Remote Sensing of Environment 194 (2017) 77–88

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Area-based vs tree-centric approaches to mapping forest carbon in Southeast Asian forests from airborne laser scanning data

David A. Coomes [a,*,1], Michele Dalponte [a,b,1], Tommaso Jucker [a,1], Gregory P. Asner [c], Lindsay F. Banin [d], David F.R.P. Burslem [e], Simon L. Lewis [f,g], Reuben Nilus [h], Oliver L. Phillips [f], Mui-How Phua [i], Lan Qie [f]

[a] Forest Ecology and Conservation Group, Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK
[b] Department of Sustainable Agro-ecosystems and Bioresources, Research and Innovation Centre, Fondazione E. Mach, Via E. Mach 1, 38010 San Michele all'Adige (TN), Italy
[c] Department of Global Ecology, Carnegie Institution for Science, 260 Panama Street, Stanford, CA 94305 USA
[d] Centre for Ecology and Hydrology, Bush Estate, Penicuik, Midlothian, EH26 0QB
[e] School of Biological Sciences, University of Aberdeen, Cruickshank Building, St Machar Drive, Aberdeen AB24 3UU, UK
[f] School of Geography, University of Leeds, Leeds LS2 9JT, UK
[g] Department of Geography, University College London, London WC1E 6BT, UK
[h] Forest Research Centre, Sabah Forestry Department, P.O. Box 1407, 90715 Sandakan, Sabah, Malaysia
[i] Faculty of Science and Natural Resources, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia

ARTICLE INFO

Article history:
Received 24 September 2016
Received in revised form 24 January 2017
Accepted 17 March 2017
Available online 28 March 2017

Keywords:
Allometry
Aboveground carbon density
Biomass estimation
Image analysis
LiDAR
Object recognition
Power-law
Tree delineation
Tropical forests

ABSTRACT

Tropical forests are a key component of the global carbon cycle, and mapping their carbon density is essential for understanding human influences on climate and for ecosystem-service-based payments for forest protection. Discrete-return airborne laser scanning (ALS) is increasingly recognised as a high-quality technology for mapping tropical forest carbon, because it generates 3D point clouds of forest structure from which aboveground carbon density (ACD) can be estimated. Area-based models are state of the art when it comes to estimating ACD from ALS data, but discard tree-level information contained within the ALS point cloud. This paper compares area-based and tree-centric models for estimating ACD in lowland old-growth forests in Sabah, Malaysia. These forests are challenging to map because of their immense height. We compare the performance of (a) an area-based model developed by Asner and Mascaro (2014), and used primarily in the neotropics hitherto, with (b) a tree-centric approach that uses a new algorithm (*itcSegment*) to locate trees within the ALS canopy height model, measures their heights and crown widths, and calculates biomass from these dimensions. We find that Asner and Mascaro's model needed regional calibration, reflecting the distinctive structure of Southeast Asian forests. We also discover that forest basal area is closely related to canopy gap fraction measured by ALS, and use this finding to refine Asner and Mascaro's model. Finally, we show that our tree-centric approach is less accurate at estimating ACD than the best-performing area-based model (RMSE 18% vs 13%). Tree-centric modelling is appealing because it is based on summing the biomass of individual trees, but until algorithms can detect understory trees reliably and estimate biomass from crown dimensions precisely, areas-based modelling will remain the method of choice.

1. Introduction

Forests are an important component of the global carbon cycle, and their future management is key to international efforts to abate climate change. During the 1990s, about 89,000 km^2 of tropical forests were lost to agriculture each year, and a further 24,000 km^2 were degraded (Nabuurs et al., 2007). Estimates of deforestation rates vary, but somewhere in the region of 230 million hectares of forest were lost per year between 2000 and 2012 (Hansen et al., 2013). Furthermore, some 30% of tropical forests were degraded by logging and/or fire during that period (Asner et al., 2009). These changes resulted in significant releases of greenhouse gases (GHGs) to the atmosphere, constituting approximately 10% of global emissions (Baccini et al., 2012) and emphasising the significance of forests in the terrestrial carbon cycle (Pan et al., 2011). Forest conservation and restoration are increasingly recognised as critical for mitigating climate change (Agrawal et al., 2011). The climate change agreement brokered at COP21 in Paris, and signed by over 200 nations, may be significant in this respect. It is now recognised that, as well as harbouring biodiversity and supporting a billion livelihoods, tropical forests are essential for climate change abatement. Even if nations de-carbonise their energy supply chains within agreed schedules, we are unlikely to avoid 2 °C global warming

* Corresponding author.
E-mail address: dac18@cam.ac.uk (D.A. Coomes).
[1] Authors contributed equally to the writing of this paper.

http://dx.doi.org/10.1016/j.rse.2017.03.017

unless 500 million ha of degraded tropical forests are protected, and land unsuitable for agriculture is afforested (Houghton et al., 2015). Forest protection can offset emissions over the next 40 years, buying time for humanity to reduce its dependency on fossil fuels (Houghton et al., 2015).

Accurate monitoring of forest extent and carbon density is essential for these renewed efforts to protect forests, because this information is the basis of programmes to reduce emissions from deforestation and forest degradation and industrial zero-net-emissions programmes, and airborne laser scanning (ALS) is widely recognised as an essential component of these projects. Regional ALS maps of carbon density are all currently generated using "area-based" approaches (Næsset, 2002). These approaches, applied in over 70 publications (Zolkos et al., 2013), relate live-wood aboveground carbon density (ACD) estimates obtained from field plots to simple summary statistics, such as mean canopy height, derived from the ALS point cloud through statistical models (Fig. 1). These approaches for mapping structural attributes of complex multi-layered forests, as outlined by Drake et al. (2002, 2003), have since been applied to carbon mapping in several tropical regions (Englhart et al., 2011; Englhart et al., 2013; Asner et al., 2012, 2013; Vincent et al., 2012; Jubanski et al., 2013; Laurin et al., 2014; Réjou-Méchain et al., 2015). However, a well-recognised problem is that many different ALS structural metrics can be used to construct the multiple regression equations, and so these models are idiosyncratic by virtue of their local fine-tuning and cannot be applied more widely than the region in which they were calibrated. An alternative approach, advocated by Asner and Mascaro (2014), uses a simple power-law function of mean top-of-canopy height (TCH) to predict carbon density within tropical landscapes. This generalised approach has obvious advantages when it comes to mapping carbon across the tropics. Yet this power-law model hinges on the assumption that (i) forest basal area and TCH are closely related and (ii) that average (i.e., 1-ha resolution) between-plot variation in basal area-weighted wood density is low at regional scales (Asner and Mascaro, 2014; Vincent et al., 2014; Duncanson et al., 2015). Currently, however, we do not have a clear understanding of situations in which these assumptions are supported.

In response to these potential issues with area-based approaches for estimating ACD from airborne laser scanning, there is current interest in developing individual-tree-based approaches to make greater use of the 3D information contained in ALS data (Fig. 1; Eysn et al., 2015; Ferraz, Saatchi, Mallet, and Meyer, 2016; Vauhkonen et al., 2012). Advances in sensor technology and computational power have generated a proliferation of approaches for detecting individual tree crowns within discrete-return ALS point clouds – including those working with the rasterized upper surface of the ALS point cloud (e.g. Hyyppa et al., 2001; Chen et al., 2006; Yu et al., 2011), and those exploiting the entire point cloud (Morsdorf et al., 2004; Reitberger et al., 2009; Duncanson et al., 2014; Ferraz et al., 2016). There are several advantages of individual-tree-based mapping compared to area-based approaches: (i) it has a strong fundamental basis because it is conceptually similar to allometric approaches used in field-based inventories; (ii) uncertainty in the estimation model is much less dependent on plot size, allowing calibration using individual trees and small plots (Dalponte and Coomes, 2016); (iii) narrow patches of forest with high conservation value, such as riparian strips, can be mapped; (iv) growth and death of individual trees can, in principle, be tracked and this information used to

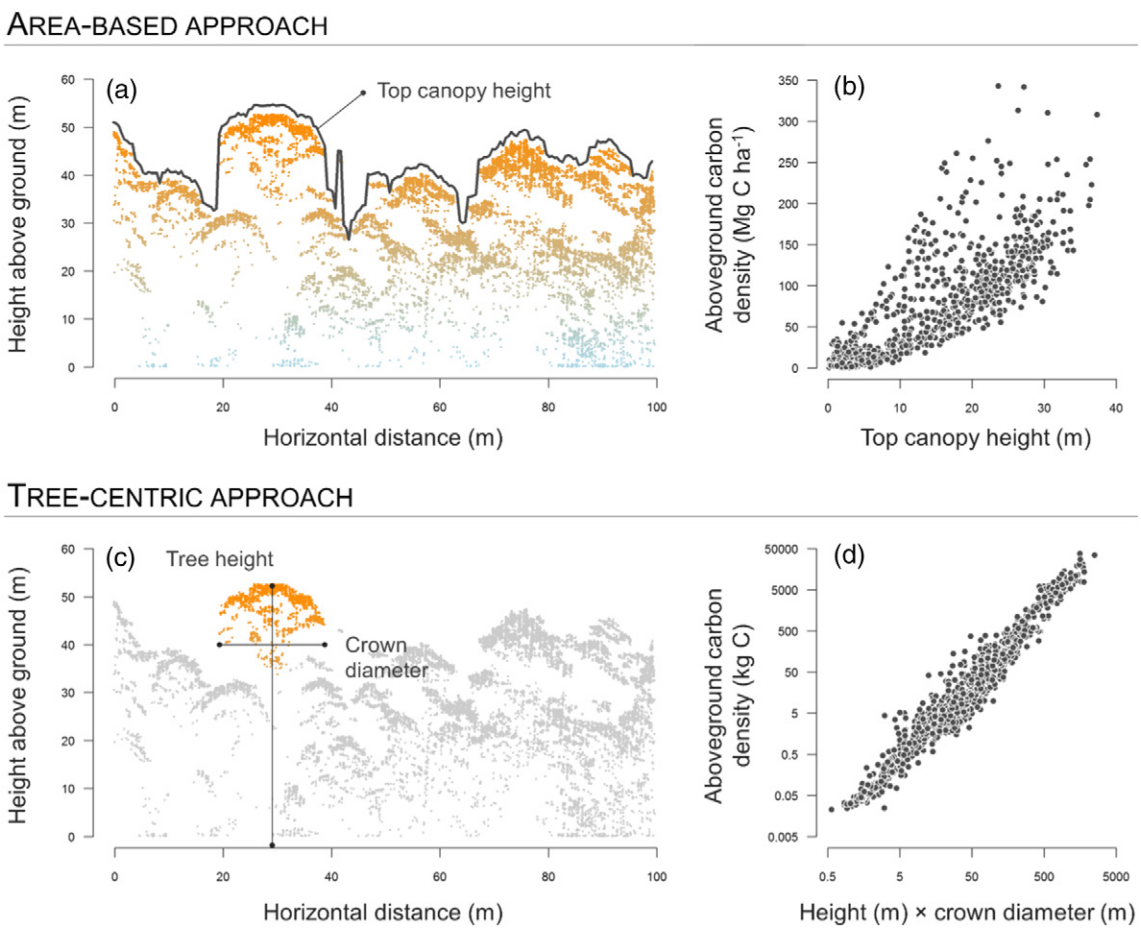


Fig. 1. Schematic diagram illustrating the key differences between area-based and tree-centric approaches used to estimate aboveground carbon density (ACD) from ALS data. Area-based approaches rely on summary statistics calculated from the ALS point cloud (e.g., top canopy height in a) to develop statistical relationships for estimating ACD (b). In contrast, tree-centric mapping aims to identify and measure the crown dimensions of individual trees within the ALS point cloud (c), and then use these to estimate their ACD (d). Data shown in panels (b) and (d) were taken from Asner and Mascaro (2014) and Jucker et al. (2016), respectively.

parameterise individual based models of forest dynamics (Shugart et al., 2015). The main disadvantages are that tree-centric approaches are computationally intensive, most delineation methods can only distinguish trees in the upper canopy, and over- or under-segmentation of trees can lead to biases in biomass estimation. Nevertheless, if these issues can be resolved, individual based modelling could mark a fundamental shift in the way forests are monitored remotely (Shugart et al., 2015). To our knowledge, only one group has used tree-centric approaches and airborne ALS data to map carbon in dense tropical forest, where obscured trees and over/under segmentation may result in substantial bias (Ferraz et al., 2016).

This paper reports on a generalised tree-centric approach for mapping the carbon density of tropical forests – i.e. one where a single model can be applied reliably across contrasting forest types. The work focusses on old-growth forests in the Sepilok Forest Reserve in lowland Sabah, Malaysian Borneo. Sabah has lost much of its lowland forests in recent decades (Gaveau et al., 2014), but has become an important testbed for global initiatives to protect and restore forests. We recently showed that tree-centric approaches to carbon mapping perform well in Alpine coniferous forests ($r^2 = 0.98$ when field- and ALS estimate of carbon stocks are compared) and that a correction factor can be applied to account for the small obscured trees that were invisible from the air (Dalponte and Coomes, 2016). However, we recognise that applying this approach to Sabah's forests is likely to be challenging, as it contains some of the tallest and densest forests in the tropics. We therefore compare the tree-centric approach with Asner and Mascaro's (2014) generalised area-based method, and critically evaluate whether the advantages of working with individual trees outweigh any disadvantages associated with the accuracy of tree detection. Sepilok Forest Reserve, within which the study focusses, contains three distinctive forest types within close proximity, providing an outstanding opportunity to test the universality of carbon mapping approaches.

2. Methods

2.1. Study site

Sepilok Forest Reserve (5°10′ N 117°56′ E) is a remnant of lowland tropical rainforest on the east coast of Sabah, Malaysia, close to the town of Sandakan (Fox, 1973). This 4294-ha protected area, which ranges in elevation from 0 to 250 m a.s.l., was founded by the Sabah Forest Department's research centre in 1931. Extensive areas have never been commercially exploited, although about 670 ha in the northeast and south of the reserve were selectively logged up until 1957, and several hundred hectares in the north had lianas and non-commercial trees thinned out in 1958 (Dent et al., 2006; DeWalt et al., 2006). Four distinct forest types are recognised within the reserve: Sandstone hill dipterocarp forest (hereafter “sandstone hill”) grows on dissected hillsides and crests; alluvial dipterocarp forest is found in the valleys (hereafter “alluvial”); heath forest is located on podzols associated with the dip slopes of cuesta (shallow-sloping slopes following bedding planes; hereafter “kerangas”); and mangroves flank the bay. The first three of these species-rich forests have been the focus of much previous research (Dent et al., 2006; DeWalt et al., 2006; Fox, 1973; Nilus et al., 2011).

2.2. Field data

There are nine permanent forest plots of 4 ha within the Sepilok reserve, three for each forest type (Alluvial, Sandstone hills and Kerangas). Three of these plots (one alluvial and two sandstone) were established in 1968 (Fox, 1973). An additional six plots were established in 2000–2001 (Nilus, 2004). All plots have been regularly re-censused since, with data standardised, quality controlled and curated within the ForestPlots.net web application and database (Lopez-Gonzalez et al., 2011). The latest census of these plots carried out between 2013 and 2015 (the “2014 resurvey”) was used for this analysis. For the purposes of this analysis, each 4 ha plot was subdivided into 1 ha subplots which were analysed separately – giving a total 36 1-ha plots. We recognise that spatial autocorrelation could impact the 1-ha plot data; however, we emphasize that the plots are not being used for landscape sampling, rather they are intended for calibration and validation purposes only. In each of the 1-ha plots, all stems ≥ 5 cm in diameter were permanently tagged, their stem diameter (D, in cm) was measured, they were identified to species (or closest taxonomic unit), and their position was marked to the nearest 10 × 10 m subplot (45,214 trees in total). The standard point of measurement (POM) for diameter was at 1.3 m from the base of the tree; for trees with buttress or deformity at 1.3 m, POM was above the non-cylindrical feature. The corners of the plots were geolocated using a Geneq Sx Blue II GPS, which differentially corrects using Space Based Augmentation and has a positional accuracy of <2 m.

2.3. ALS acquisition and processing

Airborne ALS data were acquired on 5 November 2014 using a Leica ALS50-II ALS flown at 1850 m altitude on a Dornier 228-201 travelling at 135 knots. The ALS sensor emitted pulses at 83.1 Hz with a field of view of 12.0°, and a footprint of about 40 cm diameter. The average pulse density was 7.3/m^2. The total area surveyed was of 26 km^2. The Leica ALS50-II sensor records full waveform ALS, but for the purposes of this study the data were discretised by the sensor system, with up to 4 returns recorded per pulse. Accurate georeferencing of ALS point cloud was ensured by incorporating data from a Leica base station running in the study area concurrently to the flight. The ALS data were pre-processed by NERC's Data Analysis Node and delivered in standard LAS format. All further processing was undertaken using LAStools (http://rapidlasso.com/lastools/). Points were classified as ground and non-ground, and a digital elevation model (DEM) was fitted to the ground returns using a TIN (LAStools module las2dem), producing a raster of 1 m resolution. The DEM elevations were subtracted from elevations of all non-ground returns to produce a normalised point cloud, and a canopy height model (CHM) was constructed from this on a 0.5 m raster by averaging the first returns. Finally, holes in the raster were filled by averaging neighbouring cells. Given the very high point density, this last procedure was applied to very few pixels (0.02% of the total).

2.4. Height-diameter allometry and crown area measurements

We manually delineated the crowns of 147 trees in images of the CHM, and measured their maximum heights (H, *m*) using the CHM. By walking around in the plots with print outs of the delineated CHM, it was possible to geo-locate 91 of these trees with confidence, giving us 91 trees for which we had H estimated from ALS and D measured on the ground. This dataset spanned the full range of tree sizes present at Sepilok (D: 12–165 cm; H: 16–72 m). Power-law and 3-parameter Weibull functions were fitted to the H-D relationships using maximum likelihood estimation, assuming residuals to be normally distributed and increasing with tree size and the model with the strongest statistical support was found using the Akaike Information Criterion (AIC) approach. AIC is founded on information theory and offsets the complexity costs of including additional parameters against likelihood improvements to select models (see Supporting information for code). Variation in allometry among soil types was also tested using AIC. The best-supported model (i.e., the one with the lowest AIC) was then used to estimate the heights of all trees recorded in the plots. We compared our H-D allometry with another dataset of 644 tree heights collected from a random subset of trees in the plot as part of the 2014 resurvey, using a Vertex III hypsometer (Haglöf, Sweden). For the 91 trees that we had hand delineated, crown area (*CA*, in m^2) was calculated in the imagery, and compared with *CAs* estimated on the ground from two orthogonal crown diameter measurements.

2.5. Estimating tree- and plot-level aboveground carbon density

Field data from the 36 1-ha plots were used to estimate the aboveground biomass (AGB, in kg) of each individual tree using Chave et al.'s (2014) pantropical equation:

$$AGB_{Field} = 0.0673 \times \left(WD \times D^2 \times H\right)^{0.976} \quad (1)$$

where WD is wood density (dry mass / wet volume, in g cm^{-3}), which we obtained from the global wood density database (Chave et al., 2009; Zanne et al., 2009), D is stem diameter (cm) and H is the estimated height (m) obtained from ground-based allometry, because heights were recorded for only a fraction of trees in the field survey. Trees were assigned species-specific WD values (51% of stems) or closest taxonomic unit (92% of stems matched to genus). For each 1 ha plot, aboveground carbon density (ACD, in Mg C ha^{-1}) was then estimated by summing the AGB of all trees within the plot and applying a carbon content conversion factor of 0.47 (Martin and Thomas, 2011).

The dataset of 91 manually-delineated trees was used to obtain a relationship between AGB_{Field} values and H_{ALS} and CA_{ALS} measurements. Regression of manually vs automatically delineated CA produced a near 1:1 relationship, with $R^2 = 0.85$ and RMSE = 43.3, so we regarded these as equivalent. Jucker et al. (2016) showed that the biomass of 2395 harvested trees in a global dataset was related to height and crown diameter (where $= 2 \times \sqrt{CA/\pi}$) according to a power-law function:

$$AGB = \alpha \times (H \times CD)^{\beta}. \quad (2)$$

Jucker et al. (2016) showed that this function delivered unbiased estimates of AGB across six orders of magnitude variation in tree mass, while α and β varied among plant clades, so is most accurate when fitted to local data. This function was fitted to data from the 91 segmented trees by least-squares regression of log-log transformed data. Goodness of fit was evaluated in comparison to a function based on height alone and on H × CA, the model adopted by Ferraz et al. (2016), using AIC.

2.6. Area-based approach

2.6.1. Generalised ACD equation

Asner and Mascaro (2014) proposed a generalised approach for estimating *ACD* using a single ALS metric – top canopy height (TCH, in m) – and minimal field data inputs. The method relates ACD to TCH, stand basal area (BA; in m^2 ha^{-1}) and the community-weighted mean wood density (WD; in g cm^{-3}) as follows:

$$ACD_{General} = 3.836 \times TCH^{0.281} \times BA^{0.972} \times WD^{1.376} \quad (3)$$

This function was fitted to data from the 36 Sepilok plots. TCH was the mean height of CHM pixels within each plot extracted using the raster package of R. As recommended by Asner and Mascaro (2014), least-squares regression was used to relate field-measured BA to TCH, and field-estimated WD to TCH. These two regression relationships were then included as sub-models in Eq. (3), such that ACD was predicted solely as a function of ALS derived TCH.

2.6.2. Local ACD equation

To reduce uncertainty in ACD estimates, Asner and Mascaro (2014) suggest that, when possible, regionally-calibrated equations for estimating ACD from TCH, BA and WD should be developed. The data presented in this manuscript comes from a single reserve, albeit with diverse forest types, so we refer to our models as "local" rather than "regional", which is more appropriate if plots are widely dispersed in the region of interest. Least-squares regression was used to estimate the parameters of the following power-law model relating ACD to TCH, BA and WD values in the Sepilok dataset:

$$ACD_{Local} = a \times TCH^{b} \times BA^{c} \times WD^{d} \quad (4)$$

As for the generalised ACD model, BA and WD were first estimated as a function of TCH, and these functions used as sub-models in Eq. (4).

2.6.3. Gap fraction as an estimator of basal area

Central to robustly estimating ACD using ALS data is identifying a metric which captures variation in BA among stands, which is considerable in Sepilok (Fig. S1). Asner and Mascaro's (2014) power-law models make the assumption that TCH can be used as a strong predictor of BA. However, this may not universally be the case (Duncanson et al., 2015; Spriggs, 2015). We tested whether gap fraction (GF) – the proportion of area not occupied by crowns at a given height aboveground – could be used as an alternative ALS metric for predicting BA (Fig. 2b). Gap fraction GF_h is calculated by creating a plane horizontal to the ground in the canopy height model (CHM), at height h above ground, and counting the number of pixels for which the CHM lies beneath the plane, divided by the total number of pixels in the plot. GF was calculated for heights of 1 to 50 m.

Goodness of fit of global, local and gap-fraction-modified models are compared by reporting %RMSE, calculated as $\sqrt{\sum (y-\hat{y})^2}/(N\overline{y})$, where y are the data, $\overline{y}$ is the mean, $\hat{y}$ are the model estimates, N is the number of observations. %Bias is reported as $\sum (y-\hat{y})/(N\overline{y})$.

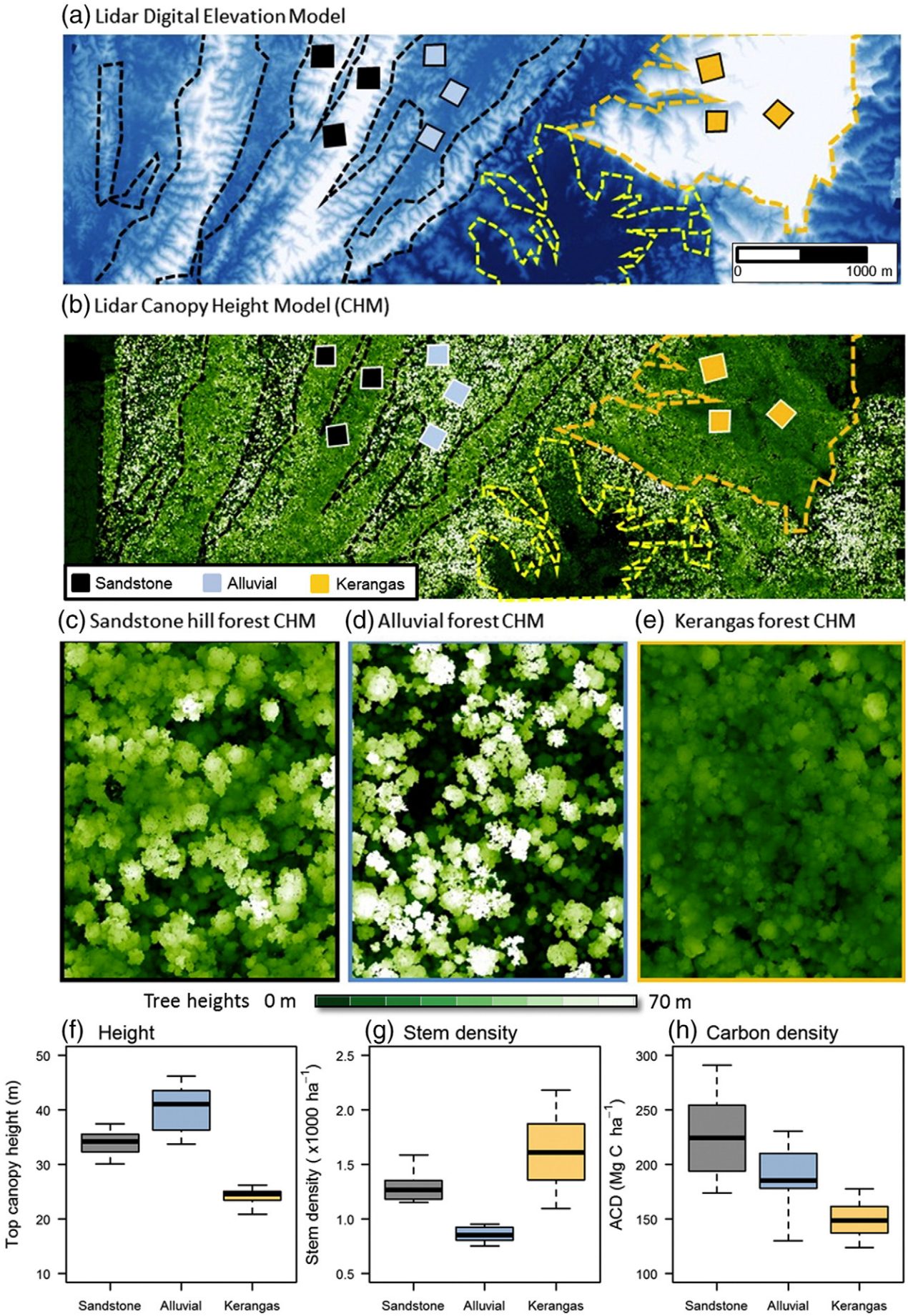
2.7. Tree-centric approach

As an alternative to area-based approaches, we tested whether ACD could be accurately estimated by summing the biomass of individual tree crowns (ITCs; Fig. 1). The *itcSegment* algorithm was used to delineate trees within each of the 36 plots and a 10 m buffer zone around them. *itcSegment* is implemented in R and made freely available on CRAN (https://cran.r-project.org/web/packages/itcSegment/index.html). It works by finding local maxima in the raster CHM, regarding these maxima as tree tops, then growing crowns around them by local searching of the raster CHM and point cloud. This approach was initially developed for coniferous forests (Dalponte et al., 2014; Eysn et al., 2015; Dalponte, 2015; Dalponte and Coomes, 2016) following the concept of Hyyppa et al. (2001). It has three stages (see Supporting information for further details):

1. *Smoothing the Canopy Height Model.* A Gaussian low-pass filter is applied to the raster CHM to smooth it, improving the reliability of the region-growing code by removing sharp changes in height within crowns that might otherwise be inferred to be crown edges.

2. *Locating local maxima in the CHM.* To function effectively in tropical forests, we found it necessary to introduce variable window sizes when searching for local maxima around the central pixel of the window. If a particular pixel in the CHM is high above the ground, then it is assumed to belong to a large tree, so a relatively large search window is needed to find the top of that tree. The dimension of the search window for a CHM pixel of elevation z is equivalent to crown radius of a wider-than-average tree of that height, ensuring that even the widest crowns in the images were not over-segmented. This crown radius was obtained by fitting an allometric function to the relationship between crown radius and height of 147 manually segmented trees, using quantile regression (*quantreg* package of R), with tau = 0.9 such that 90% of data-points sat

Fig. 2. (a) ALS-generated Digital Elevation Model (lighter shades corresponding to higher elevation) of Sepilok Nature Reserve, Sabah, showing sandstone hills (outlined in black), a cuesta supporting kerangas forest (orange), mangrove (yellow) and alluvial areas; locations of nine 4-plots ha in the sandstone hill (grey), kerangas (orange) and alluvial (light blue) forests are shown; (b) canopy height model; (c–e) magnified to show typical canopies of the three forest types; (f–g) boxplots contrasting the structural attributes of the three forest types, based on measurements taken in 36 1-ha plots.


(a) Lidar Digital Elevation Model
0
1000 m
(b) Lidar Canopy Height Model (CHM)
Sandstone
Alluvial
Kerangas
(c) Sandstone hill forest CHM
(d) Alluvial forest CHM
(e) Kerangas forest CHM
Tree heights 0 m
70 m
(f) Height
Top canopy height (m)
50
40
30
20
10
Sandstone
Alluvial
Kerangas
(g) Stem density
Stem density (x1000 ha^{-1})
2.5
2.0
1.5
1.0
0.5
Sandstone
Alluvial
Kerangas
(h) Carbon density
ACD (Mg C ha^{-1})
300
250
200
150
100
Sandstone
Alluvial
Kerangas

below the regression line (Fig. S4). A look-up table was created so that a window size could be obtained quickly for a given value of z based on this allometry.

3. *Region growing around local maxima:* The algorithm then searches iteratively around each local maximum for neighbouring pixels that are slightly lower in elevation but not greatly lower (i.e. off the crown's edge) or higher (i.e. belonging to a neighbouring tree), and regards these as being part of the same crown. The algorithm has various stopping rules to constrain the growth of each crown: crown width cannot exceed those observed in manually delineated crowns and crown depth cannot exceed a fixed percentage of height. Once regions are fully grown, the algorithm explores the ALS point cloud and uses the data to create a 2-D convex-hull polygon representing the crown.

Once trees are delineated in the imagery, the next step is to estimate their individual biomass. This was achieved by predicting biomass using Eq. (2), which was parameterised using the crown diameters and heights of the 91 trees. We then summed the AGB of segmented trees, and multiplied by a carbon density of 0.47, to obtain ACD estimates for each of the 1 ha plots. The final step is to calculate a correction factor to account for understorey trees that were not segmented because they are obscured by the upper canopy, and large trees that were over-segmented. This correction factor was obtained by fitting a regression line, passing through the origin, to the relationship between ACD_{ALS} and ACD_{Field} (Dalponte and Coomes, 2016).

The accuracy of the ITC delineation was assessed by comparing the numbers of delineated trees within stem diameter classes with the numbers observed in the field plots. To do this, the heights of the delineated trees were used to estimate stem diameters, using an allometric function derived from the 91 trees co-located on the ground and in the aerial imagery.

3. Results

3.1. Comparison of forest types

Alluvial, sandstone hill and kerangas forest types are clearly visible in the CHM and DEM of Sepilok Reserve (Fig. 2). In the three panels showing magnified-CHMs of the forest types, the alluvial forest is seen to contain the tallest trees (some reaching 70 m in height), while also exhibiting large gaps between emergent trees (Fig. 2d). The sandstone hill forest appears less tall (50–60 m) and is more densely packed (Fig. 2c); while the kerangas is much shorter with densely packed trees (Fig. 2e). Summary statistics from the plots corroborate these descriptions: the sandstone forests have the greatest above-ground carbon density, not because they are the tallest but because they have high basal areas (Fig. 2g–i), whereas the kerangas has low biomass by virtue of its height, despite having high wood densities (Fig. S1). The sandstone hill and kerangas forests have unimodal distribution of ALS returns with height, which is typical of the signal returned from forest canopies (Fig. 3). However, the alluvial has a distinctive bimodal distribution, resulting from a dense subcanopy sitting beneath the scattered emergent trees (Fig. 3). This strong variation in forest structure makes it challenging to map carbon density using a single model, as is our aim.

3.2. Allometric relationships

ALS-estimated heights (H_{ALS}) were related to field-measured diameters (D) as follows: $H_{LiDAR} = 5.29 \times D^{0.50}$ (red circles and curve; Fig. 3a). This power-law function was better-supported statistically than a 3-parameter Weibull (see Supporting information). The ALS-estimated heights are mostly within the upper quartile of data points obtained from field-measurements, and the power-law regression line through the field data is systematically different to the ALS dataset (grey circles and curves; Fig. 3a). The allometric function obtained from the field data is more similar to previous reports than our ALS-based function estimates (See Fig. 3a; Morel et al., 2011, Feldpausch et al., 2011 and Chave et al., 2014), so it seems likely that our approach for segmenting ALS imagery is preferentially sampling emergent trees that have systematically different HD allometries. For this reason, we used field-based allometries to estimate heights of trees, and generate forest-type-specific allometries (see Supporting information). If we had used the ALS-based allometry, ACD would have been 1.19× greater than the ones presented; this underscores the need to collect terrestrial LiDAR, ALS and field datasets across multiple sites in the tropics, to get a better handle on tree allometry.

As expected, a linear relationship was found between field and remotely measured crown areas, although field estimates are 30% smaller on average, with large uncertainty (Fig. 3b), probably because estimates based on field data assume that crowns are symmetrical ellipses, based on just two diameter measurements, whereas the ALS approach allows the margins to be determined more accurately (e.g. Fig. S4).

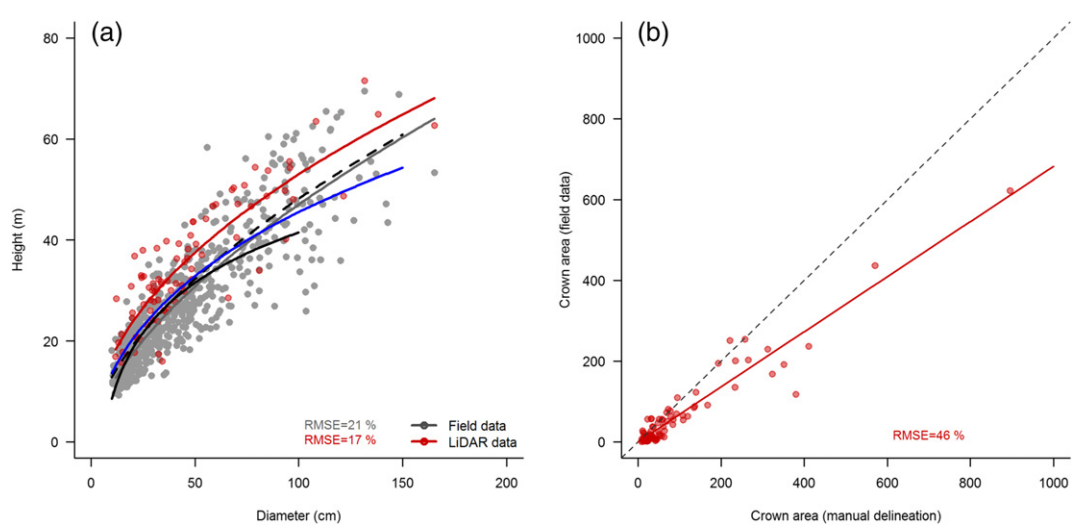


Fig. 3. Panel (a) compares the H–D relationship for trees where H was measured from the ground (grey circles) with those where ALS was used to measure height (red circles) Stem diameters (D) were measured on the ground. Power function fitted to these datasets are shown. In addition, the dashed black curve is the allometry reported from another study in lowland Sabah (Morel et al., 2011), the solid black curve is a pan-SE Asia allometry reported by Feldpausch et al. (2011), and the blue curve is calculated from Eq. (6) in Chave et al. (2014) with E = −0.147. Panel (b) shows the relationship between CA measured from the ground versus manual delineation from the ALS data.

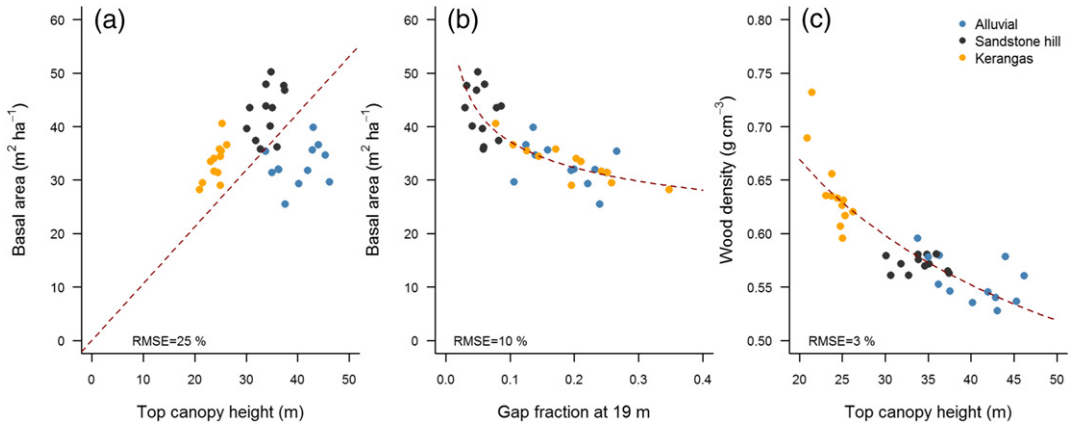


Fig. 4. Relationship between (a) basal area and top canopy height, (b) basal area and gap fraction at 19 m aboveground, and (c) basal area-weighted mean wood density and top canopy height. Regression lines fitting to all three sites are shown as dashed lines and %RMSE values provided.

3.3. Area-based approaches

3.3.1. Generalised model of Asner and Mascaro (2014)

The generalised model (Eq. (3)) gave accurate estimates of ACD when sub-models of the relationships between BA, WD and TCH were fitted to the Sepilok datasets. We obtained the following relationship between basal area and top canopy height (Fig. 4a):

$$BA = TCH \times 1.062 \tag{5}$$

Note the relationship is forced through the origin for reasons explained by Asner and Mascaro (2014). The basal area-weighted mean WD is estimated from TCH as follows (Fig. 4c):

$$WD = 1.541 \times TCH^{-0.278} \tag{6}$$

By entering Eqs. (5) and (6) into Asner and Mascaro's (2014) generalised model (Eq. (3)), we get $ACD_{generalised} = 7.37 \times TCH^{0.870}$. This function produces biased estimates of biomass when applied to the Sepilok Forest Reserve, underestimating ACD_{field} by 19% on average (Fig. 5a).

3.3.2. Local ACD equation

A local model, obtained by fitting Asner and Mascaro's (2014) power-law model to the data available from Sepilok yielded the following relationship:

$$ACD_{Local} = 0.719 \times TCH^{0.628} \times BA^{1.043} \times WD^{0.687} \tag{7}$$

where TCH was obtained from the ALS survey, and BA and WD from the field data collected from the 36 plots. Entering the BA and WD relationships into Eq. (7), gives the following relationship: ACDLocal = $1.030 \times TCH^{1.535}$. As shown in Fig. 5b, the local model is unbiased, but its accuracy remains low (%RMSE = 27%). It is clear that the sandstone hill and kerangas plots fall along a line, but that the alluvial plots deviate from the line, and that there is no correlation between field- and ALS-predicted ACD among these plots.

3.3.3. Gap fraction as an estimator of basal area

The primary reason why Eq. (7) is inaccurate in its prediction is that BA is so poorly correlated with TCH (Fig. 5a). However, we find that gap fraction at 19 m above ground (GF_{19}) is closely related to BA (Fig. 5b) as follows:

$$BA = 23.4 \times GF_{19}^{-0.20} \tag{8}$$

GF_{19} was chosen after a comparison of models including gap fraction calculated at all heights between 1 and 50 m (Fig. S3). When Eq. (8) is used as a sub-model within Eq. (7), the estimation equation becomes:

$$ACD_{Local} = 25.93\ GF_{19}^{-0.209}\ TCH^{0.437} \tag{9}$$

and a much closer relationship was obtained between field and ALS

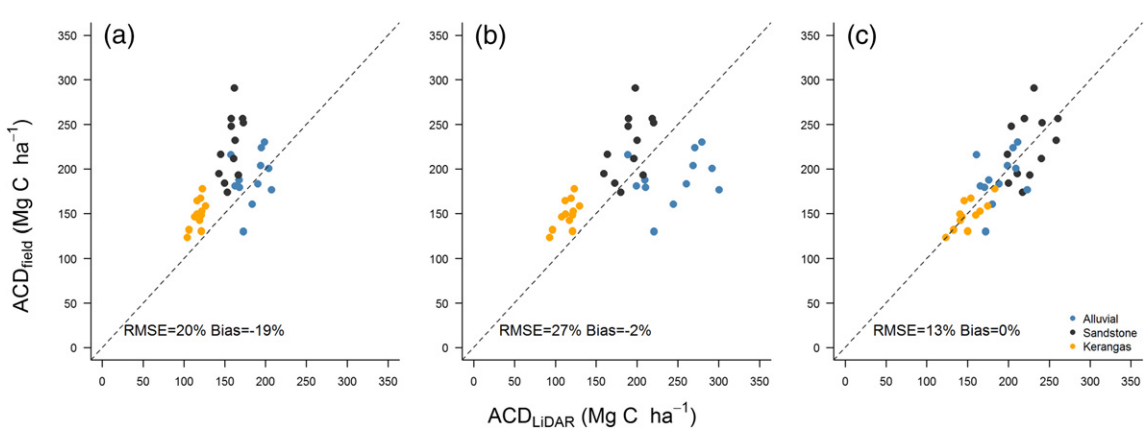


Fig. 5. Relationship between aboveground carbon density (ACD) estimated from field data and (a) ACD estimated using Asner and Mascaro's (2014) generalised model (Eq. (3)), (b) ACD estimated using a locally-calibrated model in which basal area is estimated from TCH, and (c) a locally-calibrated model in which basal area is estimated from gap fraction. The one-to-one line is given for reference in all panels (dashed black lines).

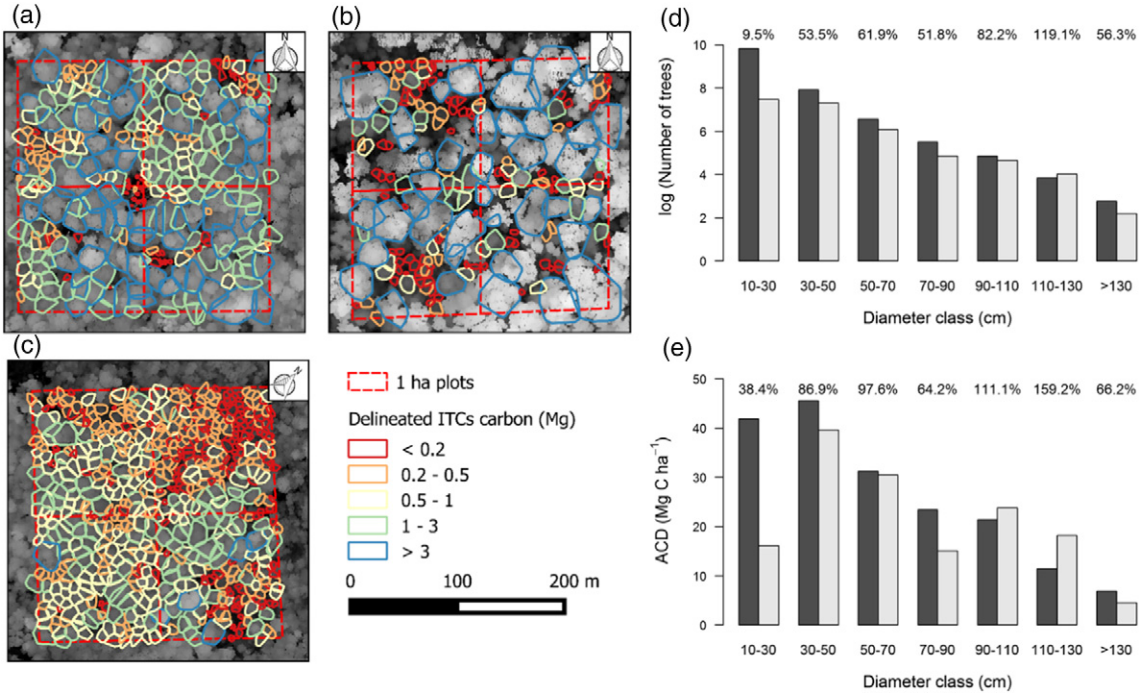


Fig. 6. Detection of individual tree crowns in ALS imagery using *itcSegment* for (a) sandstone hill, (b) alluvial and (c) kerangas forests. Within diameter size classes, the (d) numbers and (e) aboveground carbon density of trees detected by *itcSegment* are compared with values obtained in permanent plots (light and dark bars, respectively), with ITC vs field values given in percentage terms above the bars.

estimates (Fig. 5c). The RMSE was 29 Mg C ha^{-1}, corresponding to a percentage RMSE of 13%.

3.4. Individual tree crown (ITC) approach

AGB (kg) is related to the height and crown diameter of 91 delineated stems as follows:

$$AGB = 0.136 \times (H \times CD)^{1.52} \quad (10)$$

where CD is crown diameter. The allometric model has $R^2 = 0.74$ and %RMSE = 11.8% and had lower AIC than alternative models (AIC = 190, 197, 209 for H × CD, H × CA and H, respectively). Introducing variable window sizes when searching for local maxima allowed detection of both large and small trees within the rainforest canopies (Fig. 6a–c). Without this flexibility, the algorithm either omits small trees when the window was large, or over-segments large ones when the window was small. Nevertheless, because itcSegment only searches for maxima in the canopy height model, it cannot find subcanopy trees, detecting only 9.5% of stems in the 10–30 cm size class. Trees in the 10–30 cm size class contribute 23% of ACD, so represent a significant pool of carbon (Fig. 6e). Furthermore, carbon stored in very large trees was overestimated. Only 63 trees exceeded 110 cm diameter within the 36 plots, representing a tiny fraction of stems (0.2%) but a more significant percentage of carbon (10%). The ITC analysis recognised 65 trees over 100 cm diameter, but overestimated their carbon by 124%. These problems were mostly restricted to the four plots at a single alluvial site, within which the very tallest trees are located.

In our previous work with conifers in the Italian Alps, we found that ITC-based estimates of ACD were consistently less than field estimates, because obscured trees were by far the most important source of bias in the tree-centric approach (Dalponte and Coomes, 2016). This allowed us to use a single correction factor to all plots. However, here we find that ACD_{ALS} is greater than the field estimates in five of the 36 plots (Fig. 7a). The bias is removed when ACD_{ALS} values are multiplied by a single correction (1.21) but the RMSE remains high (20%). Even when

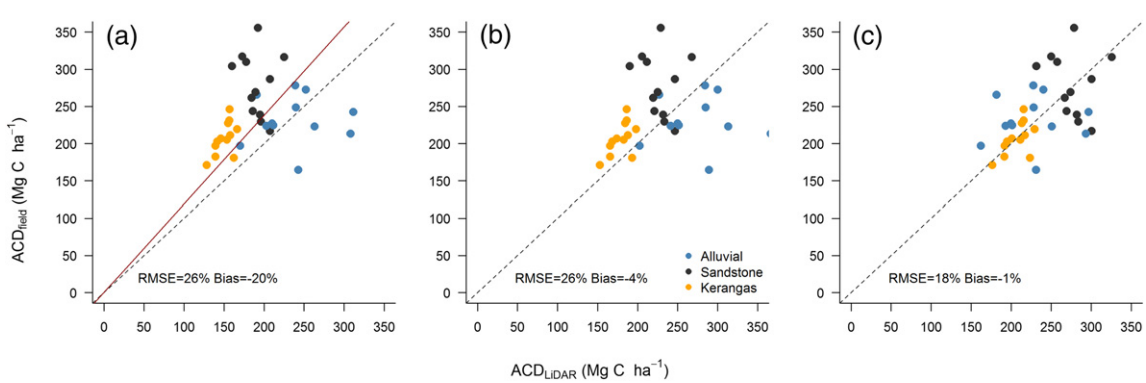


Fig. 7. Relationship between aboveground carbon density (ACD) estimated from field data and ACD estimated using the tree-centric approach; (a) without correcting for undetected subcanopy trees; (b) applying a universal correction factor; (c) applying forest-specific correction factors to the ACD estimates obtained from the tree-centric approach to account for missing stems. The one-to-one line is given for reference in all panels (dashed black lines); the solid red line in (a) is a regression line fitted to all plots from which the universal correction factor (1.21) is obtained.

separate correction factors are applied to the three forest types (alluvial = 1.07, sandstone hill = 1.30, kerangas = 1.32), the %RMSE of tree-centric approach remains greater than that obtained by area-based modelling (RMSE = 18 vs 13%). Note once again that the alluvial forest is problematic to model, because over-segmentation in one site causes the correction factor to differ greatly from the other forest types.

4. Discussion

4.1. Area-based approaches

The power-law modelling approach developed by Asner and Mascaro (2014) was effective for mapping forest carbon in the old-growth lowland rainforests of Sabah once adjustments had been made for use in this region. Their generalised model parameterised mostly from Neotropical datasets (Eq. (3)) gave biased results when applied to the Sepilok data, even when local BA-TCH and WD-TCH sub-models were used, so locally parameterised power-law functions were developed. Furthermore, basal area was better modelled as a function of gap fraction rather than top-of-canopy height.

4.1.1. Distinctiveness of SE Asian forests

Asner and Mascaro's (2014) paper shows that tropical forests from 14 regions differ greatly in structure, exemplified by the TCH-BA scatterplot in Fig. 8. Remarkably, they found that a generalised power-law relationship could be fitted that transcended all these contrasting forests types, once regional differences in structure were incorporated as sub-models relating BA and WD to TCH. However, that generalised model delivered biased results when applied to Sepilok forest datasets, underestimating ACD by 19% when the models were fitted with local data (Fig. 5a). The explanation for this bias is the dominance of lowland forests by trees in the Dipterocarpaceae, which make them distinctive from all other tropical forests (Corlett, 2009). The height-diameter allometry of these dipterocarp dominated forests are different from those found elsewhere - they have particularly narrow stems for their height (Feldpausch et al., 2011; Banin et al., 2012). The basal area of the plots in Sabah is not especially high compared with those in the Asner and Mascaro dataset (Fig. 8a), but ACD is in the upper quantiles of other forest of comparable basal area (Fig. 8a): the extraordinary height of the alluvial and sandstone hill forests gives rise to high ACDs, while high wood density in the stunted kerangas trees are the underlying cause of their high ACD. The ACD values in Sepilok Reserve (sandstone hill, alluvial and kerangas contain on average 226, 190, 150 Mg C/ha, respectively) are not dissimilar to those reported from other unlogged plots in lowland Sabah (176 Mg C/ha; Morel et al., 2011) or to regional estimates derived remotely (160 Mg C/ha Avitabile et al., 2016), but they are among the greatest ACD recorded anywhere on the planet (Avitabile et al., 2016; Sullivan et al., 2017).

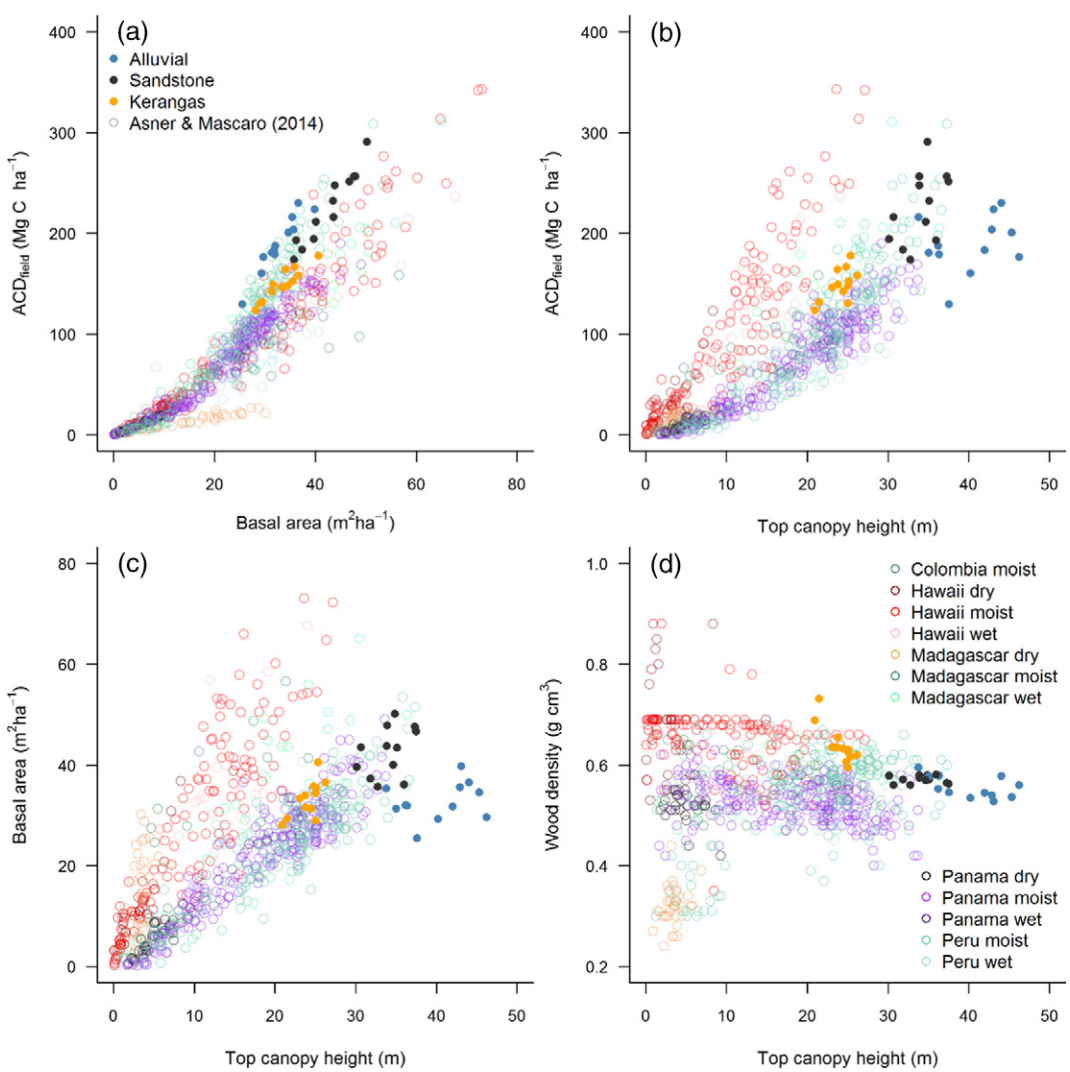


Fig. 8. Relationship between (a) field-estimated aboveground carbon density and basal area, (b) field-estimated aboveground carbon density and top canopy height, (c) basal area and top canopy height, and (d) basal-area weighted mean wood density and top canopy height in Asner and Mascaro's (2014) database (open circles) and at Sepilok (filled circles).

The kerangas forests at Sepilok contain many stems of *Tristaniopsis merguensis* and *Cotylelobium melanoxylon*, which have dense wood (0.94 and 0.83 g cm^{-3}, respectively) reflecting adaptation to the nutrient-poor soils. The distinctive structure of the Sepilok forest compared with anything used to calibrate the power-law model previously meant that it needed be custom-fitted with data only from that forest. Rainforests of SE Asia are structurally diverse (Corlett, 2009), and it remains to be seen whether a generalised model can be produced for this region, as found previously for 14 forest types in the Asner and Mascaro dataset.

4.1.2. Predicting basal area by gap fraction

ACD is more closely related to basal area than to height (Fig. 8), so it is essential that the basal-area sub-model predicts basal area accurately from ALS statistics. Here we found that gap fraction at 19 m above ground (GF_{19}) was a better predictor of basal area than TCH. This makes intuitive sense if one considers an open forest comprised of just a few trees - the crown area of each tree scales with its basal area so the gap fraction at ground level of a plot is negatively related to the basal area of its trees. A similar principle applies in denser forests, but now that canopies overlap, the best-fitting relationship between gap fraction and basal area is no longer at ground level, but is instead further up the canopy. Our study is the first to formally introduce gap fraction into the modelling framework of Asner and Mascaro (2014), but several other studies have concluded that gap fraction is an important variable to include in multiple regression models of forest biomass (Colgan et al., 2012a; Li et al., 2014; Spriggs, 2015; Singh et al., 2016). Additionally, (Colgan et al., 2012b) used fractional canopy cover ($CC = 1 - GF$) to modify the relationship between aboveground carbon density and TCH for African woodland savannas. In a follow-on paper to this one, we develop a state-wide carbon model for Sabah, including secondary as well as old-growth forests. Modelling basal area as a function of canopy cover at 19-m height is once again shown to improve goodness of fit, although the functional form is different (Coomes et al. archived). In the specific case of Sepilok, GF_{19} proved particularly effective for predicting basal area, as it was able to clearly distinguish between alluvial forests – where large emergent crown tower over a dense understory layer – and the more compact vertical distribution of crowns in sandstone hill and kerangas forests (Fig. S2). Incorporating GF_{19} into the basal area sub-model yielded the best predictions of ACD (Fig. 5c), with an RMSE of 29.0 Mg C ha^{-1}. This is slightly greater than the RMSE 24.7 Mg C ha^{-1} reported by Asner and Mascaro (2014) for their generalised model fitted to 14 forest types, but their study included secondary forest, which have lower variance in ACD (Fig. 8), so a lower RMSE was anticipated.

4.2. Tree-centric approach

ITC delineation has the potential to make much fuller use of the information within 3D ALS point clouds compared to area-based approaches, yet in the present study, the end result was less accurate ACD estimates than those achieved by area-based approaches (RMSE = 18% vs 13%). Our previous work in the Italian Alps showed tree-centric approaches to give marginally more accurate results ($R^2 = 0.98$ vs 0.96), but that paper focussed on structurally simple conifer-dominated forests. Another recent study working in dense tropical forests also found the tree-centric approach to underperform compared to area-based methods. Ferraz et al. (2016) developed an algorithm capable of detecting both sub-canopy as well as canopy crowns, yet it too delivered less accurate results than an area-based model using TCH as in input variable when applied to the 50-hectare plot on Barro Colorado island in Panama (RMSE = 13.8% vs 12.5% at the 1-hectare scale).

Three key factors contributing to uncertainty in ACD estimates are over-segmentation of large trees, incomplete detection of sub-canopy trees, and uncertainty in AGB estimates from ITC information. These are discussed in turn:

Over-segmentation of emergent dipterocarps, such as the one shown in Fig. S4, was observed. This is concerning given that 90% of biomass in alluvial plots is held in just 12% of large trees, compared to 36% in the kerangas plots and 23% in the sandstone hill plots. Yet over-segmentation is not actually such a problem, because of the $H \times CD$ term in the tree biomass function (i.e., $AGB = 0.136 \times (H \times CD^{1.52})$). To illustrate this point, imagine we split a tree in two, vertically. If we only modelled AGB as a function of H, we would severely overestimate ACD. However, by including crown diameter (CD) in the equation this error is reduced, because the two resulting trees would each have half the CD of the original. This functional form has been shown previously to provide unbiased estimates of biomass when applied to a global dataset of AGB, H and CD values (Jucker et al., 2016). In contrast, Ferraz et al. (2016) found that the sum of $WD \times \sum CA \times H$ was the function that gave the closest relationship between field- and ALS-estimates of ACD. Including crown dimensions in regression modelling may be useful when mapping regions with diverse forest types, because Blanchard et al. (2016) report that CA-D allometry is less variable among forest types than H-D relationships.

Incomplete detection of understory trees accounted for 46 Mg C ha^{-1} of missed carbon in the Sepilok forest, a significant fraction of total ACD. Detecting more understory trees will help reduce bias in carbon estimation. Recent years have seen substantial progress in segmentation, as both ALS instruments and the algorithms used to delineate trees from ALS data have improved considerably (Popescu et al., 2003; Yao et al., 2012; Duncanson et al., 2014; Paris, Valduga and Bruzzone 2016; Shendryk et al., 2016). Several recent papers point the way forward. Duncanson et al. (2014) developed an approach that first finds trees in the upper canopy, then strips points from the 3D point cloud that are associated with these trees, before searching the remaining point cloud for further trees. Ferraz et al. (2016) developed a "bandwidth model", which estimates the likely crown dimensions of trees of a certain height, based on information gained by manual delineation of crowns from the imagery, and uses this information to search the entire point cloud for trees. Paris et al. (2016) developed a segmentation method that was able to correctly delineate the crowns of 97% and 77% of canopy dominant and understorey trees, respectively, as well as accurately measuring the crown dimensions of all segmented trees. Equally promising is Shendryk et al.'s (2016) algorithm which segments trees from the bottom up, mimicking the approach used to process terrestrial laser scanning data (Calders et al., 2015). Nevertheless, our less-sophisticated approach of segmenting only the trees in the upper canopy, and applying a correction factor to account for missed carbon, may deliver similar precision provided the size distribution of forests in known, such as a power-law or Weibull distribution (Spriggs, 2015). Comparisons of the effectiveness of different algorithms when applied to the same tropical forest plots will help the refinement of approaches.

Estimates of biomass from ITC information have high variance, because the allometric relationship between tree-AGB and ($H \times CD$) has high residual variance (see also Jucker et al., 2016): our model predicts that a 50 m tall tree contains 2300 kg C with 95% confidence intervals of 1870–2920 kg C, which is very wide. One reason for this uncertainty is that ALS measures tree volumes and is generally poor at estimating wood density: we found a strong relationship between TCH and WD at plot level, but the relationship between individual H and WD values was very weak.

5. Conclusions: is it worth trying to see both the forest and the trees?

Individual-based modelling performed less well than area-based approaches, as also found by Ferraz et al. (2016) working in Panama. Given that ITC approaches are computationally demanding, is there any justification in using them? We argue that researchers should certainly persist in their development for several reasons. First, maps of trees are of practical value. For example, farm owners in several tropical countries

are legally obliged to protect strips of riparian forests on their land, but enforcing such laws has proven difficult for lack of fine-scale maps of riparian zone extents and the fate of trees within them: individual based mapping is an excellent tool for this. Similarly, the high carbon stock (HCS) approach developed in partnership with the Roundtable on Sustainable Palm Oil commits oil palm companies to protect and restore forests on their estates to make them carbon neutral (http://highcarbonstock.org/), and fine-scale mapping of carbon in these forest fragments could be valuable in this context. Secondly, there is great interest in tracking the dynamics of the largest trees in tropical forest, because they account for a large fraction of forest carbon and are known to be especially vulnerable to drought, acting as early-warning signals of the effects of climate change (Nepstad et al., 2007; Bennett et al., 2015). Repeat surveying with ALS could prove valuable for this, but only area-based approached have been used to date (e.g., Simonson et al., 2016).

Another advantage of the ITC approach is that it is directly analogous to well-established field-based approaches for measuring carbon, and so has a strong theoretical basis in terms of identifying and minimising sources of uncertainty and bias, and propagating error. For example, differences in height estimates between ALS and ground measurements (Fig. 3) shine a light on a major source of bias in ACD estimates (Hunter et al., 2013). With area-based approaches, prediction error decreases with increasing plot sizes due to reduced edge effects (Goetz and Dubayah, 2011). The same is true of ITC approaches but the errors are less dependent on plot size (Dalponte and Coomes, 2016). Intriguingly, Ferraz et al. (2016) showed the coefficients of models developed by a ITC approach were invariant of scale, whereas the coefficients of power-law models varied with scale. ITC approaches could allow a more explicit characterisation of such issues. For the moment, area-based approaches are likely to remain the method of choice for mapping tropical forest carbon, but ITC approaches are advancing rapidly and could offer a viable alternative as algorithms, sensors and analysis platforms mature.

Acknowledgement

This project was supported by a grant through the Human Modified Tropical Forests programme of NERC (NE/K016377/1). We thank members of the NERC Airborne Remote Sensing Facility and NERC Data Analysis Node for collecting and processing the data (project code MA14-14). David Coomes was supported by an International Academic Fellowship from the Leverhulme Trust. Lindsay Banin contributed field allometry data which were collected during her PhD at University Leeds, supported by NERC and a RGS Henrietta Hutton Grant. Oliver Phillips, Simon Lewis and Lan Qie provided census data collected as part of an ERC Advanced Grant (T-Forces). Thanks to Richard Bryan Sebastian for collecting field data on delineated trees.

Appendix A. Supplemenatry data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.rse.2017.03.017.

References

Agrawal, A., Nepstad, D., Chhatre, A., 2011. Reducing emissions from deforestation and forest degradation. Annu. Rev. Environ. Resour. 36:373–396. http://dx.doi.org/10.1146/annurev-environ-042009-094508.

Asner, G.P., Mascaro, J., 2014. Mapping tropical forest carbon: calibrating plot estimates to a simple LiDAR metric. Remote Sens. Environ. 140:614–624. http://dx.doi.org/10.1016/j.rse.2013.09.023.

Asner, G., Rudel, T., Aide, T., DeFries, R., 2009. A contemporary assessment of change in humid tropical forests. Conserv. Biol. 23:1386–1395. http://dx.doi.org/10.1111/j.1523-1739.2009.01333.x.

Asner, G.P., Clark, J.K., Mascaro, J., Vaudry, R., Chadwick, K.D., Vieilledent, G., ... Knapp, D.E., 2012. Human and environmental controls over aboveground carbon storage in Madagascar. Carbon Balance Manag. 7 (1). http://dx.doi.org/10.1186/1750-0680-7-2.

Asner, G.P., Mascaro, J., Anderson, C., Knapp, D.E., Martin, R.E., Kennedy-Bowdoin, T., ... Bermingham, E., 2013. High-fidelity national carbon mapping for resource management and REDD+. Carbon Balance Manag. 8:7. http://dx.doi.org/10.1186/1750-0680-8-7.

Avitabile, V., Herold, M., Heuvelink, G.B.M., Lewis, S.L., Phillips, O.L., Asner, G.P., ... Willcock, S., 2016. An integrated pan-tropical biomass map using multiple reference datasets. Glob. Chang. Biol. 22:1406–1420. http://dx.doi.org/10.1111/gcb.13139.

Baccini, A., Goetz, S.J., Walker, W.S., Laporte, N.T., Sun, M., Sulla-Menashe, D., ... Houghton, R.A., 2012. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nat. Clim. Chang. 2 (1). http://dx.doi.org/10.1038/nclimate1354.

Banin, L., Feldpausch, T.R., Phillips, O.L., Baker, T.R., Lloyd, J., Affum-Baffoe, K., ... Lewis, S.L., 2012. What controls tropical forest architecture? Testing environmental, structural and floristic drivers. Glob. Ecol. Biogeogr. 21 (12):1179–1190. http://dx.doi.org/10.1111/j.1466-8238.2012.00778.x.

Bennett, A.C., McDowell, N.G., Allen, C.D., Anderson-Teixeira, K.J., Trenberth, K.E., Nepstad, D.C., ... Haack, R.A., 2015. Larger trees suffer most during drought in forests worldwide. Nat. Plants 1 (10):15139. http://dx.doi.org/10.1038/nplants.2015.139.

Blanchard, E., Birnbaum, P., Ibanez, T., Boutreux, T., Antin, C., Ploton, P., ... Couteron, P., 2016. Contrasted allometries between stem diameter, crown area, and tree height in five tropical biogeographic areas. Trees:1–16 http://dx.doi.org/10.1007/s00468-016-1424-3.

Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., ... Kaasalainen, M., 2015. Nondestructive estimates of above-ground biomass using terrestrial laser scanning. Methods Ecol. Evol. 6:198–208. 10.1111/2041-210X.12301.

Chave, J., Coomes, D., Jansen, S., Lewis, S.L., Swenson, N.G., Zanne, A.E., 2009. Towards a worldwide wood economics spectrum. Ecol. Lett. 12 (4):351–366. http://dx.doi.org/10.1111/j.1461-0248.2009.01285.x.

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M.S., Delitti, W.B.C., ... Vieilledent, G., 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. Glob. Chang. Biol. 20:3177–3190. http://dx.doi.org/10.1111/gcb.12629.

Chen, Q., Baldocchi, D., Gong, P., Kelly, M., 2006. Isolating individual trees in a savanna woodland using small footprint lidar data. Photogramm. Eng. Remote Sens. 72 (8), 923–932.

Colgan, M.S., Asner, G.P., Levick, S.R., Martin, R.E., Chadwick, O.A., 2012a. Topo-edaphic controls over woody plant biomass in South African savannas. Biogeosciences 9 (5):1809–1821. http://dx.doi.org/10.5194/bg-9-1809-2012.

Colgan, M.S., Baldeck, C.A., Féret, J.-B., Asner, G.P., 2012b. Mapping savanna tree species at ecosystem scales using support vector machine classification and BRDF correction on airborne hyperspectral and LiDAR data. Remote Sens. 4 (11), 3462–3480.

Corlett, R.T., 2009. The Ecology of Tropical East Asia. first ed. Oxford University Press, Oxford (book).

Dalponte, M., Reyes, F., Kandare, K., Gianelle, D., 2015. Delineation of individual tree crowns from ALS and hyperspectral data: a comparison among four methods. Eur. J. Remote Sens. 48:365–382. http://dx.doi.org/10.5721/EuJRS20154821.

Dalponte, M., Coomes, D.A., 2016. Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. Methods Ecol. Evol. 7 (10):1236–1245. 10.1111/2041-210X.12575.

Dalponte, M., Ørka, H.O., Ene, L.T., Gobakken, T., Næsset, E., 2014. Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. Remote Sens. Environ. 140:306–317. http://dx.doi.org/10.1016/j.rse.2013.09.006.

Dent, D.H., Bagchi, R., Robinson, D., Majalap-Lee, N., Burslem, D.F.R.P., 2006. Nutrient fluxes via litterfall and leaf litter decomposition vary across a gradient of soil nutrient supply in a lowland tropical rain forest. Plant Soil 288:197–215. http://dx.doi.org/10.1007/s11104-006-9108-1.

DeWalt, S.S.J., Ickes, K., Nilus, R., Harms, K.E.K., Burslem, D.D.F.R.P., 2006. Liana habitat associations and community structure in a Bornean lowland tropical forest. Plant Ecol. 186:203–216. http://dx.doi.org/10.1007/s11258-006-9123-6.

Drake, J.B., Dubayah, R.O., Clark, D.B., Knox, R.G., Blair, J.B., Hofton, M.A., ... Prince, S.D., 2002. Estimation of tropical forest structural characteristics using large-footprint lidar. Remote Sens. Environ. 79 (2–3), 305–319.

Drake, J.B., Knox, R.G., Dubayah, R.O., Clark, D.B., Condit, R., Blair, J.B., Hofton, M., 2003. Above-ground biomass estimation in closed canopy Neotropical forests using lidar remote sensing: factors affecting the generality of relationships. Glob. Ecol. Biogeogr. 12 (2), 147–159.

Duncanson, L.I., Cook, B.D., Hurtt, G.C., Dubayah, R.O., 2014. An efficient, multi-layered crown delineation algorithm for mapping individual tree structure across multiple ecosystems. Remote Sens. Environ. 154, 378–386.

Duncanson, L.I., Dubayah, R.O., Cook, B.D., Rosette, J., Parker, G., 2015. The importance of spatial detail: assessing the utility of individual crown information and scaling approaches for lidar-based biomass density estimation. Remote Sens. Environ. 168: 102–112. http://dx.doi.org/10.1016/j.rse.2015.06.021.

Englhart, S., Keuck, V., Siegert, F., 2011. Aboveground biomass retrieval in tropical forests — the potential of combined X- and L-band SAR data use. Remote Sens. Environ. 115: 1260–1271. http://dx.doi.org/10.1016/j.rse.2011.01.008.

Englhart, S., Jubanski, J., Siegert, F., 2013. Quantifying dynamics in tropical peat swamp forest biomass with multi-temporal LiDAR datasets. Remote Sens. 5:2368–2388. http://dx.doi.org/10.3390/rs5052368.

Eysn, L., Hollaus, M., Lindberg, E., Berger, F., Monnet, J.M., Dalponte, M., et al., 2015. A benchmark of lidar-based single tree detection methods using heterogeneous forest data from the Alpine Space. Forests 6 (5):1721–1747. http://dx.doi.org/10.3390/f6051721.

Feldpausch, T.R., Banin, L., Phillips, O.L., Baker, T.R., Lewis, S.L., Quesada, C.A., ... Lloyd, J., 2011. Height-diameter allometry of tropical forest trees. Biogeosciences 8 (5), 1081–1106.

Ferraz, A., Saatchi, S., Mallet, C., Meyer, V., 2016. Lidar detection of individual tree size in tropical forests. Remote Sens. Environ. 183:318–333. http://dx.doi.org/10.1016/j.rse.2016.05.028.

Fox, J.E.D., 1973. A Handbook to Kabili-Sepilok Forest Reserve, Sabah Forest Record No. 9. Borneo Literature Bureau, Kuching, Sarawak, Malaysia.

Gaveau, D.L.A., Sloan, S., Molidena, E., Yaen, H., Sheil, D., Abram, N.K., ... Meijaard, E., 2014. Four decades of forest persistence, clearance and logging on Borneo. PLoS ONE 9, e101654. http://dx.doi.org/10.1371/journal.pone.0101654.

Goetz, S., Dubayah, R., 2011. Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. Carbon Manag. 2: 231–244. http://dx.doi.org/10.4155/cmt.11.18.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., ... Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. Science 342:850–853. http://dx.doi.org/10.1126/science.1244693.

Houghton, R.A., Byers, B., Nassikas, A.A., 2015. A role for tropical forests in stabilizing atmospheric CO_2. Nat. Clim. Chang. 5:1022–1023. http://dx.doi.org/10.1038/nclimate2869.

Hunter, M.O., Keller, M., Victoria, D., Morton, D.C., 2013. Tree height and tropical forest biomass estimation. Biogeosciences 10:8385–8399. http://dx.doi.org/10.5194/bg-10-8385-2013.

Hyyppa, J., Kelle, O., Lehikoinen, M., Inkinen, M., 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. Geosci. Remote Sens., IEEE Trans. 39:969–975. http://dx.doi.org/10.1109/36.921414.

Jubanski, J., Ballhorn, U., Kronseder, K., Franke, J., Siegert, F., 2013. Detection of large above-ground biomass variability in lowland forest ecosystems by airborne LiDAR. Biogeosciences 10:3917–3930. http://dx.doi.org/10.5194/bg-10-3917-2013.

Jucker, T., Caspersen, J., Chave, J., Antin, C., Barbier, N., Bongers, F., ... Coomes, D.A., 2016. Allometric equations for integrating remote sensing imagery into forest monitoring programs. Glob. Chang. Biol. http://dx.doi.org/10.1111/gcb.13388.

Laurin, G.V., Chen, Q., Lindsell, J.A., Coomes, D.A., Del Frate, F., Guerriero, L., ... Valentini, R., 2014. Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. ISPRS J. Photogramm. Remote Sens. 89:49–58. http://dx.doi.org/10.1016/j.isprsjprs.2014.01.001.

Li, W., Niu, Z., Gao, S., Huang, N., Chen, H., 2014. Correlating the horizontal and vertical distribution of LiDAR point clouds with components of biomass in a *Picea crassifolia* forest. Forests 5:1910–1930. http://dx.doi.org/10.3390/f5081910.

Lopez-Gonzalez, G., Lewis, S.L., Burkitt, M., Phillips, O.L., 2011. ForestPlots.net: a web application and research tool to manage and analyse tropical forest plot data. J. Veg. Sci. 22 (4):610–613. http://dx.doi.org/10.1111/j.1654-1103.2011.01312.x.

Martin, A.R., Thomas, S.C., 2011. A reassessment of carbon content in tropical trees. PLoS ONE 6, e23533. http://dx.doi.org/10.1371/journal.pone.0023533.

Morel, A.C., Saatchi, S.S., Malhi, Y., Berry, N.J., Banin, L., Burslem, D., Nilus, R., Ong, R.C., 2011. Estimating aboveground biomass in forest and oil palm plantation in Sabah, Malaysian Borneo using ALOS PALSAR data. For. Ecol. Manag. 262:1786–1798. http://dx.doi.org/10.1016/j.foreco.2011.07.008.

Morsdorf, F., Meier, E., Kötz, B., Itten, K.I., Dobbertin, M., Allgöwer, B., 2004. LIDAR-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management. Remote Sens. Environ. 92:353–362. http://dx.doi.org/10.1016/j.rse.2004.05.013.

Nabuurs, G.J., Masera, O., Andrasko, K., Benitez-Ponce, P., Boer, R., Dutschke, M., ... Zhang, X., 2007. Forestry. In: Metz, B., Davidson, O.R., Bosch, P.R., Dave, R., Meyer, L.A. (Eds.), Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press:pp. 541–584 http://dx.doi.org/10.2753/JES1097-203X330403.

Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens. Environ. 80:88–99. http://dx.doi.org/10.1016/S0034-4257(01)00290-5.

Nepstad, D.C., Tohver, I.M., Ray, D., Moutinho, P., Cardinot, G., 2007. Mortality of large trees and lianas following experimental drought in an Amazon forest. Ecology 88: 2259–2269. http://dx.doi.org/10.1890/06-1046.1.

Nilus, R., 2004. Effect of Edaphic Variation on Forest Structure, Dynamics, Diversity and Regeneration in a Lowland Tropical Rain Forest in Borneo PhD thesis. University of Aberdeen, UK.

Nilus, R., Maycock, C.R., Majalap-Lee, N., Burslem, D.F.R.P., 2011. Nutrient limitation of tree seedling growth in three soil types found in Sabah. J. Trop. For. Sci. 23, 133–142.

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R.A., Kauppi, P.E., Kurz, W.A., ... Houghton, R.A., 2011. A large and persistent carbon sink in the world's forests. Science 333: 988–993. http://dx.doi.org/10.1126/science.1201609.

Paris, C., Valduga, D., Bruzzone, L., 2016. A hierarchical approach to three-dimensional segmentation of LiDAR data at single-tree level in a multilayered forest. IEEE Trans. Geosci. Remote Sens. 54 (7):4190–4203. http://dx.doi.org/10.1109/TGRS.2016.2538203.

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. Can. J. Remote. Sens. 29 (5), 564–577.

Reitberger, J., Schnörr, C., Krzystek, P., Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform LIDAR data. ISPRS J. Photogramm. Remote Sens. 64: 561–574. http://dx.doi.org/10.1016/j.isprsjprs.2009.04.002.

Réjou-Méchain, M., Tymen, B., Blanc, L., Fauset, S., Feldpausch, T.R., Monteagudo, A., ... Chave, J., 2015. Using repeated small-footprint LiDAR acquisitions to infer spatial and temporal variations of a high-biomass Neotropical forest. Remote Sens. Environ. 169:93–101. http://dx.doi.org/10.1016/j.rse.2015.08.001.

Shendryk, I., Broich, M., Tulbure, M.G., Alexandrov, S.V., 2016. Bottom-up delineation of individual trees from full-waveform airborne laser scans in a structurally complex eucalypt forest. Remote Sens. Environ. 173:69–83 http://doi.org/10.1016/j.rse.2015.11.008.

Shugart, H.H., Asner, G.P., Fischer, R., Huth, A., Knapp, N., Le Toan, T., Shuman, J.K., 2015. Computer and remote-sensing infrastructure to enhance large-scale testing of individual-based forest models. Front. Ecol. Environ. 13:503–511. http://dx.doi.org/10.1890/140327.

Simonson, W., Ruiz-Benito, P., Valladares, F., Coomes, D., 2016. Modelling above-ground carbon dynamics using multi-temporal airborne lidar: insights from a Mediterranean woodland. Biogeosciences 13:961–973. http://dx.doi.org/10.5194/bg-13-961-2016.

Singh, M., Evans, D., Coomes, D.A., Friess, D.A., Suy Tan, B., Samean Nin, C., ... Liu, D., 2016. Incorporating canopy cover for airborne-derived assessments of forest biomass in the tropical forests of Cambodia. PLoS ONE 11, e0154307. http://dx.doi.org/10.1371/journal.pone.0154307.

Spriggs, R., 2015. Robust Methods for Estimating Forest Stand Characteristics Across Landscapes Using Airborne LiDAR. PhD thesis, University of Cambridge.

Sullivan, M.J.P., Talbot, J., Lewis, S.L., et al., 2017. Diversity and carbon storage across the tropical forest biome. Sci. Rep. 7, 39102.

Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkanen, J., ... Maltamo, M., 2012. Comparative testing of single-tree detection algorithms under different types of forest. Forestry 85:27–40. http://dx.doi.org/10.1093/forestry/cpr051.

Vincent, G., Sabatier, D., Blanc, L., Chave, J., Weissenbacher, E., Pélissier, R., ... Couteron, P., 2012. Accuracy of small footprint airborne LiDAR in its predictions of tropical moist forest stand structure. Remote Sens. Environ. 125:23–33. http://dx.doi.org/10.1016/j.rse.2012.06.019.

Vincent, G., Sabatier, D., Rutishauser, E., 2014. Revisiting a universal airborne light detection and ranging approach for tropical forest carbon mapping: scaling-up from tree to stand to landscape. Oecologia 175 (2):439–443. http://dx.doi.org/10.1007/s00442-014-2913-y.

Yao, W., Krzystek, P., Heurich, M., 2012. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. Remote Sens. Environ. 123:368–380. http://dx.doi.org/10.1016/j.rse.2012.03.027.

Yu, X., Hyyppä, J., Vastaranta, M., Holopainen, M., Viitala, R., 2011. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. ISPRS J. Photogramm. Remote Sens. 66:28–37. http://dx.doi.org/10.1016/j.isprsjprs.2010.08.003.

Zanne, A.E., Lopez-Gonzalez, G., Coomes, D.A., Ilic, J., Jansen, S., Lewis, S.L., ... Chave, J., 2009. Global wood density database. Dryad. 235 (Identifier: Http://hdl. Handle.net/10255/dryad).

Zolkos, S.G., Goetz, S.J., Dubayah, R., 2013. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. Remote Sens. Environ. 128:289–298. http://dx.doi.org/10.1016/j.rse.2012.10.017.