Above-ground biomass estimation in closed canopy Neotropical forests using lidar remote sensing: factors affecting the generality of relationships

JASON B. DRAKE*, ROBERT G. KNOX†, RALPH O. DUBAYAH*, DAVID B. CLARK‡, RICHARD CONDIT§, J. BRYAN BLAIR¶ and MICHELLE HOFTON*

*Department of Geography, University of Maryland, College Park, MD 20742, U.S.A.; †Biospheric Sciences Branch, NASA’s Goddard Space Flight Center, Greenbelt, MD 20771, U.S.A.; ‡Department of Biology, University of Missouri-St Louis, St Louis, MO 63121, U.S.A. and La Selva Biological Station, Costa Rica; §Center for Tropical Forest Science, Smithsonian Tropical Research Institute, Unit 0948, APO AA 34002–0948, U.S.A.; ¶Laser Remote Sensing Branch, NASA’s Goddard Space Flight Center, Greenbelt, MD 20771, U.S.A.

ABSTRACT

Aim Previous studies have developed strong, site-specific relationships between canopy metrics from lidar (light detecting and ranging) remote sensing data and forest structural characteristics such as above-ground biomass (AGBM), but the generality of these relationships is unknown. In this study, we examine the generality of relationships between lidar metrics and forest structural characteristics, including AGBM, from two study areas in Central America with different precipitation patterns.

Location A series of tropical moist forest sites in Panama and a tropical wet forest in Costa Rica.

Methods Canopy metrics (e.g. canopy height) were calculated from airborne lidar data. Basal area, mean stem diameter and AGBM were calculated from measurements taken as a part of ongoing forest dynamics studies in both areas. We examined the generality of relationship between lidar metrics and forest structural characteristics, including AGBM, from two study areas in Central America with different precipitation patterns.

Results We found that lidar metrics were strongly correlated ($R^2: 0.65–0.92$) with mean stem diameter, basal area and AGBM in both regions. We also show that the relationships differed between these regions. Deciduousness of canopy trees in the tropical moist forest area accounted for the differences in predictive equations for stem diameter and basal area. The relationships between lidar metrics and AGBM, however, remained significantly different between the two study areas even after adjusting for leaf drop. We attribute this to significant differences in the underlying allometric relationships between stem diameter and AGBM in tropical wet and moist forests.

Conclusions Important forest structural characteristics can be estimated reliably across a variety of conditions sampled in these closed-canopy tropical forests. Environmental factors such as drought deciduousness have an important influence on these relationships. Future efforts should continue to examine climatic factors that may influence the generality of the relationships between lidar metrics and forest structural characteristics and assess more rigorously the generality of field-derived allometric relationships.

Key words above-ground biomass, allometry, carbon cycle, Central America, environmental factors, forest structure, laser altimetry, lidar, remote sensing, tropical forests.

INTRODUCTION

Above-ground biomass is the total amount of biological material (usually oven-dried to remove water) present above the soil surface in a specified area. Because plant biomass is approximately 50% carbon, estimates of the total above-ground biomass in forest ecosystems are critical for carbon dynamics studies at multiple scales. These estimates provide initial conditions for ecosystem and biogeochemical models (e.g. Foley et al., 1996; Friend et al., 1997; Hurtt et al., 1998; Potter, 1999) that simulate the exchange of carbon and energy between the atmosphere and biosphere. Estimates of carbon fluxes from deforestation, land cover change and other disturbances depend on knowing the
Forest canopy structure is highly dynamic both temporally and spatially. As forests recover from past disturbance events, there are typically changes in the horizontal (e.g. increases in basal area) and vertical (e.g. an increase in stand height) distribution of forest structure that accompany an overall increase in above-ground biomass (Aber, 1979; Bormann & Likens, 1979; Oliver & Larson, 1990; Richards, 1996). Additionally, variability in climatic (e.g. temperature, precipitation), edaphic, and other environmental factors (e.g. exogenous disturbances) result in differences in the spatial distribution of above-ground biomass and vertical canopy structure (e.g. Lieberman et al., 1996; Yamakura et al., 1996; Lauarance et al., 1999; Clark & Clark, 2000). For example, in nutrient-poor areas, forests are typically lower-stature and contain less above-ground biomass than in nutrient-rich areas (Oliver & Larson, 1990; Kimmins, 1997). Furthermore, there is usually a connection between differences in vertical canopy structure and differences in biomass both through plant succession and across areas with contrasting environmental conditions.

The interconnection of vertical structure and above-ground biomass creates an opportunity to estimate above-ground biomass using lidar (light detecting and ranging) remote sensing. Lidar remote sensing has proved to be an efficient tool in the study of forest structure in a variety of forest environments (Lefsky et al., 1999, 2002; Magnusson et al., 1999; Drake et al., 2002a). Because lidar instruments sample the vertical distribution of canopy (e.g. leaves and branches) and ground surfaces (Blair & Hofton, 1999; Dubayah & Drake, 2000; Dubayah et al., 2000; Harding et al., 2001) and because of ecological and biomechanical links between biomass and vertical structure (King & Loucks, 1978; Oohata & Shinozaki, 1979; O’Neill & DeAngelis, 1981; Givnish, 1986; Franco & Kelly, 1998), several studies have found a strong correlation between lidar metrics and above-ground biomass (Nelson et al., 1988; Lefsky et al., 1999; Means et al., 1999; Drake et al., 2002a).

In a recent study, Drake et al. (2002a) found a strong, non-asymptotic relationship between canopy height metrics from lidar data and above-ground biomass in a dense, closed-canopy tropical forest. These results were encouraging because the broad-scale estimation of above-ground biomass in tropical forests has been a difficult task (Houghton et al., 2001). Previous remote sensing studies have shown that although passive optical and active radar sensors are sensitive to differences in above-ground biomass in young (0–15 years old) secondary forests, they are not as sensitive to differences in biomass in older, high-biomass primary forest areas (e.g. Luckman et al., 1997; Nelson et al., 2000; Steininger, 2000).

The airborne field experiment providing lidar data used in this study and closely related work (Blair et al., 1999; Dubayah et al., 2000; Weishampel et al., 2000; Drake et al., 2002a; Drake et al., 2002b) was scheduled originally for early in the dry season, before much leaf drop was expected in drought-deciduous trees. This would have avoided strong phenological differences among sites while also keeping flight costs manageable in otherwise persistently cloudy areas. A delay in aircraft deployment, due to changes to the schedule for the Vegetation Canopy Lidar (VCL) Mission, combined with El Niño weather patterns to produce conditions more typical of the dry season at the time of sampling in the Isthmus of Panama. In contrast, there was little noticeable leaf loss among emergent trees at the wet forest site in Costa Rica until after lidar data collection. Although this made phenological conditions inconsistent between the two study areas, it provided us with the opportunity for a more stringent test of the generality of lidar algorithms for biomass estimation. The generality of these relationships across different regions and biomes bears on how global terrestrial biomass estimates are developed using data from future airborne and spaceborne lidar instruments such as the VCL Mission’s Multi-beam Laser Altimeter (Dubayah et al., 1997) and the Ice, Cloud and Land Elevation Satellite’s Geoscience Laser Altimeter System (ICESat, Schutz et al., 1998). Because of cloud cover, lidar data from the wet season in tropical moist forests usually will be less complete or more costly to obtain than similar dry season data, so it is important to understand the limitations of using lidar data from contrasting phenological states to estimate the biomass of semi-evergreen broadleaf forests.

Our primary goal in this study was to examine the relationship between lidar metrics and above-ground biomass in closed-canopy Neotropical forest areas with different annual precipitation amounts. We focused on a tropical moist forest (sensu Holdridge et al., 1971) area in Panama, and a tropical wet forest area in Costa Rica that receives 50–75% more rainfall on average. Our first objective was to test whether the relationships between lidar metrics and allometrically estimated above-ground biomass differ between the two study areas. Although biomass must be estimated from allometric relationships, basal area and mean stem diameter are more direct summaries of field data that are themselves often used to predict forest biomass (Brown, 1997). Our second objective was to test whether relationships between lidar metrics and these simple summary statistics (e.g. basal area) differ between the two study areas. Our third objective was to examine what factors might account for differences found, and their relevance to producing global data sets from on-orbit lidar instruments.

**METHODS**

**Study areas**

This study concentrates on field and lidar data collected in two areas of Central America. The first study area is centred
on the Isthmus of Panama along the Panama Canal. This area spans a precipitation gradient ranging from approximately 2000 mm per year on the Pacific coast of Panama to 3000 mm per year on the Atlantic side (Condit et al., 2000; Pyke et al., 2002) and is classified as lowland tropical moist forest (TMF) by Holdridge et al. (1971). Within this broad area, we focused on a series of 1-ha plots distributed throughout this precipitation gradient (Pyke et al., 2002), and on the 50-ha research site on Barro Colorado Island (Condit, 1998). We refer to this study area as the TMF study area.

The second study area is the La Selva Biological Station in the Atlantic lowlands of north-eastern Costa Rica (McDade et al., 1994). La Selva is a 1540-ha research facility that is comprised of a mixture of primary and secondary tropical forest, agroforestry and current or abandoned pasture areas. This area receives approximately 4200 mm rainfall per year (Sanford et al., 1994; OTS, 2001) and is classified as tropical wet forest (TWF) (Holdridge et al., 1971). In this study we report data from several primary and secondary forest and agroforestry sites (Table 1). We refer to this study area as the TWF study area.

Field data

Field data collected as a part of different ongoing field studies at each study area were used in this work (Table 1). At the TMF study area, field data were collected in 1-ha research plots near the Panama Canal (the 'Panama Canal plots' in Table 1; Pyke et al., 2002). There were a total of 19 of these 1-ha plots that were sampled with the lidar instrument. Four of these sites are mature secondary forests, and the rest are primary forest (Table 1). The remaining field data in Panama were from the 50-ha plot on Barro Colorado Island (BCI, Condit, 1998). In this case, the 50-ha plot was first divided into 50 1-ha square plots. Next, the spatial correlation length of the lidar metrics used in this study (metrics discussed below) was determined to be approximately 90–100 m (Fig. 1) so every other 1-ha plot was discarded to maintain relative independence of the independent variable in the regression analysis. This left a total of 25 1-ha plots from BCI in a checkerboard pattern.

In the TWF study area, field data were collected in 18 0.5-ha primary forest plots (Clark & Clark, 2000), and three secondary forest areas of 14, 22 (Guariguata et al., 1997; Nicotra et al., 1999) and 31 (Pierce, 1992) years since abandonment as of March 1998. In addition, published data for six agroforestry plots (Menalled et al., 1998) were included among the TWF study sites to correspond to an earlier study (Drake et al., 2002a).

Within each of these primary and secondary forest plots at both study areas, stem diameters were measured in a marked location either at breast height or, when necessary, above buttressing where the bole is approximately cylindrical (see

<table>
<thead>
<tr>
<th>Study site</th>
<th>Land cover type</th>
<th>No. of sites</th>
<th>Mean quadrant mean stem diameter (cm)</th>
<th>Mean basal area m²/ha</th>
<th>Mean estimated AGBM (Mg ha⁻¹)</th>
<th>Related study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barro Colorado Island, Panama</td>
<td>Primary forest</td>
<td>2.5</td>
<td>1.0</td>
<td>28.16</td>
<td>26.27</td>
<td>Condit (1998)</td>
</tr>
<tr>
<td>Panama Canal Plots</td>
<td>Primary forest</td>
<td>15</td>
<td>1.0</td>
<td>26.69</td>
<td>25.23</td>
<td>Condit (2000)</td>
</tr>
<tr>
<td></td>
<td>Secondary forest</td>
<td>4</td>
<td>1.0</td>
<td>24.35</td>
<td>23.52</td>
<td>Pyke et al. (2002)</td>
</tr>
<tr>
<td></td>
<td>31-year secondary forest</td>
<td>1</td>
<td>0.5</td>
<td>20.76</td>
<td>20.69</td>
<td>Condit et al. (2000)</td>
</tr>
<tr>
<td></td>
<td>22-year secondary forest</td>
<td>1</td>
<td>0.3</td>
<td>22.24</td>
<td>21.80</td>
<td>Nicotra et al. (1999)</td>
</tr>
<tr>
<td></td>
<td>14-year secondary forest</td>
<td>1</td>
<td>0.25</td>
<td>14.28</td>
<td>13.89</td>
<td>Nicotra et al. (1999)</td>
</tr>
<tr>
<td></td>
<td>Agroforest</td>
<td>6</td>
<td>0.12</td>
<td>9.03</td>
<td>9.03</td>
<td>Menalled et al. (1998)</td>
</tr>
<tr>
<td>La Selva Biological Station, Costa Rica</td>
<td>Primary forest</td>
<td>18</td>
<td>0.5</td>
<td>20.76</td>
<td>20.69</td>
<td>Condit (1998)</td>
</tr>
<tr>
<td></td>
<td>Secondary forest</td>
<td>31</td>
<td>0.25</td>
<td>14.28</td>
<td>13.89</td>
<td>Condit et al. (2000)</td>
</tr>
<tr>
<td></td>
<td>22-year secondary forest</td>
<td>1</td>
<td>0.3</td>
<td>12.85</td>
<td>12.46</td>
<td>Nicotra et al. (1999)</td>
</tr>
<tr>
<td></td>
<td>14-year secondary forest</td>
<td>1</td>
<td>0.25</td>
<td>10.46</td>
<td>10.07</td>
<td>Nicotra et al. (1999)</td>
</tr>
</tbody>
</table>

* Estimated above-ground biomass (Mg/ha) using general equation for tropical moist forests in Brown (1997). ** Including plots with trees whose diameters are greater than the largest tree used to develop the regression in Brown (1997). *** Estimated above-ground biomass (Mg/ha) using general equation for tropical wet forests in Brown (1997). **** Data collected as part of 1998 pre-launch VCL field campaign at La Selva.
Methods in Condit, 1998; Clark & Clark, 2000). In the present study all stems with diameters greater than or equal to 10 cm were measured in each plot.

Stem diameter measurements were used to estimate above-ground biomass values for each measured tree using general allometric equations (Brown, 1997) for TWFs (eqn 1) at the TWF study area, and for TMFs (eqn 2) at the TMF study area. Stem diameters were also used to calculate quadratic stem diameter (QMSD, eqn 3) and basal area (eqn 4) for each plot:

\[
AGBM_s = 21.297 - 6.953(D) + 0.740(D^2) \quad \text{(eqn 1)}
\]

where \( D \) is the stem diameter in cm, and \( AGBM_s \) is the estimated oven-dried \( AGBM \) for the stem in kg:

\[
AGBM_s = \exp(-2.134 + 2.530 + \ln(D)) \quad \text{(eqn 2)}
\]

where ‘exp’ means ‘\( e \) raised to the power of’;

\[
QMSD = \sqrt{\frac{\sum D^2}{n}} \quad \text{(eqn 3)}
\]

where \( n \) is the number of stems in the plot and \( QMSD \) is the quadratic mean stem diameter in cm;

\[
BA = \frac{\sum \pi (D/2)^2}{A} \quad \text{(eqn 4)}
\]

where \( BA \) is the basal area (m\(^2\)) measured at breast height or above buttressing within a given area (A).

Plot-level values of estimated above-ground biomass were then calculated by summing all estimated stem-level above-ground biomass values and converting to mass per unit area (Mg/ha). Note that the plot-level structure metrics are not perfectly correlated, although they are all functions of the enumerated population (n) and the measured diameters (Table 2). Sixteen plots at the TMF study area contained stems whose diameters were larger than the maximum diameter used to develop the original allometric equation (eqn 2), therefore only plots that contained stems within this regression range (< 150 cm) were used in the regression analysis comparing lidar metrics with estimated above-ground biomass (EAGB). However, all plots were used in the regression analysis involving lidar metrics and QMSD or basal area.

**Lidar data**

Lidar data were collected over both study areas in March 1998 by the Laser Vegetation Imaging Sensor (LVIS, Blair et al., 1999; Dubayah et al., 2000). LVIS is an airborne scanning laser altimeter (Fig. 2) designed and developed at

<table>
<thead>
<tr>
<th>n</th>
<th>QMSD</th>
<th>Basal area</th>
<th>AGBM</th>
</tr>
</thead>
<tbody>
<tr>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>QMSD</td>
<td>—</td>
<td>0.76</td>
<td>0.92</td>
</tr>
<tr>
<td>Basal area</td>
<td>—</td>
<td>—</td>
<td>0.90</td>
</tr>
<tr>
<td>AGBM</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

* Not significant (\( P = 0.42 \)).
NASA’s Goddard Space Flight Center. LVIS measures the roundtrip time for pulses of near-infrared laser energy to travel to the surface and back. The incident energy pulse interacts with canopy (e.g. leaves and branches) and ground features and is reflected back to a telescope on the instrument. Unlike most other laser altimeters, LVIS digitizes the entire time-varying amplitude of the backscattered energy (in 30-cm vertical bins). This yields a ‘waveform’ or profile related to the vertical distribution of intercepted surfaces from the top of the canopy to the ground (see Fig. 2 and Blair et al., 1999; Dubayah & Drake, 2000; Dubayah et al., 2000).

In this study, a footprint size of approximately 25 m diameter was used. This exceeds the average crown diameter of large emergent trees in closed-canopy tropical forests (King, 1996; Richards, 1996) and thereby consistently allows lidar energy to reach the ground through intercrown gaps (Dubayah et al., 1997). LVIS scanned across a swath of approximately 1 km with a 50% overlap of footprints across swath, and contiguous along-track footprint spacing (Fig. 2). LVIS footprints can be geolocated to within 2 m (Blair & Hofton, 1999). At both study areas, only LVIS footprints that were entirely coincident with field plots were included.

For the LVIS observations that fell within each field plot some were eliminated according to two different filtering rules. First, if the total energy received at the instrument was less than 10% of the mean total energy for all shots, and the shot elevation was greater than 200 m higher than the lowest elevation for the plot, then the shot was eliminated. This filter was used to remove only shots that were reflected off clouds. Secondly, if the last portion of the waveform did not return to the background noise level for the shot (plus 1 SD) the shot was eliminated. This eliminated shots whose ground returns were obviously not recorded. Together these filters removed approximately 30% of all LVIS shots.

Two metrics were derived from the LVIS waveforms. First, canopy height was calculated by identifying the location within the waveform where the signal initially increases above the mean background noise level (the canopy top). Next the ground return was identified as the centre of the last Gaussian pulse. The canopy height was then the distance between these two locations.

The other lidar metric, height of median energy (HOME), is calculated by first identifying the location of the median of the entire signal (i.e. above the noise level), including the energy from both canopy and ground surfaces (Drake et al., 2002a). This location is then referenced to the ground to derive a height. The HOME metric is therefore influenced by both the vertical distribution of canopy elements (Drake et al., 2002b) and the canopy cover because in more open canopies a greater proportion of the lidar energy is reflected from the understory and ground, thus lowering the HOME metric. Plot-level means for canopy height and HOME were then calculated for all shots that fell within each plot.

**Adjustment of lidar HOME for deciduousness in the TMF study area**

The TWF study area receives approximately 4200 mm rain/year and is in the TWF life zone (Holdridge et al., 1971). The leaf loss from canopy trees at the TWF study area was minimal when the lidar flights occurred in mid-March 1998. In contrast, the TMF study area spans a rainfall gradient from approximately 2000–3000 mm rain/year. When the lidar flights occurred in late March 1998, this area was at the end of its dry season and leaf loss from canopy-forming trees was pronounced in some areas. As a result, less of the lidar energy was reflected from the upper canopy, thereby reducing the lidar HOME metric relative to the TWF study area.

To compensate for this effect, a simple adjustment for the proportional reduction of the HOME metric was made to all lidar data from the TMF study area based on the estimated fraction of crown area that was deciduous (FCAD) in each plot. This involved a simple two-step approach. First, we estimated the FCAD in each plot based on average annual precipitation values in each plot. We assumed that in areas that receive less rainfall and have a longer dry season the leaf loss from canopy trees will be greater. This general trend has been qualitatively described for forests throughout the tropics (see Reich, 1995) and has recently been quantitatively examined in our TMF study area (Condit et al., 2000). We interpolated linearly between data points for precipitation and the fraction of crown area deciduous as listed in Condit et al. (2000). The relationship (eqn 5) was then used to estimate FCAD from average precipitation values for all TMF field plots. Although this relationship is based on only three points, the plots in the Condit et al. (2000) study span the rainfall gradient in our study.

\[
\text{FCAD} = \left(\frac{-0.02 \times \text{Rainfall} + 60.27}{100}\right)
\]  

where Rainfall = mm/year and \(\text{FCAD} = \text{fraction of crown area deciduous}\) (between 0 and 1, equation developed from data in Condit et al., 2000)

The second step was to adjust the lidar HOME values based on the estimated FCAD in each plot in our TMF study area. In tropical forests, leaf loss is primarily from large canopy-forming trees (Reich, 1995; Condit et al., 2000). In areas with greater leaf loss from canopy trees we assume that more lidar energy will be reflected from the understory and ground, thus reducing the waveform median (HOME). To compensate for this effect we developed a relatively simple model (eqn 6) that adjusts for the proportional reduction of HOME based on canopy deciduousness without requiring assumptions about canopy density or reflectivity. We feel that this model is a reasonable preliminary approach for closed canopy forests that are deciduous from the top of the canopy down (Reich, 1995).
\[ HOME' = \text{HOME}/(1 - \text{FCAD}) \]  
(eqn 6)

where \( \text{HOME} \) = lidar height of median energy (m) and \( \text{FCAD} \) = fraction of crown area deciduous.

**Data analysis**

A linear regression analysis was used to develop relationships between plot-level averages of lidar metrics and field-derived QMSD, basal area and EAGB for each study area. The lidar metric that was the best single predictor was then identified. For each forest structural characteristic (e.g. QMSD), an analysis of covariance (ANCOVA) was then performed to test for significant differences in the slope and intercept of the relationships (Zar, 1996) developed for each site. For the TMF data this process was used for both the normal and deciduous adjusted HOME metrics.

**RESULTS**

**General site characteristics**

Primary and secondary forest sites in the TMF study area have a larger quadratic mean stem diameter (QMSD), basal area and estimated above-ground biomass on average than the corresponding primary and secondary field sites in the TWF study area (Table 1). Of these forest structural characteristics, the difference in mean basal area between the two sites is the least, at approximately 10% for primary forest plots, followed by an approximately 30% larger QMSD in the plots at the TMF study area. The largest difference is for estimated above-ground biomass, which is approximately 70% larger in the primary TMF plots that fall within the range of the general allometric equation (eqn 2), and 95% larger for all the TMF plots (including those outside the range) than primary forest plots in the TWF study area. The mean estimated above-ground biomass value for BCI that we calculated using eqn 2 (287 Mg/ha) corresponds well with a mean of 290 Mg/ha reported in Chave et al. (in review) using an equation that includes stem diameter, height and wood density.

The differences between lidar metrics from both study areas (Table 3) were minimal. Lidar canopy heights were on average approximately 5% larger in primary forest plots in TMF compared to the TWF plots. The differences between HOME values were on average less than 2%, as both were approximately 20 m.

**Relationship between lidar HOME and allometrically estimated above-ground biomass**

The lidar metric HOME is highly correlated with EAGB in both study areas (Fig. 3). In the TMF study area, the \( R^2 \) value is 0.66 for plots whose tree diameters are all within the range of the general allometric equation (eqn 2), and 0.82 for all TMF plots, with RMSE values of 31.52 Mg/ha and 39.10 Mg/ha, respectively. For the TWF relationship, the \( R^2 \) value is 0.89 and the RMSE is 22.54 Mg/ha.

Although the \( R^2 \) value is greater for the TWF relationship, the ‘relative regression error’ (RMSE divided by mean) is slightly lower for the TMF relationship compared to the TWF relationship (11.5% and 14.06%, respectively, Table 4). Thus, both the TMF and TWF regression relationships are similar in terms of proportional errors in above-ground biomass estimates. In addition, the relative regression errors for both relationships are at or below the coefficient of variation (CV) for EAGB (Table 4).

In other regards, however, there is a great deal of divergence between the relationships for the two study areas (Fig. 3). The slope is much greater in the TMF relationship (22.33) compared to the TWF relationship (6.33). The TMF regression equation also has a negative intercept, probably the result of only sampling within relatively high biomass areas. An analysis of covariance (ANCOVA) shows that the

<table>
<thead>
<tr>
<th>Study site</th>
<th>Land cover type</th>
<th>No. of sites</th>
<th>Plot size (ha)</th>
<th>Mean no. of shots per plot</th>
<th>Mean lidar height (m)</th>
<th>Mean height of median energy (HOME) (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barro Colorado Island, Panama</td>
<td>Primary forest</td>
<td>25</td>
<td>1.0</td>
<td>29</td>
<td>34.93</td>
<td>20.29</td>
</tr>
<tr>
<td>Panama Canal Zone Plots</td>
<td>Primary forest</td>
<td>15</td>
<td>1.0</td>
<td>13</td>
<td>31.28</td>
<td>19.16</td>
</tr>
<tr>
<td>Secondary forest</td>
<td>4</td>
<td>1.0</td>
<td>12</td>
<td>29.77</td>
<td>18.41</td>
<td></td>
</tr>
<tr>
<td>La Selva Biological Station, Costa Rica</td>
<td>Primary forest</td>
<td>18</td>
<td>0.5</td>
<td>9</td>
<td>31.33</td>
<td>20.14</td>
</tr>
<tr>
<td>31-year secondary forest</td>
<td>1</td>
<td>0.3</td>
<td>8</td>
<td>25.48</td>
<td>16.65</td>
<td></td>
</tr>
<tr>
<td>22-year secondary forest</td>
<td>1</td>
<td>0.25</td>
<td>6</td>
<td>17.68</td>
<td>11.32</td>
<td></td>
</tr>
<tr>
<td>14-year secondary forest</td>
<td>1</td>
<td>0.25</td>
<td>7</td>
<td>23.67</td>
<td>9.21</td>
<td></td>
</tr>
<tr>
<td>Agroforestry</td>
<td>6</td>
<td>0.12</td>
<td>3</td>
<td>9.47</td>
<td>1.68</td>
<td></td>
</tr>
</tbody>
</table>

© 2003 Blackwell Publishing Ltd, Global Ecology & Biogeography, 12, 147–159
slopes and intercepts of these two equations are significantly different \( P < 0.01 \).

The regression lines from Fig. 3 show that there is a great deal more EAGB in the TMF study area for any lidar HOME above 9 m. To some degree this is expected, because although mean EAGB is approximately 70% higher in TMF primary forest plots, the lidar HOME is essentially the same (Tables 1 and 3).

However, the differences in mean QMSD and basal area between the two study areas are smaller than the differences between mean above-ground biomass. We next test the generality of the relationships between lidar HOME and these direct structural summaries.

### Relationship between lidar metrics and direct structural indices

Lidar HOME is strongly correlated with quadratic mean stem diameter (QMSD) in both study areas (Fig. 4). The level of variation in QMSD explained by the HOME metric (i.e., the \( R^2 \) value) is approximately 92% in the TWF study area compared to 66% in the TMF study area. However, the

---

**Table 4** Coefficients of variation in forest structural characteristics and the relative regression errors from lidar equations for TMF and TWF study areas

<table>
<thead>
<tr>
<th>Forest structural characteristic</th>
<th>TMF study area CV</th>
<th>TMF study area relative regression error*</th>
<th>TWF study area CV</th>
<th>TWF study area relative regression error*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quadratic mean stem diameter</td>
<td>11.83</td>
<td>6.94 ((6.59^{**}))</td>
<td>9.60</td>
<td>8.77</td>
</tr>
<tr>
<td>Basal area</td>
<td>16.44</td>
<td>8.95 ((9.89^{**}))</td>
<td>11.94</td>
<td>12.87</td>
</tr>
<tr>
<td>Estimated above-ground biomass</td>
<td>19.25 ((28.94^{***}))</td>
<td>11.50 ((12.27^{*<strong>})) ((11.36^{</strong>}))</td>
<td>14.01</td>
<td>14.06</td>
</tr>
</tbody>
</table>

\* Relative regression error = RMSE/mean. \** With leafdrop-modified lidar height of median energy. \*** Including plots with trees whose diameters are greater than the largest tree used to develop the regression in Brown (1997).

---

© 2003 Blackwell Publishing Ltd, Global Ecology & Biogeography, 12, 147–159
RMSE from the TMF relationship (1.89 cm) is somewhat lower than the value from the TWF relationship (2.09 cm). In addition, the relative regression errors for both relationships are smaller than the CV in QMSD for the respective sites (Table 4).

HOME is also strongly correlated with basal area in both areas (Fig. 5). In this case the level of variation explained is approximately equal for both areas (~70%). The relative regression error (Table 4) from the TMF relationship (9%) is smaller than in the TWF relationship (12.9%). In this case the relative regression error from the TMF relationship is smaller than the CV for basal area (16.4%); however, the relative regression error from the TWF relationship is slightly larger than the CV for basal area (11.9%).

The relationships between lidar HOME and QMSD (Fig. 4) and between HOME and basal area (Fig. 5) are not as divergent between the two study areas as are the relationships between HOME and EAGB (Fig. 3). The slopes of the relationships between HOME and QMSD are similar at both sites, and were not found to be significantly different ($P = 0.55$) in an ANCOVA analysis. Intercepts for the HOME–QMSD relationships, however, were found to be significantly different ($P < 0.01$), indicating that the relationships are not equivalent between study sites. Similarly, both the slopes and the intercepts of the HOME–basal area relationships were found to be significantly different between the two study areas. However, it should be noted that the y-intercept term in the TMF linear regression relationship was not significantly different from zero ($P = 0.19$).

The differences in these relationships show that the TMF plots have a greater QMSD and more basal area for the same median height (lidar HOME). Again, from the field and lidar summary statistics (Tables 1 and 3) this is anticipated, because although the mean basal area and QMSD are higher in the TMF plots, HOME is approximately the same.

These differences may be the result of the canopy deciduousness at the TMF study area, especially in drier sites where leaf loss from canopy trees is greater (Condit et al., 2000). The HOME metric is determined by the vertical distribution of canopy elements (e.g. leaves and branches) and the canopy cover. Leaf loss from canopy-forming trees will result in less energy reflection from the upper canopy thereby lowering the HOME value. We therefore examine next the effect of this deciduousness on the relationship between HOME and forest structural summaries.

Relationships between deciduous-adjusted lidar HOME values and direct structural indices

The adjustment of HOME for the effect of leaf loss of canopy trees preserved the strength of the relationships between HOME and QMSD (Fig. 6) and between HOME and basal area (Fig. 7). Whereas the level of variation in basal area that is explained with HOME decreased slightly (from 70% to 65%), the $R^2$ value for the HOME–QMSD relationship increased slightly (from 0.66 to 0.7). More importantly, after adjustment the relationships are much more similar between the two study areas.
The differences between the slope and the intercept from both HOME–QMSD relationships are not significantly different using an ANCOVA test (\( P = 0.85 \) and 0.21, respectively). Similarly, the difference in slope from the HOME–basal area relationships in the two study areas is smaller and not significantly different (\( P = 0.06 \)). The intercepts in the HOME–basal area relationships were found to be significantly different (\( P < 0.01 \)). None the less, where the data cover the same range of HOME and basal areas, the two point clouds now overlap more completely and appear similar (Fig. 7). Thus, adjustment of HOME by simple fraction of crown area deciduous values eliminated much of the difference between HOME–basal area and HOME–QMSD relationships at both TWF and TMF plots.

**Relationship between deciduous-adjusted lidar HOME and allometrically estimated above-ground biomass**

As with the direct structural characteristics above, the adjustment in HOME also did not greatly affect the strength of the relationship between HOME and EAGB at the TMF study area (Fig. 8). The \( R^2 \) and the RMSE stayed approximately the same (66% and ~31 Mg/ha, respectively) after HOME values were adjusted. In contrast, although the adjustment did slightly reduce the slope of the TMF relationship (from 22.33 to 21.46), the relationships from the two study areas were still significantly different in both slope and the intercept (\( P < 0.01 \), from ANCOVA).

**DISCUSSION**

Lidar remote sensing has shown promise for the estimation of above-ground biomass in a variety of temperate and tropical forests (e.g. Means et al., 1999; Drake et al., 2002a). In these studies, different metrics from large-footprint lidar waveforms (normal or transformed to adjust for attenuation and absorption of lidar energy through the canopy) have been used to predict above-ground biomass. However, the generality of relationships between lidar metrics and above-ground biomass remained untested. The present paper provides an initial test of the generality of the relationships between lidar metrics and above-ground biomass in closed-canopy Neotropical forests.

Our first objective was to test whether the relationships between lidar metrics and allometrically estimated above-ground biomass differ between the TMF and TWF study areas. Our results show that relationships between a simple lidar metric (height of median energy) and EAGB are significantly different between the TMF and TWF study areas.

In response to our second objective, we found that relationships between lidar HOME and direct structural summaries
The remaining divergence in the HOME–EAGB relationships is most probably the result of differences in the underlying biophysical and allometric relationships for the two study areas. If the same allometric equation (e.g., the TWF equation; eqn 1) is applied in both study areas, the relationships between HOME and EAGB are no longer divergent at the two study areas (Fig. 9). Estimated above-ground biomass values for the TMF plots fall entirely within the data envelope of the relationship between HOME and EAGB for the TWF plots. This suggests that the differences in the lidar–biomass relationships at the two study areas are primarily the result of the different allometric relationships between stem diameter and above-ground biomass in TWF and TMF areas.

This led us naturally to question whether the differences in the general allometric relationships from TMFs and TWFs are significant. To address this we went back to the original measurements of stem diameters and above-ground biomass for harvested trees in Brown (1997). We performed a jackknife analysis to estimate the range of uncertainty in the parameters of the TMF allometric equation (Fig. 10a). Based on the trees harvested in TWF and TMF areas (listed in Brown, 1997), there do appear to be significant differences in allometric relationships between stem diameter and AGBM. The uncertainty envelope around the TMF relationship is distinct from the TWF relationship for all trees > 50 cm diameter (Fig. 10a). As a result, trees with the same diameter are predicted to have a larger above-ground biomass in TMFs compared to TWFs. We caution, however, that these allometric relationships are based only on stem diameter. At least one of the trees in the Brown (1997) TMF dataset was an emergent tree from Indonesia with a diameter of 130.7 cm and a height of 70.7 m (Yamakura et al., 1986). We feel that more thorough examination of these and other general biomass allometric equations, especially accounting for variability in wood density, height and form factor (Cannell, 1984), would greatly improve our current understanding.

We next scaled up the tree-level allometric uncertainty to the plot level by applying the upper and lower jack-knife equations to all trees sampled in the TMF study area. This process was necessary because the number of trees in each plot also influences total plot-level biomass. We found that the plot-level relationship between field-measured basal area and EAGB from the TWF site does not fall within the uncertainty envelope from the TMF study area (Fig. 10b). In other words, for the same basal area there will be a larger predicted above-ground biomass in our TMF area compared to the TWF area.

Several related studies support the fidelity of the general allometric equations used in our work. A recent study found that biomass estimates from the general TMF equation (eqn 2) were similar to estimates from locally derived equations in another TMF in Brazil (Keller et al., 2001). In another study, biomass estimates from a general TMF equation (similar to
eqn 2) varied by less than 10% from estimates from a site-specific allometric equation (Brown et al., 1995). Further, our mean EAGB value at BCI (287 Mg/ha) is equivalent to a mean of 290 Mg/ha reported by Chave et al. (in press) using an equation that includes height and wood density. This lends support for the use of the general TMF equation when local equations (e.g. based on locally harvested trees) are unavailable, as is the case in our TMF study area. Although we are unaware of any studies that directly assess the accuracy of the general TWF equation, Clark & Clark (2000) aptly point out that many of the TWF trees (57%) were harvested within 20 km of La Selva. As a result we feel that these two equations are reasonably accurate for our two study areas.

Our results could have significant implications for how global observations from future spaceborne lidar instruments (e.g. VCL) should be used to produce global estimates of terrestrial above-ground biomass. Our results show that it will probably be necessary to develop a series of relationships between lidar metrics and above-ground biomass in different bioclimatic life zones. Future research should also test the applicability of these new relationships at other study areas within the same bioclimatic life zone. For example, does the relationship from our TMF study area also apply to TMF areas in Brazil?

These results illustrate the importance of climatic variables for developing general algorithms for the estimation of above-ground biomass in different tropical areas using lidar data. For example, average rainfall data and other environmental factors (e.g. soil type) can be used to estimate leaf loss in canopy-forming trees during the dry season, which, as we have shown, will affect the generality of the relationship between lidar metrics and tropical forest structural summaries. At a minimum, if leaf phenology data and models are unavailable, estimates could be developed from lidar data collected while the canopy is fully leafed out.

This study only begins the exploration of the generality of the relationships between lidar metrics and above-ground biomass in closed-canopy Neotropical forests, but it illustrates that this research area holds great potential. The strong correlation of lidar metrics with above-ground biomass in a variety of tropical forests is an improvement over many existing remote sensing techniques that are currently not able to estimate reliably biomass in older secondary and primary forests (Luckman et al., 1997; Nelson et al., 2000; Steininger, 2000). Although relationships between lidar metrics and biomass may differ, the geographical regions where these differences occur appear to be distinct and may be predicted using climatic variables such as temperature and rainfall.

Future work in other tropical and extra-tropical forest environments may reveal that it is possible to develop a relatively simple algorithm or model to estimate terrestrial above-ground biomass globally from a suite of lidar and climatic metrics. In more open tropical woodlands, it is likely that additional lidar metrics such as canopy top height and a canopy cover index will be necessary to estimate above-ground biomass accurately. We also expect that the fusion of lidar data with high spatial and temporal satellite imagery will further extend the utility of these data.
ACKNOWLEDGMENTS

We thank Nancy Casey, David Rabine and Dave Kendig for their help in processing the LVIS data. We thank especially Robin Chazdon for the use of her secondary forest field data at La Selva. We are grateful to Leonel Ramos, William Miranda, Marcos Molina, Jeanette Panaiagua, Braulio Vilchez, Robin Chazdon, Laura Chazdon, Birgit Peterson, Steve Prince and John Weishampel for valuable help in collecting TWF field data. We also sincerely thank S. Hubbell and R. Foster for initiating the BCI 50-ha plot, and A. Hernandez, S. Aguilar and R. Perez for mapping and identifying trees in the 1-ha plots in Panama. Many thanks to Stephanie Bohlman, Chis Pyke and Jerome Chave for sending copies of their manuscripts and for providing helpful comments. Our special thanks also go to the La Selva Biological Station, the Organization for Tropical Studies, the Carbone Project (funded by DOE, NSF and the Andrew W. Mellon Foundation), the Center for Tropical Forest Science of the Smithsonian Tropical Research Institute, the Wallops Flight Facility Aircraft Programs Office, the National Geographic Institute of Costa Rica and the Governments of Costa Rica and Panama. This project is funded by a NASA contract to the University of Maryland for the implementation and execution of the Vegetation Canopy Lidar Mission under the Earth System Science Pathfinder program. In addition, J. Drake was funded through a NASA Earth System Science Fellowship.

REFERENCES


**BIOSKETCH**

Jason Drake’s primary research interest is in the field of forest ecology. He is particularly interested in exploring the link between forest canopy structure and function. In his research he uses field data, remote sensing, geographical information systems, spatial analysis techniques and ecological models to work across a variety of scales ranging from individual plants to landscape and regional levels.