A critique of general allometry-inspired models for estimating forest carbon density from airborne LiDAR

Short title: LiDAR estimation of forest carbon density

Rebecca A. Spriggs[1, ¶,*], Mark C. Vanderwel[2], Trevor A. Jones[3], John P. Caspersen[4], David A. Coomes[1, ¶,*]

[1] Department of Plant Sciences, University of Cambridge, Cambridge, UK

[2] Department of Biology, University of Regina, 3737 Wascana Parkway, Regina, Saskatchewan, Canada

[3] Ontario Ministry of Natural Resources and Forestry, Forest Research and Monitoring Section, 1235 Queen Street East, Sault Ste. Marie, Ontario, Canada

[4] Faculty of Forestry, University of Toronto, 33 Willcocks Street, Toronto, Ontario, Canada

[¶] These authors contributed equally to this work.

[*] Corresponding authors e-mail: ras212@cam.ac.uk; dac18@cam.ac.uk

Abstract

There is currently much interest in developing general approaches for mapping forest aboveground carbon density using structural information contained in airborne LiDAR data. The most widely utilized model in tropical forests assumes that aboveground carbon density is a compound power function of top of canopy height (a metric easily derived from LiDAR), basal area and wood density. Here we derive the model in terms of the geometry of individual tree crowns within forest stands, showing how scaling exponents in the aboveground carbon density model arise from the height−diameter (H−D) and projected crown area−diameter (C−D) allometries of individual trees. We show that a power function relationship emerges when the C−D scaling exponent is close to 2, or when tree diameters follow a Weibull distribution (or other specific distributions) and are invariant across the landscape. In addition, basal area must be closely correlated with canopy height for the approach to work. The efficacy of the model was explored for a managed uneven−aged temperate forest in Ontario, Canada within which stands dominated by sugar maple (*Acer saccharum* Marsh.) and mixed stands were identified. A much poorer goodness−of−fit was obtained than previously reported for tropical forests (R^2 = 0.29 vs. about 0.83). Explanations for the poor predictive power on the model include: (1) basal area was only weakly correlated with top canopy height; (2) tree size distributions varied considerably across the landscape; (3) the allometry exponents are affected by variation in species composition arising from timber management and soil conditions; and (4) the C-D allometric power function was far from 2 (1.28). We conclude that landscape heterogeneity in forest structure and tree allometry reduces the accuracy of general power-function models for predicting aboveground carbon density in managed forests. More studies in different forest types

are needed to understand the situations in which power functions of LiDAR height are appropriate for modelling forest carbon stocks.

Keywords: biomass, temperate forest carbon, airborne LiDAR, scaling relationships, crown area−diameter allometry, heterogeneity

Introduction

Aboveground carbon density (ACD) is an important forest property to map in the context of the global carbon cycle [1-3]. Classically, ACD has been estimated using tree size measurements recorded from networks of forest plots, with generalised or species−specific allometries used to convert field measures of diameter and height into tree biomass estimates, and then into ACD estimates [4, 5]. More recently, methods using remote sensing technologies have been developed to complement these plot networks: airborne or spaceborne LiDAR sensors have proven to be particularly effective for estimating ACD because they provide detailed information about forest structure, which is in turn closely related to ACD [6].

There is currently much interest in developing a general method for predicting ACD from LiDAR [7, 8]. A common approach has been to estimate ACD in field plots and then use regression to relate these measurements to various LiDAR metrics [9]. This approach can deliver accurate estimation models within sampling regions, but the models lack physical underpinnings because they are purely empirical. Consequently, they either need to be re-parameterized for each new site, or generalised by estimating how parameters vary geographically. Asner and Mascaro [8] have developed a General Model (henceforth AM−GM) for predicting ACD, which uses measures of the top canopy height derived from LiDAR (H_L), along with local relationships predicting basal area (B_P) and basal−area−weighted mean wood density ($\bar{\rho}_P$):

$$ACD = aH_L{}^{b_1}B_P{}^{b_2}\bar{\rho}_P{}^{b_3} \tag{1}$$

where a, b_1, b_2 and b_3 are parameters estimated by regression using the log−transformed function. Note that subscript L denotes a LiDAR−based measurement, and subscript P a plot−based measurement. Asner and Mascaro [8] argue that this model is analogous to the allometric formula used to calculate an individual tree's biomass from its height H_i, diameter D_i and wood density ρ_i measurements, namely $aH_i^bD_i^c\rho_i^d$ where a, b, c and d vary with forest type [10] and *i* denotes measurements on an individual tree. Fitting the AM−GM to data from four contrasting tropical forests, Asner et al. [7] found that a single, universally fitted relationship reduced model accuracy by no more than 1% relative to regional−specific models. Furthermore, the accuracy was only slightly diminished by replacing plot−level measurements of B_P and $\bar{\rho}_P$ with regional averages and, as a result, the major benefit of their approach is that it requires less additional field data to calibrate than traditional regressions [11].

A key reason why the AM−GM has worked well, where it has, is that basal area and top-of-canopy height were closely correlated in the forests investigated. Asner and Mascaro [8] showed that – for the four tropical forests studied – the AM−GM could be calibrated simply by generating a local relationship estimating B_P from LiDAR and finding a regional $\bar{\rho}_P$ estimate. Others have questioned the generality of the approach [12,13]. In some forest types the correlation between forest height and basal area is weak, especially for mature stands. In these situations two stands can have the same top-of-canopy height, but quite different basal area [14,15].

The problem is that the carbon density of a plot is obtained by summing the biomass of individual trees, but because a tree's biomass is non-linearly related to its dimensions (height, stem diameter), this summation is only exact under certain conditions that we explain below. Although Asner et al. [7] did not claim that the AM−GM could be applied outside the tropics, testing the accuracy of the model across different forest types is important to understanding the applicability and limitations of the general model. For example, tropical and temperate forests have contrasting size structures: rain forests contain shade-tolerant species that develop a dense understory beneath the upper canopy (i.e. stands contain many small trees and few large trees), while temperate forests often lack dense understories and can have unimodal size-frequency distributions [16]. Perhaps for this reason the AM−GM had low goodness−of−fit when applied to broadleaf and coniferous forests in the USA [13], but this has yet to be evaluated critically. Vincent et al. [12] suggest that forests should first be delineated into homogenous regions with respect to the relationships between forest structure and LiDAR data to improve model performance. Unfortunately, this requirement would severely limit the generality of the model.

The aim of this study is to derive the AM−GM from first principles using the geometry of individual trees and, by doing so, to improve understanding of when the AM−GM is likely to yield accurate predictions (i.e., have high goodness−of−fit when applied to data from the field and from LiDAR scanners). Our individual−tree−based general model (ITB−GM) has the same functional form as the AM−GM (1), but its parameters are derived from individual tree allometries and other assumed scaling relationships. We fit the AM−GM to data from an uneven−aged forest in central Ontario, Canada and compare the parameter estimates with those obtained from tree−based measurements using the ITB−GM. By doing so, we explore why the AM−GM has poor predictive ability in this temperate forest. We then examine whether fitting separate models for two forest types within the Canadian dataset leads to significant improvements in goodness−of−fit. Finally, we outline forest conditions that determine the accuracy of the AM−GM.

Theory: An individual−tree−based general model

Consider a tree with stem diameter D_i (in cm), height H_i (in m), vertically projected crown area C_i (in m^2) and wood density ρ_i (in g/cm^3) growing in a plot with an area A_P (in ha). The tree's aboveground biomass can be modelled as $a_1 \pi {D_i}^2 H_i \rho_i$ where a_1 is a species−specific coefficient that depends on crown and stem form. The total aboveground biomass of the plot is found by summing the biomasses of all N_P trees in the plot. ACD is calculated by dividing this biomass value by A_P and multiplying by carbon content a_0 (typically 0.5):

$$ACD = a_2 \sum_{i=1}^{N_p} a_0 a_1 \rho_i {D_i}^2 H_i \qquad (2)$$

where $a_2 = \pi / A_P$. For ease of presentation, the limits of summations are dropped in subsequent equations, but remain the same throughout.

Assuming that a tree's height is related to its diameter by a power function ($H_i = a_H {D_i}^{k_H}$), we get:

$$ACD = a_2 \sum a_0 a_1 a_H \rho_i {D_i}^{2+k_H} \qquad (3)$$

We can use individual tree heights and crown areas to estimate the average top canopy height H_P: this is calculated by summing the crown top height of all trees in the plot, weighted by their crown areas, $H_P = (\sum a_3 C_i H_i)/C_P$ where the canopy area of the plot is $C_P = \sum_{j=1}^{N_p} C_j$ and a_3 is a multiplier that takes into account that the average height of each tree's crown is some fraction of that tree's maximum height [15]. Assuming that crown area is also a power function of stem diameter ($C_i = a_C D_i^{k_C}$), and that $H_i = a_H D_i^{k_H}$ as before, we get:

$$H_P = \frac{1}{C_P} \sum a_3 a_H a_C D_i^{k_C + k_H} \tag{4}$$

Our aim is to substitute (4) into (3) to remove the D_i terms, so that ACD is expressed in terms of H_P, B_P and $\bar{\rho}_P$. However, this is not straightforward for two reasons. The first problem is that a_0, a_1, a_3, a_H, a_C and ρ are inside the summations, but cannot necessarily be moved outside the summations because they are species−specific variables. As an approximation, we represent them by tree−volume−weighted mean values and take them outside of the summation [12] to give $ACD \approx a_2 \bar{a}_0 \bar{a}_1 \bar{a}_H \bar{\rho}_P \sum {D_i}^{2+k_H}$ and $H_P \approx \bar{a}_3 \bar{a}_H \bar{a}_C {C_P}^{-1} \sum D_i^{k_C + k_H}$. The second problem is that D_i is raised to different exponents inside the two summations (except when k_C= 2). In order to progress, we need to assume that the two summations are themselves related by a scaling function: $\sum D_i^{2+k_H} \approx a_D \left(\sum D_i^{k_C + k_H}\right)^{k_D}$; we call this the volume summation scaling relationship. The canopy area can be substituted with basal area by assuming a second scaling function: $C_P \approx a_B {B_P}^{k_B}$; we call this the canopy area scaling relationship. Making these substitutions, we obtain an individual-tree-based general model (ITB−GM):

$$ACD \approx a_4 (H_P)^{k_D} (B_P)^{k_D k_B} \bar{\rho}_P \text{ where } a_4 \approx \bar{a}_0 \bar{a}_1 \bar{a}_H a_2 a_D \left(\frac{a_B}{\bar{a}_3 \bar{a}_H \bar{a}_C}\right)^{k_D} \tag{5}$$

This equation is analogous to the AM−GM, given in (1), with $a = a_4$, $b_1 = k_D$, $b_2 = k_D k_B$ and $b_3 = 1$, but it has more parameters and so is less powerful for predictions.

Our derivation based on tree allometries shows that certain parameters in the AM−GM depend on the exponents of the volume scaling relationship and canopy area scaling relationship. It is important to realise that it would be impossible to derive a function having the form of the AM−GM unless these scaling relationships are valid. In the Supporting Information (S1 Text) we show that these relationships are mathematically valid when tree sizes are precisely power-law or Weibull distributed. If the tree size distributions of all stands across a forest follow one of these functions (with identical parameters), the summation can be replaced by an integral that has an analytical solution. Specifically, if a large number of diameters (D_i,…D_N) are drawn from $p(D)=\alpha D^{-\beta}$ (where α is a normalising constant), then a given power function summation can be approximated by:

$$\sum_{i=1}^{N} D_i^{\gamma} \approx N \int_{D_{min}}^{D_{max}} D^{\gamma} p(D)\, dD = \alpha N \int_{D_{min}}^{D_{max}} D^{\gamma-\beta}\, dD \tag{6}$$

which can in turn be solved to give:

$$\sum_{i=1}^{N} D_i^{\gamma} \approx \frac{\alpha N}{\gamma-\beta+1}\left[{D_{max}}^{\gamma-\beta+1} - {D_{min}}^{\gamma-\beta+1}\right] \tag{7}$$

A similar property holds for a Weibull distribution of tree diameters (see S1 Text). If the power or Weibull distribution is identical across stands, it can be shown that $k_D = k_B = 1$ and a_D and a_B are both predictable.

We now compare the performance of the AM−GM and ITB−GM using data from a temperate forest, to gain a better understanding of when these models are appropriate for estimating ACD from LiDAR data.

Materials and methods

Study area and inventory dataset

We used datasets from Haliburton Forest and Wildlife Reserve in central Ontario, Canada (45º13’N, 78º35’W). The forest is managed using selection silviculture and consists mostly of uneven−aged stands [17]. Sugar maple (*Acer saccharum* Marsh.) is the most prevalent species, but a number of other species are common, including eastern hemlock (*Tsuga canadensis* (L.) Carrière), balsam fir (*Abies balsamea* (L.) Mill.) and American beech (*Fagus grandifolia* Ehrh.). There were 154 circular plots inventoried across the forest each with an area of 2500 m^2. The plot locations were chosen to stratify the variation across the forest. The stem diameters of all trees with a stem diameter equal to or greater than 8 cm were recorded along with their species identity. The plots were randomly split into a calibration (114 plots) and a validation dataset (40 plots). The calibration dataset was used for fitting the models and relationships, whilst the validation dataset was reserved for assessing model performance.

ACD was estimated for each plot using species−specific allometric equations developed for Canadian inventories, which relate stem diameter to aboveground tree biomass [18, 19]. Species−specific equations were used for the seven most prevalent species and then generic conifer and broadleaf equations were used for all remaining species (~ 17% of total trees). The individual tree aboveground biomasses were summed for each plot and converted to a per hectare estimate; this aboveground biomass estimate was then multiplied by the carbon content of wood (0.5; [20]) to estimate ACD. Wood density estimates were extracted from [21] and represent the oven dry mass divided by green volume. To parameterise the LiDAR models (AM−GM and ITB−GM), wood density was summarised as a volume−weighted average for each plot ($\bar{\rho}_P$). Finally, we succinctly described the tree size distribution of each plot by calculating the quadratic mean diameter (QMD) as $200\sqrt{A_P B_P/(\pi N_P)}$, and by fitting a Weibull distribution to the list of stem diameters.

Airborne LiDAR

The LiDAR data were collected using an Optech ALTM 3100 four–pass system flown in August 2009 (altitude = 1500 m; pass overlap = 30%; pulse density = 2 pulses/m^2). The dataset consisted of x, y and z coordinates (converted to the height above the ground by subtracting the digital elevation model) with up to four returns recorded from a single pulse. We used discrete–return airborne LiDAR data clipped in ArcGIS 10 to overlay the inventoried plots, which had been georeferenced to sub–metre accuracy using a Trimble Geo XH 6000. The LiDAR metrics used in the analyses were H_L and gap fraction (G_L) (

Table **1**). We split each plot into 1 m by 1 m tiles and extracted the maximum recorded height of pulses in each of those tiles. H_L was calculated as the mean of the tile heights that were recorded at 2 m and above, which excluded the tiles where LiDAR pulses were not intercepted by the canopy. G_L was calculated as the proportion of first returns recorded at a height less than 2 m above the ground.

Table 1 Definitions of all terms and parameters in the AM−GM and ITB−GM.

Term	**Definition**	**Units**
Lidar metrics		
H_L	Top canopy height	m
G_L	Gap fraction	No units
Tree level measurements		
ρ_i	Wood density	$Mg\ m^{-3}\ \equiv\ g\ cm^{-3}$
D_i	Diameter	cm
B_i	Basal area	m^2
H_i	Stem height	m
C_i	Crown area	m^2
Plot based measurements		
ACD	Aboveground carbon density	$Mg\ C\ ha^{-1}$
$\bar{\rho}_P$	Mean wood density (weight by relative abundances of species)	$Mg\ m^{-3}\ \equiv\ g\ cm^{-3}$
C_P	Canopy area ($C_P = \sum_{j=1}^{N} C_j$)	m^2
A_P	Plot area	ha
N_P	Total number of stems in a plot	No units
H_P	Average top canopy height	m
B_P	Basal area	$m^2\ ha^{-1}$
QMD	Quadratic mean diameter	cm
Model parameters		
a_0	Carbon content of trees	
a_1	Coefficient related to crown and stem form	
a_2	Factor scaling stem diameter to plot level basal area	
a_3	Average crown height as a proportion of tree height	
a_4	Coefficient in final ACD equation which amalgamates other coefficients	
$\bar{a}_0, \bar{a}_1, \bar{a}_3$	Means of a_0, a_1, and a_3, weighted by tree volumes	
a_H, k_H	Coefficient and exponent of scaling relationship between stem diameter and height (H−D)	
a_C, k_C	Coefficient and exponent of scaling relationship between stem diameter and crown area (C−D)	
a_D, k_D	Coefficient and exponent of scaling relationship between two summations of stem diameter raised to different powers (volume scaling relationship)	
a_B, k_B	Coefficient and exponent of scaling relationship between canopy area and basal area (canopy area scaling relationship)	

Forest types from aerial photography

The study area was classified into two forest types using aerial photographs (captured by an ADS52 Leica camera). The photographs were manually delineated into 42 forest types using standard methods developed by Ontario's Forest Resources Inventory programme [22]. We reduced the number of forest types to just two according to estimated species composition: stands dominated by sugar maple, and mixed stands that contained a significant coniferous component alongside sugar maple (see [23] for further details on the method used).

Fitting the AM−GM to the Canadian data

The log−transformed AM−GM was fitted using least squares regression to ACD measured in the calibration plots:

$$\ln ACD = \ln a + b_1 \ln H_L + b_2 \ln B_P + b_3 \ln \bar{\rho}_P \tag{8}$$

Predicted ACD values included a $e^{MSE/2}$ multiplier (where MSE is the mean square error of the regression) to correct for a bias introduced by the log transformation [24]. B_P and $\bar{\rho}_P$ were estimated from relationships with LiDAR so that the model could be used to predict ACD outside of the measured plots. We compared the accuracy of models based on LiDAR estimates of B_P and $\bar{\rho}_P$ against models where B_P and $\bar{\rho}_P$ were ground measurements, to quantify the loss in accuracy as a result of this estimation approach.

We measured the accuracy of the 40 validation plot predictions of the ACD model and the B_P and $\bar{\rho}_P$ equations using the coefficient of determination (R^2):

$$R^2 = 1 - \frac{\sum_{j=1}^{40}(P_j - O_j)^2}{\sum_{j=1}^{40}(O_j - \bar{O})^2} \tag{9}$$

where the observed and predicted value for each plot is denoted by O_j and P_j, respectively, and the overall mean observed value is denoted by $\bar{O}$. We compared model support using the Akaike information criterion (AIC) where k is the number of estimated parameters and L is the maximised likelihood function:

$$AIC = 2k - 2\ln(L) \tag{10}$$

We also calculated the percentage root mean square error (% RMSE) which is normalised using the mean of the observed values:

$$\% \, RMSE = \frac{100}{\bar{O}} \sqrt{\frac{\sum_{j=1}^{40}(P_j - O_j)^2}{40}} \tag{11}$$

Estimating the parameter values of the ITB−GM from tree level information

Exponents k_B and k_D of the ITB−GM equation ($ACD \approx a_4 {H_L}^{k_D} {B_P}^{k_D k_B} \bar{\rho}_P$) are derived from the volume summation and canopy area scaling relationships. To estimate these, we first estimated allometric scaling exponents k_H and k_C from dimensional measurements of 5436 trees at a site 230 km from the study area [25]. We calculated the relative abundances of species within the 114 calibration plots (Table S1), then drew 500 trees at random from the height and crown radius dataset such that the species composition of the sample was the same as observed in the plots. Power functions were then fitted to the height vs. diameter and crown area vs. diameter relationships for these 500 trees. The fitted power functions gave values for k_H and k_C that were representative of the species composition in our study area. Exponent k_D (of the volume scaling relationship) was estimated by calculating $\log(\sum D_i^{2+k_H})$ and $\log(\sum D_i^{k_C+k_H})$ for each of the 114 calibration plots, and then fitting a power function through these data. Similarly, exponent k_B of the canopy area scaling relationship was estimated by calculating $\log(C_P)$ and $\log(B_P)$ for each of the 114 calibration plots, and then fitting a power function through these data. Theoretically, a_4 in the ITB−GM could be calculated as $\bar{a}_0\bar{a}_1\bar{a}_H a_2 a_D {a_B}^{k_D}(\bar{a}_3\bar{a}_H\bar{a}_C)^{-k_D}$ but in practice several of these variables are hard to determine. For this reason, a_4 was estimated by linear regression: we fit log(ACD) as a linear function of log H_L, log B_P and log $\overline{\rho}_P$ with the coefficients associated with these explanatory variables fixed at the values calculated from individual−tree−based information, such that only a_4 was estimated.

Testing whether forest type information improves model accuracy

To explore whether incorporating forest type information improved the predictive power of the estimation model, we split the plots into sugar maple and mixed stands using the aerial photographs and repeated the same procedures as above for fitting AM−GM and ITB−GM. Forest type was incorporated into both of these models and into the equations estimating B_P and $\overline{\rho}_P$ from H_L and G_L.

Results

Predicting temperate forest biomass using general power-law models

A summary of the coefficients and goodness−of−fit estimates of the AM−GM (1) fitted to the Canadian temperate forest dataset are provided in Table 2. The coefficient of the log($\bar{\rho}_P$) term was not significantly different from zero, so we set the power (b_3) to 1 to match the ITB−GM. The resulting model performed relatively poorly, as the R^2 of the fit to the validation plots was only 0.18. Fitting the model with ground−measured B_P and $\bar{\rho}_P$ increased the R^2 to 0.41, but unfortunately B_P was poorly predicted from LiDAR estimates of H_L and G_L (R^2 = 0.09; Table 3), and $\bar{\rho}_P$ was unrelated to the LiDAR metrics (Figs 1, 2). As a result, we found that ACD could be estimated using the AM−GM with relatively low accuracy (22.5% RMSE; equivalent to a RMSE of 15.7 Mg C ha^{-1}; Fig 3).

Table 2. Aboveground carbon density (ACD) estimation models fit to a Canadian temperate forest dataset containing sugar maple and mixed broadleaf-conifer stands. Parameters shown in bold were estimated from individual tree data, while all other parameters were estimated using least-squares regression of calibration plot data. The AIC gives the relative performance of the models and the R^2 denotes the fit to the validation plots: 1) using ground measured B_P and $\bar{\rho}_P$ and 2) using LiDAR estimated B_P and $\bar{\rho}_P$.

Model type	ACD estimation equation	AIC	1) ground B_P and $\bar{\rho}_P$ R^2	2) LiDAR B_P and $\bar{\rho}_P$ R^2
Asner and Mascaro's General Model (AM−GM)				
All stands	$5.11 H_L^{0.271} B_P^{0.808} \bar{\rho}_P$	947.8	0.405	0.179
Sugar maple stands	$2.99 H_L^{0.258} B_P^{0.991} \bar{\rho}_P$	944.4	0.453	0.292
Mixed stands	$10.1 H_L^{0.258} B_P^{0.616} \bar{\rho}_P$			
Individual Tree Based General Model (ITB−GM)				
All stands	$0.285\, H_P^{\mathbf{1.24}} B_P^{\mathbf{0.870}} \bar{\rho}_P$	1009.6	−0.111	−0.213
Sugar maple stands	$0.552\, H_P^{\mathbf{1.15}} B_P^{\mathbf{0.729}} \bar{\rho}_P$	1002.5	−0.088	−0.330
Mixed stands	$0.314\, H_P^{\mathbf{1.22}} B_P^{\mathbf{0.867}} \bar{\rho}_P$			

Table 3. Basal area and wood density estimation equations obtained by least squares regression. Explanatory variables were LiDAR metrics top canopy height (H_L) and gap fraction (G_L) and forest type derived from aerial photographs in the sugar maple and mixed stand specific equations. The AIC gives the relative performance of the models and the R^2 denotes the fit to the validation plots.

Response variable	**Estimation equations**	**AIC**	**R^2**
Basal area			
B_P (all stands)	$14.2 + 0.871\ \mathrm{H_L} - 29.4\ \mathrm{G_L}$	728.5	0.093
B_P (sugar maple stands)	$4.83 + 1.21\ \mathrm{H_L} - 20.3\ \mathrm{G_L}$	666.2	0.286
B_P (mixed stands)	$12.5 + 1.21\ \mathrm{H_L} - 20.3\ \mathrm{G_L}$		
Volume−weighted mean wood density			
$\bar{\rho}_P$ (all stands)	0.533	−307.0	−0.022
$\bar{\rho}_P$ (sugar maple stands)	0.576	−364.3	0.188
$\bar{\rho}_P$ (mixed stands)	0.497		

Fig 1. Relationships between field-measured biophysical properties (basal area B_P and wood density $\bar{\rho}_P$) and LIDAR metrics (top-of-canopy height H_L and gap fraction G_L). The lines are predictions from multiple regression analyses of data from all sites (solid), mixed stands (dashed) and sugar maple (dotted). For panels (a) and (c), the predicted lines are obtained by holding G_L constant at its mean value, whilst for panels (b) and (d) the value of H_L was held at its mean value.

Fig 2. Predictions made for the validation plots by multiple-regression models for basal area (left column) and volume weighted wood density (right column) with: a) no species information and b) forest types.

Fig 3. LiDAR vs ground estimated ACD in 40 validation plots, where LiDAR estimates are based on Asner and Mascaro's general model (AM−GM; first column) and the individual tree based general model (ITB−GM; second column). The first row gives the fit of the AM−GM and ITB−GM to the 40 validation plots (AM−GMa and ITB−GMa) and the second row gives the fit of the models fitted separately to 24 sugar maple and 16 mixed−species stands (AM−GMb and ITB−GMb). The overall RMSE value for each model version is given in the bottom right corner of the plot and the individual RMSE for the sugar maple (Mh) plots and mixture plots (Mix) are given above the plot.

Including forest type into the B_P and $\bar{\rho}_P$ estimation models led to increased goodness−of−fit (R^2 rose from 0.09 to 0.29 in the B_P models and from −0.02 to 0.19 in the $\bar{\rho}_P$ equations; Table 3) and was strongly supported by AIC (B_P: Δ = 62.3; $\bar{\rho}_P$: Δ = 57.3). The % RMSE of the B_P estimator fell from 23.3 to 20.7% and that of $\bar{\rho}_P$ from 11.7 to 10.4% (Fig. 2b). The mixed−forest plots had higher basal area and lower wood density than the sugar maple plots (Fig 1). Incorporating forest type improved overall performance of the AM−GM with the R^2 rising from 0.18 to 0.29 (RMSE: 20.9 vs. 22.5%), with moderate AIC support (Δ = 3.4).

Estimating the exponents of individual-tree-based generalised model (ITB−GM)

The ITB−GM model, which fixed the values of model parameters based on the field-measured allometries of individual trees, performed less well than the Asner-Mascaro model in which the parameters were estimated by regression. The exponents of ITB−GM estimated from the fitted allometric powers of the H−D and C−D relationships are presented in Table 4 and the fitted relationships are presented in Fig 4. For all stands, height and crown area were fitted as power functions of diameter, with exponents of 0.521 and 1.28 respectively. The log−log regression relationship between summed stem volume ($\sum D_i^{2+k_H}$) and the maximum canopy volume ($\sum D_i^{k_C+k_H}$) had a higher goodness−of−fit (R^2 = 0.814) than the log−log regression relationship between canopy area (C_P) and basal area (B_P) (R^2 = 0.654) indicating that the volume scaling relationship was better supported than the canopy area scaling relationship.

Table 4. Estimates of power function parameters of relationships between (a) height vs diameter; (b) crown area vs diameter; (c) summed diameters raised to 2 different powers (see text; crown volume scaling relationship); (d) basal area vs canopy area (canopy area scaling relationship).

Model version	**(a) H_i vs D_i**			**(b) C_i vs D_i**			**(c) $\sum D_i^{2+k_H}$ vs $\sum D_i^{k_C+k_H}$**		**(d) C_P vs B_P**	
	a_H	k_H	R^2	a_C	k_C	R^2	k_D	R^2	k_B	R^2
All stands	3.26	0.521	0.593	0.465	1.28	0.419	1.24	0.814	0.701	0.654
Sugar maple stands	3.89	0.476	0.634	0.898	1.10	0.431	1.15	0.659	0.632	0.503
Mixed stands	3.73	0.466	0.503	0.397	1.29	0.378	1.22	0.813	0.711	0.676

Fig 4. Height−diameter power relationships are given in the left panel whilst the crown area−diameter power relationships are given in the right panel. The exponents from these fitted power functions are used to estimate the powers in the ITB−GM model (Table 4): top row for all stands, middle row for sugar maple stands and the bottom row for mixed stands.

Both scaling relationships contained residual error and had exponent values different from 1 because our set of plots did not follow a single diameter distribution (Fig 5). Although the Weibull distributions that we fit showed that stem diameters were monotonically decreasing in most calibration plots, quadratic mean diameter ranged from 13 to 33 cm across the plots. Plots with a higher QMD generally had a higher top canopy height as measured by LiDAR. In the Supporting Information (S1 Text, Figs S1-S3), we provide a comprehensive analysis of how variation in tree diameter distributions affects model fit for a range of different H−D and C−D scaling relationships.

Fig 5. Weibull distributions of tree diameters in each calibration plot. The rug plot along the x-axis shows the quadratic mean diameter of each plot, coloured according to top canopy height. The left panel represents sugar maple stands; the right panel represents mixed stands.

The ITB−GM, with exponents fixed at their theoretical values and a_4 fitted by linear regression is given in Table 2. The exponent associated with B_P was similar in the two models (AM−GM: 0.81 vs ITB−GM: 0.87), but the exponent associated with H_L differed greatly (0.271 vs 1.24). The ITB−GM model had a RMSE of 27.3%, indicating it is less able to explain variance in biomass than the AM−GM (22.5%).

The best predictions were obtained by using the AM−GM and including forest type information (RMSE: sugar maple: 15.8%; mixture: 25.5%). The exponent of the H_L term in the AM−GM was unaffected by forest type, but the B_P exponent of mixed stands was much lower than the sugar maple exponent (0.616 vs 0.991; Table 2). Including forest type led to greater improvements in the fit of the ITB−GM than that of the AM−GM (ΔAIC = 7.1 vs. ΔAIC = 3.4). However, the predictions to the validation plots of the ITB−GM were slightly less accurate (RMSE: no forest types = 27.3%; forest type = 28.6%). In all versions of the model, the sugar maple plots were predicted more accurately than the mixture plots.

Discussion

Deriving the AM−GM from individual tree measurements has revealed the origins of its parameters, the assumptions behind the power function formula, and the situations in which it is unlikely to make accurate predictions. Below, we explore specific explanations for low goodness−of−fit, including that (1) the basal area and wood density of plots are not closely correlated with top canopy height or gap fraction as measured by LiDAR; (2) tree size distributions are not conserved across the landscape; and (3) the exponents of the allometries are affected by systematic changes in species composition, and the exponent of the crown area allometry deviates from 2. Our findings suggest that among–stand variability in structure and composition are key factors in determining the accuracy of the AM−GM.

Basal area is weakly correlated with height

Basal area is a key element of allometry-inspired models for estimating forest carbon. It is not directly measured from LiDAR, but instead is inferred indirectly from other height metrics [23]. The goodness–of–fit of the AM−GM for this Canadian forest was substantially reduced when ground–measured B_P and $\bar{\rho}_P$ were replaced with LiDAR estimates (R^2 = 0.41 vs 0.18) and therefore LiDAR offered a poor substitute for ground data on these quantities. Predictions of B_P from LiDAR metrics were weak in our study area (R^2 = 0.09; Fig 1) compared with that reported by Asner et al. [7] for tropical forests ($R^2 \geq 0.55$), although the fit was improved by splitting the plots into two forest types (R^2 = 0.29). Improving the accuracy of the LiDAR–based models of B_P may therefore require other metrics than H_L and G_L to be included in regression relationships [9, 12] or the application of individual-based approaches [14].

Tree size distributions vary across the landscape

Basal area is fundamentally linked with the stem diameter distribution, and variability in this distribution weakens the correlation between canopy height and basal area. When the stem diameter distribution follows either a power- or truncated-Weibull function and is conserved across a landscape, then the volume summation and crown area scaling relationships are exact and the exponents of the AM−GM all reduce to 1 (see S2 Fig**.**1 and derivation in S1 Text). However, when the underlying diameter distributions vary among stands, the exponents relating these quantities will deviate from 1 and the accuracy of the relationships will decrease (Fig S3).

The AM−GM is therefore likely to be less accurate in forests where there is large variability in tree size distributions.

Why are size distributions more variable in temperate forests than in natural tropical forests?

Size distributions of forests are linked to size−dependent growth and mortality [26], and can be similar across forested landscapes if these demographic functions remain constant over space and time [27, 16]. This may be a reasonable assumption in old-growth tropical forests where size distributions are often close to power functions with exponents of roughly −2 [28, but see 27]. Temperate forests are often managed and comprise a patchwork of stands at different stages of recovery following disturbance (natural or human). Temperate forest size distributions tend to be more variable [29] and are often modelled by a Weibull distribution with the flexibility to fit both unimodal and power function−type distributions [16]. The selection−managed forests considered here are uneven−aged, and exhibit varying tree size distributions as a legacy of their management history. Our analyses suggest that assumptions of the AM−GM are compromised in structurally heterogeneous forests, and that this model is not expected to produce high goodness−of−fits in such areas. In our particular study area, changing management practices over time have produced a wide range of diameter distributions, which in turn have weakened the accuracy of the AM−GM.

Wood density is very weakly correlated with LiDAR-measured height

LiDAR and RADAR measure forest structure, but not wood density. Predictive models can give rise to markedly different maps of ACD depending on the assumed spatial variation in wood density [36]. Wood density ($\bar{\rho}_P$) was even less well predicted ($R^2 = -0.02$) from LiDAR than basal area, but was improved by separating the landscape into forest types (R^2 = 0.19) because conifer and broadleaf species vary in wood density. There is no evidence in our derivation, or from previous work [10, 38], that $\bar{\rho}_P$ should have an associated power in the AM−GM, even though the model has commonly been fitted with an $\bar{\rho}_P$ exponent included [7, 10]. Consistent with theory, we found that including the $\bar{\rho}_P$ exponent (b_3) did not lead to significant improvements in model fit in our temperate data.

Influences of crown area allometry on goodness of fit

The exponent of the C−D relationship, k_C, can also affect accuracy. When k_C = 2, the powers in the ITB−GM all reduce to 1, total stem volume is directly proportional to the maximum canopy volume and canopy area is directly proportional to basal area. The AM−GM is therefore most accurate when k_C = 2; conversely, the further k_C departs from 2, the more inaccurate the volume and crown area summation scaling relationships become (see S1 Text for a detailed exploration). Even with variable size distributions, the goodness−of−fit of the total stem volume vs canopy volume relationship is high (R^2 > 0.8) when k_C is greater than 1.3. There is a sharp drop off in the accuracy of the volume scaling relationship if the C−D exponent is less than 1.3 (Fig S1), and the AM−GM is expected to perform poorly in forests with variable size distributions when the C−D exponent has a lower value. Since k_C was 1.28 for the Canadian temperate forest, the crown area allometry also contributed to low model accuracy.

We lack a clear picture of how k_C varies globally, but there is some evidence that values are lower for temperate forests. Classical self−thinning theory was based on an assumption of an exponent of 2 [30, 31], whereas metabolic scaling theory predicts an exponent of 4/3 [32], both above the threshold of 1.3 below which accuracy deteriorates. An average value of k_C = 1.36

was found for tropical forests [30], whereas a wide range of k_C values have been reported for temperate forests (0.85 for Virginia, USA, [33]; 1.19 for European beech, [34]; 2.16 for New Zealand mountain beech, [31]). Competition amongst the trees becomes an important feature determining crown shape and the C-D exponent [31] and that too varies at different scales. The goodness−of−fits of the C−D power functions in our analyses were low ($R^2 < 0.45$), suggesting that uneven−aged stands may require a variable relationship between height and diameter, which would consequently require an alternative formulation of the AM−GM. Dietze et al. [35] found that the C−D scaling relationship was more variable than the H−D relationship for two managed temperate forest sites in North Carolina, USA.

The H−D scaling exponent, k_H, has less influence on the ITB−GM than k_C, as it only contributes to the volume scaling relationship and appears on both sides of this equation. The magnitude of k_H affects the accuracy of the power function by influencing the relative magnitude of the summations; increasing k_H would mitigate the effects of k_C deviating from 2 (Fig S1).

Influences of forest composition on power-law exponents and goodness of fit

Changes in forest composition within a landscape can have major effects on ACD estimates if those changes are associated with systematic variation in crown geometry and wood density [12, 36]. In our study area, the model was not substantially improved when forest type was accounted for (Fig 3), but an examination of its assumptions highlighted some combinations of H−D and C−D exponents where forest type could influence the generality of the model (Fig S1). Given that the AM−GM is based on scaling relationships of individual trees (H−D and C−D), it is clear that species composition may be important if it results in changes to these allometric functions across the landscape. Previous studies indicate that H−D and C−D power functions vary with site and species, suggesting that AM−GM exponents will vary across heterogeneous landscapes. The inclusion of forest type improved the ACD predictions of the sugar maple stands more than the mixed stands. Delineation of the sugar maple forest type, which essentially represents a single species, may therefore have been beneficial because there is expected to be more variation in allometry between species than within species. Lines *et al.* [37] noted that the H−D relationships of Spanish conifer species had exponents close to 2/3 (the value predicted by biomechanical theory), but those of broadleaf species were much more variable and often less than 2/3 [34]. Such differences between conifers and broadleaves could result in different AM−GM exponents across forests with shifting species dominance.

Conclusion

The allometry-inspired AM-GM model appears to predict forest carbon more reliably in tropical forests than in temperate ones. Asner and Mascaro [8] achieved a goodness−of−fit of $R^2 = 0.83$ compared with $R^2 = 0.18$ in this study, even though the models were identical (Table 2). Their RMSE was 9% of the mean ACD compares with 23% for our models (Fig 3). Duncanson et al. [14] also observed poor model performance when testing the AM−GM in two out of three temperate forest sites in the USA ($R^2 = 0.13$, 0.18 and 0.73).

A key issue is that stand basal area is weakly correlated with canopy height in temperate landscapes comprised of patchworks of stands at various stages of succession/development after disturbance. Selection management created a variety of structural conditions in the Canadian forests studied here, whereas in natural temperate forests variation in stand structure is induced by disturbance from wind, disease, fire and pests. Variability in regeneration, growth and mortality among these stands leads to weak correlations between basal area and height – whereas these are closely coupled in many tropical forests [7]. The allometry-inspired model is reliant on predicting basal area from height, which is a particular problem in heterogeneous landscapes.

Deriving the AM−GM from individual tree information further underscores the importance of variability in size distributions across landscapes. Given that a tree's biomass is obtained by multiplying its wood volume by its wood density (and assuming conical form), the values of b, c and d in the individual biomass model function $aH_i^b\ D_i^c\ \rho_i^d$ should be close to 1, 2 and 1, respectively [10, 38]. By analogy we would expect b_1, b_2 and b_3 to all be approximately 1 in the AM−GM if the summation had no effect on exponents; however, two of the exponents are far from 1 for the tropical forests analysed by Asner and Mascaro [8] (b_1 = 0.28, b_2 = 0.97 and b_3 = 1.38). Non-linearities in the process of scaling from trees to stands are clearly influential in determining these exponents. This also explains why our ITB-GM was ineffective.

This paper has described the theoretical basis of the AM-GM, demonstrating that the reliability of the approach is dependent on having invariant size distributions across landscapes and on the crown area-diameter power relationship of individual trees. Landscape heterogeneity in these attributes resulted in the poor performance of the AM-GM in a managed temperate system compared with species-rich tropical forests. Model performance is improved by stratification into forest types, but this does not address the issue of varying size distributions. More studies into the spatial variability of tree size distribution are needed to understand when allometry-inspired general models can be reliably used to predict forest aboveground carbon stocks.

References

1. Avitabile V, Herold M, Heuvelink GBM, Lewis SL, Phillips OL, Asner GP, et al. An integrated pan-tropical biomass map using multiple reference datasets. Glob Chang Biol. 2016;22: 1406–1420.
2. Mitchard ETA. The tropical forest carbon cycle and climate change. Nature. 2018;559: 527–534.
3. Pan Y, Birdsey RA, Fang J, Houghton R, Kauppi PE, Kurz WA, et al. A Large and Persistent Carbon Sink in the World's Forests. Science. 2011;333: 988–994.
4. Chave J, Réjou-Méchain M, Búrquez A, Chidumayo E, Colgan MS, Delitti WBC, et al. Improved allometric models to estimate the aboveground biomass of tropical trees. Glob Chang Biol. 2014;20: 3177–3190.
5. Smith JE, Heath LS, Woodbury PB. How to estimate forest carbon for large areas from inventory data. J Forest. 2004; 102(5): 25–31.
6. Wulder MA, White JC, Nelson RF, Næsset E, Ørka HO, Coops NC, et al. Lidar sampling for large-area forest characterization: A review. Remote Sens Environ. 2012;121: 196–209.
7. Asner GP, Mascaro J, Muller-Landau HC, Vieilledent G, Vaudry R, Rasamoelina M, et al. A universal airborne LiDAR approach for tropical forest carbon mapping. Oecologia. 2012; 168: 1147–1160.
8. Asner GP, Mascaro J. Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. Remote Sens Environ. 2014; 140: 614–624.
9. Bouvier M, Durrieu S, Fournier RA, Renaud J-P. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. Remote Sens Environ. 2015; 156: 322–334.

10. Chave J, Andalo C, Brown S, Cairns MA, Chambers JQ, Eamus D, et al. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia. 2005; 145: 87–99.

11. Kalliovirta J, Laasasenaho J, Kangas A. Evaluation of the laser-relascope. Forest Ecol Manag. 2005; 204: 181–194.

12. Vincent G, Sabatier D, Rutishauser E. Revisiting a universal airborne light detection and ranging approach for tropical forest carbon mapping: scaling-up from tree to stand to landscape. Oecologia. 2014; 175: 439–43.

13. Jucker T, Asner GP, Dalponte M, Brodrick PG, Philipson CD, Vaughn NR, et al. Estimating aboveground carbon density and its uncertainty in Borneo's structurally complex tropical forests using airborne laser scanning. Biogeosciences. 2018; 15: 3811-3830.

14. Duncanson LI, Dubayah RO, Cook BD, Rosette J, Parker G. The importance of spatial detail: Assessing the utility of individual crown information and scaling approaches for lidar-based biomass density estimation. Remote Sens Environ. 2015; 168: 102–112.

15. Kent R, Lindsell J, Laurin G, Valentini R, Coomes D. Airborne LiDAR detects selectively logged tropical forest even in an advanced stage of recovery. Remote Sens. 2015; 7: 8348–8367.

16. Coomes DA, Allen RB. Mortality and tree-size distributions in natural mixed-age forests. J. Ecol. 2007; 95, 27-40.

17. Vanderwel MC, Thorpe HC, Shuter JL, Caspersen JP, Thomas SC. Contrasting downed woody debris dynamics in managed and unmanaged northern hardwood stands. Can J For Res. 2008; 38: 2850–2861.

18. Lambert M, Ung C, Raulier F. Canadian national tree aboveground biomass equations. Can J For Res. 2005; 35: 1996–2018.

19. Ung CH, Bernier P, Guo XJ. Canadian national biomass equations: new parameter estimates that include British Columbia data. Can J For Res. 2008; 38: 1123–1132.

20. Intergovernmental Panel on Climate Change (IPCC). Vol. 4: Agriculture, forestry and other land use. In: Eggleston S, Buendia L, Miwa K, Ngara T, Tanabe K. Guidelines for National Greenhouse Gas Inventories. Institute for Global Environmental Strategies; 2006.

21. Gonzalez JS. Wood Density of Canadian Tree Species. Information Report – Northwest Region, Forestry Canada. 1990; NOR-X-315.

22. Ontario Ministry of Natural Resources. Ontario Forest Resources inventory photo interpretation specifications. 2009. Available from: http://www.ontario.ca/environment-and-energy/forest-resources-inventory

23. Spriggs RA, Vanderwel MC, Jones TA, Caspersen JP, Coomes DA. A simple area-based model for predicting airborne LiDAR first returns from stem diameter distributions: an example study in an uneven-aged, mixed temperate forest. Can J For Res. 2015; 45: 1338–1350.

24. Baskerville GL. Use of logarithmic regression in the estimation of plant biomass. Can J For Res. 1972; 2: 49–53.

25. Caspersen JP, Vanderwel MC, Cole WG, Purves DW. How stand productivity results from size- and competition-dependent growth and mortality. PLoS ONE. 2011; 6: e28660.

26. Coomes DA, Duncan RP, Allen RB, Truscott J. Disturbances prevent stem size-density distributions in natural forests from following scaling relationships. Ecol Lett. 2003; 6: 980–989.

27. Muller-Landau HC, Condit RS, Harms KE, Marks CO, Thomas SC, Bunyavejchewin S, et al. Comparing tropical forest tree size distributions with the predictions of metabolic ecology and equilibrium models. Ecol Lett. 2006; 9: 589–602.
28. West GB, Enquist BJ, Brown JH. A general quantitative theory of forest structure and dynamics. Proc Natl Acad Sci USA. 2009; 106: 7040–7045.
29. Duncanson LI, Dubayah RO, Enquist BJ. Assessing the general patterns of forest structure: quantifying tree and forest allometric scaling relationships in the United States. Glob Ecol Biogeogr.; 2015;24: 1465–1475.
30. Muller-Landau HC, Condit RS, Chave J, Thomas SC, Bohlman SA, Bunyavejchewin S, et al. Testing metabolic ecology theory for allometric scaling of tree size, growth and mortality in tropical forests. Ecol Lett. 2006; 9: 575–588.
31. Coomes DA, Holdaway RJ, Kobe RK, Lines ER, Allen RB. A general integrative framework for modelling woody biomass production and carbon sequestration rates in forests. J Ecol. 2012; 100: 42–64.
32. Enquist BJ, West GB, Brown JH. Extensions and evaluations of a general quantitative theory of forest structure and dynamics. Proc Natl Acad Sci USA. 2009; 106: 7040–7045.
33. Anderson-Teixeira KJ, McGarvey JC, Muller-Landau HC, Park JY, Gonzalez-Akre EB, Herrmann V, et al. Size-related scaling of tree form and function in a mixed-age forest. Funct Ecol. 2015; 29: 1587-1602.
34. Pretzsch H, Dieler J. Evidence of variant intra- and interspecific scaling of tree crown structure and relevance for allometric theory. Oecologia. 2012; 169: 637–649.
35. Dietze MC, Wolosin MS, Clark JS. Capturing diversity and interspecific variability in allometries: A hierarchical approach. Forest Ecol Manag. 2008; 256: 1939–1948.

36. Mitchard ETA, Feldpausch TR, Brienen RJW, Lopez-Gonzalez G, Monteagudo A, Baker TR, et al. Markedly divergent estimates of Amazon forest carbon density from ground plots and satellites. Glob Ecol Biogeogr. 2014; 23: 935–946.

37. Lines ER, Zavala MA, Purves DW, Coomes DA. Predictable changes in aboveground allometry of trees along gradients of temperature, aridity and competition. Glob Ecol Biogeogr. 2012; 21: 1017–1028.

38. Schumacher FX, Hall FDS. Logarithmic expression of timber-tree volume. J Agric Res. 1933; 47: 719–734.

Supporting information captions

S1 Text. Assessing the validity of the volume and canopy area scaling relationships.

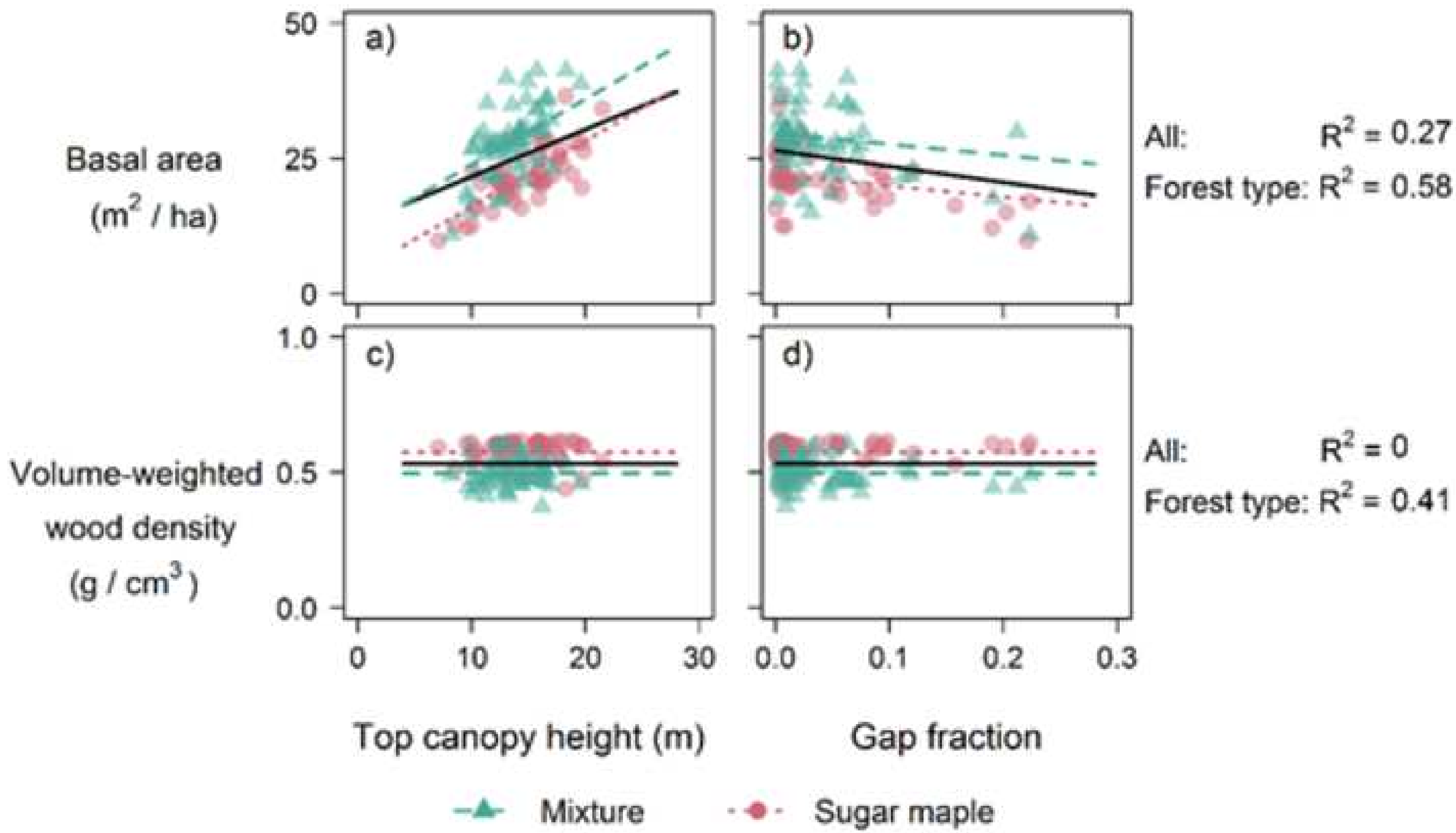
S1 Table. Species compositions extracted from the 114 calibration plots. Trees from the tree height and crown area dataset were sampled to match these compositions.

S1 Fig. Goodness−of−fit with different values of k_H and k_C when substituting top canopy height (H_L) and basal area (B_P) into the ITB−GM (5) using the relationships modelled by the volume and canopy area scaling relationships. Parameter values for the modelled relationships are also given. For particular values of k_H and k_C, each matrix cell represents a relationship fitted to the 114 calibration plots. The square matrices give the power (k_D), coefficient (a_D) and R^2 of the relationship in the volume scaling relationship, whilst the bars give the equivalent (k_B, a_B and R^2) for the canopy area scaling relationship. In the square matrices, both k_H and k_C vary, whilst only the latter affects the bars. Points represent the values of k_H and k_C estimated from allometric data (Table 4).

S2 Fig. Goodness−of−fit of the scaling relationships when underlying size distributions follow a power function or Weibull distribution. The square matrices represent R^2 values for the volume scaling relationship and the bars represent R^2 values for the canopy area scaling relationship as the exponent parameters of H−D and C−D are varied. The leftmost and centre panels represent pseudo−data plots that exhibit a power function and a Weibull distribution, respectively. The rightmost panels show the difference in R^2 for each combination of $\mathrm{k_H}$ and $\mathrm{k_C}$.

S3 Fig. Exponent values and goodness−of−fit of the volume summation and crown area assumptions as the H−D and C−D relationships are varied and as the Weibull stem diameter distributions become more variable. The exponent of the volume scaling relationship when the Weibull parameters were changed to produce low and high variance is given in the top row. The difference in R^2 between each of the variable Weibull datasets and the fixed Weibull is given in the bottom matrices, where the square matrices correspond to the volume scaling relationship and the bars correspond to the canopy area scaling relationship.

S1 Dataset. Plot data used for main analyses.


Basal area
(m² / ha)
50
25
0
a)
b)
All: R² = 0.27
Forest type: R² = 0.58
1.0
0.5
0.0
Volume-weighted
wood density
(g / cm³)
c)
d)
All: R² = 0
Forest type: R² = 0.41
0
10
20
30
0.0
0.1
0.2
0.3
Top canopy height (m)
Gap fraction
Mixture
Sugar maple

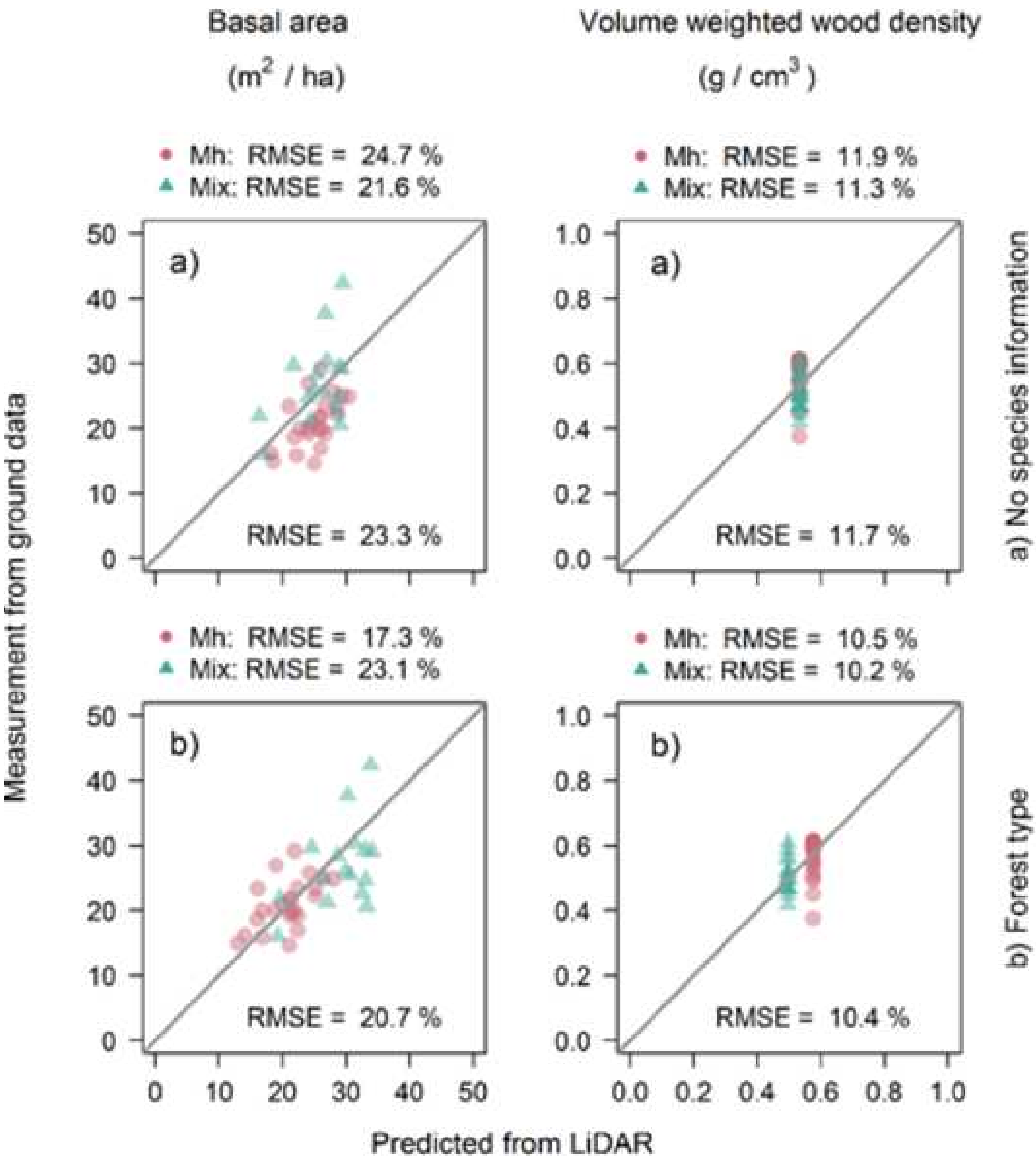
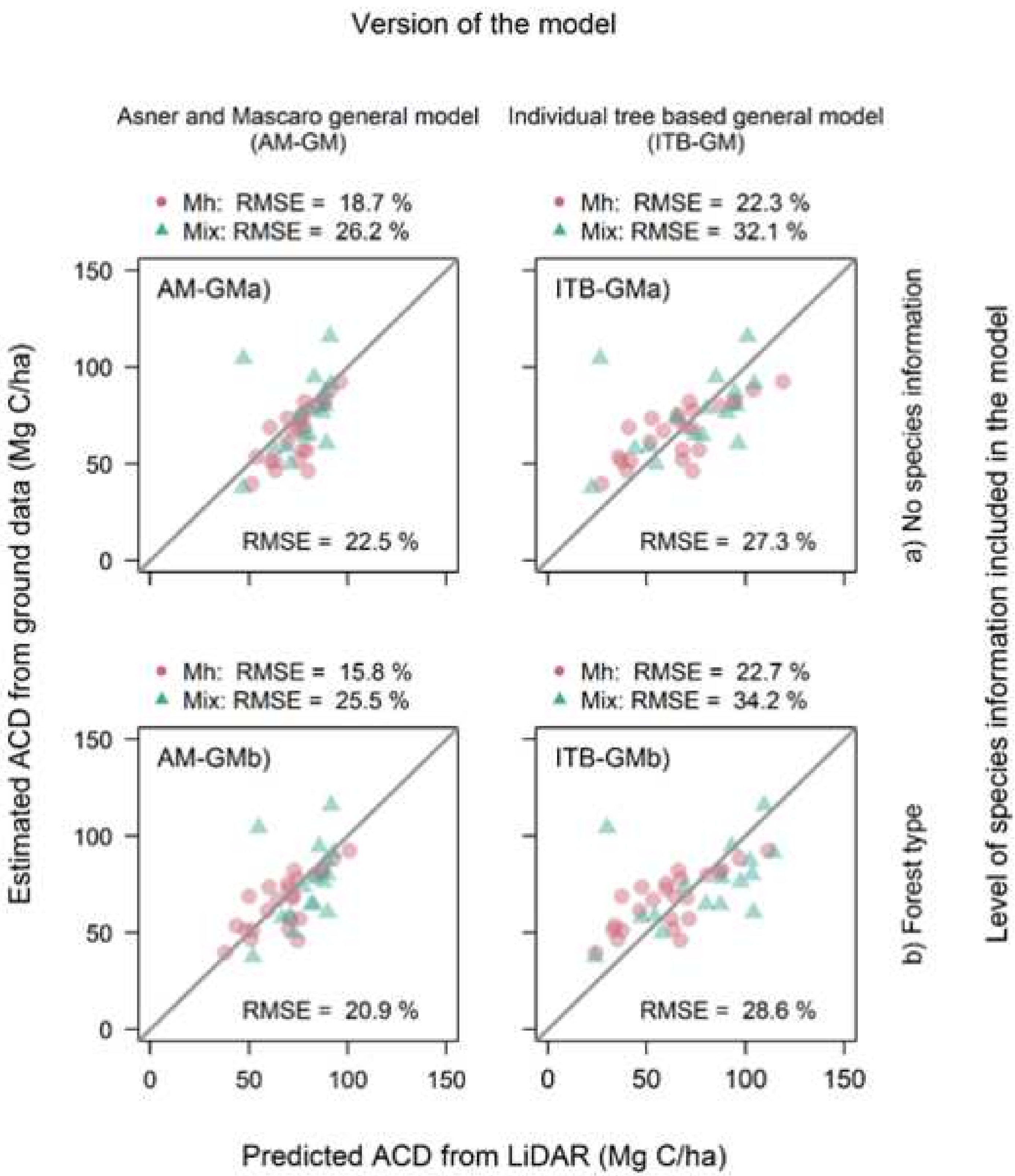
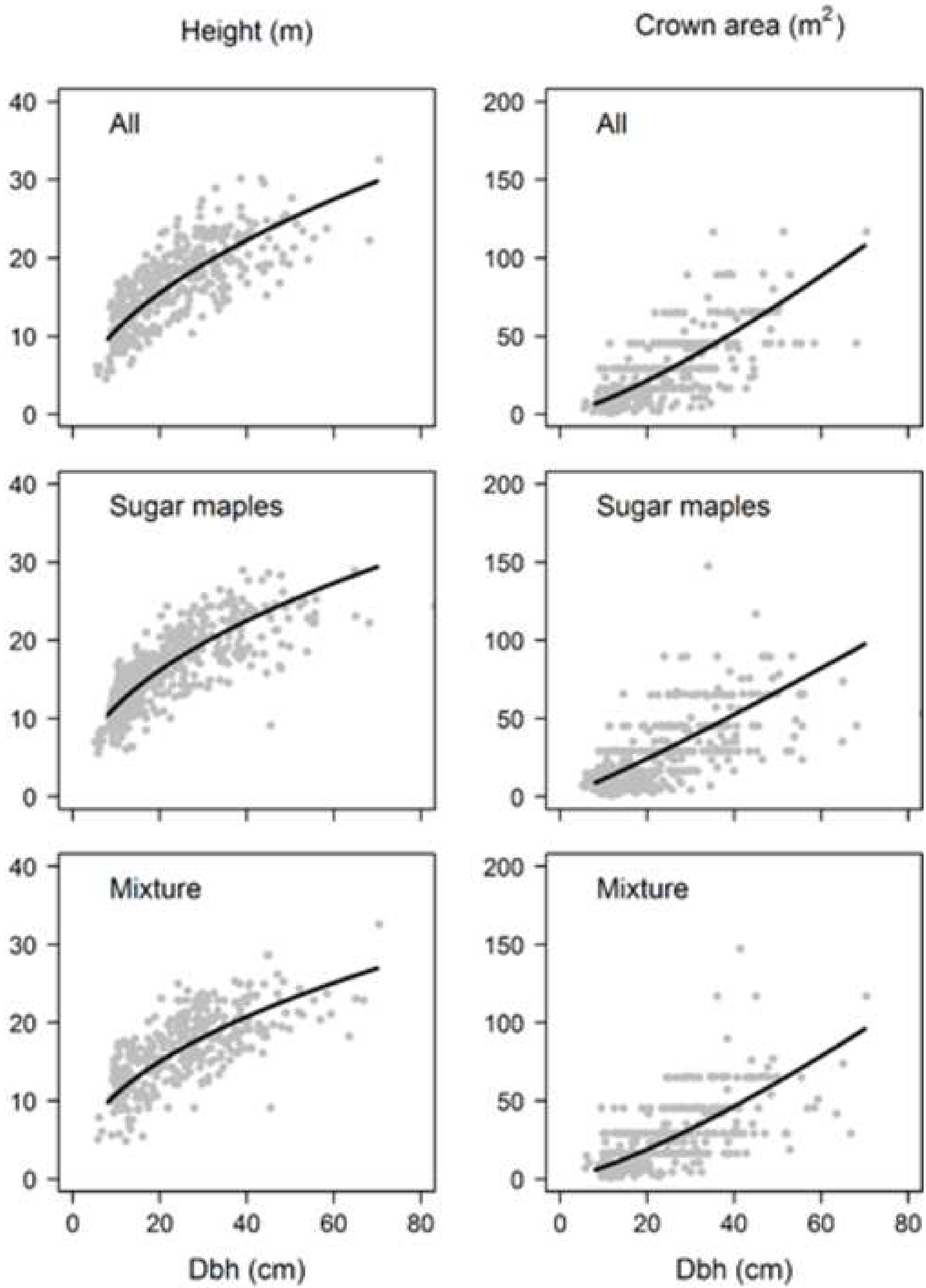

Basal area
(m² / ha)
Volume weighted wood density
(g / cm³)
Mh: RMSE = 24.7 %
Mix: RMSE = 21.6 %
a)
RMSE = 23.3 %
Mh: RMSE = 11.9 %
Mix: RMSE = 11.3 %
a)
RMSE = 11.7 %
a) No species information
Mh: RMSE = 17.3 %
Mix: RMSE = 23.1 %
b)
RMSE = 20.7 %
Mh: RMSE = 10.5 %
Mix: RMSE = 10.2 %
b)
RMSE = 10.4 %
b) Forest type
Measurement from ground data
Predicted from LiDAR

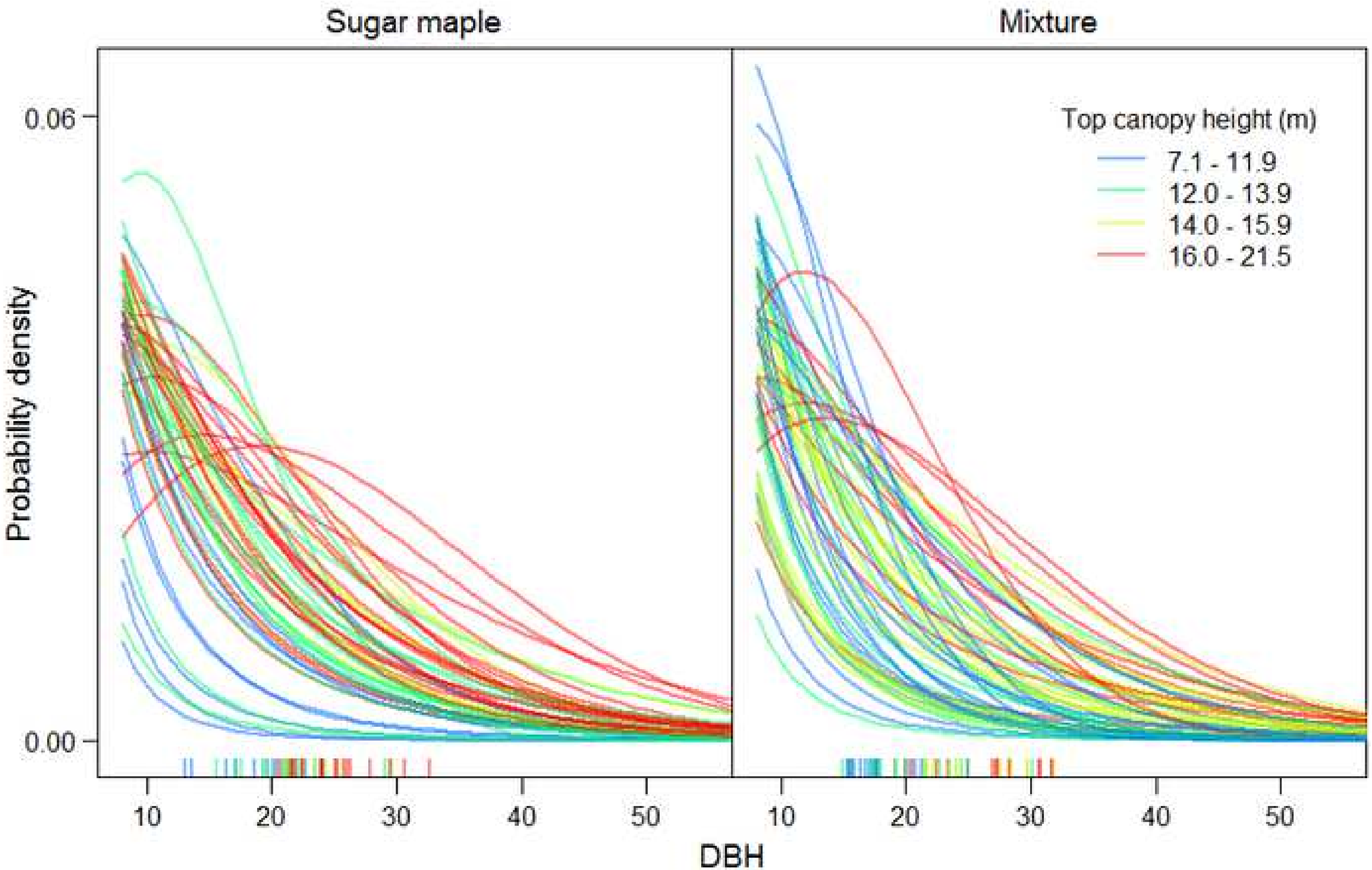
Figure Click here to access/download;Figure;Fig 3.png



Figure

Click here to access/download;Figure;Fig 4.png




Sugar maple
Mixture
Top canopy height (m)
7.1 - 11.9
12.0 - 13.9
14.0 - 15.9
16.0 - 21.5
0.06
0.00
Probability density
10
20
30
40
50
DBH

Figure

Click here to access/download
Supporting Information
S1 Text.docx

Click here to access/download
Supporting Information
Table S1.docx