Available online at www.sciencedirect.com
SCIENCE DIRECT®
Remote Sensing of Environment 91 (2004) 68–89
Remote Sensing of Environment
www.elsevier.com/locate/rse

Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape

Matthew L. Clark[a,*], David B. Clark[b], Dar A. Roberts[a]

[a] Department of Geography, University of California, Santa Barbara, Santa Barbara, CA 93106, USA
[b] University of Missouri-St. Louis, St. Louis, MO, USA, and La Selva Biological Station, Puerto Viejo de Sarapiquí, Costa Rica

Received 15 April 2003; received in revised form 23 February 2004; accepted 28 February 2004

Abstract

Meso-scale digital terrain models (DTMs) and canopy-height estimates, or digital canopy models (DCMs), are two lidar products that have immense potential for research in tropical rain forest (TRF) ecology and management. In this study, we used a small-footprint lidar sensor (airborne laser scanner, ALS) to estimate sub-canopy elevation and canopy height in an evergreen tropical rain forest. A fully automated, local-minima algorithm was developed to separate lidar ground returns from overlying vegetation returns. We then assessed inverse distance weighted (IDW) and ordinary kriging (OK) geostatistical techniques for the interpolation of a sub-canopy DTM. OK was determined to be a superior interpolation scheme because it smoothed fine-scale variance created by spurious understory heights in the ground-point dataset. The final DTM had a linear correlation of 1.00 and a root-mean-square error (RMSE) of 2.29 m when compared against 3859 well-distributed ground-survey points. In old-growth forests, RMS error on steep slopes was 0.67 m greater than on flat slopes. On flatter slopes, variation in vegetation complexity associated with land use caused highly significant differences in DTM error distribution across the landscape. The highest DTM accuracy observed in this study was 0.58-m RMSE, under flat, open-canopy areas with relatively smooth surfaces. Lidar ground retrieval was complicated by dense, multi-layered evergreen canopy in old-growth forests, causing DTM overestimation that increased RMS error to 1.95 m.

A DCM was calculated from the original lidar surface and the interpolated DTM. Individual and plot-scale heights were estimated from DCM metrics and compared to field data measured using similar spatial supports and metrics. For old-growth forest emergent trees and isolated pasture trees greater than 20 m tall, individual tree heights were underestimated and had 3.67- and 2.33-m mean absolute error (MAE), respectively. Linear-regression models explained 51% (4.15-m RMSE) and 95% (2.41-m RMSE) of the variance, respectively. It was determined that improved elevation and field-height estimation in pastures explained why individual pasture trees could be estimated more accurately than old-growth trees. Mean height of tree stems in 32 young agroforestry plantation plots (0.38 to 18.53 m tall) was estimated with a mean absolute error of 0.90 m ($r^2 = 0.97$; 1.08-m model RMSE) using the mean of lidar returns in the plot. As in other small-footprint lidar studies, plot mean height was underestimated; however, our plot-scale results have stronger linear models for tropical, leaf-on hardwood trees than has been previously reported for temperate-zone conifer and deciduous hardwoods.

Keywords: Tropical environment; Airborne laser scanner; Lidar methods; Digital terrain models; Geostatistics; Kriging; Tree height estimation; Digital canopy model

1. Introduction

Tropical rain forests (TRF) represent some of the most biologically diverse, structurally complex and carbon-rich ecosystems on the planet. Our ability to understand the complexity of these forests, and to generalize findings from plot-scale measurements, is greatly aided by the meso-scale (1 to 100 km^2) to global-scale perspective afforded by remote sensing. Digital terrain models (DTMs) and canopy-height estimates are two important remote sensing products for studies of TRF ecology and management. DTMs describe the variation of elevation across a landscape and have been used in applications including mapping drainage basin geomorphology (Yin & Wang, 1999), flood modeling (Bates & De Roo, 2000), calculation of biophysical controls on vegetation distribution (e.g., temperature, solar radiation) (Dymond & Johnson, 2002), and spatial analysis of soil properties (Gessler et al., 2000).

* Corresponding author. Tel.: +1-805-893-4434; fax: +1-805-893-3146.
E-mail address: mclark@geog.ucsb.edu (M.L. Clark).

0034-4257/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2004.02.008

Canopy-height estimates from remote sensing technology have a variety of potential applications in TRF, such as calculating surface roughness for atmosphere–land interaction models (Raupach, 1994), spatial analyses of forest dynamics, such as canopy gap formation, distribution and turn-over (Birnbaum, 2001), identifying plant species (Brandtberg et al., 2003; Holmgren 2003), mapping of wildlife habitat (Hinsley et al., 2002), modeling canopy rain interception (Herwitz & Slye, 1995), and modeling light penetration (Clark et al., 1996; Montgomery and Chazdon, 2001). Vegetation height is allometrically related to forest structure parameters, such as estimated aboveground biomass (Brown et al., 1995; Clark et al., 2004). Because roughly half of biomass is composed of carbon, improvements in our ability to map biomass through remote sensing will translate into better estimates of carbon stocks and flux at broad scales. Such advances are particularly important in tropical forests, which contain a large proportion of terrestrial carbon, and consequently have the greatest potential to increase atmospheric carbon dioxide from deforestation (Dixon et al., 1994). Although passive optical and active synthetic aperture radar (SAR) signals and associated metrics are sensitive to forest aboveground biomass variation (Ranson et al., 1997; Sader et al., 1989), biomass estimates from these sensors tend to saturate at the high biomass levels typically found in tropical forests (Imhoff, 1995; Luckman et al., 1998; Steininger, 2000).

Airborne scanning light detection and ranging (lidar) sensors, also referred to as airborne laser scanners (ALS), have emerged as the premier instrument for both the generation of detailed DTMs (Hodgson et al., 2003; Kraus & Pfeifer, 1998; Petzold et al., 1999) and the estimation of individual and plot-scale tree height (Brandtberg et al., 2003; Magnussen & Boudewyn, 1998; Næsset, 1997, 2002; Persson et al., 2002). These active remote sensing systems record the round-trip time of a laser light pulse to leave and return to the sensor after interacting with the surface target (Baltsavias, 1999a). The recorded time is converted to distance using the speed of light, and the height above an ellipsoid of the target is measured. Terrestrial lidar systems typically operate in the near-infrared portion of the spectrum (900 to 1064 nm). Commercial airborne sensors are generally flown at altitudes between 20 m (helicopter) to 1000 m (airplane) and record discrete returns of light within fine-scale 0.25- to 0.60-m footprints (Baltsavias, 1999a; Huising & Gomes Pereira, 1998). Older sensors (e.g., FLI-MAP used in this study) record only the first and/or last returns, while newer terrain-mapping sensors record multiple discrete returns (e.g., 5 or more). More experimental sensors, such as Laser Vegetation Imaging Sensor (LVIS), are flown at higher altitudes (5 to 10 km), cover larger extents, and digitize detailed vertical profiles of lidar returns (i.e., waveforms) from relatively coarse-scale, 1- to 80-m (typically $\sim$ 20 m) footprints (Blair et al., 1999; Harding et al., 2001; Lefsky et al., 2002).

There have been relatively few published applications that have used small-footprint lidar sensors in tropical rain forest environments (Blair & Hofton, 1999; Hofton et al., 2002). However, studies using the large-footprint LVIS sensor in Costa Rica have shown that there is immense potential of lidar technology for TRF research and monitoring efforts (Blair & Hofton, 1999; Drake et al., 2002a,b; Hofton et al., 2002; Weishampel et al., 2000). For example, Drake et al. (2002a) showed that lidar metrics applied to large-footprint waveforms can accurately predict aboveground biomass over a wide range of tropical forest conditions without saturation.

The objective of this study was to assess small-footprint lidar technology for the estimation of ground elevation and tree heights in tropical landscapes. Lidar data from a small-footprint, first-return ALS sensor were used to generate a DTM and DCM, from which elevation and tree-height estimates were derived, respectively. The study site is considered an extreme test of lidar technology because: (1) several classes include forests with dense, structurally complex canopy that severely restricts ground-level light transmission and lidar returns (Clark et al., 1996; Drake et al., 2002a; Montgomery and Chazdon, 2001), (2) all vegetation classes are in ''leaf-on'' conditions, exacerbating ground-retrieval difficulties, and (3) it is situated over a range of terrain conditions. We begin with a review of applying lidar technology to sub-canopy elevation, forest canopy height and forest structure estimation, with an overall emphasis on commercially available, small-footprint lidar systems. In this study, original lidar footprints were interpolated into a digital surface model (DSM). Next, a DSM ground-retrieval scheme was developed for interpolating a DTM using geostatistical techniques. The accuracy of DTM-derived elevation estimates was rigorously tested with a comparison to in situ field-survey points. We also discuss differences in DTM accuracy related to terrain slope and land-use factors. The DSM and the derived DTM were then used to calculate a map of canopy height, i.e., DCM, and the accuracy of DCM-derived estimates of stem heights at individual tree and plot scales was assessed with reference to comparable field measurements.

1.1. Digital terrain models and lidar

Photogrammetry has been the traditional source of broad-scale DTMs throughout the world. There has been an increase in the use of interferometric synthetic aperture radar (IfSAR) and lidar active sensors as more economic alternatives for producing DTMs (Hodgson et al., 2003; Petzold et al., 1999). In many developing countries, where most tropical forests occur, the best publicly available topographic information is from the Shuttle Radar Topography Mission (SRTM), which used IfSAR technology to produce a global terrestrial DTM with a 90-m horizontal resolution (Rabus et al., 2003). However, the relative vertical accuracy of this product at the 50–100 km scale

is roughly 6 m due to errors from several systematic and random factors that can only be partially reduced by data post-processing.

Lidar-derived DTMs estimated in open areas or under areas with low vegetation can have vertical accuracies ranging from 0.06 to 0.61-m root-mean-square error (RMSE) (Cobby et al., 2001; Huising & Gomes Pereira, 1998). Trunks, branches and leaves in dense vegetation tend to cause multiple-scattering reflections or absorption of the emitted laser energy so that fewer backscattered returns are reflected directly from the ground (Harding et al., 2001; Hofton et al., 2002). This effect increases with more canopy closure, canopy depth (or volume) and structural complexity (Harding et al., 2001; Hodgson et al., 2003; Hofton et al., 2002), and it is expected to be more severe for first-return only lidar systems because recorded returns generally come from the canopy (Magnussen & Boudewyn, 1998). The result is that the RMSE between the lidar-derived DTM and reference elevation data tends to increase in areas of dense vegetation because: (1) there are fewer ground-return samples for DTM surface interpolation, and (2) those samples selected as ground may actually be understory vegetation, logs or rocks (Cobby et al., 2001; Hodgson et al., 2003; Raber et al., 2002). When non-ground samples are included in the interpolation, mean-signed residual error will be positive (lidar-reference), i.e., the lidar DTM overestimates the reference elevation. When actual ground samples are mistakenly filtered out of the data before DTM interpolation, as discussed below, results are less predictable and depend on local topographic curvature. DTM peaks may be clipped or valleys filled due to inadequate retrieval of ground samples, resulting in the underestimation or overestimation of local elevation, respectively. Additional sources of error in lidar-derived DTMs include vertical and horizontal error in positioning the laser platform, laser scan angle, surface reflectivity, and slope of the terrain, all of which combined can add 0.20–2.00 m of error to an elevation estimate (Baltsavias, 1999b; Hofton et al., 2002; Huising & Gomes Pereira, 1998; Kraus & Pfeifer, 1998).

The filtering of vegetation from sub-canopy ground returns for the interpolation of DTMs, or "bald Earth," has been an active area of research. However, most algorithms for small-footprint lidar data are proprietary, and reported instrument and DTM accuracies are often poorly documented and generally assumed to be measured under optimal conditions—flat areas with no vegetation (Baltsavias, 1999a; Huising & Gomes Pereira, 1998). Some vegetation-filtering techniques include morphological filters and statistical analyses of heights in a neighborhood, and may be fully automated or involve some human interpretation (Huising & Gomes Pereira, 1998). Kraus and Pfeifer (1998) used an automated, iterative technique that interpolated a mean surface from the lidar cloud of *xyz* points and then successively removed or down-weighted points with residuals higher than a specified threshold. A relatively simple approach is to find local-minima relative to neighboring samples at a specified scale and/or search configuration (Cobby et al., 2001; Petzold et al., 1999). Resulting ground samples (i.e., local minima) must then be interpolated to form a surface.

There have been relatively few studies that provide rigorous accuracy assessments of lidar-derived DTMs under dense forest canopy with leaves, be they simple or composite leaves in hardwood forests or needles in conifer forests. Working with last-return lidar data flown over a leaf-on pine/deciduous forest landscape, Hodgson et al. (2003) identified ground points through a combination of proprietary software and human interpretation. A comparison of DTM elevation against 1470 survey-grade field measurements had an overall RMSE of 0.93 m. DTM error differed significantly by land use. Although RMSE was 0.33 m for low grass, it increased to 1.22 and 1.53 m for the more structurally complex scrubs/shrub and deciduous vegetation types, respectively. Furthermore, these researchers found that in the dense, multi-layered shrub/scrub class, there was a highly significant increase in DTM error of roughly 2 m from lowest (0–2°) to steepest (6–8°) slopes, which the authors attributed to vertical inaccuracies over relatively short horizontal distances under complex canopy. Cobby et al. (2001) developed an automated ground-retrieval scheme for a floodplain environment that included deciduous forests with leaf-on conditions. An initial DTM was interpolated from local-minima cells retrieved from non-overlapping, 5×5-pixel windows (10-m side) overlaid on a last-return DSM (2-m support). The final DTM was achieved by tailoring the ground-retrieval algorithm to short and tall vegetation classes. While terrain under short vegetation could be predicted with a 0.17-m RMSE ($n = 5$), the RMSE was 3.99 m ($n = 12$) under deciduous forests on steeper slopes (10–15°). Using last-returns filtered with proprietary methods, Reutebuch et al. (2003) found that a DTM under conifer plantations had a 0.32-m RMSE and a +0.22-m mean-signed error (overestimated, $n = 347$ reference points). There was only a slight 0.15-m increase in mean error between clearcut and uncut forest classes.

Hodgson et al. (2003) used a triangulated irregular network (TIN) to interpolate the DTM from ground points, while Cobby et al. (2001) used bilinear interpolation. Additional interpolation techniques include covariance linear weighting approaches, such as inverse distance weighting described below (Reutebuch et al., 2003), which can help smooth the high-frequency effects of spurious vegetation points, especially when applied in iterative filtering schemes (Lohmann & Koch, 1999).

1.2. Geostatistical methods for DTM generation

We used geostatistical techniques to interpolate a DTM from 0.33-m support ground cells derived from a vegeta-

tion-filtering algorithm. In the interpolation process, cells were treated as *xyz*-coordinate samples (i.e., points). Two common geostatistical interpolation algorithms, inverse distance weighted (IDW) (Bartier & Keller, 1996; Isaaks & Srivastava, 1989) and ordinary kriging (OK) (Goovaerts, 1997; Isaaks & Srivastava, 1989; Lloyd & Atkinson, 2002a,b), were considered in this study. Both techniques involve a weighted linear combination of neighboring data samples in estimating values at unsampled locations. In the case of IDW, weighting of neighboring samples is based solely upon an estimation location-to-sample inverse-distance function, along with a user-specified power weight factor (Bartier & Keller, 1996). The further away a sample is from the estimation location, and the greater the power weight factor, the less influence the sample will have on the estimate value. OK uses a distance-covariance model to weight samples based on their distance from the estimation location, as well as to down-weight samples that are clustered in space—an effect which tends to smooth the variance in the surface (Isaaks & Srivastava, 1989). The kriging covariance model is generally developed with reference to an empirical semivariogram (Goovaerts, 1997; Isaaks & Srivastava, 1989). Recently, Lloyd and Atkinson (2002a,b) reported that the more sophisticated OK technique was more accurate than IDW when interpolating DTMs from photo-interpreted (Lloyd & Atkinson, 2002a) or lidar-derived (Lloyd & Atkinson, 2002b) elevation points.

1.3. Lidar vegetation height and forest structure estimation

With discrete-return small-footprint systems, vegetation height is calculated as the difference between the original footprint heights and the bald-Earth DTM. The result is a set of estimated canopy heights with a footprint-scale support. Alternatively, a DSM is interpolated from footprint heights and subtracted from the DTM, thereby creating a canopy surface with height values recorded in square pixels or cells, i.e., a digital canopy model (DCM).

Small-footprint lidar technology now permits the detection and segmentation of individual tree crowns from fine spatial resolution DCMs (Brandtberg et al., 2003; Persson et al., 2002). Metrics applied to either DCM cells or footprint heights from within crown segments have been used to estimate the height and structure of individual crowns (Brandtberg et al., 2003; Gaveau & Hill, 2003; Næsset & Økland, 2002; Persson et al., 2002; Popescu et al., 2003). Such measures also hold promise for tree species classification (Brandtberg et al., 2003; Holmgren & Persson, in press). Persson et al. (2002) estimated individual conifer tree heights from DCM-cell maxima with a model r^2 of ~ 0.98 and RMS error of 0.63 m. Næsset and Økland (2002) estimated conifer tree height with the maximum of first-return footprints within crowns. The regression model explained 75% of the variance with a prediction RMS error of 0.23 m. Tree height was slightly overestimated by 0.18 m (3.15-m S.D.) in the work of Næsset and Økland, yet underestimated by 1.13 m in the work of Persson et al.. Næsset and Økland concluded that to reduce underestimation, high footprint density is needed to increase the probability of detecting the tops of conifer crowns. We know of only two studies that estimated heights for hardwood individuals. Working in West Virginia with forests in leaf-off winter conditions, Brandtberg et al. (2003) estimated hardwood tree heights with first-return footprints from within crown segments. A regression model explained 69% of the variance (RMSE not reported). There was an underestimation of height for taller trees, which the authors attributed to the low probability of the laser detecting the maximum crown height with leaf-off conditions, as well as random error in field measurements and inaccuracy in ground retrieval. In the United Kingdom, Gaveau and Hill (2003) estimated leaf-on hardwood tree and shrub heights from a first-return DCM with a model that explained 95% of the variance (1.89-m RMSE). The authors also presented strong evidence that an underestimation of tree heights (-2.12-m mean-signed error) resulted from laser pulses penetrating into the crown before reflecting a detectable first-return signal. The depth of signal penetration depended on variation of foliage and branches in the crown at the fine scale of a lidar footprint.

Stand-scale height of conifer-dominated forests have been predicted with varying levels of success from small-footprint lidar metrics (Næsset, 1997, 2002; Næsset & Økland, 2002). Several authors have shown that laser underestimation can be minimized by comparing a quantile of ground measurements (i.e., mean, maximum height) to a certain quantile of upper-most canopy returns in the stand (Magnussen & Boudewyn, 1998; Næsset, 1997, 2002). For example, Næsset (1997) found that the mean of lidar maxima (footprints) from a grid-overlay of square cells (i.e., 15×15 m) estimated average conifer stand height with a mean-signed error of -0.40 to 1.90 m (8–20-m tree height; 1.5-ha stands; RMSE not reported). For mixed conifer and hardwood forests with 6 to 29-m tall trees, Næsset (2002) used multiple-regression analyses to relate several lidar metrics to mean and dominant tree height at plot (0.02 ha) and stand (average 1.6 ha) scales. Lidar data included first and last pulses and metrics included the maximum, mean, density and various percent-quantile measurements within the plots. Model cross-validation RMS errors for young to mature forest plots ranged from 0.05 to 0.07 m ($r^2 = 0.82$ to 0.95) and 0.07 to 0.08 m ($r^2 = 0.74$ to 0.93) for mean and dominant plot-scale height, respectively. Stand-scale heights were estimated as the mean of plot-scale lidar estimates, and resulting models explained 92% and 87% of the variance for mean and dominant tree height, respectively (RMSE not reported).

We are not aware of any studies that have used small-footprint lidar to estimate stand-scale height for areas composed of evergreen, tropical tree species (i.e., hardwood trees, palms). However, one recent study by Lim et al. (2003) focused on leaf-on hardwood forests in Ontario,

Canada. Plot-scale (0.04 ha) Lorey's mean height was estimated using either the mean or maximum of lidar pulses (first and last return). Regression models explained 66% and 86% of the variance, and it was found that tree height was underestimated (RMSE not reported).

Fine-scale lidar data has also been used to estimate stand-scale biophysical properties, such as volume, biomass, crown diameter and diameter breast height (Lim et al., 2003; Popescu et al., 2003). Advances in analyzing large-footprint waveforms have produced metrics that can estimate biophysical parameters over dense conifer and broad-leaf forests with considerable accuracy [see Lefsky et al. (2002) for a recent review of the literature]. Most notable for this study, Drake et al. (2002a) used multiple-regression and metrics from LVIS waveforms to predict plot-scale (025 to 0.5 ha) basal area, aboveground biomass, and quadratic-mean stem diameter over a range of tropical forest types at the La Selva Biological Station, Costa Rica; their resulting models explained 93%, 72% and 93% of the variance, respectively, and models did not saturate with increasing forest height and complexity.

2. Methods

2.1. The La Selva Biological Station study site

This study was conducted at the La Selva Biological Station of the Organization for Tropical Studies (OTS). Located in the Atlantic lowlands of north–east Costa Rica 84°00′ 13.0 W, 10°25′ 52.5 N), the 1614-ha reserve contains a mixture of old-growth terra firme, swamp, secondary and selectively logged forests, as well as agroforestry plantations, developed areas with buildings and mowed grass, and abandoned pastures with large grasses, shrubs and remnant trees (Fig. 1). This site receives on average 4 m of precipitation annually, and has a short dry season (December through April) that is rarely long or severe. The old-growth forest is classified as a tropical wet forest in the Holdridge Life Zone System and is characterized by a species-rich, multi-layered community composed of trees, palms, lianas, and other terrestrial and epiphytic plants (Hartshorn & Hammel, 1994). There are 323 species of hardwood trees and leaf morphology includes simple or compound arrangements, large or small sizes and variable thickness. Although some trees are deciduous in the dry season, the canopy as a whole can be considered evergreen (Fetcher et al., 1994). The geomorphology of the landscape is structured by two main features: (1) highly eroded lava flows that contain a system of alternating ridges and stream valley-bottoms separated by steep slopes, and (2) flat to gently undulating alluvial terraces (Sanford et al., 1994). The reserve is covered by a 50 × 100-m grid, oriented 32° from North, with permanent markers at each grid intersection. The local grid coordinate system permits researchers to accurately geo-locate field data in a local Geographic Information System (GIS) Cartesian coordinate system.

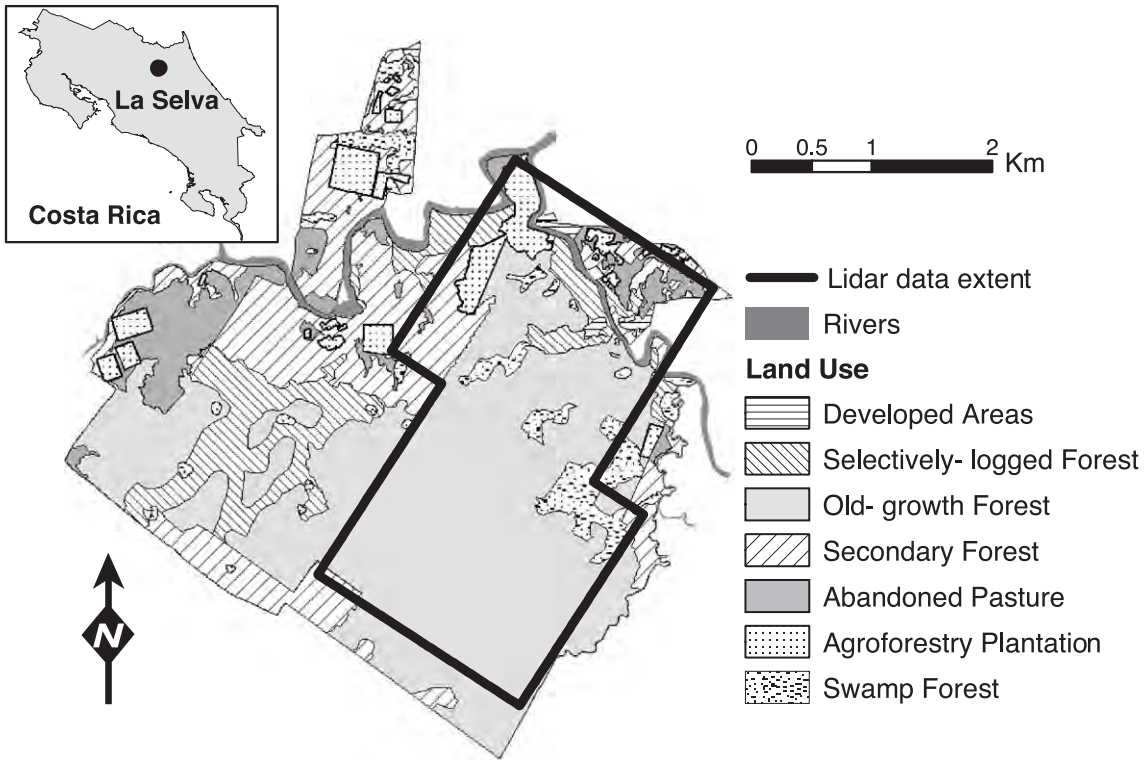

Fig. 1. La Selva Biological Station study area land-use and lidar data extent.

Table 1
Topographic reference data

	Count	Min	Max	Mean	Median
Elevation (m)	3859	38.78	141.37	70.93	62.38
Slope (deg)[a]	932	0.00	44.00	12.94	12.00

[a] Measured only in old-growth forest.

2.2. Topographic reference data

The DTM elevation was compared to 3859 in situ elevation points (Table 1) that were surveyed in 1991 using optical-leveling techniques (Hofton et al., 2002). In the comparison, co-located survey points and DTM cells were determined by linking points to the nearest 1-m DTM cell centroid based on *x*,*y* coordinates. Survey data cover the range of land-use and geomorphic conditions found in the reserve. A total of 1321 of these points were measured at grid intersections, while the remaining points do not have permanent markers and are located off the grid, mainly in the northwest to southeast direction between grid intersections. All survey points were transformed to the Universal Transverse Mercator (UTM), WGS-84 datum coordinate system from the local La Selva coordinate system using a least-squares affine transformation based on five differential-GPS points, as used by Hofton et al. (2002). DGPS measurements were taken from permanent towers or in open areas to avoid obstruction by vegetation. Differential corrections were performed using data from a base station at La Selva. The transformation had an overall RMSE of 0.57 m ($X_{RMS}=0.48$ m, $Y_{RMS}=0.31$ m). The vertical values of the survey points were shifted from the local reference mean sea level (MSL) to the WGS-84 ellipsoid MSL using the constant 11.44-m offset published by Hofton et al.. This offset was verified by selecting the four to five nearest FLI-MAP (discussed below) lidar returns around three grid monuments in open areas with very flat, mowed lawn. The average vertical offset between surveyed and lidar-derived elevations was also found to be 11.44 m. Neither the FLI-MAP data nor the reference data were referenced relative to a geoid, and so the elevation offset between the datasets is constant across the study extent.

DTM error was analyzed relative to slope and land-use factors. A total of 932 survey points at grid intersections in old-growth forest were further classified into one of four slope classes based on field measurements from a separate forest structure study (Clark et al., 1999; field data summarized in Table 1). Slope classes included: 0–3°, 3–10°, 10–20° and >20°. A total of 2060 survey points were classified into one of seven land-use categories based on an overlay operation with an existing year 2000 land-use map (Fig. 1) derived from historic aerial photographs, IKONOS imagery and survey maps (OTS, unpublished data). The land-use categories, from short-simple to tall-structurally complex vegetation, were: Developed Areas, Pastures, Plantations (abandoned and current), Secondary forests (1 to 34 years old), Selectively logged Forest (including 50-year-old, abandoned agroforestry areas), Swamp forests and Old-growth forests. To limit the confounding effect of slope on DTM error (discussed in Section 3), only survey points on slopes less than 10° were considered in the land-use analysis. A mask delineating areas with slopes less than 10° was created by sampling the DTM-derived slope (Arc/Info 8.0.1, Environmental Systems Research Institute, Redlands, CA) after filtering with an averaging 3 × 3-cell filter.

2.3. Reference field heights

2.3.1. Individual tree heights

In the field, the maximum height of a crown was estimated with a handheld laser range-finder (Impulse-200LR, Laser Technology, Englewood, CO) for individual trees taller than 20 m. Several readings of the crown apex (i.e., highest foliage) were taken per individual from different locations (if possible) and trees were included in analyses if their measurement standard deviation was less than 1 m. The mean of the readings for an individual established the tree's height. Only the >20-m height class was considered because field measurements were taken 3–4 years after the FLI-MAP overflight, and rapid tree growth in smaller tree-height classes during this time would confound the analysis.

The heights of 21 trees in pasture with isolated crowns were measured in the field during July 2001 (~ 9 species, Table 2). The crown center (centroid) for each pasture tree was estimated using DGPS readings from a Trimble Pro XL GPS and on-site base station (Trimble Navigation Limited, Sunnyvale, CA) and subsequent visual spatial adjustment of the DGPS points with reference to the DCM.

Between February and October 2000, the heights of 59 old-growth forest emergent trees were measured (11 species, Table 2) in the field and their trunks were geo-located relative to the closest grid marker using a tape measure and compass. Trunk locations were then transformed to UTM coordinates (Hofton et al., 2002) and crown centroids were identified through visual adjustment of trunk points with reference to the DCM. We selected canopy-emergent trees in old-growth forest for analysis because (1) they are major components of overall forest biomass (Clark & Clark, 1996), (2) they were easy to locate unequivocally in the DCM, and (3) their height growth in the 3 years between field measurement and the lidar overflight was expected to be minimal.

Table 2
Height reference data for individual trees and plots (units in meters)

	Count	Min	Max	Mean	Median	S.D.
Old-growth emergent trees	59	31.00	56.39	45.04	44.34	5.76
Pasture trees	21	20.01	51.19	31.64	27.54	10.73
Plots (all stems)	32	0.38	17.53	6.19	4.70	5.27
Plots (tree stems)	32	0.38	18.53	7.28	4.70	6.17

2.3.2. Plot-scale stem heights

For plot-scale reference data, we used stem heights from a 1997 census of agroforestry plantations within La Selva (Menalled et al., 1998). The 32 plantation plots considered in this study were on 1-year ($n=9$), 4-year ($n=9$) and 16-year ($n=14$) cutting cycles and had stems spaced 2 m apart at roughly equal densities at the time of the census. Each plot had a total area of 0.04, 0.08, and 0.13 ha, respectively, and the oldest trees in a plot were 6 years old in 1997. All trees were one of three species: *Hyeronima alchorneoides* (*Euphorbiaceae*), *Cedrela odorata* (*Meliaceae*), or *Cordia alliodora* (*Borginaceae*). Seven of the fourteen 16-year rotation plots included a mixture of one of the three dominant tree species as well as a sub-canopy palm *Euterpe oleracea* (mean height 7.13 m) and a smaller understory shrub *Heliconia imbricata* (mean height 1.53 m). The two-dimensional area of each field plot (i.e., a rectangle polygon) was digitized in a GIS and located with visual reference to the DCM.

In the field, height measurements were recorded with a laser range-finder for every individual stem in the plot. Plot mean height (Table 2) was calculated in two ways: (1) as the average of heights for all stems in the plot (termed "all stems"), and (2) the average of heights for trees in the plot (termed "tree stems"). Note that the difference between these two metrics lies in the inclusion of a mix of canopy (tree) and sub-canopy (palm and shrub) stem heights in the plot mean-height calculation for 7 of the 32 plots; the mean height of the other 25 plots, which did not include the palms and shrubs, did not differ between the two methods.

2.3.3. Accuracy of field-height measurements

We established the vertical accuracy of the laser range-finder in two ways. In one case, five height estimates at 10-, 20-, 30- and 40-m distances from a 15.00-m pole in an open area were found to have an overall vertical mean error of $+0.18 \pm 0.14$ S.D. m (slight overestimation). There was no significant effect on the mean height measured with distance (Kruskal–Wallis non-parametric ANOVA, $\alpha=0.5$). Next we analyzed the stem height of actual trees. Crown apices were more difficult to identify from the ground than the top of the 15-m pole. Four trees in open areas had range-finder height estimates made before being cut down. The actual height of the tree was then measured from its horizontal position on the ground, including the stump. The four trees had an average ground-measured height of 27.3 m and a range-finder estimate mean error of $+1.45 \pm 1.67$ S.D. m (overestimation). In summary, reference measurements using the laser range-finder have their own associated error, mostly due to the human observer's inability to identify the apex of the tree crown, rather than distance from the target. The overall bias appears to be an overestimation of tree height by the laser range-finder technique.

2.4. Small-footprint lidar data

The lidar data used in this study is from the FLI-MAP system (John E. Chance and Associates, Lafayette, LA), which was flown in a helicopter over La Selva in October 1997. The sensor is a small-footprint, first-return 0.9-μm laser sensor that has a 8000-Hz pulse rate, 30° scan angle, 2-mrad beam divergence, typically 9-point/m^2 sampling density, 30-cm footprint spacing, and a rated vertical accuracy of ~ 10 cm (Blair & Hofton, 1999; Huising & Gomes Pereira, 1998). The 10-cm footprints (Hofton et al., 2002) were converted to a Triangular Irregular Network (TIN) (John E. Chance and Associates), and the final product delivered for analysis by the data distributor (Fig. 2A) was a raster digital surface model (DSM) interpolated from the TIN with 0.33-m cell support (each containing a height).

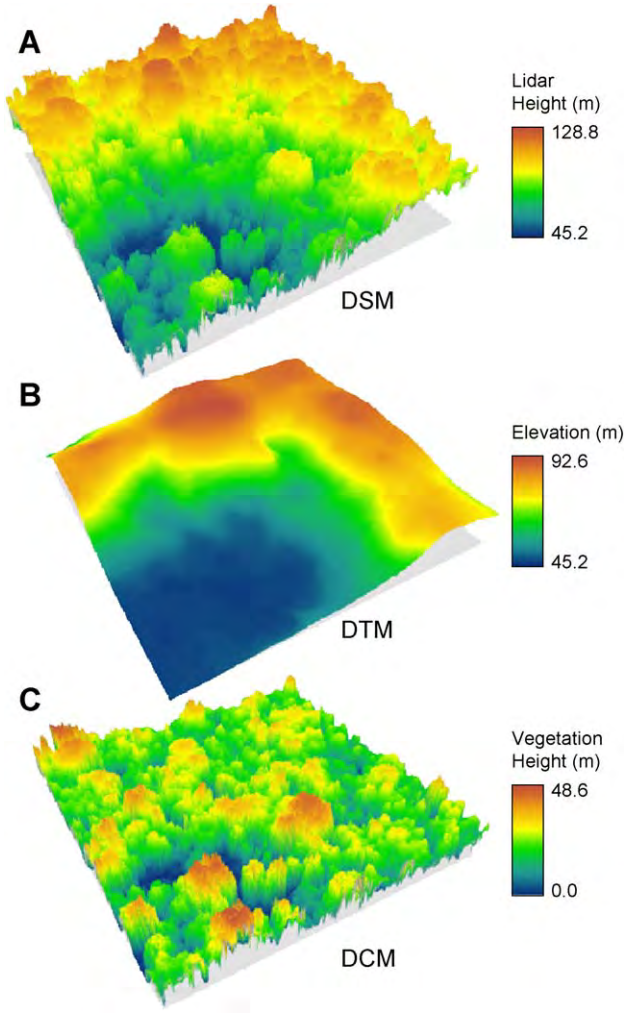

Fig. 2. A three-dimensional perspective of a 250 × 250-m subset of the lidar raster products, all covering the same geographic extent in an old-growth forest: (A) the unprocessed lidar height surface (i.e., digital surface model, DSM), (B) the estimated sub-canopy elevation surface (i.e., digital terrain model, DTM), and (C) the estimated vegetation height surface (i.e., digital canopy model, DCM) resulting from the subtraction of the DTM (B) from the DSM (A). In (C), canopy emergent trees are red-tone, concave domes.

The DSM covered a 754-ha area of the reserve (Fig. 1). Previous lidar research that used this dataset include analyses of LVIS pseudo-waveforms synthesized from the DSM cells (Blair & Hofton, 1999) and a comparison of LVIS to FLI-MAP sub-canopy elevation retrieval (Hofton et al., 2002). In the latter study, FLI-MAP elevation readings over areas of structurally complex vegetation (e.g., old-growth forest) were found to be mostly from canopy, not the ground, as is to be expected from a first-return sensor. However, in a 25-m diameter circular area, there were at least some FLI-MAP hits that coincided with survey elevations (Hofton et al., 2002), which may be due in part to sensor's relatively high sampling density (Huising & Gomes Pereira, 1998).

2.5. Ground retrieval and DTM interpolation

To create a DTM, FLI-MAP data processing involved two major steps: (1) identifying ground-return cells in the DSM, and (2) subsequent geostatistical interpolation of ground cells to form a DTM. We developed two simple ground-retrieval algorithms that sought a balance between processing speed and the minimization of elevation error over the range of vegetation and terrain conditions found at La Selva.

2.5.1. Local-minima ground-retrieval scheme

The local-minima algorithm proceeded as follows: a grid of non-overlapping, square cells was overlaid on top of the original DSM. Within each grid cell, one local-minima DSM cell (0.33-m support) was selected and identified as a ground return. This procedure resulted in a population of ground-return cells for each of the five grid scales considered independently: 5, 10, 15, 20 and 30 m. As previously discussed in Section 2.4, at a relatively coarse scale (e.g., a 25-m diameter circle), there is expected to be at least one lidar pulse that penetrates to or near the ground surface. This pulse would register as a cell of relatively low height in the DSM; and thus, the above local-minima scheme is analogous to selecting the lowest return in a square footprint of a specified scale (i.e., 5, 10 m, etc.). Ground-return cells identified at each scale were then used in separate geostatistical interpolation schemes (described below) that generated DTMs with a 1-m cell size. Samples from each DTM were compared to 3859 co-located reference points (Section 2.2). The overall RMS errors of the resulting DTMs were used as the basis for the selection of a final ground-retrieval/interpolation scheme.

2.5.2. DTM interpolation schemes

The inverse distance weighted (IDW) and ordinary kriging (OK) geostatistical techniques were assessed for the interpolation of the final DTM (Section 1.2). In both procedures, DSM ground-return cells (0.33-m support) were treated as representing a cloud of *xyz*-coordinate points, with the cell's centroid defining the *x* and *y* coordinates and the cell's DSM height defining the *z* value. Note that these points have less variance than they would if derived directly from the original lidar footprints, instead of from the DSM. This is because DSM cell values were smoothed from TIN-interpolation and subsequent rasterization.

IDW interpolation of ground points was conducted using Arc/Info 8.0.1 (Environmental Systems Research Institute, Redlands, CA). The IDW interpolation moved in a circular window of 50-m radius and included at least four data points. The IDW power parameter specifies how sample points are weighted with distance from the interpolation node. A lower power relaxes the weighting of near points, causing more smoothing in the interpolated surface (Isaaks & Srivastava, 1989). Initial experiments based on minimizing RMSE identified that a value of 2 was the most appropriate IDW power for this dataset.

For OK, an isotropic normal-score transform variogram (Goovaerts, 1997) for each set of ground points (i.e., 5, 10, etc. scales) was manually fitted using Variowin v2.21 (Pannatier, 1996). These models contained a short- and long-range nested structure and a very small nugget effect (Goovaerts, 1997; Isaaks & Srivastava, 1989). Ordinary kriging of the normal-score data points was performed using the GSLIB v2.0 software package (Deutsch & Journel, 1992) with a 250-m search radius, and the inclusion of 4 to 12 data points at each interpolation node.

2.5.3. Iterative-addition ground-retrieval scheme

An additional ground-retrieval technique was tested that iteratively added in ground points at successively finer scales (Fig. 3). Point selection proceeds through three iterations, each with three steps (in Fig. 3; Iterations 1–3 are rows and steps are columns). In the first step of the first iteration (Fig. 3; row 1, column 1), a coarse-scale grid (i.e., 20-m cells) is overlaid on the original DSM and one 0.33-m minima cell is selected from within each 20 × 20-m cell. In Step 2 (Fig. 3; row 1, column 2), DSM minima cells are converted to *xyz* points. In Step 3 (Fig. 3; row 1, column 3), a first-pass DTM is created from IDW interpolation with a 5-m support. The first iteration is essentially the local-minima scheme as described above, but with a 5-m IDW interpolation.

For Step 1 of Iterations 2 and 3 (Fig. 3; rows 2 and 3, column 1), the coarse-scale grid is reduced by 5-m (i.e., from 20 to 15 m, or 15 to 10 m) and overlaid on the DSM; again, minima cells are identified, converted to *xyz* points, and then added to the population of minima points from the previous iteration. In Step 2 of Iterations 2 and 3 (Fig. 3; rows 2 and 3, column 2), minima-point *z* values are subtracted from the DTM resulting from the previous iteration (row 1 or 2, column 3), and individual points 0.25-m higher than the DTM (marked with "+" in Fig. 3) are removed from the population of minima points (Fig. 3; row 2 or 3, column 2). In Iteration 2, Step 3 (Fig. 3; row 2, column 3) another preliminary DTM is interpolated with IDW for use in the third iteration. At the end of the Iteration 3, a final population of ground points is interpolated with IDW or Ordinary Kriging (OK) with a 1-m support (Fig. 3; row 3, column 3).

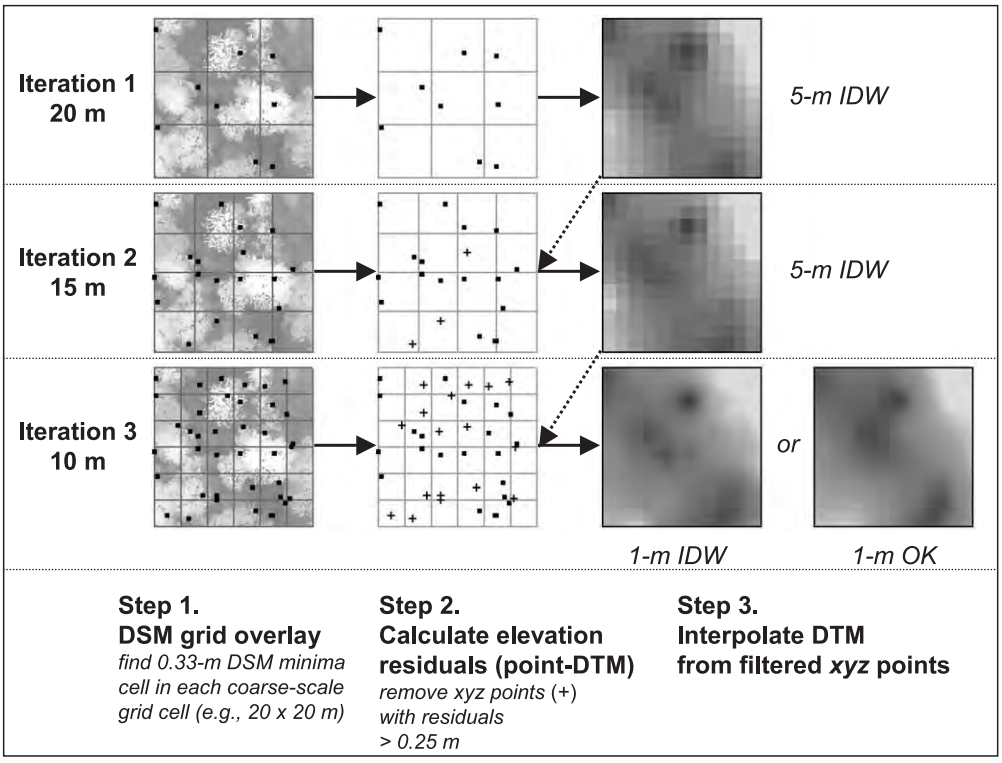

Fig. 3. Conceptual flow of the iterative-addition ground-retrieval scheme (see Section 2.5 for details). Abbreviations are: digital surface model (DSM), digital terrain model (DTM), inverse distance weighed interpolation (IDW), and ordinary kriging interpolation (OK). Note: to create the figure, the same 60 × 60-m extent was placed on top of the actual DSM, and ground samples outside the extent were included in the DTM interpolation (Step 3, all iterations).

The effect on the mean-signed error of the final DTM using a 0.25-, 0.50- and 1.0-m residual threshold was evaluated; and subsequently, 0.25 m was deemed a conservative residual threshold in that it allowed finer-scale variation to be added back to the population of ground-return points, while it minimized the risk of including understory vegetation points (and subsequently inflating the mean-signed error). To minimize processing time, an IDW interpolator with 5-m support was chosen for generating intermediary DTMs (Fig. 3; row 1 or 2 only, column 3); future research should test the impact of using a finer-scale interpolation (i.e., 1 m) or an ordinary kriging interpolator on the accuracy of the final DTM product.

2.5.4. Sink removal

Roughly 0.5% of the DSM-derived, *xyz* points from the ground-retrieval schemes contained very low, spurious points that created local sinks in the interpolated DTM. We prevented some sinks from entering the population of ground points by imposing two restrictions on DSM local-minima cells: (1) that the minima cell not lie 60-m lower than its highest neighboring DSM cell (no trees or other objects in the scene are expected to be taller than 60-m high), and (2) the cell not lie lower than the minimum study site elevation (30 m). Sink *xyz* points that persisted through the ground-retrieval process were automatically removed by calculating their sink depth from a first-pass DTM interpolation. All sink points greater than 5-m in depth were deleted from the population of ground points, and the DTM surface was re-interpolated.

2.6. Digital terrain model (DTM) accuracy assessment

For each DTM interpolated, error was calculated by subtracting DTM elevation from the elevation measured at co-located field-survey points (DTM-reference; sensu Hodgson et al., 2003). Calculated error statistics included mean-signed error, mean absolute error (MAE), and RMSE (Isaaks & Srivastava, 1989).

The final DTM with the lowest RMSE was submitted to a more rigorous analysis that sought to understand the variation of error across the TRF landscape. We asked the following questions:

(1) Does DTM error differ among slope inclination categories in old-growth forest?
(2) Does DTM error differ among land-use categories?

As reviewed in Section 1.1, relatively dense and structurally complex vegetation decreases DTM accuracy. Lidar range error also increases non-linearly with greater surface-inclination angle (Baltsavias, 1999a); and, when dense vegetation types occur in areas of steep terrain, these combined vertical errors decrease the accuracy of the interpolated surface over relatively short horizontal distances (Hodgson et al., 2003).

The square-root transformation of the mean absolute error (SqrtMAE) was used in an ANOVA to test for statistical significance of the above hypotheses (sensu Hodgson et al., 2003), with $\alpha = 0.05$. The square-root transformation of MAE was chosen to adjust the positively skewed distribution of MAE to a normal distribution, as required by ANOVA. Since the minimum distance between reference points was 0.28 m, the classical ANOVA assumption of independence of residuals was suspect; therefore, we opted to use the generalized least-squares ANOVA with a residual variance–covariance weighting, formulated by Gotway and Cressie (1990). The weightings take into consideration residual autocorrelation and reduce chances of committing Type I errors. This technique requires a variance–covariance model of the SqrtMAE residuals (SqrtMAE minus a trend component). The trend component of the SqrtMAE points was estimated using ordinary kriging of the local mean (Goovaerts, 1997). An isotropic variogram of the resulting SqrtMAE residuals was modeled, and this model in conjunction with the data-to-data distances of the survey points were used to estimate the SqrtMAE residual variance–covariance sub-matrices for each class (e.g., Old-growth, Developed Areas, etc.), as required by the spatial ANOVA (Gotway & Cressie, 1990).

2.7. Stem height accuracy assessment

A digital canopy model (DCM) was calculated by subtracting the final 1-m support DTM, bald-Earth surface from the lidar DSM (i.e., DCM = DSM – DTM, Fig. 2). The DCM maintained the 0.33-m support of the DSM. Individual tree and plot-scale stem heights were estimated from the DCM using metrics that had comparable spatial supports and calculations as those used in the field (sensu Drake et al., 2002b; Næsset, 1997).

Plot-scale DCM metrics considered the same two-dimensional area as field plots, i.e., same spatial support. In field measurements, plot mean height was calculated as the average height of stems spaced 2-m apart (Section 2.3). Following Næsset (1997), we devised a comparable lidar metric that averaged the DCM local-maxima (i.e., maximum height) from each 2 × 2-m cell in a grid overlaying each plot. Alternatively, we also calculated the average of all heights (i.e., all DCM cells) in the plot extent. These two lidar metrics are referred to as $\text{Mean}_{2\times 2}$ and Mean_{ALL}, respectively.

Field measures of individual tree height located the highest leaf or twig in a crown (i.e., crown apex). Similarly, we based DCM metrics for estimating tree height on a sample of cells from within the crown. All DCM cells within 5 horizontal meters of the crown centroid were sampled and tree height was estimated as the maximum value of the cells (referred to as ''Maximum''). Since crowns had multiple peaks of high foliage, field measures may have missed the highest crown apex. In consideration of this potential field error, we devised another metric that averaged DCM cells whose values were above the 95% quantile (sensu Ritchie et al., 1993; referred to as Mean_{95}).

At individual tree and plot scales of analysis, height error was calculated as the difference between the DCM and field-height metrics (DCM – field, sensu Næsset, 1997). These errors, which are residuals from a 1:1 relationship, were summarized with the statistics mean-signed error, MAE and standard deviation (Isaaks & Srivastava, 1989).

We also evaluated the linear relationship between lidar and field-height metrics with regression analyses, which provided a coefficient of determination (r^2) value and a model RMSE (referred to as RMSEm). RMSEm was calculated from prediction residuals (sensu Drake et al., 2002a; Næsset, 2002; Næsset & Økland, 2002) and is the general error expected if the regression model were applied to DCM-derived estimates, i.e., calibration using ground measurements. Because we did not have an independent set of field data for model validation, model prediction residuals needed for calculating RMSEm were acquired through a common cross-validation procedure (Drake et al., 2002a; Næsset, 2002; Næsset & Økland, 2002; Popescu et al., 2003).

3. Results

3.1. DTM generation and accuracy assessment

The five local-minima ground-retrieval scales tested in this study (i.e., 5, 10, 15, 20 and 30 m) had overall RMSE

Table 3
Error of digital terrain model (DTM) elevation estimates ($n = 3859$, units in meters)[a]

Interpretation[b] (ground retrieval[c])	RMSE[d]	MAE[e]	Mean	S.D.	Median	Min	Max
IDW (local minima)	2.47	1.78	+0.68	2.37	0.69	−14.80	17.64
IDW (iterative-addition)	2.47	1.75	+0.08	2.47	0.29	−16.29	16.54
OK (local minima)	2.39	1.69	+1.10	2.13	0.94	−14.77	18.62
OK (iterative-addition)	2.29	1.60	+0.97	2.07	0.83	−17.14	16.89

[a] Error is calculated as DTM–field-survey elevation (deviation from 1:1 relationship).
[b] Interpolation schemes are: IDW = inverse distance weighted, or OK = ordinary kriging.
[c] Ground-retrieval schemes are: local-minima = minimum DSM cell in a 20 × 20-m grid, or iterative-addition = iteratively adding local-minima from 20-, 15- to 10-m scales.
[d] RMSE is root mean-square error.
[e] MAE is mean absolute error.

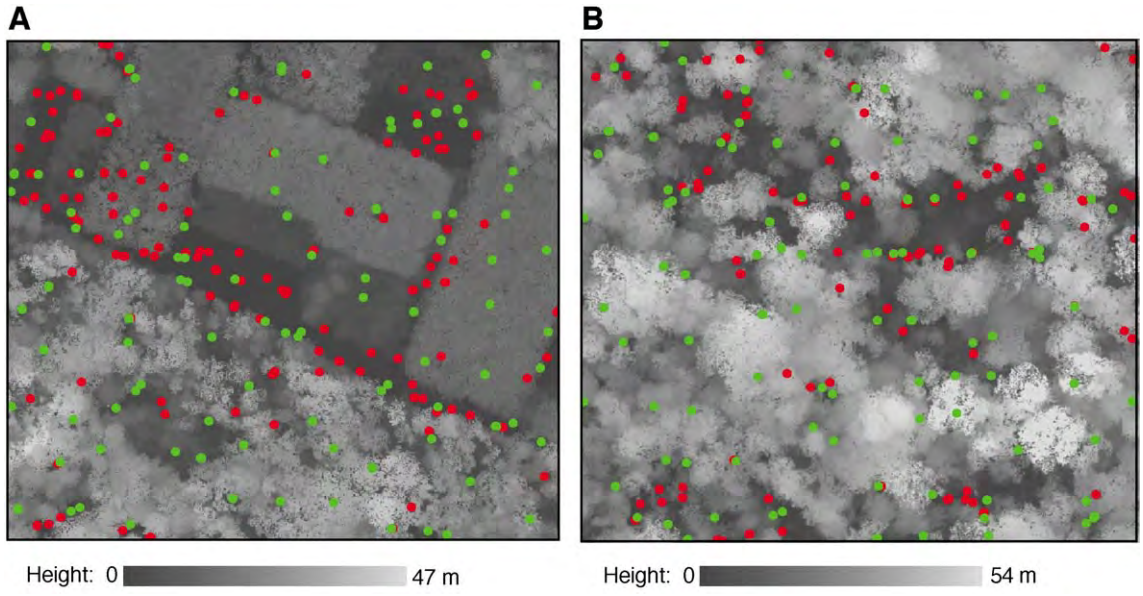

Fig. 4. An example of ground-retrieved samples (i.e., *xyz* points) for (A) agroforestry plantations, and (B) an old-growth forest with an open swamp. Points shown are for the 20-m local-minima retrieval scheme (green dots) and the iterative-addition scheme spanning 20-, 15- and 10-m scales (green and red dots combined). The underlying gray-scale surface is from the final DCM (Figs. 2c and 8). Extent is 210×180 m.

errors ranging from 2.29 to 5.09 m, using either IDW or OK for surface interpolation (data not shown). The scale with the lowest RMSE for both interpolation methods was found to be 20 m. For this tropical landscape and lidar sampling density, 20 m appears to be the near-optimum scale to identify ground returns with the local-minima approach, and so only the results from this scale will be discussed (Table 3). This optimal scale is likely determined by the average crown dimensions and canopy gap characteristics in old-growth forest, which comprises 69% of the study area. By visual interpretation of colorized 1-m IKONOS imagery, Clark et al. (2004) measured mean maximum crown diameter to be 19.6 m for all trees within old-growth plots. In this study, mean maximum crown diameter of the very largest, canopy-emergent trees in old-growth forests were measured from the ground to be 27.2 m ($n = 36$). At 15-m scales or finer, there was a higher probability that a DSM local-minima cell could lie completely within one large crown or many connected crowns. If these crowns are vertically dense, or have understory trees underneath, the

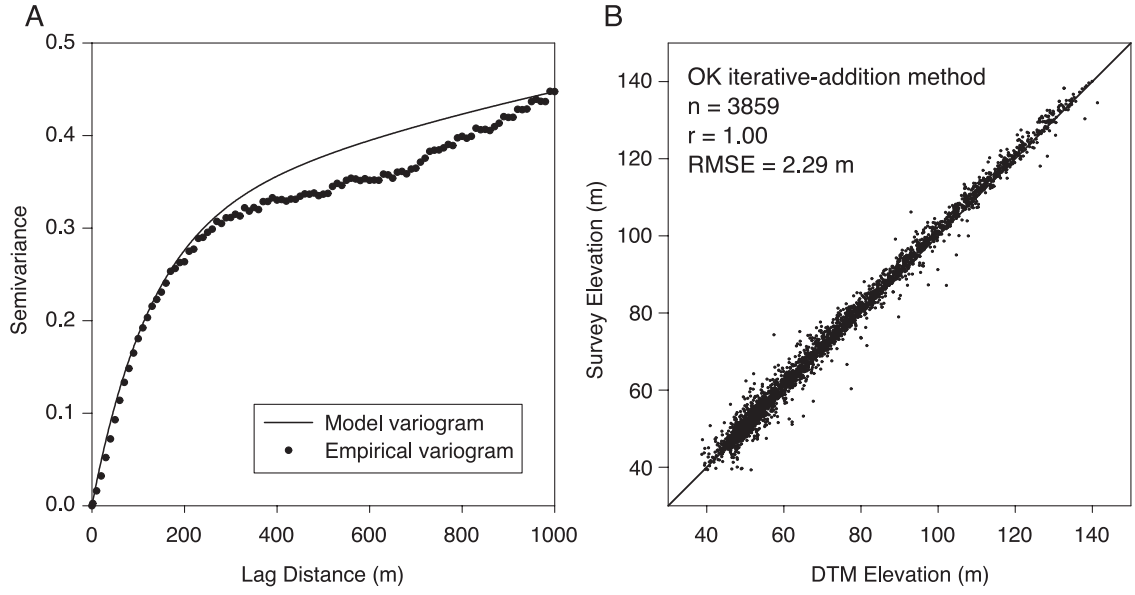

Fig. 5. (A) Empirical and modeled variograms from elevation points retrieved from the iterative-addition algorithm (20-m starting scale). (B) Relationship between lidar-estimated ground elevation (sampled from the final DTM (Fig. 2b)) and field-surveyed elevation.

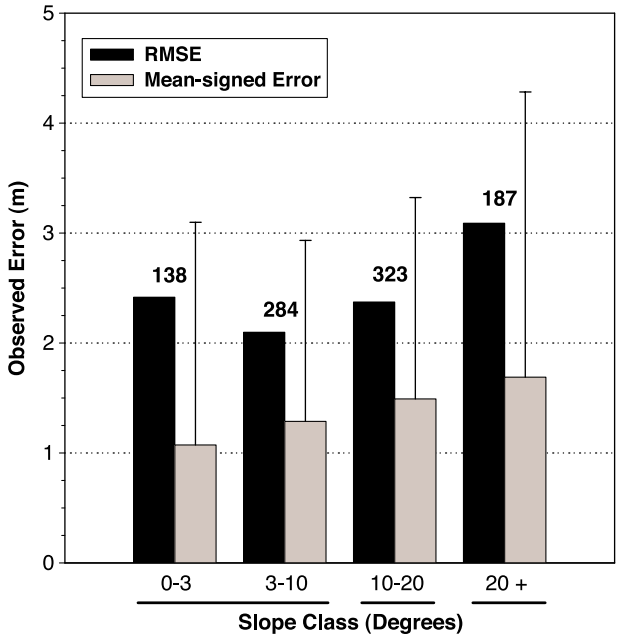

Fig. 6. Root-mean-square error (RMSE) and mean-signed error distribution by slope class within old-growth forest. Numbers above the bars indicate the number of reference points for each category. Lines connecting classes indicate groups of homogenous mean absolute error ($\alpha = 0.05$; *t*-tests account for heteroscedasticity (Gotway & Cressie, 1990)).

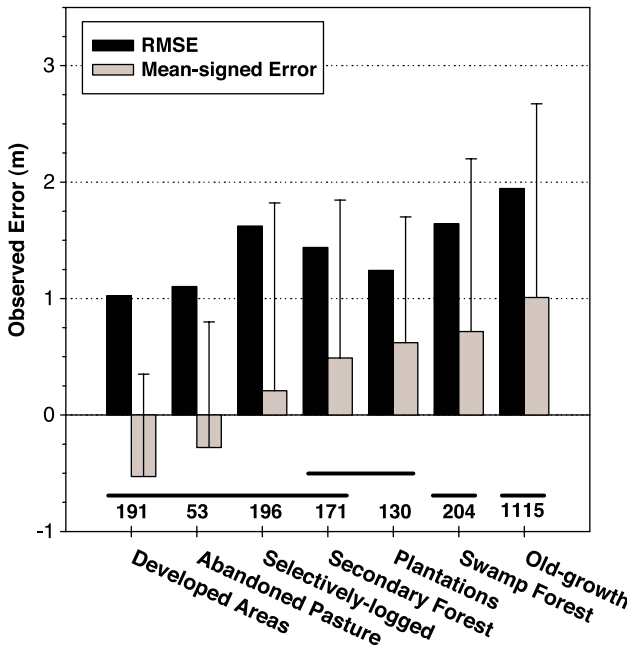

Fig. 7. Root-mean-square error (RMSE) and mean-signed error distribution by land-use class for all samples with slopes $\leq 10°$. Numbers below the bars indicate the number of reference points for each category. Lines connecting classes indicate groups of homogenous mean absolute error ($\alpha = 0.05$; *t*-tests account for heteroscedasticity (Gotway & Cressie, 1990)).

DSM local minima would likely be from upper- or lower-canopy leaves or branches, not the ground. At 30-m scales or greater, there was a higher chance of selecting a canopy-gap DSM cell with a height from or near the ground; however, a trade-off exists in that local minima at coarser scales fail to sample fine-scale topography. As shown in Fig. 4 (green points), the 20-m local-minima scheme identified DSM-cell minima (subsequently converted to *xyz* points) on the periphery of dense vegetation, such as in the shortest plantation plots, on the edges of emergent tree crowns and in more open swamps. The 20-m scale, iterative-addition scheme could identify more of these local-minima cells (*xyz* points), while still avoiding the areas of densest vegetation (Fig. 4, green and red points together).

The population of *xyz* points for each ground-retrieval scheme was used to model variograms needed for OK interpolation. The normal-score variogram from the 20-m scale, iterative-addition dataset (method used for final DTM interpolation, Fig. 5A) was found to have a zero nugget effect, exponential short-range structure of 380-m range with a sill of 0.31, and an additional long-range spherical structure spanning to a 2000-m range, comprising an additional 0.20 of the total sill (maximum sill modeled in study extent was 0.51).

Considering all 3859 survey points distributed throughout the landscape, the correlation (*r*) between DTM and reference elevation was +0.99 ($p < 0.0001$) for IDW interpolation and +1.00 ($p < 0.0001$) for OK interpolation, using either ground-retrieval scheme. In terms of RMSE, OK also performed better than IDW for DTM interpolation. There was a 0.18-m RMSE difference between the best IDW and OK interpolated surfaces (Table 3). The inclusion of finer-scale local-minima points using the iterative-addi-

Table 4
DTM error on slopes $\leq 10°$ summarized by land use (in meters)[a]

Land-use class	Mean ± S.D. (RMSE)
Developed areas	−0.53 ± 0.88 (1.02)
Abandoned pastures	−0.28 ± 1.08 (1.10)
Selectively-logged Forest	+0.21 ± 1.61 (1.62)
Secondary Forest	+0.49 ± 1.36 (1.44)
Agroforestry plantations	+0.62 ± 1.08 (1.24)
Swamp Forest	+0.72 ± 1.48 (1.64)
Old-growth Forest	+1.01 ± 1.66 (1.95)
All classes	+0.66 ± 1.59 (1.72)

[a] DTM interpolation from OK, iterative-addition scheme.

Table 5
Canopy-surface height summarized by land use[a]

Land-use class	No. of cells	Area (Ha)	Mean (m)	S.D. (m)
Abandoned pastures	2,126,838	23.6	3.3	5.0
Developed areas	826,562	9.2	8.0	8.2
Agroforestry plantations	3,287,534	36.5	12.8	9.0
Secondary Forest	5,395,687	60.0	12.9	8.2
Old-growth Forest	43,675,481	485.3	20.5	8.8
Swamp Forest	4,210,094	46.8	20.5	11.1
Selectively-logged Forest	3,271,990	36.4	20.8	11.0

[a] Calculated by from the digital canopy model (DCM), 0.33-m cells.

tion scheme (i.e., iteratively adding local-minima from 20,15 to 10-m scales) improved the OK-interpolation RMSE by 0.10 m, resulting in an overall RMSE of 2.29 m (Fig. 5B; Table 3). All DTMs had a positive mean-signed error, and so they tended to overestimate elevation. Although this overestimation was up to 0.87 m higher with OK relative to IDW interpolation, the mean absolute error (MAE) using OK was lower (Table 3). Indeed, the OK DTM was significantly more accurate than IDW interpolation, using either ground-retrieval technique ($n = 3859$, paired t-test of sqrtMAE, $p < 0.0001$ for both comparisons). As expected, OK was found to reduce the variance of errors (Table 3, standard deviation). This smoothing of the variance across space tends to minimize the influence of spurious understory vegetation or downed trunks that are inevitably included in the DTM interpolation. In contrast, IDW does not exhibit this desirable smoothing effect as strongly as OK. The OK iterative-addition method was selected to generate the final DTM because the surface had a relatively low error variance and the lowest overall RMSE and MAE. This DTM was free of obvious underestimation (i.e., sinks) or overestimation (i.e., peaks) errors and revealed remarkable fine-scale detail of geomorphology (Fig. 2B).

As expected, elevation error in the final DTM was not distributed evenly across the landscape. Mean absolute error (MAE) was significantly different between slope inclination classes in old-growth forest ($p < 0.0001$). Mean-signed error was positive (overestimation of DTM elevation) and increased 0.62 m from the lowest to the steepest slope classes (Fig. 6, +1.07-m [0–3°] vs. +1.69-m [>20°] mean-signed error). The difference in MAE between slopes 0–3° and 0–10° was not significantly different (paired t-test; $p = 0.68$); however, the MAE on slopes ≤10° was significantly less

Table 6
Estimation error of individual tree heights[a]

	MAE[b]	Mean	S.D.	Min	Max
Old-growth trees[c] *(n = 59)*					
Maximum[d]	3.67	−2.11	4.11	−5.02	11.59
$Mean_{95}$[d]	3.94	−2.72	4.15	−4.40	12.10
Pasture trees[c] *(n = 21)*					
Maximum[d]	2.33	−1.58	2.40	−5.11	2.68
$Mean_{95}$[d]	2.84	−2.30	2.46	−5.64	2.35

Units are in meters.

[a] Error is calculated as: lidar (footnote d) − field (footnote c) height metric (deviation from 1:1 relationship).

[b] MAE is mean absolute error.

[c] Field metric is the maximum height in an individual crown (old-growth emergent or isolated pasture trees).

[d] Lidar metrics are: Maximum = maximum value of DCM cells; $Mean_{95}$ = mean of cells above 95% quantile.

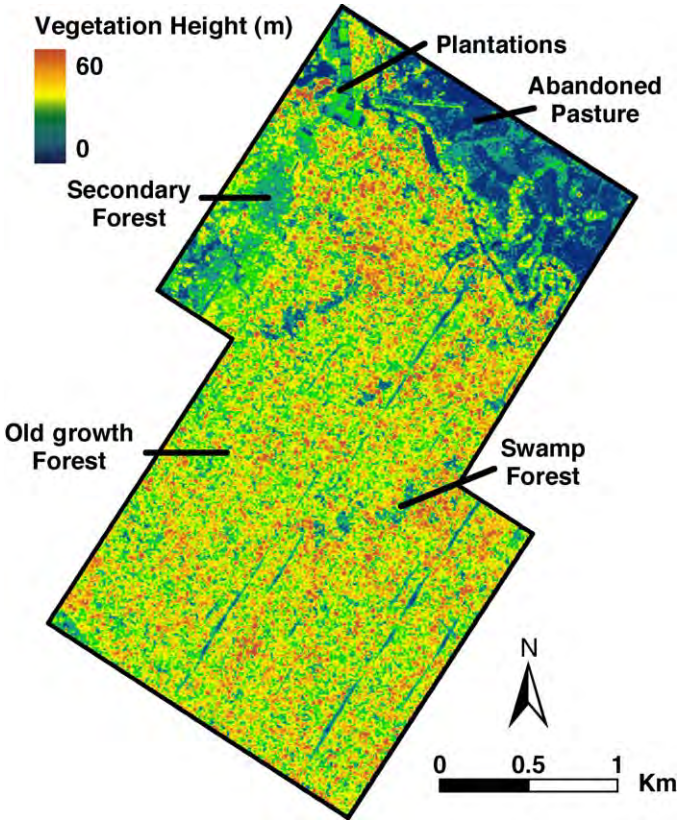

Fig. 8. Landscape-scale overview of the final digital canopy model.

than on slopes 10–20° and >20° (Fig. 6, lines connecting classes; $p < 0.004$ both comparisons). Under old-growth forest, RMS error was lowest on slopes $\leq 10°$ (2.21 m, 0–3° and 3–10° combined) and highest on slopes greater than 20° (3.09 m; Fig. 6), indicating that very steep slopes had the largest overestimation bias and error variance (Isaaks & Srivastava, 1989).

We next assessed DTM elevation error relative to land use. The slope error analysis indicated that slopes $\leq 10°$ had statistically similar DTM errors, so we limited land-use analyses to survey points on slopes $\leq 10°$ to avoid the confounding effects of slope error.

For all land-use classes on slopes less than 10°, the overall RMSE for the DTM was 1.72 m (Table 4). There was a highly significant difference in mean absolute error between the land-use classes ($p < 0.0001$). The classes with relatively few trees or shrubs, Developed Areas and Abandoned Pastures, had a slight elevation underestimation (mean-signed error) of -0.58 and -0.28 m, respectively (Fig. 7; Table 4). As expected, DTM elevation was overestimated (positive mean-signed error) under forest canopies (Fig. 7; Table 4). DTM error was most severe in old-growth forests, which had extremely dense, multi-layered canopies. Old-growth forests had significantly higher MAE (pair-comparisons with other classes, $p < 0.05$), the largest mean-signed error (+1.01-m overestimation), and highest RMSE (1.95 m) relative to all other classes (Fig 7; Table 4). In contrast, the classes with very little to no canopy cover,

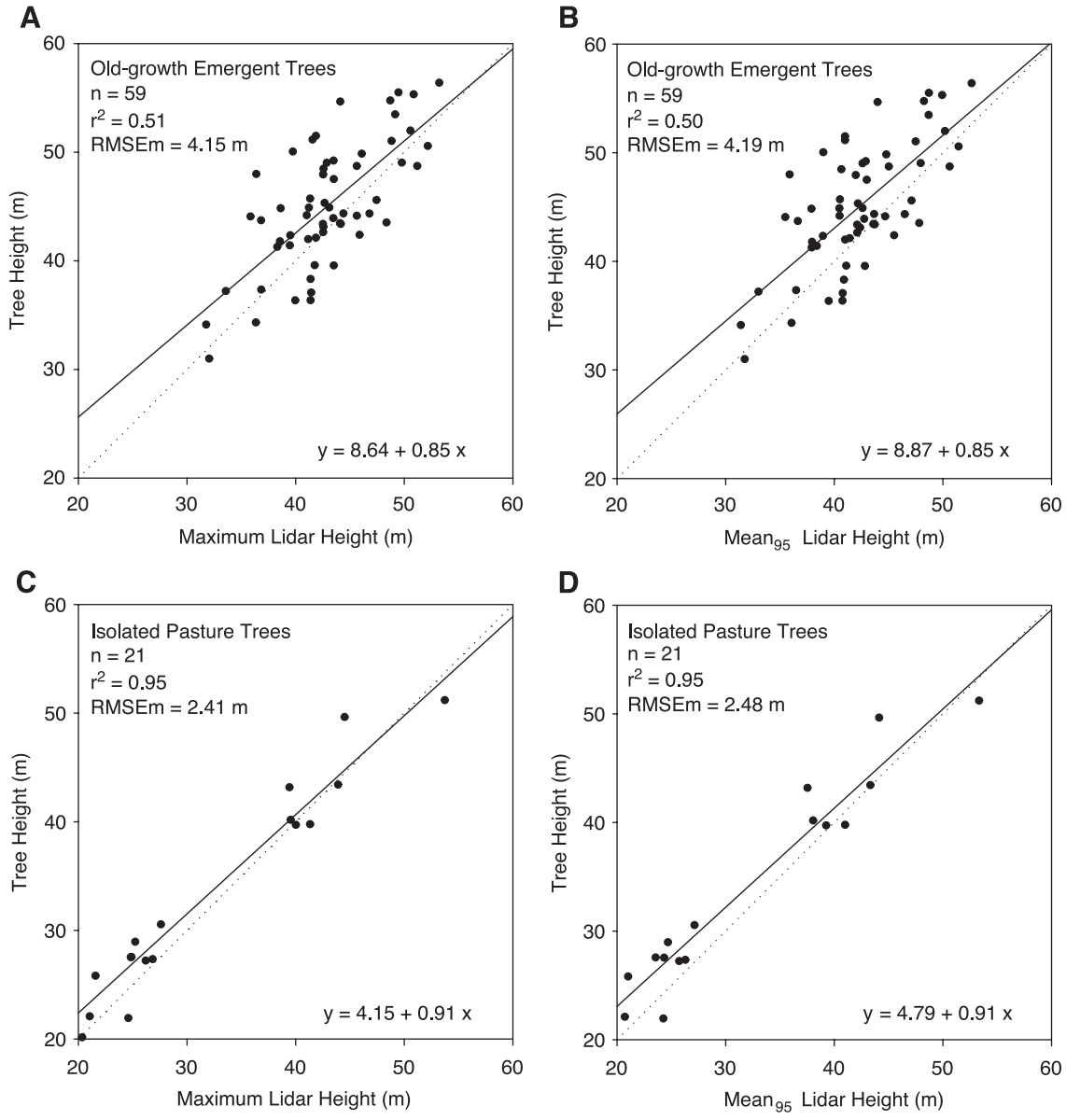

Fig. 9. Individual-tree height regression models for old-growth emergent trees (A and B) and isolated pasture trees (C and D). Reference data are the maximum height within a tree's crown. Lidar metrics include Maximum (A and C) or $Mean_{95}$ (B and D). Solid lines are the fitted models and dotted lines are the 1:1-relationships. RMSEm is the root-mean-square error of the model (i.e., cross-validation prediction residuals).

Developed Areas and Abandoned Pastures, had the lowest RMS errors (1.02 and 1.10 m, respectively). Interestingly, classes with continuous canopies of relatively low tree height, Secondary forests and Agroforestry Plantations had statistically similar absolute errors (Fig 7, lines).

3.2. Landscape view of canopy-surface heights

Differences in land use depicted in Fig. 1 are clearly observable as mean height and textural variations in the DCM (Fig. 8). As expected, mean canopy-surface height is greatest for old-growth, swamp and selectively logged forests, which have relatively low human disturbance (Table 5). Several linear strips of low elevation running southwest to northeast indicate areas where no lidar footprints were recorded due to gaps in the flight-line (Fig. 8). These areas were avoided in tree-height analyses.

3.3. Individual-tree height estimation

In estimating old-growth tree height, the Maximum lidar metric had a significantly lower mean absolute error than the $Mean_{95}$ lidar metric (Table 6; $p < 0.001$, t-test). The relative MAE error for Maximum and $Mean_{95}$, respectively, was 8.1% and 8.7% of the mean height for individual emergent trees measured in the field (Tables 2 and 6). Mean-signed error was negative, indicating that the DCM metrics underestimate actual emergent tree heights in old-growth forest. Linear-regression models for predicting old-growth tree height explained 51% and 50% of the variance (RMSEm of 4.15 and 4.19 m) for Maximum and $Mean_{95}$, respectively (Fig. 9A and B). Lidar metric underestimation can be seen as points clustering above the 1:1-relationship line in the scatter plot (Fig. 9A and B).

As with old-growth heights, pasture tree heights were underestimated by the DCM metrics (Table 6, negative mean-signed error). However, DCM height estimates of pasture trees had a lower MAE, with 2.33 and 2.84 m for Maximum and $Mean_{95}$, respectively (Table 6). Mean absolute errors using Maximum and $Mean_{95}$ were 7.4% and 9.0%, respectively, of the mean field height of pasture trees (Tables 2 and 6). For pasture trees, the Maximum lidar metric had a significantly lower mean absolute error than the $Mean_{95}$ lidar metric (Table 6; $p < 0.001$, t-test). Height errors were also less variable than those of old-growth trees (Table 6, standard deviations; Fig. 9, 1:1-relationship line). For pasture trees, the linear models relating lidar-derived to field-derived height explained 95% of the variance (using either lidar metric) and relationships were stronger than those observed for old-growth tree heights (Fig. 9C and D vs. Fig. 9A and B). Model RMS errors for pasture trees (2.41 and 2.48 m for Maximum and $Mean_{95}$, respectively) were roughly half of those expected in applying old-growth height models (Fig. 9).

In summary, direct estimation of individual tree height from the DCM was more accurate (i.e., less absolute error) when using the Maximum lidar metric, which calculated height as the maximum DCM cell within 5-m radius of the crown centroid. Furthermore, pasture height estimates using either DCM metric had less absolute error and stronger linear relationships with field measurements than those from old-growth forest (Table 6; Fig. 9A and C).

3.4. Plot-scale height estimation

In terms of mean-signed error, MAE and standard deviation of errors, the $Mean_{ALL}$ lidar metric out-performed the $Mean_{2\times2}$ method in estimating plot mean height, calculated from all or tree-only stems (Table 7). Considering tree-only stems, mean-signed error was slightly negative (-0.36 m) when using the mean of all DCM cells in the plot (i.e., $Mean_{ALL}$ metric), indicating a slight underestimation of plot mean height. In contrast, we found that the $Mean_{2\times2}$ lidar metric overestimated mean stem height ($+1.55$ mean-signed error). Mean absolute error was significantly less for $Mean_{ALL}$ relative to $Mean_{2\times2}$ when considering the plot mean height of all or tree-only stems ($p < 0.0001$, paired t-tests).

All plot-scale regression models were highly significant ($p < 0.001$). Using the $Mean_{ALL}$ lidar metric, the mean height of tree stems in plantation plots was predicted with a model r^2 of 0.97 and RMSEm of 1.08 m (Fig. 10B—trees-only plots), whereas this relationship dropped to an r^2 of 0.84 and RMSEm rose to 2.26 m when predicting mean height for all canopy and sub-canopy stems in plots (Fig. 10D—all stem plots). A similar pattern was observed in using the $Mean_{2\times2}$ lidar metric (Fig. 10A and C). Height was overestimated for the seven all-stem plots, and so those points fall below the regression line (Fig. 10C and D—open circles) and decrease the overall strength of the models. In general, the best plot-scale relationship was between the $Mean_{ALL}$ lidar metric and the trees-only field metric. Errors between this combination of lidar and field metrics had the lowest 1:1-relationship bias (-0.36-m mean-signed

Table 7
Estimation error of plot-mean stem heights[a]

	MAE[b]	Mean	S.D.	Min	Max
Tree stems[c] *(n = 32)*					
$Mean_{2\times2}$[d]	1.74	+1.55	1.55	−0.52	4.92
$Mean_{ALL}$[d]	0.90	−0.36	1.09	−3.69	1.60
All stems[c] *(n = 32)*					
$Mean_{2\times2}$[d]	2.82	+2.63	2.78	−0.52	8.08
$Mean_{ALL}$[d]	1.64	+0.73	2.32	−3.69	5.91

Units are in meters.

[a] Error is calculated as: lidar (footnote d) − field (footnote c) height metric (deviation from 1:1 relationship).

[b] MAE is mean absolute error.

[c] Field metrics are mean of individual stem heights for tree species only (Tree stems) or trees and other species (All stems).

[d] Lidar metrics are: $Mean_{2\times2}$ = mean of maximum DCM values in a 2 × 2-m grid overlaid on plot; $Mean_{ALL}$ = mean of all DCM values in plot.

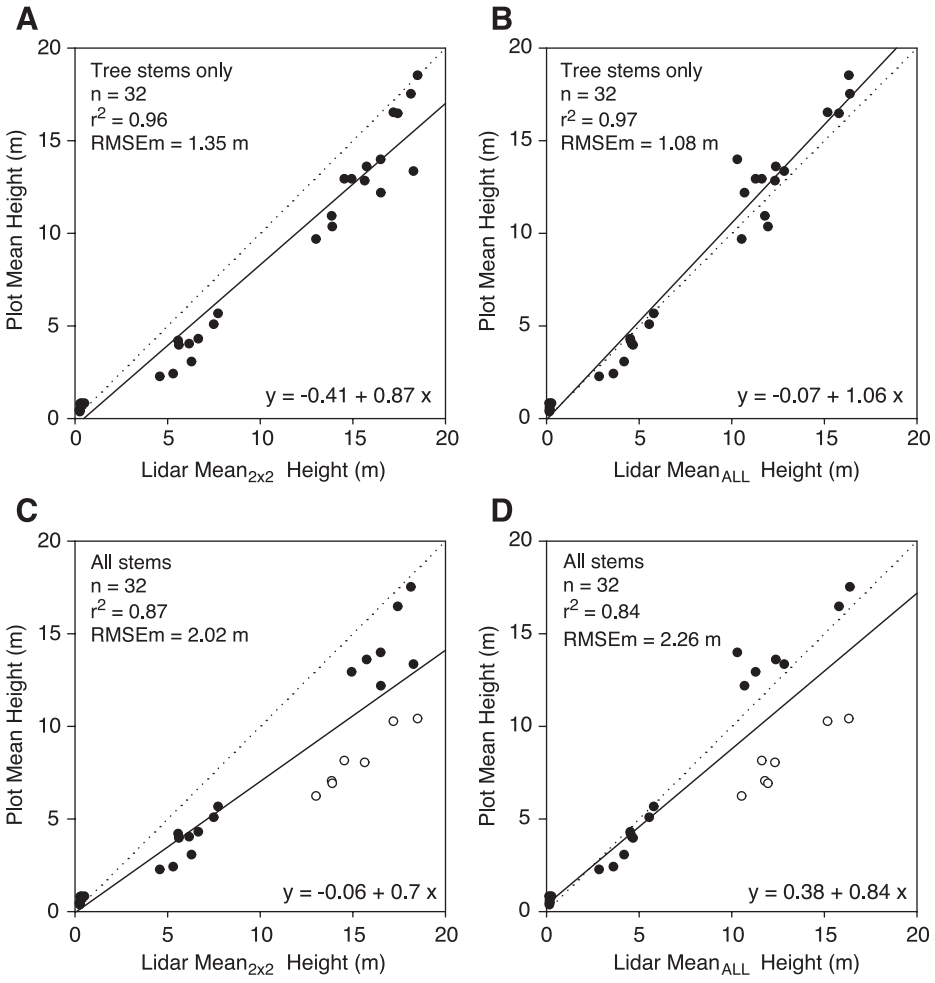

Fig. 10. Plot mean height regression models. Reference data are the plot mean of stem heights computed from only tree stems (A and B; filled circles) or from trees, sub-canopy palms and shrub stems (C and D; open circles). Plot-scale lidar metrics include $Mean_{2\times2}$ (A and C) or $Mean_{ALL}$ (B and D). Solid lines are the fitted models and dotted lines are the 1:1-relationships. RMSEm is the root-mean-square error of the model (i.e., cross-validation prediction residuals).

error), strongest linear relationship ($r^2 = 0.97$) and lowest model RMSE (1.08 m).

4. Discussion

4.1. Lidar ground retrieval in a tropical landscape

The overall 2.29-m RMSE of the DTM generated using an iterative ground-retrieval and ordinary kriging interpolation scheme (Table 3) fell short of the decimeter accuracies reported for other laser sensors (Cobby et al., 2001; Huising & Gomes Pereira, 1998). However, those accuracies were achieved only under ideal conditions favoring retrieval of ground points, such as in areas with relatively flat terrain, without vegetation or with deciduous vegetation in a leaf-off state. In our tropical wet forest landscape, areas exhibiting these characteristics are rare. The "bald Earth" geomorphology in the study area is not flat due to its volcanic history and leaf-on canopy persists throughout the entire year. In comparing areas of similar slope ($< 10°$), our DTM error is 0.79 m greater than the error (1.72 vs. 0.93-m RMSE) observed by Hodgson et al. (2003) when analyzing a lidar-derived DTM from a temperate-zone, deciduous (leaf-on) landscape. Our study used a fully automated technique to retrieve ground *xyz* points from a DSM, which was interpolated from the original lidar footprints. In contrast, Hodgson et al. used proprietary software to identify ground-return footprints and then used a visual-interpretation step to remove spurious *xyz* points prior to DTM interpolation. Due to our first-return lidar data, DSM-interpolation smoothing, fully automated ground retrieval, and extremely dense vegetation, we would expect our study DTM to have relatively greater overall error than the Hodgson et al. study. Despite the many challenges in using lidar over a TRF landscape, our observed DTM error is well below the USGS maximum-permitted RMSE of 7 m for Level-3 DTMs, the highest-quality nationwide products publicly offered in the United States (Hodgson et al., 2003); and furthermore, this lidar-generated DTM is a vast improvement over previously available topographic data at this important tropical research site.

We found that DTM accuracy followed a gradient in human-disturbance intensity. The DTM had the greatest RMS error (1.95 m) in old-growth forests. The structure of these forests is composed of multi-layered leaves and branches maintained by a fine-scale, tree-fall disturbance regime. In effect, this structure is a dense media which acts to absorb or multi-scatter photons of near-infrared laser light, thereby reducing the probability that photons reach the ground and return to the sensor; consequentially, our ground-retrieval algorithm had a greater chance of mistaking sub-canopy returns for ground returns in old-growth forests, and the accuracy of the interpolated DTM decreased. In contrast, DTM accuracy was 0.93-m greater in developed areas of the reserve (1.02-m RMSE), which had scattered overstory trees and shrubs, mowed grass and cobble roads. Even under these more ideal ground-retrieval conditions, the algorithm failed to entirely separate ground from overlying vegetation height. When we considered just those survey points in developed areas with mowed grass or roads, overstory vegetation >3-m away, and on flat terrain, RMSE was 0.58 m and mean-signed-error was -0.49 ± 31 S.D. m ($n = 20$). This error can be considered the elevation error related to sensor artifacts when using first-return laser pulses, i.e., footprint *xyz* positioning, scan angle (Baltsavias, 1999b; Huising & Gomes Pereira, 1998). A last-return, small-footprint system would likely provide better ground-retrieval performance in this densely vegetated landscape. Indeed, last-return systems are generally deployed for DTM-generating campaigns (Hodgson et al., 2003; Lefsky et al., 2002) because there is a higher probability that laser returns will come from ground reflections. Flying a lidar sensor with higher sample density (i.e., more frequent postings) is also expected to increase the chances of detecting the ground (Hodgson et al., 2003).

In structurally complex old-growth forest, we found vertical errors to increase 0.67 m on the steepest slopes relative to flattest areas (Fig. 6, >20° vs. 0–3° class RMSE). In comparison, Hodgson et al. (2003) found a 2-m greater RMSE on flat vs. steep slopes under structurally complex, leaf-on vegetation (0–2° vs. 6–8° classes). Observed DTM vertical error has both horizontal and vertical error components (Hodgson et al., 2003), such as footprint positioning and instrument range errors, respectively. As mentioned above, these factors combined may contribute 0.58 m to DTM elevation error. On steep slopes, vertical errors between DSM-derived ground points translate into large DTM interpolation errors relative to flat areas, which have more points of similar elevation that act to down-weight the impact of a spurious point during interpolation. Also, most DTM error analyses assume that errors in reference data are negligible. In this study, field-survey data suffer from horizontal, planimetric errors (~ 0.57 m, Section 2.2) due to the transformation between the local and UTM WGS-84 coordinate and datum systems. On the steepest slopes [44° (Table 1)], planimetric error could introduce up to 0.40-m vertical error in reference data. When we considered only reference points on slopes $\leq 10°$, RMS error decreased from 2.29 to 1.72 m (all land use). This 0.57-m difference in DTM error can easily be explained by greater reference error combined with instrument-related vertical error on steep slopes. However, our data do not permit us to rigorously assess these differences in lidar- and reference-related vertical errors.

Compared to previously published accuracies achieved with the large-footprint LVIS sensor at our tropical study site (Hofton et al., 2002), FLI-MAP elevation estimation had better overall performance across the entire landscape (5.64-m [LVIS] vs. 2.29-m [FLI-MAP] RMSE). Also, the DTM from this study performed better than LVIS on flatter areas under the range of vegetation conditions in the study area (1.72-m [FLI-MAP, slopes $\leq 10°$] RMSE vs. 2.42-m [LVIS, slopes $\leq 3°$]). In the LVIS study, the estimated elevation at centroids of 25-m diameter footprints were compared directly to the nearest reference point, and so steep slopes compounded the vertical error associated with planimetric discrepancies between centroids and survey points. In contrast, the FLI-MAP DSM had a 0.33-m spatial support that permitted more precise planimetric location of ground-retrieved samples. In addition, fine spatial support provided denser ground sampling relative to the LVIS footprints. Due to these benefits of using a small-footprint sensor, our observed elevation error was relatively low compared to the Hofton et al. (2002) study, even on the steepest slopes and under the densest vegetation in the landscape (2.42 to 3.09-m RMSE from flat to steep slopes in old-growth forest).

As was found by Lloyd and Atkinson (2002b), ordinary kriging was a better DTM interpolator of lidar ground-retrieved samples than the more conventional IDW technique. Although the mean-signed error was smaller for IDW, OK interpolations had smaller mean absolute errors, lower RMS errors, and smaller error variance. The variance-dampening effect afforded by OK is desirable because it tends to reduce the impact of spurious sub-canopy vegetation samples that have inadvertently passed through the vegetation filter. There are other variants of kriging that could be useful to DTM interpolation of lidar data. For example, co-kriging (Goovaerts, 1997) the ground samples with a co-varying variable, such as land-use type, optical reflectance data or a textural variable computed from the original lidar data, could provide a robust means of adjusting the data covariance weighting to various landscape units.

4.2. Lidar estimation of tropical vegetation height

Individual tree heights were underestimated by the lidar metrics (Table 6, negative mean-signed error). These results concur with those by Brandtberg et al. (2003), Gaveau and Hill (2003), and Persson et al. (2002). The FLI-MAP sensor used in this study detected first-return signals from densely sampled small footprints, and so it is unlikely that this

underestimation error is entirely from the laser missing the upper-most reaches of the crown (Næsset & Økland, 2002). Based on recent findings by Gaveau and Hill for leaf-on hardwood trees, we believe underestimation is partly caused by laser pulses penetrating below crown surfaces until inner-crown materials reflect a detectable first-return signal.

Some of the discrepancy between lidar and field estimates is also related to DTM error. In abandoned pastures, the DTM tends to slightly underestimate elevation by 0.28 m on average and RMS error was 1.10 m (Table 4). In contrast, the DTM error under old-growth forests was significantly greater. The DTM was overestimated by 1.01 m on average and RMS error increased to 1.95 m (Table 4). In abandoned pastures, broad areas without tree canopy permitted more laser energy to reach the ground and return to the sensor, resulting in better DTM accuracy. In old-growth forest, dense canopy causes DTM overestimation that effectively "clips" trees at their base when the DCM is calculated. In terms of RMSE, DTM-related error can account for up to 47% and 42% of the model error observed for old-growth and pasture trees, respectively (Fig. 10, Maximum). We found that DTM error increases on steep slopes in old-growth forests; however, slope was not linearly related to the absolute difference between lidar- and field-derived old-growth tree heights [Maximum metric (Fig. 9A), $r^2 = 0.003$; $p = 0.68$]. We therefore conclude that slope-related effects on tree height estimates are not severe.

Another source of random error in the predictive models is from field reference data (Brandtberg et al., 2003; Gaveau & Hill, 2003; Persson et al., 2002). In our study, field measurements of individual tree heights were taken 3 years after the lidar mission, and many trees could have grown higher in that time. This growth would result in the field estimates being higher than the lidar measurements, thus adding to the underestimation bias seen here. Although our field data included only trees with precise measurements (i.e., measurement S.D. ≤ 1 m), high precision does not ensure absolute accuracy. In this study, we found that directly measured stem height was overestimated by the laser range-finder. Range-finder estimation precision was between 0.14 and 1.67 m (Section 2.3.3, S.D. of errors). Again, this field measurement error would create an additional source of apparent underestimation in lidar-derived estimates and may account for up to 69% and 40% of model RMSE for pasture and old-growth tree heights, respectively (Fig. 9, Maximum).

Given the difficulties in estimating elevation and field stem heights in old-growth forest, it is not surprising that regression models had a stronger linear relationship for pasture trees than old-growth trees (r^2 0.95 vs. 0.51; Fig. 9A and C, Maximum;). The strength of the relationship for pasture trees is greater than that observed by Brandtberg et al. (2003) in estimating deciduous trees in leaf-off conditions (r^2 0.69), and the same as that observed by Gaveau and Hill (2003) in estimating leaf-on hardwood trees and shrubs (r^2 0.95). All studies used the first-returns from small-footprint lidar data to estimate vegetation height and observed an underestimation for tall trees. Our study and the Brandtberg et al. study used the local-maxima of samples from within a crown, while Gaveau and Hill compared DCM cells directly to field points measured over crowns with high *xyz*-coordinate precision. The two-return lidar data used by the other two studies, as well as the leaf-off conditions in the Brandtberg et al. study, likely permitted better ground retrieval and higher DTM accuracy relative to that which can be expected with a first-return lidar flown over an evergreen TRF landscape. Relatively high DTM error in old-growth forests likely explains why old-growth TRF tree heights were not as reliably predicted as those of tropical pasture trees (this study) and temperate-zone hardwood trees (Gaveau & Hill, 2003). Despite the limitations of our elevation and field-height estimates, our pasture tree calibration model was surprisingly strong relative to both the Brandtberg et al. and Gaveau and Hill models. Hardwood trees in leaf-on conditions were underestimated by -1.58 m and -2.12 (mean-signed error) in this study and that of Gaveau and Hill, respectively. Exposed leaves in hardwood trees favor the detection of the crown surface because they readily transmit and multiple-scatter near-infrared laser light (Grant, 1987). In contrast, the Brandtberg et al. model had weaker accuracy because leaf-off crowns expose only twigs and branches to the sensor and allow greater laser penetration into the canopy before returning a signal.

Regression models for estimating individual conifer tree height had RMS errors of 0.23 and 0.63 m (r^2 0.75 and $\sim$ 0.98) in the works of Næsset and Økland (2002) and Persson et al. (2002), respectively. In this study, we observed a 2.41-m model RMSE for pasture trees (r^2 0.95; Fig 9C, Maximum) and Gaveau and Hill (2003) observed a 1.98-m model RMSE for deciduous hardwoods and shrubs (r^2 0.95). The two conifer studies used first and last return, small-footprint lidar data. Assuming DTM error is equal between the studies, it thus appears that lidar calibration models (i.e., linear-regression models) have a lower RMS error in estimating conifer tree heights than they do for leaf-on, hardwood trees. However, more comparative studies between tropical and temperate species, tree conditions (i.e., stress, senescence), lidar sensors and associated analytical methods are needed to confirm this accuracy assessment.

The linear correlation between lidar and reference measurements of plot mean height of tree stems was very strong (Fig. 10B). The 0.97 r^2 exceeded the 0.38–0.49, 0.91, 0.82–0.95, and 0.91 r^2 values reported for conifer-dominated stands by Magnussen and Boudewyn (1998), Næsset (1997, 2002), and Næsset and Økland (2002), respectively, and the 0.68 r^2 value for leaf-on hardwood stands (Lim et al., 2003) As was found in this and other lidar studies (Lim et al., 2003; Magnussen & Boudewyn, 1998; Næsset, 1997), mean tree height at the plot scale was slightly underestimated by the Mean_{ALL} lidar metric (Table 7). In this study, this underestimation (negative mean-signed error) was 0.36

m, while it was 2.1 to 4.1 m and 0.70 m in the Næsset and Magnussen and Boudewyn studies, respectively. At the stand scale, conifer underestimation likely results because too few footprints detect the upper-most twigs and branches of conical trees (Magnussen & Boudewyn, 1998). In our study, laser penetration into the hardwood tree canopy is likely an issue (Gaveau & Hill, 2003), and square DSM cells may not always record the height of the highest surface material within the cell—an underestimation that would persist in DCM heights. Furthermore, our field metric considered the mean heights of tree stems (i.e., maximum height of individual crowns), whereas the $Mean_{ALL}$ lidar metric averaged all canopy-surface heights (i.e., DCM cells) in the plot, not just DCM crown-maxima, and this lidar measurement is expected to be low relative to the field measurement (Magnussen & Boudewyn, 1998).

Although the $Mean_{2\times2}$ metric was designed to isolate DCM crown-maxima (i.e., stem heights of trees spaced 2 m apart), the metric overestimated mean tree height by an average of 1.55 m (Table 7). These results are contrary to those of Næsset (1997), who found that a local-maxima grid-overlay metric reduced mean-signed error to insignificant levels. However, the Næsset and other "grid-overlay" studies have used last-return lidar systems with relatively sparse sampling densities, properties which combined lower the probability of recording a return from the upper-most canopy. This problem may have been compounded by the geometry of conifer crowns in the previous studies (Magnussen & Boudewyn, 1998; Næsset, 1997, 2002). In contrast to TRF tree crowns, which are broad and generally hemispherical in shape, conical crowns expose much less upper-most canopy material to a lidar sensor (Magnussen & Boudewyn, 1998). A grid-overlay lidar metric can compensate for missing the crown by selecting only the highest heights in the dataset. In the case of FLI-MAP, which has both a high sampling density and first-returns, there is less chance of missing the upper-canopy, especially over these agroforestry plots that contained even-aged stands with a relatively flat upper-canopy layer of broad leaves. In this case, the grid-overlay metric needlessly over-compensates the plot-height estimate upward relative to the reference height.

Given the dense upper-canopy layer in the older plots, FLI-MAP first-return signals were not expected to be sensitive to sub-canopy stem heights. This sub-canopy insensitivity was revealed clearly in linear-regression analyses. When considering all stems in plots, regression models had weaker relationships and higher model RMS errors than those built when considering just tree-stem heights (Fig. 10). The averaging of canopy and sub-canopy stem heights from the field lowered the mean height for seven of the reference plots, and since FLI-MAP was sensitive mainly to the upper-canopy trees, the lidar metric estimation was high relative to the field measurement.

As was found by Næsset and Økland (2002) for conifer trees, the estimation of height for individual hardwood trees (Table 6) was not as accurate as that found for plot-scale estimation (Table 7). This finding is not surprising, because at the plot scale, both lidar and field measurement errors are minimized by averaging many co-located lidar and reference data points.

5. Conclusions

In this study, we found that a small-footprint, first-return lidar sensor can be used to predict sub-canopy elevation with 2.29-m accuracy in a tropical landscape. The accuracy of the elevation surface was significantly affected by vegetation cover, with largest errors detected in areas with structurally complex old-growth forests, especially on steep slopes. The digital terrain model that resulted from this analysis had a 1-m spatial support, which makes it an ideal input for other ecological or management applications such as modeling of inundation zones, hill-slope processes, and habitat associations.

We showed that small-footprint lidar systems have potential for the estimation of individual tree height of tropical, hardwood species in leaf-on conditions. For old-growth forest emergent trees and isolated abandoned-pasture trees greater than 20 m tall, individual tree heights could be estimated directly from a lidar canopy-height surface (i.e., DCM) with mean absolute errors that were 8.1% and 7.4% of mean field heights, respectively. Models for individual old-growth and pasture trees explained 59% and 95% of the variance, with model RMS errors of 2.41 and 4.15 m, respectively; however, as was found in other studies (Brandtberg et al., 2003; Persson et al., 2002), it was difficult to separate lidar-related error from reference measurement error.

Plot-scale, mean stem height for trees within plantation stands was estimated from the DCM with much greater accuracy than for individual tree heights. The best plot-scale models explained 87% of the variance when estimating the mean height of all canopy and sub-canopy stems in plots, while 97% of the variance when considering just canopy-tree stems. These results are encouraging in that a stand (i.e., a plot in this study) is often an important scale for many broad-scale studies and management decisions. Given that stand height has an allometric relation to tree diameter and biomass, it is expected that fine-scale, lidar-derived DCMs can be used to predict carbon stocks and forest structure parameters (Lim et al., 2003; Popescu et al., 2003). Such a strong relationship to plantation and old-growth forest structure has been found for a large-footprint, waveform lidar system at this study site (Drake et al., 2002a,b). However, a small-footprint sensor provides a fine-scale perspective on forest structure that may be more appropriate for certain kinds of ecological research, such as tracking tree-fall gap dynamics, identifying wildlife habitat, or scaling point-based data (e.g., tower eddy-flux carbon

measurements) to a broader scale. Future research will assess whether the stand-scale findings for structurally simple plantations can be extended to the estimation of stand-scale biomass in plantations as well as in structurally complex, old-growth forest.

The fine-scale DCM from this study is also an important layer for fusion with other types of remote sensing data. For example, Hudak et al. (2002) used regression and geostatistical techniques to scale estimated canopy height from a small-footprint sensor with limited spatial extent, to a much larger, continuous spatial extent using relatively inexpensive Landsat ETM+ multispectral data. Furthermore, textural measures from the DCM surface can aid image segmentation and/or enhance class statistical separation in image classification schemes (e.g., Raber et al., 2002). The fine-scale DCM from this study is also an excellent data source for the delineation of individual tree crowns for forest inventory and species identification applications (e.g., Brandtberg et al., 2003; Holmgren, 2003). In conjunction with high spatial resolution hyperspectral imagery, future research will investigate the utility of DCM-derived texture and crown morphology information in the identification of individual tree species within old-growth tropical rain forest landscapes.

Acknowledgements

This work was supported by NASA Headquarters under the Earth System Science Fellowship Grant NGT5-30436. The authors would like to thank Leonel Campos, William Miranda and students from Organization for Tropical Studies (OTS) graduate tropical ecology course 00-1 for their invaluable field assistance. Old-growth tree height and slope inclination data used in this study was based upon work supported by the National Science Foundation under Grant DEB-0129038. Plot stem-height measurements from agroforestry plantations were provided by Dr. Jack Ewel, whose work is supported by National Science Foundation award DEB-9975235 and by the Andrew W. Mellon Foundation. FLI-MAP data was graciously donated to OTS by the U.S. Army Corps of Engineers Topographic Engineering Center. OTS provided logistic support for the lidar data acquisition and support for the field component of this research. We would also like to thank Dr. Jason Drake and two anonymous reviewers for their useful comments that greatly improved this manuscript.

References

Baltsavias, E. P. (1999a). Airborne laser scanning: Basic relations and formulas. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*, 199–214.

Baltsavias, E. P. (1999b). Airborne laser scanning: Existing systems and firms and other resources. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*, 164–198.

Bartier, P. M., & Keller, C. P. (1996). Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). *Computers & Geosciences*, *22*(7), 795–799.

Bates, P. D., & De Roo, A. P. J. (2000). A simple raster-based model for flood inundation simulation. *Journal of Hydrology*, *236*(1–2), 54–77.

Birnbaum, P. (2001). Canopy surface topography in a French Guiana forest and the folded forest theory. *Plant Ecology*, *153*(1–2), 293–300.

Blair, J. B., & Hofton, M. A. (1999). Modeling laser altimeter return waveforms over complex vegetation using high-resolution elevation data. *Geophysical Research Letters*, *26*(16), 2509–2512.

Blair, J. B., Rabine, D. L., & Hofton, M. A. (1999). The laser vegetation imaging sensor: A medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*, 115–122.

Brandtberg, T., Warner, T. A., Landenberger, R. E., & McGraw, J. B. (2003). Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America. *Remote Sensing of Environment*, *85*(3), 290–303.

Brown, I. F., Martinelli, L. A., Thomas, W. W., Moreira, M. Z., Ferreira, C. A. C., & Victoria, R. A. (1995). Uncertainty in the biomass of Amazonian forests—an example from Rondonia, Brazil. *Forest Ecology and Management*, *75*(1–3), 175–189.

Clark, D. B., & Clark, D. A. (1996). Abundance, growth and mortality of very large trees in neotropical lowland rain forest. *Forest Ecology and Management*, *80*(1–3), 235–244.

Clark, D. B., Clark, D. A., Rich, P. M., Weiss, S., & Oberbauer, S. F. (1996). Landscape scale evaluation of understory light and canopy structure: Methods and application in a neotropical lowland rain forest. *Canadian Journal of Forest Research*, *26*(5), 747–757.

Clark, D. B., Palmer, M., & Clark, D. A. (1999). Edaphic factors and the landscape-scale distributions of tropical rain forest trees. *Ecology*, *80*, 2662–2675.

Clark, D. B., Read, J. M., Clark, M. L., Murillo Cruz, A., Fallas Dotti, M., & Clark, D. A. (2004). Application of 1-m and 4-m resolution satellite data to studies of tree demography, stand structure and land-use classification in tropical rain forest landscapes. *Ecological Applications*, *14*(1), 61–74.

Cobby, D. M., Mason, D. C., & Davenport, I. J. (2001). Image processing of airborne scanning laser altimetry data for improved river flood modelling. *ISPRS Journal of Photogrammetry and Remote Sensing*, *56*, 121–138.

Deutsch, C. V., & Journel, A. G. (1992). *GSLIB: Geostatistical software library and user's guide*. New York: Oxford University Press.

Dixon, R. K., Brown, S., Houghton, R. A., Solomon, A. M., Trexler, M. C., & Wisniewski, J. (1994). Carbon pools and flux of global forest ecosystems. *Science*, *263*, 185–190.

Drake, J. B., Dubayah, R. O., Clark, D. B., Knox, R. G., Blair, J. B., Hofton, M. A., Chazdon, R. L., Weishampel, J. F., & Prince, S. D. (2002a). Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment*, *79*, 305–319.

Drake, J. B., Dubayah, R. O., Knox, R. G., Clark, D. B., & Blair, J. B. (2002b). Sensitivity of large-footprint lidar to canopy structure and biomass in a neotropical rainforest. *Remote Sensing of Environment*, *83*, 378–392.

Dymond, C. C., & Johnson, E. A. (2002). Mapping vegetation spatial patterns from modeled water, temperature and solar radiation gradients. *ISPRS Journal of Photogrammetry and Remote Sensing*, *57*(1–2), 69–85.

Fetcher, N., Oberbauer, S. F., & Chazdon, R. L. (1994). Physiological ecology of plants. In L. A. McDade, K. S. Bawa, H. A. Hespenheide, & G. S. Hartshorn (Eds.), *La Selva: Ecology and natural history of a neotropical rain forest* (pp. 128–141). Chicago: University of Chicago Press.

Gaveau, D. L. A., & Hill, R. A. (2003). Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. *Canadian Journal of Remote Sensing*, *29*(5), 650–657.

Gessler, P. E., Chadwick, O. A., Chamran, F., Althouse, L., & Holmes, K. (2000). Modeling soil–landscape and ecosystem properties using terrain attributes. *Soil Science Society of America Journal*, *64*(6), 2046–2056.

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. New York: Oxford University Press.

Gotway, C. A., & Cressie, N. A. C. (1990). A spatial analysis of variance applied to soil–water infiltration. *Water Resources Research*, *26*(11), 2695–2703.

Grant, L. (1987). Diffuse and specular characteristics of leaf reflectance. *Remote Sensing of Environment*, *22*, 309–322.

Harding, D. J., Lefsky, M. A., Parker, G. G., & Blair, J. B. (2001). Laser altimeter canopy height profiles: Methods and validation for closed-canopy, broadleaf forests. *Remote Sensing of Environment*, *76*, 283–297.

Hartshorn, G. S., & Hammel, B. E. (1994). Vegetation types and floristic patterns. In L. A. McDade, K. S. Bawa, H. A. Hespenheide, & G. S. Hartshorn (Eds.), *La Selva: Ecology and natural history of a neotropical rain forest* (pp. 73–89). Chicago: University of Chicago Press.

Herwitz, S. R., & Slye, R. E. (1995). 3-Dimensional modeling of canopy tree interception of wind-driven rainfall. *Journal of Hydrology*, *168* (1–4), 205–226.

Hinsley, S. A., Hill, R. A., Gaveau, D. L. A., & Bellamy, P. E. (2002). Quantifying woodland structure and habitat quality for birds using airborne laser scanning. *Functional Ecology*, *16*(6), 851–857.

Hodgson, M. E., Jensen, J. R., Schmidt, L., Schill, S., & Davis, B. (2003). An evaluation of LIDAR- and IFSAR-derived digital elevation models in leaf-on conditions with USGS Level 1 and Level 2 DEMs. *Remote Sensing of Environment*, *84*, 295–308.

Hofton, M. A., Rocchio, L. E., Blair, J. B., & Dubayah, R. (2002). Validation of vegetation canopy lidar sub-canopy topography measurements for a dense tropical forest. *Journal of Geodynamics*, *34*, 491–502.

Holmgren, J. (2003). Estimation of Forest Variables using Airborne Laser Scanning. *Doctoral Dissertation. Department of Forest Resource Management and Geomatics. Umeå, Sweden: Swedish University of Agricultural Sciences* (41 pp.) http://diss-epsilon.slu.se/archive/00000334

Hudak, A. T., Lefsky, M. A., Cohen, W. B., & Berterretche, M. (2002). Integration of lidar and Landsat ETM plus data for estimating and mapping forest canopy height. *Remote Sensing of Environment*, *82*, 397–416.

Huising, E. J., & Gomes Pereira, L. M. (1998). Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, *53*, 245–261.

Imhoff, M. L. (1995). Radar backscatter and biomass saturation-ramifications for global biomass inventory. *IEEE Transactions on Geoscience and Remote Sensing*, *33*(2), 511–518.

Isaaks, E. H., & Srivastava, R. M. (1989). *An introduction to applied geostatistics*. New York: Oxford University Press.

Kraus, K., & Pfeifer, N. (1998). Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *53*, 193–203.

Lefsky, M. A., Cohen, W. B., Parker, G. G., & Harding, D. J. (2002). Lidar remote sensing for ecosystem studies. *Bioscience*, *52*(1), 19–30.

Lim, K., Treitz, P., Baldwin, K., Morrison, I., & Green, J. (2003). Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. *Canadian Journal of Remote Sensing*, *29*(5), 658–678.

Lloyd, C. D., & Atkinson, P. M. (2002a). Non-stationary approaches for mapping terrain and assessing prediction uncertainty. *Transactions in GIS*, *6*(1), 17–30.

Lloyd, C. D., & Atkinson, P. M. (2002b). Deriving DSMs from LiDAR data with kriging. *International Journal of Remote Sensing*, *23*(12), 2519–2524.

Lohmann, P., & Koch, A. (1999). Quality assessment of laser-scanner-data. In Proceedings of the International Society of Photogrammetry and Remote Sensing. (ISPRS) workshop "Sensors and Mapping from Space 1999", Hannover, Germany, September 27–30, 1999.

Luckman, A., Baker, J., Honzák, M., & Lucas, R. (1998). Tropical forest biomass density estimation using JERS-1 SAR: Seasonal variation, confidence limits, and application to image mosaics. *Remote Sensing of Environment*, *63*, 126–139.

Magnussen, S., & Boudewyn, P. (1998). Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research*, *28*, 1016–1031.

Menalled, F. D., Kelty, M. J., & Ewel, J. J. (1998). Canopy development in tropical tree plantations: A comparison of species mixtures and monocultures. *Forest Ecology and Management*, *104*, 249–263.

Montgomery, R. A., & Chazdon, R. L. (2001). Forest structure, canopy architecture, and light transmittance in tropical wet forests. *Ecology*, *82*(10), 2707–2718.

Næsset, E. (1997). Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *52*, 49–56.

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

Næsset, E., & Økland, T. (2002). Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment*, *79*(1), 105–115.

Pannatier, Y. (1996). *VARIOWIN: Software for spatial data analysis in 2D*. New York: Springer-Verlag.

Persson, A., Holmgren, J., & Söderman, U. (2002). Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing*, *68*(9), 925–932.

Petzold, B., Reiss, P., & Stössel, W. (1999). Laser scanning—surveying and mapping agencies are using a new technique for the derivation of digital terrain models. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*, 95–104.

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

Raber, G. T., Jensen, J. R., Schill, S. R., & Schuckman, L. (2002). Creation of digital terrain models using an adaptive lidar vegetation point removal process. *Photogrammetric Engineering and Remote Sensing*, *68*(12), 1307–1315.

Rabus, B., Eineder, M., Roth, A., & Bamler, R. (2003). The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. *ISPRS Journal of Photogrammetry and Remote Sensing*, *57*(4), 241–262.

Ranson, K. J., Sun, G., Weishampel, J. F., & Knox, R. G. (1997). Forest biomass from combined ecosystem and radar backscatter modeling. *Remote Sensing of Environment*, *59*, 118–133.

Raupach, M. R. (1994). Simplified expressions for vegetation roughness length and zero-plane displacement as functions of canopy height and area index. *Boundary - Layer Meteorology*, *71*(1–2), 211–216.

Reutebuch, S. E., McGaughey, R. J., Anderson, H. E., & Carson, W. W. (2003). Accuracy of a high-resolution lidar terrain model under a conifer forest canopy. *Canadian Journal of Remote Sensing*, *29*(5), 527–535.

Ritchie, J. C., Evans, D. L., Jacobs, D., Everitt, J. H., & Weltz, M. A. (1993). Measuring canopy structure with an airborne laser altimeter. *Transactions of the ASAE*, *36*(4), 1235–1238.

Sader, S. A., Waide, R. B., Lawrence, W. T., & Joyce, A. T. (1989). Tropical forest biomass and successional age class relationships to a vegetation index derived from Landsat TM data. *Remote Sensing of Environment*, *28*, 143–156.

Sanford Jr., R. L., Pabby, P., Luvall, J. C., & Phillips, E. (1994). Climate,

geomorphology, and aquatic systems. In L. A. McDade, K. S. Bawa, H. A. Hespenheide, & G. S. Hartshorn (Eds.), *La Selva: Ecology and natural history of a neotropical rain forest* (pp. 19–33). Chicago: University of Chicago Press.

Steininger, M. K. (2000). Satellite estimation of tropical secondary forest aboveground biomass: Data from Brazil and Bolivia. *International Journal of Remote Sensing*, *21*(6), 1139–1157.

Weishampel, J. F., Blair, J. B., Knox, R. G., Dubayah, R., & Clark, D. B. (2000). Volumetric lidar return patterns from an old-growth tropical rainforest canopy. *International Journal of Remote Sensing*, *21*(2), 409–415.

Yin, Z. Y., & Wang, X. H. (1999). A cross-scale comparison of drainage basin characteristics derived from digital elevation models. *Earth Surface Processes and Landforms*, *24*(6), 557–562.