Remote Sensing of Environment 113 (2009) 182–196

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers

Kaiguang Zhao [a,*], Sorin Popescu [a], Ross Nelson [b]

[a] Spatial Sciences Lab, Dept. of Ecosystem Science and Management, Texas A&M University, College Station, Texas 77801, USA
[b] Biospheric Sciences Branch Code 614.4, NASA's Goddard Space Flight Center Greenbelt, Maryland 20771, USA

ARTICLE INFO

Article history:
Received 30 December 2007
Received in revised form 30 July 2008
Accepted 16 September 2008

Keywords:
Lidar
Biomass
Carbon
Scale
Scale-invariant
Airborne laser
Forest inventory
Functional regression

ABSTRACT

Researchers in lidar (Light Detection And Ranging) strive to search for the most appropriate laser-based metrics as predictors in regression models for estimating forest structural variables. Many previously developed models are scale-dependent that need to be fitted and then applied both at the same scale or pixel size. The objective of this paper is to develop methods for scale-invariant estimation of forest biomass using lidar data. We proposed two scale-invariant models for biomass: a linear functional model and an equivalent nonlinear model that use lidar-derived canopy height distributions (CHD) and canopy height quantile functions (CHQ) as predictors, respectively. The two models are called functional regression models because the predictors CHD and CHQ are themselves functions or functional data. The model formulation was justified mathematically under moderate assumptions. We also created a fine-resolution biomass map by mapping individual tree component biomass in a temperate forest of eastern Texas with a lidar tree-delineation approach. The map was used as reference data to synthesize training and test datasets at multiple scales for validating the two scale-invariant models. Results suggest that the models can accurately predict biomass and yield consistent predictive performances across a variety of scales with an R^2 ranging from 0.80 to 0.95 (RMSE: from 14. 3 Mg/ha to 33.7 Mg/ha) among all the fitted models. Results also show that a training data size of around 50 plots or less was enough to guarantee a good fitting of the linear functional model. Our findings demonstrate the effectiveness of CHD and CHQ as lidar metrics for estimating biomass as well as the capability of lidar for mapping biomass at a range of scales. The functional regression models of this study are useful for lidar-based forest inventory tasks where the analysis units vary in size and shape. They also hold promise for estimating other forest characteristics such as below-ground biomass, timber volume, crown fuel weight, and Leaf Area Index.

1. Introduction

Accurate estimation of biomass is essential for better understanding the carbon cycles over terrestrial ecosystems where forests serve as a primary reservoir of terrestrial carbon (Houghton, 2005). Monitoring changes in forest biomass over time has both environmental and economical significance as exemplified by the establishment of the Kyoto Protocol. In the context of global warming and change, quantification of carbon sequestration in forests provides insights into the relevant biogeochemical processes such as vegetation responses to elevated CO_2 levels and environmental impacts of anthropogenic activities (Houghton et al., 2001). A growing need for spatially-explicit mapping of forest biomass is being partially compensated by recent advances in remote sensing technologies. In particular, Light Detection And Ranging (lidar), with the capability of directly measuring vegetation structure, has brought a breakthrough in remotely inventorying forest resources, and therefore provides a superior choice for remote sensing of forest above-ground biomass (AGB) as compared to optical sensors that suffer from saturation in spectral response to dense canopies with high biomass (Lefsky et al., 2002).

Discrete-return lidar data acquired by small-footprint airborne laser scanners (ALS) consist of a collection of spatially-distributed points, each representing an intercepted target along the illumination path of laser pulses. Each lidar return (hit) is stored as a coordinate triple of (x, y, z) and often tagged with auxiliary variables such as the intensity and scan-angle. Thus far, the mainstream lidar research has focused on the use of lidar coordinate triple information while only few studies partially investigated the utility of lidar auxiliary variables (Brandtberg et al., 2003; Holmgren, 2004; Popescu and Zhao, 2008). Previous efforts demonstrate the potential of lidar for estimating forest structural attributes at different levels (Næsset, 1997; Nelson et al., 2003a; Omasa et al., 2003). At individual tree levels, many algorithms were proposed to delineate single trees from lidar data for estimating tree height and/or crown dimensions. Most results suggest that lidars tend to underestimate tree height due to a large probability of missing treetops even with a high sampling density (Popescu et al.,

* Corresponding author.
E-mail addresses: zhaokg@tamu.edu (K. Zhao), s-popescu@tamu.edu (S. Popescu), Ross.F.Nelson@nasa.gov (R. Nelson).

0034-4257/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2008.09.009

2002; Yu et al., 2004; Chen et al., 2006). Plot- or stand-level estimates of forest characteristics from lidar (e.g., basal area, AGB, and timber volume) have also been reported for various forest types and conditions (Næsset, 2004; Riano et al., 2004; Andersen et al., 2005; Maltamo et al., 2006; Mutlu et al., 2008; Zhao, Popescu, & Nelson, 2008a).

Standard practice in establishing lidar-based models for estimating plot-level forest attributes involves the use of regression analysis for relating some carefully-selected lidar metrics to the spatially coincident in-situ measurements that are often prorated if not temporally concomitant with laser data. Upon validation, these regressed models will be applied to the rest of the lidar data for prediction (Næsset and Bjerknes, 2001). The use of lidar for AGB generally follows this two-stage procedure where ground reference biomass is obtained by destructive sampling or more often with the recourse to allometric equations. Past work demonstrates promising results in estimating AGB with lidar (Lefsky et al., 1999; Means et al., 1999; Nelson et al., 2004). However, the lidar metrics and regression models reported in these studies often lack commonalities. The reported lidar metrics mainly include mean, maximum and median canopy height, quadratic mean canopy height, quantile heights and etc. These metrics have been used alone or combined in linear models or nonlinear models (Lim et al., 2003). Nelson et al. (1988) found, for example, that the logarithmic model with mean height from profiling laser data as predictor accounted for the most variation in ground-measured AGB. Popescu et al. (2003) reported two plot-level linear biomass models for pine and deciduous plots, respectively; the significant predictors of the pine model are the lidar-derived mean and maximum crown diameter, and those of the deciduous model are the lidar-derived maximum tree height and mean crown diameter. Also, using ALS data, Lim and Treitz (2004) explored laser canopy-based quantiles in a log-transformed single-predictor linear model for AGB, and they concluded that all laser-based quantiles are equally capable of predicting biomass if the tree allometry in a forest remains the same.

The importance of scale in remotely sensing the surface biophysical parameters has long been recognized, partly due to the incapability of sensors to directly measure the heterogeneity at sub-pixel scales (Strahler et al., 1986; Woodcock and Strahler, 1987). From either an applied or philosophical perspective, scale issues arise virtually from all disciplines of the earth sciences (Oreskes et al., 1994). As far as remote sensing theories and applications are concerned, scale problems often reveal themselves by a range of incompatibilities among relevant scales such as those at which sensors make measurements, those at which variables of interest are defined physically and observed *in-situ*, those at which remote sensing application models are formulated and applied, and those at which retrieval algorithms together with the involved remote sensing forward and inversion models are formulated and implemented. Moreover, ad-hoc image processing procedures such as resampling and filtering often introduce some compounding scale effects (Turner et al., 1989; Marceau et al., 1994; Marceau, 1999). To compensate the scale discrepancies, practical applications often require devising some viable scaling-up or scaling-down schemes. In addition, for a specific application there often exists an optimal scale that may not be valid to other applications (Marceau et al., 1994).

In the context of lidar remote sensing of forest structures, most estimation models in previous studies are likely to be not only site- or species-specific but also scale-dependent, which indicates that the models are fitted using data collected at a given plot size and should be applied at a scale or pixel (cell) size commensurate to the plot size used in the model fitting (Næsset, 2002). Patenaude et al. (2004) explicitly noted the scale-dependence of lidar approaches in a ground carbon study by observing that using the same exponential model with the 80th lidar height percentile as the predictor, the total AGB calculated by applying the model to each cell is different from the value obtained by using the single 80th percentile of the whole area. They attributed this discrepancy to the nonlinearity of the exponential model (Patenaude et al., 2004). To be more precise, even if the model is linear with a lidar percentile height as the only predictor, such scaling effect could still exist due to the nonadditivity of the percentiles from two canopy height distributions. In general, the scale-dependence of lidar-based prediction models is revealed by a simple observation that the aggregation of predicted values at individual cells is not equal to the single prediction based on the single predictor over the aggregated cells. This scaling effect may result from one or more of the nonadditive properties of predictors, models, and response variables when changing the scale (Oreskes et al., 1994).

One way to quantify AGB using lidar with minimal scaling effects is to inventory forest biomass at individual tree levels (Popescu, 2007). AGB at scales above tree levels will be immediately available by integrating tree-level results up to the desired scale. Apparently, such a method relies on the successful delineation of trees. Considering the powerful capacities of modern computers, computational demands by tree delineation should be of little concern even for extensive areas. However, a possible concern is the incapability and inefficiency of many tree-segmentation algorithms to identify understory and suppressed trees, or to delineate trees under certain canopy conditions such as dense forests and grouped trees (Maltamo et al., 2000). This limitation probably explains why little literature reports on the use of the tree segmentation and integration approach for mapping plot or stand-level AGB over large areas. In some cases, this approach can be effective because the main contributor to AGB under certain forest conditions is those dominant trees that could be identified from lidar by tree-delineation algorithms (Zavitkovski, 1976).

The main objective of this study is to explore the possibility of developing scale-invariant prediction models of AGB using lidar data. To this end, we proposed two scale-invariant biomass models: a linear functional model that uses the lidar-derived canopy height distribution (CHD) as the predictor, and a nonlinear functional model equivalent to the first one that instead uses canopy height quantile function (CHQ) as the predictor. The two models are called functional regression models because the predictors CHD and CHQ are themselves functions or functional data. We provided a theoretical justification for formulating these biomass models, and evaluated the models based on a series of multi-scale training and test data sets that were synthesized from a realistic fine-resolution biomass map over a forested area in eastern Texas, USA with a lidar individual tree-based method.

2. Materials

2.1. Study area

A 4800-ha forested region in eastern Texas, USA, was chosen for this study (Fig. 1). The airborne laser coverage consists of pine plantations in various developmental stages, old growth pine stands in the Sam Houston National Forest, many of which with a natural pine stand structure, and upland and bottomland hardwoods. Much of the southern U.S. is covered by forest types similar to those of our study area, with similar species, productivity, and land use/change patterns.

2.2. Field measurements

Fieldwork was undertaken during May–Jun. 2004 on 62 randomly sampled circular plots, including twenty-six 0.01-ha plots within young unthinned pine plantations and thirty-six 0.1-ha plots. A total of 1004 trees were tallied with respect to height, crown width, crown base height (CBH), diameter at breast height (DBH), species, and crown class (Kraft class). Tree height was measured using a Vertex Forester hypsometer; DBH was measured using a diameter tape; crown width was calculated as the average of four perpendicular radii measured from the bole of a tree towards each cardinal direction, and crown class was determined as one of four categories, i.e., dominant,

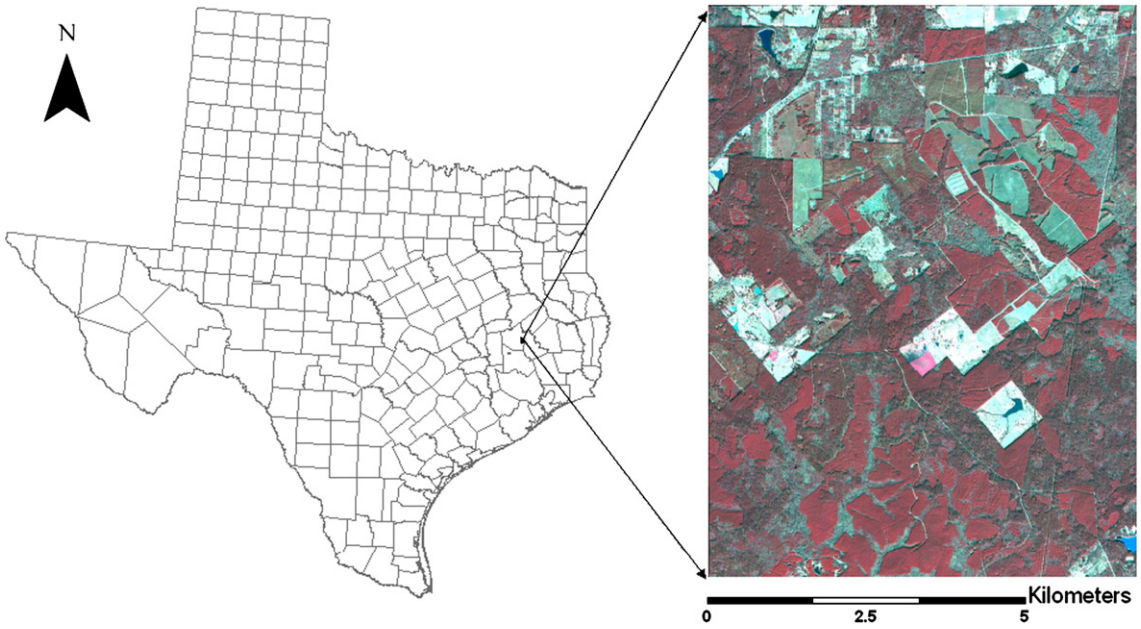


Fig. 1. The study area (right: the Quickbird image) is a forested region located in eastern Texas, USA (left).

co-dominant, intermediate, and overtopped (USDA Forest Service FIA National Core Field Guide, 2005, p. 78). The 261 overtopped trees were excluded for this study because they only intercepted a limited number of laser pulses and it is difficult to reconstruct their structures from lidar data. Of the 743 trees left, 504 are Loblolly pines (*Pinus taeda* L.), and 239 are deciduous trees such as water oak (*Quercus nigra* L.), red oak (*Quercus falcata* Michx), sweetgum (*Liquidambar styraciflua* L.), and post oak (*Quercus stellata* Wangenh.).

Tree locations were indirectly calculated according to trigonometric relations using the geographic coordinates of plot centers and the azimuth and distance of tree boles relative to the plot center that have been measured by a Differential GPS, a Suunto compass (KB-14), and the Vertex hypsometer, respectively. Thus, the mapped tree locations refer to boles, not treetops that may deviate from tree boles. All of these factors made it difficult to match field-measured trees with lidar-detected trees identified by the tree segmentation algorithm. As a result, only 117 out of 743 field trees, including 94 pines and 23 deciduous trees were matched to lidar-detected trees with high confidence (Popescu and Zhao, 2008). The descriptive statistics of the matched trees are summarized in Table 1.

2.3. Lidar data and multispectral imagery

ALS data were collected with a Leica-Geosystems ALS40 during the leaf-off season in March 2004 by M7 Visual Intelligence of Houston, Texas. The lidar system was operated to record two returns per pulse, i.e., first and last, with a reported accuracy of 20–30 cm and 15 cm for horizontal and vertical positioning, respectively, and was configured to scan +/−10° from nadir. The dataset features a full coverage from two perpendicular directions with 19 and 28 flight lines in the north–south and the east–west direction, respectively. The average point density is 2.6 laser hits per m^2. A digital elevation model (DEM) derived from the lidar data using a proprietary package was also delivered by the data vendor. A Quickbird image acquired in the leaf-off season of 2004 is available for the study area. The image, with a spatial resolution of 2.5 m, was classified to differentiate between pines and deciduous trees as well the non-forest class with an overall accuracy of 89.7%.

Table 1
Descriptive statistics of the 117 field-measured trees that have been correctly matched with lidar-detected trees

	DBH (cm)	Height (m)	Crown diameter (m)	Height to crown base (m)
Minimum	5.23	8.83	0.83	3.2
Maximum	78.49	37.49	13.27	25.9
Range	73.26	28.66	12.44	22.7
Standard deviation	17.21	6.91	2.68	4.45
Average	30.99	19.98	5.94	11.76

3. Methodology

3.1. Overview

The primary goal of this study is to formulate lidar-based scale-invariant biomass models that are fitted at a given scale but could be applied at the same or different scales with minimal scaling effects (Fig. 2). We also expect that these models, either during fitting or application, are less sensitive to shapes of analysis units (e.g., plot shape). Obviously, the validation of such models entails a data set of ground-reference biomass at continuous, or at least, multiple scales (plot sizes). Such type of field data, however, is not directly available in most studies including this investigation.

To obtain such reference data of biomass at various scales for model evaluation, we resorted to a lidar individual tree-based approach to create a 0.5×0.5 m fine-resolution biomass map that served as a reliable substitute to field data. This AGB map covers the full extent of the study area and can be used to synthesize reference AGB values at a desired scale (plot size) by properly aggregating up to that scale. These synthesized plot-level AGB data were used to train and test (i.e., calibrate and validate) the scale-invariant biomass models.

For clarification purposes, we summarize the reasons why we prefer such a fine-resolution map to the *in-situ* measurements for model evaluation.

(1) All of the field plots are fixed in size and are circular in shape, thus failing to provide ground-reference biomass at various scales or to study the effects of shapes of analysis units, while the lidar-derived AGB map allows for simulating the biomass observations at a variety of plot sizes or for different plot shapes.
(2) The use of lidar-derived AGB map allows generating a considerable number of reference values at a given scale,

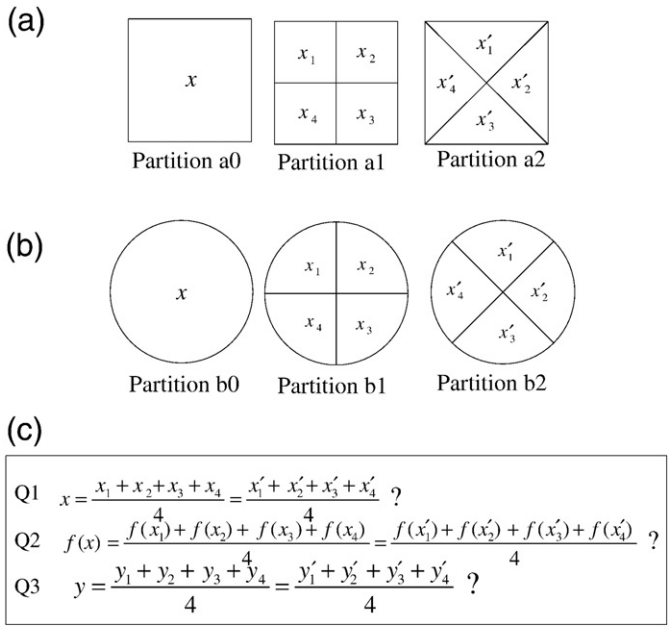


Fig. 2. Potential limitations of remote sensing-based retrieval models ($\hat{y}=f(x)$) used for estimating biophysical variables y from remotely-sensed variables x on a per-analysis-unit basis (e.g., pixel): Two hypothetical study areas, (a) one square and (b) another circular, have been partitioned with different types of analysis units (AU). Often enough, x (e.g., some lidar metrics) is non-additive in that the equalities of Q1 in (c) may not hold due to scale-dependence (e.g., P. a0 vs. P. a1, or P. b0 vs. P. b1), shape-dependence (e.g., P. a1 vs. P. a2 where AUs of the two partitions are equal in size but different in shape), or translation-dependence (P. b1 vs. P. b2 where AUs of P. b1 are translated or rotated in the angular direction to arrive at P. b2). In terms of their physical definitions, most biophysical variables y such as LAI and biomass are additive so that the equalities of Q3 hold, but some canopy attributes such as mean tree height and Lorey's height are not additive when scaling. Depending on the properties of x, $f(\cdot)$ and y, the remote sensing-based estimates ($\hat{y}$) may be scale-, shape- or translation-dependent. Note that Partitions a2, b1 and b2 are given here only for pedagogic purposes.

which, therefore, produces enough data points for model validation as opposed to a limited number of field plots (i.e., 62 plots).

(3) Lidar metrics that will be related to AGB in the scale-invariant models are extracted from the same raw lidar data source as that of the lidar-derived biomass map. This allows for seamless co-registration between independent and dependent variables, thereby, avoiding the mis-registration problem that usually is a concern when pairing field-geolocated plots with airborne-geolocated lidar canopy height model (CHM) especially at individual tree levels.

(4) The training and test data for model evaluation were generated in the same manner through the random splitting of the synthesized data. This consistency makes model validation more viable because the uncertainties caused by disparities of data sources between training and testing are precluded.

The rest of the "Methodology" section is organized as follows: First, we briefly describe the individual tree-based approach to creating the reference fine-resolution AGB map (Fig. 3); next, we formulate two scale-invariant AGB models and also justify them within a theoretical and methodological framework; finally, we describe practical schemes for fitting the models.

3.2. Mapping individual tree biomass

Individual tree biomass was mapped using a lidar-based individual tree delineation approach together with regression analysis (upper panel, Fig. 3). The purpose here is two-fold: (1) to introduce this individual tree-based method as an effective alternative to map AGB, and (2) more importantly, to generate a fine-resolution AGB map for synthesizing reference biomass values to fit and test the scale-invariant models proposed later.

3.2.1. Tree location, height and crown diameter from CHM

The raw lidar data were first processed to create a digital surface model (DSM) by selecting the highest lidar hit within each 0.5×0.5 m cell and interpolating the selected hits into a regular grid of 0.5×0.5 m. Then, a CHM was generated by computing the pixel-wise difference between DSM and DEM. Next, we applied the TreeVaw software to the CHM to identify individual trees (Popescu and Wynne, 2004). The software implements a local maximum filtering with a variable circular window. The window diameter is locally determined and sized adaptively according to the CHM height at the window center, assuming that a taller tree has a wider crown. The output from this software is a list of identified trees that includes tree location, tree height and crown width.

3.2.2. Individual tree crown base height from lidar height-bin products

Lidar height-bin data were a multi-band product derived by first discretizing the 3D spatial domain into a 3D array of voxels with a prescribed horizontal and vertical (height-bin) resolutions and then counting the laser hits that fall within each voxel. Unlike CHM, the height-bin product makes use of the full set of laser hits, and proves useful for analyzing 3D forest structures. Moreover, height-bin data can be analyzed as multi-band images where each horizontal slice at a given height-bin is treated as a single-band image. In this investigation, we used a voxel dimension of 0.5×0.5×1.0 m ($x\times y\times z$) with up to 31 height-bins starting from 0.0 m.

Individual-tree CBH was estimated using the height-bin data. Specifically, for each TreeVaw-identified tree a vertical profile of laser hits versus height-bin was derived by vertically cutting out of the height-bin data a cylinder that is centered at the tree location and that has a diameter equal to crown width. We calculated CBH as the height at which an abrupt drop in the vertical profile occurs. Interested readers are referred to Popescu and Zhao (2008) for more details.

3.2.3. Fine-resolution AGB map

DBH is a common surrogate for biomass through the use of allometric equations. Lidars, however, do not directly measure DBH. Therefore, regression analysis was used to estimate DBH from the aforementioned lidar-derived tree dimension variables including tree height, crown width and CBH; for this purpose, two linear models, one for pines and another for the deciduous, were developed by relating the field-measured DBH to the lidar-derived tree dimension variables for the 117 matched trees (Table 2). The lidar-estimated DBH by the

Table 2
Results of lidar-derived individual tree dimension variables, DBH and biomass for the 117 matched trees

	Species[a]	R^2	RMSE	Method or model[a]
Height (m)	Pines	0.96	1.38	Directly measured by using the variable-
	Deciduous	0.90	2.15	window filtering algorithm known as
Crown width	Pines	0.57	1.68	"TreeVaw"
(m)	Deciduous	0.59	2.08	
Crown base	Pines	0.79	2.03	Height-bin approach
height (m)	Deciduous	0.74	1.88	
DBH (cm)	Pines	0.87	5.62	dbh = −11.2 + 1.52LH + 0.21LCW + 1.06LCBH
	Deciduous	0.89	4.80	dbh = −0.95 + 0.70LH + 3.14LCW + 0.37LCBH
Biomass (kg)	Pines	0.80	237	bm = exp(−2.5356 + 2.4349 * ln(dbh))
	Deciduous	0.88	138	bm = exp(−2.48 + 2.4835 * ln(dbh))

[a] For brevity, LH, LCW, and LCBH in the regression models represent lidar-derived tree height, crown width and crown base height, respectively. The biomass equations are the national-scale generic models for USA that have been developed by Jenkins et al. (2003).

two models was then incorporated into two species-specific dbh-based general allometric equations as compiled in Jenkins et al. (2003) to compute total AGB for each identified tree. Tree species, either pine or deciduous (hardwood), was determined by referring to the classified Quickbird image. Furthermore, the estimated total AGB of each tree was apportioned into component biomass according to the ratio equations of Jenkins et al. (2003) in order to generate a spatially-explicit biomass map at a 0.5 m spatial resolution. The generation process proceeded tree by tree: the bole biomass was assigned to the pixel at the tree location; the foliage and branch biomass was distributed with a 2-D Gaussian distribution over the pixels covered by the crown.

3.3. Scale-invariant biomass models

To formulate effective models for estimating forest attributes of interest, specifically biomass, we chose the lidar-derived CHD or, its equivalent, CHQ as predictors (Fig. 4). Because CHD and CHQ are themselves perceived as functions or functional data, a model using CHD or CHQ as predictor is considered as a functional (i.e., a function of functions), which makes this study distinct from most previous research that used various lidar-based individual statistics from ALS data as predictors in linear or nonlinear regression models.

To clarify, let us first address one of the two predictors, i.e., CHD. We define CHD as the probability of observing the top canopy surface at a height h or, equivalently speaking, the proportion of the ground surface obscured by the canopy surface of height h when viewed downward from the above (see a more rigorous definition in the next section). Obviously, CHD is a function with respect to h; denote it by $p(h)$. Such a definition of $p(h)$ is made more compatible with the lidar-derived CHM (Fig. 4). Thus, $p(h)$ or, more precisely, its estimate $\hat{p}(h)$ can be easily derived from the CHM or the raw lidar data. Below, we introduce two functional models with CHD and CHQ as predictors, respectively, for estimating biomass.

3.3.1. A linear functional model

The first AGB model is a linear functional that uses CHD as the predictor. In addition to preserving more information than the commonly used lidar metrics (Fig. 4), the predictor $p(h)$ takes on a benign property of being linearly scalable, which means that if $p_1(h)$ and $p_2(h)$ are two respective CHDs defined for two contiguous regions with areas of a and b, the CHD for the combined region, $p_{12}(h)$, is an area-weighted linear combination of $p_1(h)$ and $p_2(h)$:

$$p_{12}(h) = \frac{a \cdot p_1(h) + b \cdot p_2(h)}{a+b}. \tag{1}$$

For an AGB model $f[\cdot]$ to be scale-invariant, it must satisfy,

$$f\left[\frac{a \cdot p_1(h) + b \cdot p_2(h)}{a+b}\right] = \frac{a \cdot f[p_1(h)] + b \cdot f[p_2(h)]}{a+b} \tag{2}$$

where a, b, $p_1(h)$ and $p_2(h)$ assume certain arbitrariness. Therefore, $f[\cdot]$ should be a linear functional of $p(h)$; a natural choice is $f[p(h)] = \langle p(h), K(h) \rangle$ where $\langle \cdot, \cdot \rangle$ denotes an inner product in a function space.

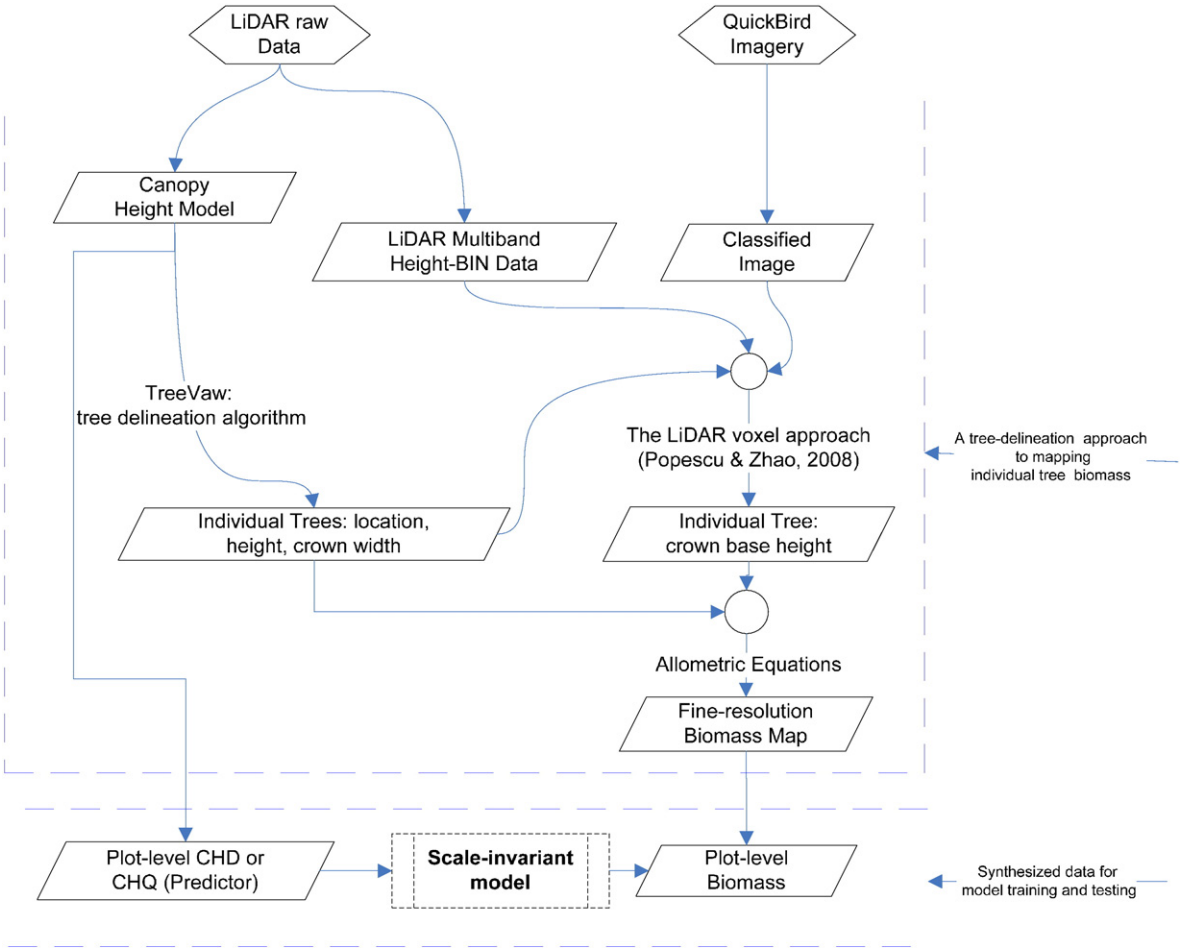
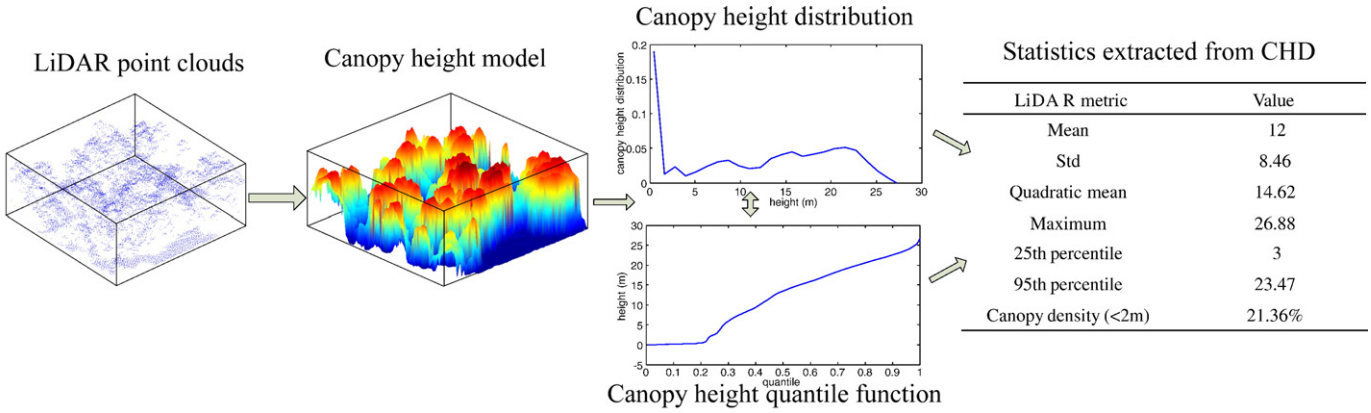


Fig. 3. A flow chart summarizing the procedures of the individual tree-based approach to mapping per-tree biomass in order to simulate reference data for training and validating the proposed scale-invariant biomass models: In the synthesized reference data, the predictors are canopy height distributions (CHD) or canopy height quantile functions (CHQ) on plots of a given size, and the dependent variable is biomass on the coincident plots.



LiDA R metric	Value
Mean	12
Std	8.46
Quadratic mean	14.62
Maximum	26.88
25th percentile	3
95th percentile	23.47
Canopy density (<2m)	21.36%

Fig. 4. The flow of diminishing information contents: raw lidar returns on a square plot, the canopy height model (CHM) in a 3-D perspective, the canopy height distribution (CHD) derived from the CHM and the associated canopy height quantile function (CHQ), and a table of several height statistics extracted from the CHD. Notice how information is being lost progressively from the left to the right. Previous studies used the raw lidar data or CHM for individual tree identification, or used lidar metrics similar to those in the table as predictors for stand characteristics. As a compromise, this study focuses on the use of CHD or CHQ for estimating biomass with functional regression models.

For simplicity, we chose the most common inner product and proposed the following scale-invariant AGB model,

$$M = f_K[p(h)] = \langle p(h), K(h) \rangle = \int_{h \geq 0} p(h) \cdot K(h) \mathrm{d}h \tag{3}$$

where M is the AGB per unit area; $K(h)$ is an unknown but fixed function that parameterizes the model $f_K[\cdot]$. The functional form of $K(h)$ needs to be inferred, e.g., from a training dataset.

Mathematically, $p(h)$, as a function or curve, is defined continuously over $h \geq 0$, but in practice, it has to be extracted from lidar data and its function values are available only at a discrete set of points as depicted in Fig. 8, i.e., $\{p(h_i)\}_{i=1}^{k}$, $h_i = (i-1)\cdot\Delta h$. Accordingly, $f_K[\cdot]$ of Eq. (3) can be approximated by,

$$M \approx f_K\left[\{p(h_i)\}_{i=1}^{k}\right] = \sum_{i=1}^{k} p(h_i) K(h_i) \Delta h \tag{4}$$

where Δh is a height-bin used to sample $p(h)$. In essence, the scale-invariant model of Eq. (4) is a linear one that has $\{p(h_i)\}_{i=1}^{k}$ as predictors with the corresponding coefficients being $K(h_i)\cdot\Delta h$.

3.3.2. An equivalent nonlinear functional model

Previous work suggested that canopy height quantile-based lidar metrics are effective predictors for estimating biomass (Lim and Treitz, 2004). An equivalent biomass model using CHQ as predictor can be derived directly from the linear functional model of Eq. (3) based on the equivalence between CHQ and CHD. Mathematically, a quantile function (i.e., CHQ) and its associated probability distribution function (i.e., CHD) can be converted to each other with no loss of information. However, unlike CHD, an undesirable feature of CHQ is its nonadditive scaling property, which means that the overall CHQ over a region cannot be generally aggregated as a linear combination of the multiple CHQs over the sub-regions that partition the region. This scale-dependence indicates that linearity and scale-invariance are unlikely to be achieved at the same time for a model using CHQ as the predictor for biomass. Hence, a scale-invariant model with CHQs should be a nonlinear functional.

Denote CHQ by $Q(q)$ where $0 \leq q \leq 1$. Considering the equivalence of $Q(q)$ to $p(h)$ (Fig. 4), $Q(q)$ is derived by first obtaining the cumulative CHD $F(h) = \int_0^h p(h')dh'$ and then inverting $F(h)$, that is, $Q(q) = F^{-1}(q)$. Using this CHD–CHQ relationship as well as the integration-by-parts theorem, Eq. (3) is re-written as,

$$M = K(h_{\max}) - \int_0^1 q \cdot \frac{dK(Q(q))}{dq} dq \tag{5}$$

where $h_{\max} = \min\{h;\ F(h) = 1\}$ is the maximum canopy height on the analysis unit. Eq. (5) is an AGB model using $Q(q)$ as the predictor, and again $K(\cdot)$ is the same model "parameter" as that of Eq. (3) and needs to be estimated by resorting to either training data or theoretical calculation.

As with $p(h)$, the observed data of CHQ $Q(q)$ are only available at a discrete set of q's, i.e., $\{h_i = Q(q_i)\}_{i=1}^{k+1}$ where $q_i = (i-1)/k$. Accordingly, we have the following approximation of Eq. (5),

$$M \approx \sum_{i=1}^{k} K(h_{i+1})/k \tag{6}$$

where $h_i = Q(q_i)$ is the $(i-1)/k$ quantile of CHD. This model appears nonlinear with respect to h_i due to the potential nonlinearity of $K(h)$. Again, it is noted that this nonlinear model is essentially equivalent to the linear one in Eq. (4).

3.4. Theoretical basis

Interesting enough, theoretical evidences exist to justify the two aforementioned AGB models and are detailed below. Readers, however, may skip Section 3.4 without too much loss of the understanding of the whole paper.

3.4.1. Modeling canopy height distributions (CHD)

This sub-section aims to model the relation between CHDs and the distributions of dbh at plot levels with the aid of tree allometric relationships. Of particular note is that dbh dictates biomass through the allometry equations while CHD is measured by lidar. Rather than to be mathematically rigorous, the modeling purpose is to provide some theoretical justification for the scale-invariant biomass models proposed above.

Before proceeding, it is beneficial to clarify some variables and notations.

The following variables or their distributions refer to the properties of the forest over an analysis unit (e.g., a 30 m × 30 m pixel or 1.0-ha circular plot):

h "Canopy height" which is the height of the top surfaces visible from above, and refers to both open grounds ($h = 0$) and canopy surfaces ($h > 0$).

$p(h)$ "Canopy height distribution" (CHD), i.e., the probability distribution of h on a unit forested area, which refers to the probability of observing the portion of the unit area that has a height of h, and consists of two terms,

i.e., one for the open ground and another for the canopy (Fig. 4).

c "Percent crown cover" which defines the fraction of a unit area obscured by vertically projected crowns when assuming that crowns are opaque.

n "Stem density" which is the number of trees per unit area.

The following notations refer to the properties of single trees:

H, D, h_{cbh} and α Tree dimension variables which are tree height, diameter, crown base height, and crown projection area (CPA), respectively; CPA refers to the vertically projected area of a single tree crown onto the ground.

$p_H(H)$, $p_D(\mathrm{D}h)$, $p_{h_{\mathrm{cbh}}}(h_{\mathrm{cbh}})$ and $p_\alpha(\alpha)$ The corresponding distributions of the tree dimension variables for all trees within an analysis unit.

$p_{\mathrm{tr}}(h|H, \alpha, h_{\mathrm{cbh}})$ "Per-tree crown height distribution" conditioned on tree dimension parameters H, α, and h_{cbh}. It is defined for a tree concerning the canopy height only over its CPA α and is affected by the geometry of the tree.

In the above, $p(h)$ is defined in a way that is compatible with the lidar CHM or the first laser hits/returns, and is slightly different from the height distribution of all laser returns because some returns, e.g., last ones, could be reflected from inside crowns. Other relevant variables will be introduced when necessary.

The variables and distributions defined above are related to each other through allometric equations. First, the percent crown cover c of a unit forested area can be modeled by the Boolean model (Li and Strahler, 1992),

$$c = 1 - \exp(-n \cdot E[\alpha]) \tag{7}$$

where $E[\alpha]=\int \alpha p_\alpha(\alpha)\mathrm{d}\alpha$ is the mean per-tree CPA. To simplify illustrations, but without losses of generality, we assume the first-order approximation of Eq. (7), i.e., $c \approx n \cdot E[\alpha]$, which corresponds to the case of non-overlapping crowns.

Next, given the percent crown cover c, the CHD is composed of two terms,

$$p(h) = (1-c) \cdot \delta(h) + c \cdot p_c(h) \tag{8}$$

where $\delta(h)$ is the Dirac function that is infinity at $h=0$ and is 0 elsewhere. The first term $(1-c)\cdot\delta(h)$ indicates that the open ground is seen with a probability of $(1-c)$, i.e., the fraction of open ground; the second term $c \cdot p_c(h)$ means that when viewed vertically downward, crowns can be seen with a probability of c, i.e., the proportion of the aggregated CPAs over the analysis unit. As such, $p_c(h)$ is deemed as the seen CHD over the canopy only ($h>0$), and it can be elucidated as a CPA-weighted sum of crown height distributions of individual trees. For example, if the stem density n is relatively large, $p_c(h)$ is approximated by,

$$p_c(h) \approx \int p_{\mathrm{tr}}(h|D) \frac{\alpha(D)}{E[\alpha]} p_D(D) dD \tag{9}$$

where $p_{\mathrm{tr}}(h|D)$ is the per-tree crown height distribution conditional on the dbh D; $\alpha(D)=E[\alpha|D]$ is the mean per-tree CPA given D, which is dictated by the allometric relationship.

To model $p_{\mathrm{tr}}(h|D)$ of Eq. (9), let us first address $p_{\mathrm{tr}}(h|H, \alpha, h_{\mathrm{cbh}})$, which can be explicitly modeled if the crown geometry is known. Also, CPA α can be explicitly computed from crown diameters if crown shapes take some known geometry primitives, e.g., ellipsoid and cone. For some crown geometries, it holds that $p_{\mathrm{tr}}(h|H, \alpha, h_{\mathrm{cbh}})=p_{\mathrm{tr}}(h|H, h_{\mathrm{cbh}})$. Meanwhile, allometric relationships, whether deterministic or probabilistic, can be used to derive the distributions of H, α and h_{cbh} conditioned on D such as $p_H(H|D)$, $p_\alpha(\alpha|D)$, $p_{h_{\mathrm{cbh}}}(h_{\mathrm{cbh}}|D)$, or $p(H, h_{\mathrm{cbh}}, \alpha|D)$ (e.g., Tewari and Gadow, 1999; Li et al., 2002). As a result, we obtain the per-tree crown height distribution conditioned on D,

$$p_{\mathrm{tr}}(h|D) = \int p_{\mathrm{tr}}(h|H, \alpha, h_{\mathrm{cbh}}) p(H, h_{\mathrm{cbh}}, \alpha|D) \mathrm{d}H \cdot \mathrm{d}\alpha \cdot \mathrm{d}h_{\mathrm{cbh}} \tag{10}$$

which generally depends on tree species. Nevertheless, if there are multiple species, e.g., $p_{\mathrm{tr}}(h|D, i)$, $i=1,\ldots, m$, they could be assimilated into an overall distribution that is unconditioned on species by assuming that the occurrence probabilities of each species, r_i, could be elicited using some prior knowledge $p_r(r_1, r_2\ldots, r_m)$ where $\sum r_i=1$. Often, it is convenient to model $p_r(r_1, r_2\ldots, r_m)$ with a multinomial distribution. Marginalizing over $r_1, r_2\ldots, r_m$ yields

$$p_{\mathrm{tr}}(h|D) = \int \left(\sum_{i=1}^{m} p_{\mathrm{tr}}(h|D, i) \cdot r_i \right) \cdot p_r(r_1, r_2, \ldots, r_m) \mathrm{d}r_1 \mathrm{d}r_2 \ldots \mathrm{d}r_m. \tag{11}$$

To finalize the model description, we substitute Eqs. (7) and (9) into Eq. (8) and obtain an expression for CHD,

$$p(h) = (1 - n \cdot E[\alpha]) \cdot \delta(h) + n \cdot \int p_{\mathrm{tr}}(h|D) \alpha(D) p_D(D) \mathrm{d}D \tag{12}$$

which establishes a link between CHD ($p(h)$) and the distribution of dbh ($p_D(D)$) through the known crown geometry and allometric relationships such as those in Eqs. (9)–(11).

On the other hand, M, the AGB over the unit forested area, is readily calculated from $p_D(D)$ and n with the recourse to the dbh-based biomass allometric equation $B(D)$,

$$M = n \int B(D) \cdot p_D(D) \mathrm{d}D. \tag{13}$$

3.4.2. Deriving biomass models from the CHD model

The two biomass models proposed above can be directly uncovered from the CHD model formulated in Eqs. (7)–(12). In the following, we focus only on the derivation of the linear model of Eq. (3) because the two models are equivalent except for the use of different predictors.

It becomes clear from Eqs. (12) and (13) that CHD is related with biomass through $n \cdot p_D(D)$. Once an estimate of CHD, $\hat{p}(h)$, is extracted from lidar data, the inversion of Eq. (12), if possible at all, provides knowledge about $n \cdot p_D(D)$ that then can be incorporated into Eq. (13) to compute M. As per this logic, we compare Eqs. (8) and (12) to isolate the canopy part of CHD, i.e., the second term of $p(h)$, as given below,

$$c \cdot p_c(h) = \int p_{\mathrm{tr}}(h|D) \cdot \alpha(D) \cdot n \cdot p_D(D) \mathrm{d}D \tag{14}$$

which essentially is a Fredholm integral equation of the first kind with an aim to solve out $n \cdot p_D(D)$ given $c \cdot p_c(h)$ and $p_{\mathrm{tr}}(h|D) \cdot \alpha(D)$ (Aster et al., 2004). In practice, $c \cdot p_c(h)$ may be estimated from lidar CHM by discarding the ground component from $p(h)$, and $p_{\mathrm{tr}}(h|D) \cdot \alpha(D)$ may be formulated with empirical allometric relationships such as height–dbh equations.

To derive the linear AGB model, we present a symbolic scheme for solving Eq. (14), although this is of little practical value due to its numerical unstability of an ill-posed problem. First, for notational convenience, denote $p_{\mathrm{sec}}(h)=c \cdot p_c(h)$; then discretize the integral equation using $\{h_i=(i-1)\cdot \Delta h\}_{i=1}^{k}$ for h and $\{D_j=(j-1)\cdot \Delta D\}_{j=1}^{l}$ for D. As a result, Eq. (14) reduces to,

$$\mathbf{p}_{\mathrm{sec},(k\times 1))} = \mathbf{A}_{(k\times l)} \cdot \mathbf{p}_{n,D,(k\times l)} \tag{15}$$

where $\mathbf{p}_{\mathrm{sec},\,(k\times 1)}=[p_{\mathrm{sec}}(h_1),\ldots, p_{\mathrm{sec}}(h_k)]^T$, $\mathbf{p}_{n,D,(l\times 1)}=[n \cdot p_D(D_1),\ldots, n \cdot p_D(D_l)]^T$, and $\mathbf{A}_{(k\times l)}$ is a $k\times l$ matrix with its (i, j) element being $p_{\mathrm{tr}}(h_i|D_j) \cdot \alpha(D_j) \cdot \Delta D$. Applying $\mathbf{A}^{-1}_{(l\times k)}$, i.e., the pseudoinverse of $\mathbf{A}_{(k\times l)}$, to Eq. (15), $\mathbf{p}_{n,D,(l\times 1)}$ is obtained as,

$$n \cdot p_D(D_j) = \sum_{i=1}^{k} a^{-1}_{ji} \cdot p_{\mathrm{sec}}(h_i) \tag{16}$$

where a_{ji}^{-1} is the (j,i) element of $\mathbf{A}_{(l\times k)}^{-1}$. Combining Eqs. (13) and (16), AGB can be estimated by,

$$M \approx \sum_{i=1}^{k} p(h_i) \cdot \sum_{j=1}^{l} a_{ji}^{-1} \cdot B(D_j) \cdot \Delta D \tag{17}$$

which is a linear model of the same form as in Eq. (4). Hence, the linear model of Eq. (4) is derived from the CHD–dbh relationship of Eqs. (7)–(12). Comparing Eqs. (4) and (17) further reveals the following relationship,

$$K(h_i) = \sum_{j=1}^{l} a_{ji}^{-1} \cdot B(D_j) \cdot \frac{\Delta D}{\Delta h}. \tag{18}$$

3.5. Inference and validation of biomass models

3.5.1. K-function

The newly introduced function $K(h)$, acting like a weighting function in the integral of Eq. (3), represents the contribution to biomass per unit area by the portion of canopies that has a height falling within/around h. In general, $K(h)$ is affected by the tree allometry and crown geometries as indicated by Eq. (18); therefore, strata-specific models should be used under various forest types or conditions.

The functional models need to be fitted by specifying the parameterizing function $K(h)$, in the same spirit as fitting linear regression models by specifying the relevant coefficients. Mathematically, $K(h)$ can be solved out by inverting the following Fredholm integral equation of the first kind,

$$B(D) = \int_{h\geq 0} K(h) \cdot p_{\mathrm{tr}}(h|D)\alpha(D)\mathrm{d}h \tag{19}$$

which is derived by substituting $p(h)$ of Eq. (12) into Eq. (3) and then relating Eq. (3) with Eq. (13). Given the allometric relationships including $B(D)$ and $p_{\mathrm{tr}}(h|D)\alpha(D)$, the inversion of Eq. (19) for $K(h)$ is an ill-posed problem that often requires regularization schemes for numerical solutions. A symbolic solution to Eq. (19) for $K(h)$ is the same as that to Eq. (14).

In this work, we attempted not to infer $K(h)$ by directly solving the integral equation Eq. (19) because this involves the explicit formulation of $p_{\mathrm{tr}}(h|D)$. Instead, we turned to a constrained regression technique to fit the linear functional model for estimating $K(h)$ in light of training data sets, as detailed below.

3.5.2. Constrained least-squares method for model fitting

The model parameter, $K(h)$, cannot be an arbitrary function. It is beneficial to impose certain constraints on the form of $K(h)$ according to some intuitive prior knowledge: First, we assumed that the AGB becomes zero if no trees are present (i.e., $M(\delta(h))=\int\delta(h)K(h)\mathrm{d}h=0$ for $p(h)=\delta(h)$), which indicates $K(0)=0$. Second, we assumed that $K(h)$ is non-decreasing with respect to h, considering that the larger the percentage of tall trees, the higher the AGB.

For simplicity, the discretized linear model of Eq. (4) is used to infer the coefficients $K(h_i)$ (i.e., the discretized version of $K(h)$), although other complicated or possibly more effective quadrature methods may be considered for discretizing the integral of the biomass model in Eq. (3) (Aster et al., 2004). However, even with Eq. (4), common regression procedures fail to work out partly due to the constraints on $K(h)$. Alternatively, we used a constrained least-square procedure, which employs an active set method as implemented in Matlab (*lsqlin* function, Mathworks, Inc.), to iteratively search for optimal values of $K(h_i)$ in terms of minimizing the mean square error with regard to the synthesized training data. The constraints take both inequality and equality forms, including $K(h_i)\geq K(h_{i-1})$ for $i>1$ and $K(h_0=0)=0$.

3.5.3. Synthesized training and testing datasets

Reference datasets were synthesized for fitting and testing the AGB functional models. In the datasets, the predictors (independent variables) include CHDs $\{p(h_i)\}_{i=1}^{k}$ and CHQs $\{h_i=Q(q_i)\}_{i=1}^{k+1}$ for the linear and nonlinear models, respectively, and were extracted from the lidar CHM. The response variable (dependent variable) is the AGB obtained by aggregating the fine-resolution AGB map on the coincident plots. In the synthesis process, we investigated two plot shapes, i.e., circular and square, as well as ten plot sizes of $100i^2$ m^2 with i ranging from 1 to 10, thus resulting in a total of 20 shape-by-size combinations. For each combination of a given shape and plot size, we systematically sampled a total of 2000 synthesized plots (data points) that would be randomly split into two subsets, one with 100 plots for training and another with the remaining 1900 for testing.

4. Results

4.1. Fine-resolution biomass map with the individual tree-based approach

Table 2 reports the results of tree dimension variables, DBH and biomass that were obtained with the individual tree-based approach for the 117 matched trees. Fig. 12 depicts the fine-resolution AGB map where the detailed spatial pattern of AGB is evident. Reference data synthesized from this map were used in the subsequent experiments to evaluate the functional AGB models.

4.2. Evaluating the linear functional model

In the synthesized data for evaluating the linear biomass model of Eq. (4), we first used a 0.25 m height-bin (sampling interval) to discretize the CHDs up to 35 m, therefore, resulting in a predictor

Table 3
Coefficients of determination (R^2) and RMSEs of the linear AGB model trained and tested across a range of plot sizes (scale) for both square and circular plot shapes

Plot size (ha)	Square plots				Circular plots			
	Training		Testing		Training		Testing	
	R^2 (min,max)	RMSE	R^2 (min,max)	RMSE	R^2 (min,max)	RMSE	R^2 (min,max)	RMSE
.01	.832 (.804, .855)	33.3	.825 (.813, .830)	34.2	.831 (.805, .854)	33.7	.823 (.815, .828)	34.6
.04	.892 (.877, .908)	24.3	.889 (.885, .892)	24.7	.892 (.874, .909)	24.3	.887 (.882, .891)	25.0
.09	.912 (.897, .922)	21.0	.910 (.907, .913)	21.3	.912 (.897, .928)	19.9	.911 (.907, .913)	21.3
.16	.921 (.910, .933)	19.2	.919 (.917, .921)	19.5	.921 (.909, .936)	19.1	.919 (.916, .921)	19.6
.25	.926 (.914, .934)	18.0	.925 (.923, .927)	18.2	.925 (.914, .934)	18.1	.924 (.923, .927)	18.3
.36	.931 (.919, .943)	16.9	.929 (.927, .931)	17.2	.932 (.921, .945)	16.9	.929 (.927, .930)	17.3
.49	.933 (.923, .943)	16.2	.932 (.930, .934)	16.4	.932 (.920, .943)	16.3	.932 (.930, .934)	16.5
.66	.937 (.925, .948)	15.4	.934 (.933, .936)	15.7	.935 (.924, .946)	15.6	.934 (.932, .936)	15.8
.81	.937 (.928, .948)	15.0	.937 (.934, .938)	15.1	.937 (.927, .947)	15.0	.936 (.935, .938)	15.2
1	.940 (.928, .952)	14.3	.938 (.936, .940)	14.6	.938 (.928, .946)	14.6	.938 (.937, .940)	14.7

Note that, in each case, R^2 is reported as the mean, minimum and maximum over 100 random runs, and RMSE is reported as the mean over the 100 runs. The predictor CHD is sampled with a 0.25 m height-bin.

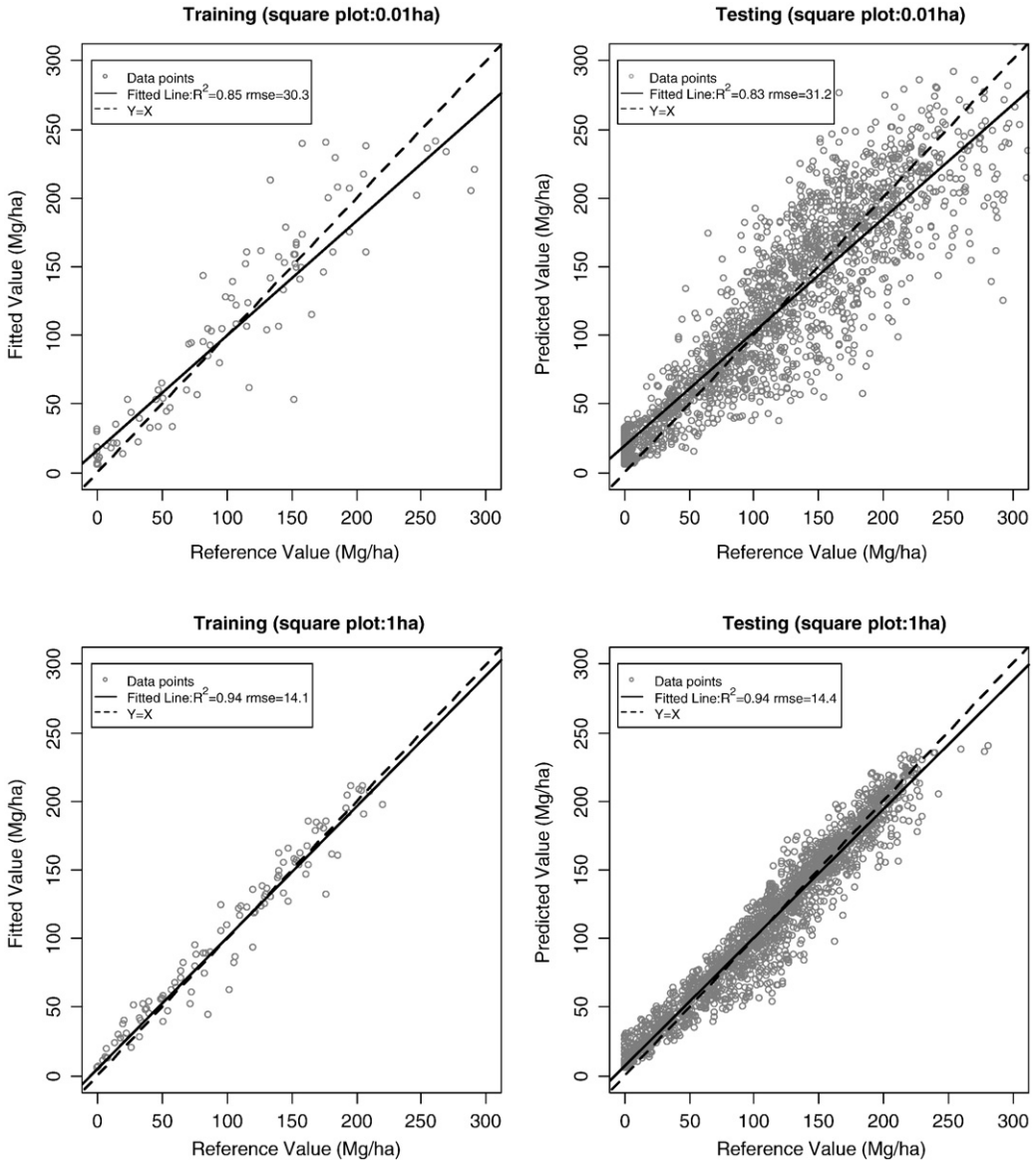


Fig. 5. Scatterplots of predicted vs. reference biomass for the training (the left two graphs) and the testing (the right two graphs), respectively, and also over the plot sizes of 0.01 ha (the upper two graphs) and 1 ha (the lower two graphs), respectively. The model evaluated here is the linear functional AGB model with its predictor CHD discretized using a 0.25 m height-bin.

vector of CHD that contains 140 scalar values. For a given group of 2000 plots of a given size and shape, we randomly split it 100 times to generate 100 datasets, each comprising 100 training and 1900 test plots, for assessing the consistency of the models across different training/test datasets.

Regression results for the linear model are shown in Table 3, which lists the mean, minimum and maximum coefficients of determination (R^2) as well as the mean root mean square errors (RMSE) over each group of 100 trials for a given plot size and shape. It is observed that for all plot sizes, plot shapes had no significant impacts on the model inferences in terms of both R^2 and RMSE (all p-values $\ll 0.01$). Overall, R^2 values increase with an associated decrease in RMSE as the plot size varies from 0.01 ha to 1.0 ha. For example, in a random trial using the 0.01-ha square-plot data, R^2 values for training and testing are 0.85 and 0.83, respectively, with the associated RMSEs of 30.3 Mg/ha and 31.2 Mg/ha, respectively (Fig. 5); in the case of 1-ha square-plot data, R^2 for training and testing are both about 0.94, with a RMSE of 14.1 and 14.4, respectively (Fig. 5). In all the 2000 trials of model fitting and testing (two plot shapes by ten plot sizes by 100 random runs), performances of the linear functional model in training are consistent

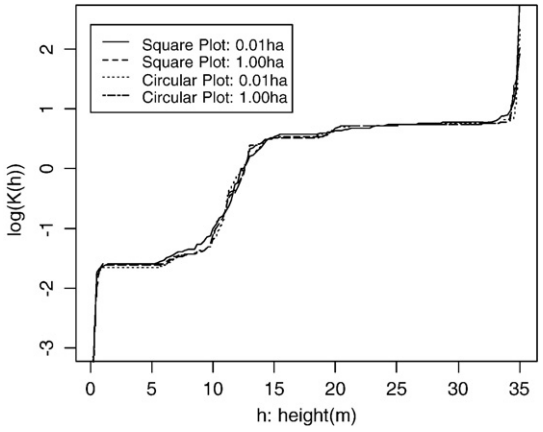


Fig. 6. *K*-functions: The estimated coefficients of the linear model are plotted as curves with respect to h. Each curve is the average over the respective 100 runs.

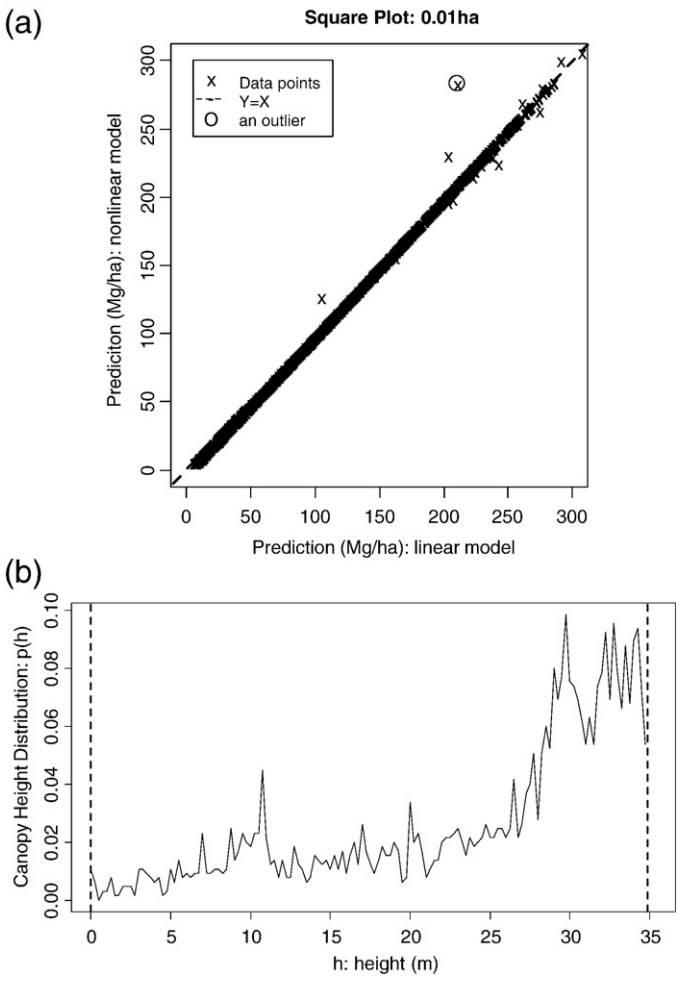


Fig. 7. (a) Scatterplot of predicted AGB values by the nonlinear model versus those by the linear model when using 0.01 ha circular plot data. (b) A truncated canopy height distribution (CHD) where portion of the canopy above 35 m are not taken into account due to the discretization scheme: This example corresponds to the outlier circled in (a).

with those in testing as revealed by a close agreement of R^2 or RMSE (all p-values $\ll 0.01$) between the fitting and testing (Table 3). Such consistency partially validates the proposed linear model, and also suggests its good generalization ability.

Four examples of the regressed $K(h)$, obtained as the series of optimized coefficients of the linear model, are plotted in Fig. 6 where each of the four curves is the average of the estimated $K(h)$'s over 100 random trials of a given shape and plot size. Across all the 20 combination of plot-related factors considered, i.e., two shapes by ten sizes, there are no significant differences between any two average $K(h)$ curves as depicted in Fig. 6 (all p-values $\ll 0.01$), although minor differences may be observed especially at the end of $K(h)$ curves or between the estimated curves from individual trials even for the same plot size.

4.3. Evaluating the nonlinear functional model

It is not an easy endeavor to estimate $K(h)$ by directly fitting the nonlinear model of Eq. (6). Therefore, we applied the $K(h)$ estimated for the linear model to the nonlinear model to show the equivalence of the two models. The predictor of the nonlinear model consists of quantile-based heights associated with CHQs that were discretized at $q_i=(i-0.5)/100\%$, $i=1,2,\ldots,100$, as well as 99.6%, 99.7%, 99.8% and 99.9%. Because the values of $K(h)$ for the linear model were estimated only at discrete heights (i.e., $h=\Delta h\cdot j$), we interpolated these estimates to compute the value of $K(h)$ at an arbitrary height h for being used in the nonlinear model. To assess the equivalence of the nonlinear model to the linear model, we compared the AGB prediction of the two models in the form of scatterplots where, for brevity, only the case for the 0.01 ha square plots is shown (Fig. 7a). The two models matched each other exactly for all but several testing data points (plots) (Fig. 7a). Most exceptions are due to truncation errors in the CHDs of the linear model that have been discretized only up to 35 m, as illustrated in Fig. 7b.

4.4. Model analysis

The predictor CHD, perceived as a curve, needs to be discretized with a height-bin when applying the linear functional model. To provide practical guides, we examined the effects of both the length of height-bin and the training data size on performances of the linear model.

4.4.1. Effects of height-bin

The choice of a height-bin (an interval used to sample CHDs) not only affects the characterization of CHDs but also determines the dimension of the CHD predictor with a larger height-bin corresponding to a lower dimension in the same way as that of a spectral

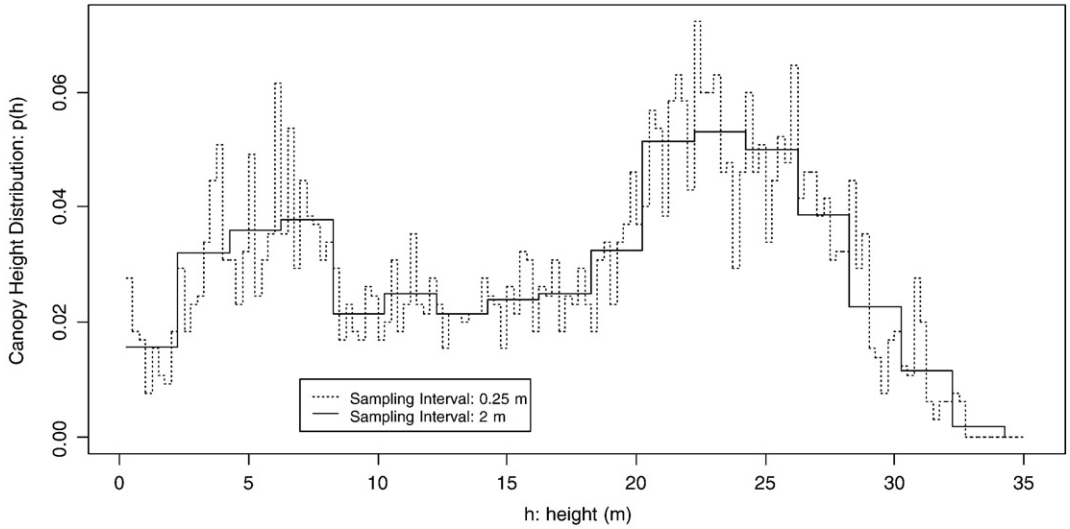


Fig. 8. A canopy height distribution (CHD) extracted from lidar data was discretized using two different height-bins, with a larger height-bin (2.0 m) leading to a coarser representation. Conceptually, CHD is similar to the canopy density metrics in Næsset (2002) and the height-bin products in Popescu and Zhao (2008).

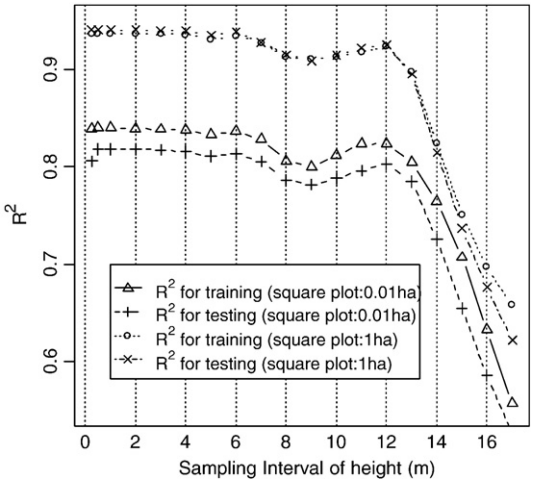


Fig. 9. R^2 vs. height-bin (sampling interval of height): R^2's for training and testing change as a function of height-bin that is used to discretize the canopy height distributions.

signature which can be characterized with different spectral resolutions (Fig. 8). Using a training data size of 100 and a test data size of 1900 for all the 20 combinations of plot shapes and sizes, we evaluated the linear model by employing a series of height bins ranging from 0.25 m to 20 m that corresponds to a predictor dimension ranging from 140 to 2. Typical results of R^2 in training (fitting) and testing are presented in Fig. 9 for the 0.01-ha and 1.0-ha square plots, respectively. It is found that the model performances showed no degrading trends until the height-bin was increased above around 6.0 m and that the R^2 dropped significantly when using a height-bin larger than 12.0 m. This observation suggests that it is practical and effective to choose a relatively large height-bin below a critical value with no great loss of model performances. The critical value may vary with respect to specific applications, e.g., around 5 to 6 m in this work.

4.4.2. Effects of training data size

We also varied the training data sizes from 10 to 500 with an increment of 10 to evaluate the effects of training data sizes on the performances of the linear AGB model. The models fitted with these different training data sizes were tested using an independent set of 1500 data points. Meanwhile, we also varied the CHD dimension from 7 to 140 that corresponds to a height-bin ranging from 5 m to 0.25 m in discretizing the CHDs. In this analysis, all evaluations were based on 0.49-ha square-plot data because plot shapes and sizes have no or little influences on the model inference. To avoid a plethora of figures, we reported the changes of R^2 with training data sizes only for two extreme models, i.e., one with a predictor dimension of 140 (height-bin: 0.25 m) and another with a predictor dimension of 7 (height-bin: 5 m) as shown in Fig. 10. We found that, for all the models considered, the addition of new data into the training datasets with a size of above 70 contributed no significant gain in terms of the R^2 for testing, although the R^2 for the training may fluctuate a little bit which suggests the possibility of slightly overfitting in some cases. The result indicates that, in our cases, a training size of 50 is typically enough for the purposes of fitting the model and making prediction. However, in practice caution should be exercised such that training data should be collected over a wide range of biomass values to reduce the risk of extrapolation.

4.5. Scaling-up for total biomass prediction

The linear functional model was further assessed in an extreme case by treating the whole study area as a large analysis unit from which a single CHD was extracted as an overall predictor to estimate the total biomass of the study area. Only the models fitted from square-plot data were examined. For each plot size, there were 100 trained models resulting from 100 random trials. These overall AGB estimates are presented in the box-whisker plots where each column summarizes the estimates from 100 models fitted with the 100 random training datasets at a given plot size (Fig. 11). On average, all the models overestimated AGB in comparison to the reference value of 109.0 Mg/ha.

5. Discussion

We employed two independent procedures for mapping biomass at different levels with lidar, i.e., a hybrid individual tree-based approach for individual tree component biomass, and two scale-invariant regression models for plot-level biomass. The emphasis of this study was on the regression models because the individual tree-based approach requires tree-segmentation algorithms that may be not available to some practitioners or not applicable to certain forest conditions (Zhao and Popescu, 2007). The individual tree-based

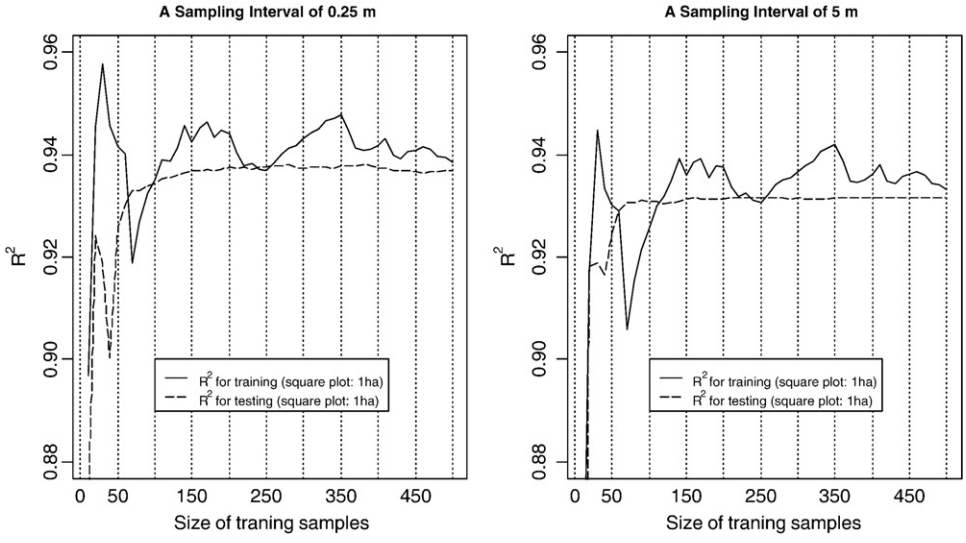


Fig. 10. The effects of training sample sizes on the R^2 for two different height-bins: 0.25 m (the left) and 5.0 m (the right). The dataset used is the 1.0 ha square-plot data.

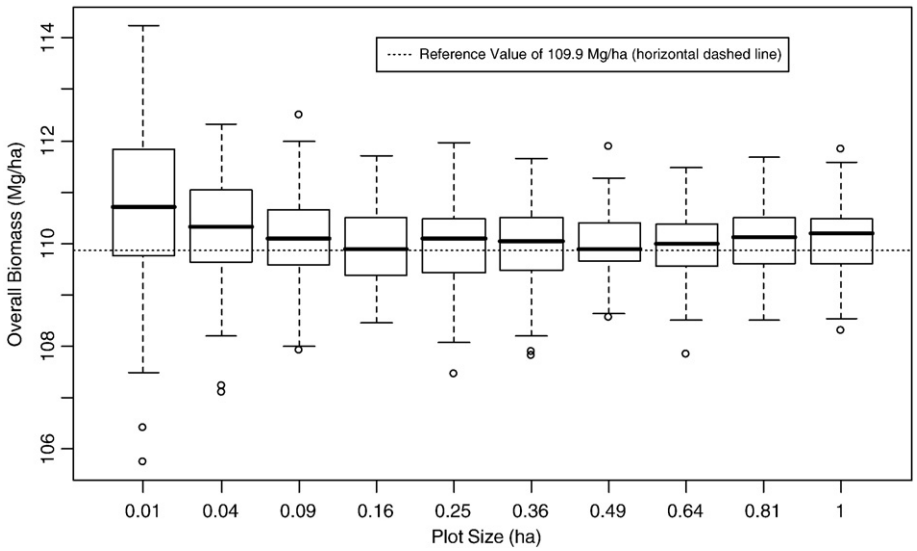


Fig. 11. The box-whisker plot of the overall biomass estimated by using the single overall CHD of the whole study area. For each plot size, there are 100 models obtained respectively from 100 runs with different random training sets, and the 100 models were applied to the overall CHD, thus producing 100 estimates of the overall biomass at each plot size.

approach allows accurately mapping tree parameters and helped to create a fine-scale raster map of biomass (Fig. 12). To our best knowledge, this map is the first of its kind, and it provides a realistic and detailed characterization of spatial distribution of biomass. Of particular note, the use of our functional models for practical applications is similar to that of the regression models of previous studies (Næsset, 1997; Nelson et al., 2004). Our models by no means rely on the fine-resolution biomass map. The sole role of this map in this study is for evaluating the performance of the functional regression models, and in this regard, results of this study imply that our functional models are at least as effective as the individual tree-based approach.

Lidar provides direct measurements of canopy height but not of dbh, which is a common proxy for biomass. Researchers have been striving to search for the appropriate lidar metrics and effective model forms to develop regression models for estimating biomass or other forest structural variables of interest. In this study, CHD and CHQ were chosen as the respective predictors for two functional models. A major tenet behind such choices is that although the CHD and CHQ contain no horizontal information (x and y coordinates) of lidar hits, they can be perceived as a function/curve with respect to height (the z coordinates), thus being able to retain more information of the original lidar data than any statistics extracted from CHDs such as mean height, truncated mean height, and quantile-based height (Fig. 4). Conceptually, CHD is similar to the set of canopy density metrics examined in Næsset (2002). In this regard, our models have potential for making the most use of information inherent in lidar data in the same spirit as the lidar voxel-based height-bin approach in Popescu and Zhao (2008), and they also have the advantage of being scale-invariant and mathematically justifiable. Our models rely on a newly introduced K function that has a specific physical meaning. The K function is site-specific or strata-specific due to the varying allometry across different forest types. However, the model formulation is independent of forest types. Thus, our models are applicable to various forest regions, but they need to be fitted and applied for the regions of similar forest conditions.

Unlike previous studies that extracted lidar-based statistics directly from lidar raw laser hits or fully-digitized waveforms (Næsset, 2004; Riano et al., 2004; Andersen et al., 2005), this study extracted the lidar predictors, i.e., CHDs or CHQs, from the lidar-derived CHM. However, certain implicit relationships exist between raw laser hits and CHMs, and such relationships can be modeled so that CHDs and the height distributions of raw laser hits could be inferred from each other under certain assumptions (Sun and Ranson, 2000; Kotchenova et al., 2003). Therefore, it is suggested that the proposed models be extended to use the height distributions of original laser hits as predictors that may include first, last or all lidar returns as deemed appropriate by the analysts. By doing so, one may characterize understory structures. However, a side effect is that the use of intermediate or last laser returns should invalidate the scale-invariance of the models due to the nonlinear scalability of the height distributions derived from such laser returns. In such a situation, the scale-dependence is partly caused by the non-uniform sampling pattern of laser hits. Another foreseen side effect using original returns is the models' dependence on ALS instruments because different ALS systems could respond differently to the same forest (Næsset and Gobakken, 2008). Nevertheless, future studies may investigate the utility and applicability of lidar height distributions of original laser hits as the predictor in our functional models for estimating biomass or other forest attributes.

No explicit form is assumed for the allometry equation $B(D)$ in Eq. (12) during the derivation of the models. Therefore, although they are intended for estimating biomass, the models can be equally applied for other attributes such as basal area, stem density, crown fuel weight, canopy base height and timber volume if these structural variables use dbh as a proxy in their allometry equations. For example, assuming $B(D)=\pi D^2/4$, we get a model for basal area. As an extreme case, assuming $B(D)=1$, we get a model for stem density. Besides, we strongly suggest that future studies should examine the utility of the linear functional models with CHD as the predictor to estimate those forest structural variables such as Leaf Area Index (LAI), which may have no explicit functional relationship with dbh, mainly for two major reasons: first, structural variables such as LAI are linearly scalable when changing scales, and the linear model we proposed can guarantee this scalable property; second, a CHD or lidar height distribution contains much more information than any subset of its extracted statistics such as mean and quantile-based heights.

Concomitant with the scale-invariance, our models feature shape- and translation-invariance, by which we mean that the model output for a forested region remains the same irrespectively of the choices of modifiable analysis units that partition the region as well as the origins where these units are placed. However, we suspect that most previously developed models are subject to not only scale-dependence but also shape- and translation-dependence. A typical

manifestation of shape-dependence is that the division of a study area into square pixels in two different map projection systems cannot establish correspondence because square pixels become parallelograms when transformed from one projection to another. Shape- and translation-dependence have not been fully examined in this study, but they could also be caused by the model nonlinearity and the non-scaling property of predictors. As such, our scale-invariant models hold great promises for more effective and wider uses in multi-scale forest inventory on variable analysis units that have been obtained by segmenting CHM as reported in van Aardt et al. (2006) and Zhao and Popescu (2007). For example, an extreme case in which our models are applicable is to estimate total AGB using profiling lasers where the analysis units are 1-dimensional lines or transects (Nelson et al., 2003b; Zhao et al., 2008a).

Two assumptions, i.e., non-negativity and non-decreasing, have been made about the K function. However, more stringent constraints can be imposed on K. For example, we can assume that K takes certain parametric forms. A simple choice is $K(h)=a\cdot h+b$ that gives a familiar model,

$$M = \int (a \cdot h + b) p(h) \mathrm{d}h = a\overline{h} + b \tag{20}$$

where $\overline{h}$ is the mean canopy height. Other rather complicated parametric forms, e.g., $K(h)=a\ \cdot[\exp(b\cdot h)-1]$ or $a\cdot[h^b-1]$, may be explored in further studies. In such cases, the inference of K essentially becomes the estimation of only its functional parameters such as a and b of the above examples. In addition, we have presumed that $K(0)=0$ such that the biomass on a plot empty of trees is zero. This assumption makes our models different from those that produce non-zero biomass for non-tree plots (e.g., zero canopy height). Patenaude et al. (2004) explained that such initial values at zero lidar-recorded height represent the biomass in ground vegetation and litter compartments. In such a sense, the model of this study only considers the biomass present in "standing" trees. However, as to those models of non-zero initial biomass for non-tree plots, a possible concern arises that the prediction for zero or small canopy heights may risk anomalous biomass values associated with extrapolation because these low-canopy plots (e.g., clear-cut) are sometimes excluded from the model fitting, e.g., for numerical reasons (Lim and Treitz, 2004). To this end, the models proposed by us have the advantage of being applicable across a wide range of biomass values, and yields exactly zero biomass value for open grounds where no trees are present.

This study used a constrained regression procedure to infer the K function during model fitting. Alternatively, we recommend that future studies investigate the usefulness of advanced data mining techniques such as support vector machines (SVM) and Gaussian Processes (GPs) to uncover the possibly nonlinear relationships between AGB and CHDs or CHQs because CHDs or CHQs are often high-dimensional functional data resulting from the discretization of curves and previous studies reported the successful applications of these techniques for supervised learning with high-dimensional data (Durbha et al., 2007; Zhao, Popescu, & Zhang, 2008b). We also strongly suggest the use of these machine learning techniques for estimating other forest structural attributes from lidar data.

Another important aspect about scale or plot size concerns our model formulation. By analogy to the formulation of radiative transfer theory in forest canopies (Ross, 1981), the mathematical derivation of our models involves some contradictory assumptions: On the first hand, we expect that the number of trees within a plot is small enough so that not too much overlapping exists to validate the first-order approximation of percent crown cover in the Boolean model; on the other hand, we assume that there is a large sample of trees for justifying continuous distributions of relevant variables. The number of trees largely depends on the plot size that is used to define the problem. Nevertheless, it is advantageous in practice to train a model on large plots, but the choices of cell size for prediction make no difference as far as the total biomass in the study area is concerned. However, be aware that creating a biomass map with our approach at an extremely-fine scale, e.g., 0.1 × 0.1 m, makes no physical sense at all. Our models are expected to predict biomass at least above individual tree levels.

The increase in R^2 with plot sizes as noted in Table 3 is likely due to plot-edge effects: Because tree stems contain a high percentage of total tree biomass and the methods for extracting predictor and

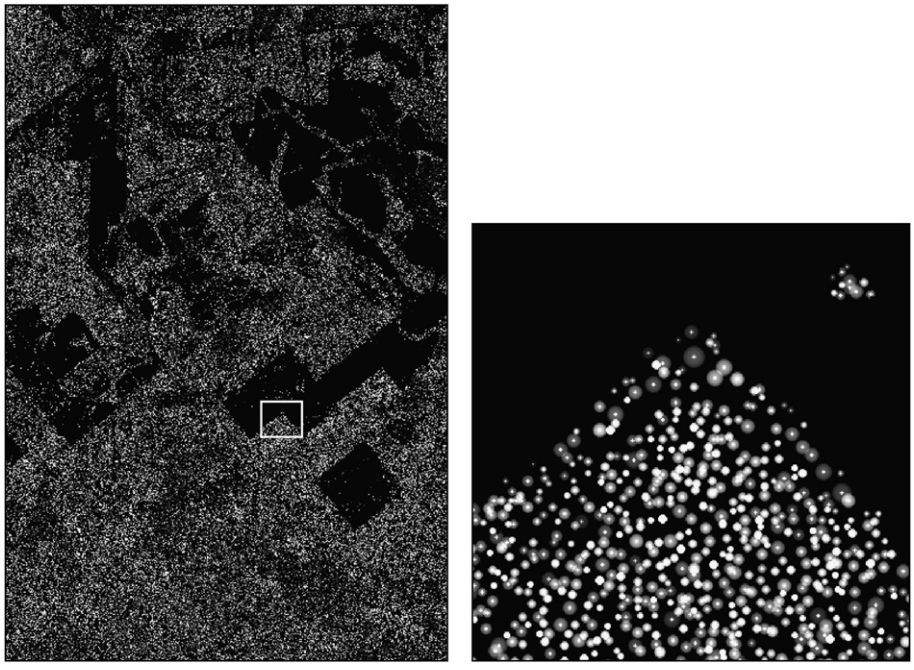


Fig. 12. The 0.5 × 0.5 m fine-resolution biomass map derived by the combined use of lidar scanner data and Quickbird image with the individual tree-based approach (left). A close-up of the subset highlighted by the white rectangle is also displayed (right). The map uses a gray scheme such that brighter pixels indicate higher biomass.

dependent variables count only the components that fall into a plot, discrepancies occur whenever a tree stem falls outside the plot while a large portion of the tree crown falls inside the plot or vice versa. The edge effect is more pronounced for small plots (Andersen et al.; 2005). However, the edge effect has minor influence on the model parameter inference (i.e., the K function) as demonstrated by the statistically indistinguishable mean functions of K that were estimated for different plot sizes (Fig. 6). In practice, it may be desirable to undertake field-work on larger plots while keeping a reasonable cost. If the inventory budget is limited, an alternative is to first construct a computer-based canopy simulator based on the limited ground data and then run the simulator to generate synthesized observations for model development (Nelson, 1997).

The overestimation in the overall biomass may be attributed to two factors. First, the reference value of total biomass, obtained by averaging the lidar-derived biomass map, only captures the biomass in those trees that were identified from CHM by the tree-delineating algorithm. Due to omission errors of the algorithm, the reference value represents an underestimate of the true biomass in the vegetative components captured by the CHM. This point is illustrated in the scatterplots of Fig. 5, the lower left corners of which show a plot-level reference value of zero due to absence of identified trees, but the predicted values are small yet non-zero due to the presence of some canopies in the CHM. This discrepancy also leveraged all the fitted lines in the scatterplots of prediction vs. reference to tilt upward near zero biomass as shown by the solid lines compared with the 1:1 lines (Fig. 5). Second, although we attempted to mask out the non-forest area by referring to the Quickbird image, not all the non-vegetative components are removed from CHM by doing so. Therefore, the remaining man-made structures in the masked CHM contributed to the overestimation of biomass.

6. Conclusions

This work adds to the extant evidence that lidar, with carefully designed models, offers an effective means for remotely measuring forest biomass at a range of scales from individual tree, plot, stand, local or even up to regional levels. To predict above-ground biomass at various scales using the same model, we proposed the use of lidar-derived CHDs and CHQs as independent variables, and accordingly presented two scale-invariant functional models. In addition to scale-invariance, a major consideration for choosing these predictors is to incorporate as much information inherent in lidar data as possible into the models. To provide theoretical justification for the model formulation, we developed an accompanied mathematical framework based on several moderate assumptions. To obtain reference data for model evaluation, we derived a fine-resolution biomass map with an individual tree-based approach.

Results of the individual tree-based approach demonstrate the utility of lidar for detecting single trees and measuring various tree biophysical variables such as height, crown width, crown base height, and component biomass. Results from the intensive model evaluations based on the synthesized realistic datasets provide initial proofs that our models can accurately predict biomass from low to high levels and that our models have consistent predictive performances across a variety of scales. Since the predictor CHD is obtained in the form of normalized histograms with a specified sampling height-bin, we found that models with a carefully selected large height-bin (e.g., less than 5.0 m) can be as effective as those with finer height bins. The results also show that in our experiments, a training sample size of around 50 or less is enough to guarantee a good fitting of the linear functional model as long as the training sample is representative of the forest conditions that the model will be applied to. No attempt was made to directly fit the nonlinear model, and future studies may examine the effectiveness of advanced nonlinear regression or machine learning techniques such as SVM and Gaussian Processes in regressing biomass on the high-dimensional data of the discretized CHQ. For further research, it is expected that the proposed models can be used to predict forest structural variables other than AGB, such as stem density, basal area, timber volume, crown fuel weight and LAI, using either CHDs or lidar height distributions of original laser hits as predictor. Future research should also be carried out to apply the models of this study in forest inventory tasks where analysis units vary in size and shape. To augment the applicability of our models, it is suggested that auxiliary information from multispectral imagery should be integrated with lidar ranging data, e.g., for developing strata-specific models.

Acknowledgments

I thank my collaborators Emm. Vassilakis, I. Papanikolaou and L. We gratefully acknowledge the support provided by the Texas Forest Service (award #: 02-DG-11083148-050) and a NASA New Investigator Program in the earth science. We thank Curt Stripling, all forestry personnel of the Texas Forest Service, Muge Mutlu and Alicia Griffin for their help with field data collection, Jared Stukey for his patient proofreading of the first draft of the paper. We are also greatly appreciative of the insightful comments from the three anonymous reviewers. Kaiguang Zhao would like to thank Drs. X. Ben Wu and Marian Eriksson for the helpful discussion.

References

Andersen, H. E., McGaughey, R. J., & Reutebuch, S. E. (2005). Estimating forest canopy fuel parameters using lidar data. *Remote Sensing of Environment*, *94*, 441–449.

Aster, R., Borchers, B., & Thurber, C. (2004). *Parameter estimation and inverse problems.* New York: Academic Press 230 pp.

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

Chen, Q., Baldocchi, D., Gong, P., & Kelly, M. (2006). Isolating individual trees in a savanna woodland using small footprint lidar data. *Photogrammetric Engineering & Remote Sensing*, *72*(8), 923–932.

Durbha, S. S., King, R. L., & Younan, N. H. (2007). Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sensing of Environment*, *107*, 348–361.

Holmgren, J. (2004). Prediction of tree height, basal area and stem volume using airborne laser scanning. *Scandinavian Journal of Forest Research*, *19*, 543–553.

Climate change 2001: The scientific basis. Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Xiaosu, D., Maskell, K., & Johnson, C. A. (Eds.). (2001). *Contribution of working group I to the third assessment report of the intergovernmental panel on climate change* Cambridge: UK7 Cambridge University Press 944 pp..

Houghton, R. A. (2005). Aboveground forest biomass and the global carbon balance. *Global Change Biology*, *11*, 945–958.

Jenkins, J. C., Chojnacky, D. C., Heath, L. S., & Birdsey, R. A. (2003). National-scale biomass estimators for United States tree species. *Forest Science*, *49*, 12–35.

Kotchenova, S. Y., Shabanov, N. V., Knyazikhin, Y., Davis, A. B., Dubayah, R., & Myneni, R. B. (2003). Modeling lidar waveforms with time-dependent stochastic radiative transfer theory for remote estimations of forest structure. *Journal of Geophysical Research*, *108*(D15), 4484.

Lefsky, M. A., Cohen, W. B., Acker, S. A., Parker, G. G., Spies, T. A., & Harding, D. (1999). Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sensing of Environment*, *70*, 339–361.

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

Li, X., & Strahler, A. H. (1992). Geometric-optical bidirectional reflectance modeling of the discrete crown vegetation canopy: Effect of crown shape and mutual shadowing. *IEEE Transaction on Geosciences and Remote Sensing*, *30*, 276–292.

Li, F., Zhang, L., & Davis, C. J. (2002). Modeling the joint distribution of tree diameters and heights by bivariate generalized beta distribution. *Forest Science*, *48*(1), 47–58.

Lim, K. S., & Treitz, P. M. (2004). Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. *Scandinavian Journal of Forest Research*, *19*(6), 558–570.

Lim, K., Treitz, P., Baldwin, K., Morrison, I., & Green, J. (2003). Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. *Canadian Journal of Remote Sensing*, *29*, 648–678.

Maltamo, M., Eerikäinen, K., Packalén, P., & Hyyppä, J. (2006). Estimation of stem volume using laser scanning-based canopy height metrics. *Forestry*, *79*(13), 217–229.

Maltamo, M., Kangas, A., Uuttera, J., Torniainen, T., & Saramäki, J. (2000). Comparison of percentile based predicted methods and Weibull distribution in describing diameter distribution of heterogeneous Scots pine stands. *Forest Ecology and Management*, *133*, 263–274.

Marceau, D. J. (1999). The scale issue in social and natural sciences. *Canadian Journal of Remote Sensing*, *25*, 347–356.

Marceau, D. J., Howarth, P. J., & Gratton, D. J. (1994). Remote sensing and the measurement of geographical entities in a forested environment. 1: The scale and spatial aggregation problem. *Remote Sensing of Environment*, *49*(2), 93–104.

Means, J. E., Acker, S. A., Harding, D. J., Blair, J. B., Lefsky, M. A., Cohen, W. B., et al. (1999). Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the western Cascades of Oregon. *Remote Sensing of Environment*, *67*, 298–308.

Mutlu, M., Popescu, S. C., Stripling, C., & Spencer, T. (2008). Assessing surface fuel models using lidar and multispectral data fusion. *Remote Sensing of Environment*, *112*, 274–285.

Næsset, E. (1997). Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment*, *61*, 246–253.

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

Næsset, E. (2004). Practical large-scale forest stand inventory using a small footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, *19*, 164–179.

Næsset, E., & Bjerknes, K. -O. (2001). Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. *Remote Sensing of Environment*, *78*, 328–340.

Næsset, E., & Gobakken, T. (2008). Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment*, *112*, 3079–3090.

Nelson, R. (1997). Modeling forest canopy heights: The effects of canopy shape. *Remote Sensing of Environment*, *60*, 327–334.

Nelson, R., Krabill, W., & Tonelli, J. (1988). Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment*, *24*, 247–267.

Nelson, R., Short, A., & Valenti, M. (2004). Measuring biomass and carbon in Delaware using airborne profiling Lidar. *Scandinavian Journal of Forest Research*, *19*, 500–511.

Nelson, R., Valenti, M. A., Short, A., & Keller, C. (2003a). A multiple resource inventory of Delaware using airborne laser data. *Bioscience*, *53*, 981–992.

Nelson, R., Parker, G. G., & Horn, M. (2003b). A portable airborne laser system for forest inventory. *Photogrammetric Engineering and Remote Sensing*, *69*, 267–273.

Omasa, K., Qiu, G. Y., Watanuki, K., Yoshimi, K., & Akiyama, Y. (2003). Accurate estimation of forest carbon stocks by 3-D remote sensing of individual trees. *Environmental Science and Technology*, *37*, 1198–1201.

Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the Earth Sciences. *Science*, *263*, 641–646.

Patenaude, G., Hill, R. A., & Milne, R. (2004). Quantifying forest above ground carbon content using lidar remote sensing. *Remote Sensing of Environment*, *93*, 368–380.

Popescu, S. C. (2007). Estimating biomass of individual pine trees using airborne Lidar. *Biomass and Bioenergy*, *31*, 646–655.

Popescu, S. C., & Wynne, R. H. (2004). Seeing the trees in the forest: Using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photogrammetric Engineering & Remote Sensing*, *70*, 589–604.

Popescu, S. C., Wynne, R. H., & Nelson, R. H. (2002). Estimating plot-level tree heights with LIDAR: Local filtering with a canopy-height based variable window size. *Computers and Electronics in Agriculture*, *37*(1–3), 71–95.

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

Popescu, S. C., & Zhao, K. G. (2008). A voxel-based lidar method for estimating crown base height for deciduous and pine trees. *Remote Sensing of Environment*, *112*(3), 767–781.

Riano, D., Chuvieco, E., Condés, S., Gonzalez-Matesanz, J., & Ustin, S. L. (2004). Generation of crown bulk density for *Pinus sylvestris* L. from lidar. *Remote Sensing of Environment*, *92*, 345–352.

Ross, J. K. (1981). *The radiation regime and architecture of plant stands.*The Hague: Dr W. Junk Publishers 391 pp.

Strahler, A. H., Woodcock, C. E., & Smith, J. A. (1986). On the nature of models in remote sensing. *Remote Sensing of Environment*, *20*, 121–139.

Sun, G., & Ranson, K. J. (2000). Modeling lidar returns from forest canopies. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(6), 2617–2626.

Tewari, V. P., & Gadow, K. V. (1999). Modelling the relationship between tree diameters and heights using SBB distribution. *Forest Ecology and Management*, *119*, 171–176.

Turner, M. G., Dale, V. H., & Gardner, R. H. (1989). Predicting across scales: Theory development and testing. *Landscape Ecology*, *3*, 245–252.

van Aardt, J. A. N., Wynne, R. H., & Oderwald, R. G. (2006). Forest volume and biomass estimation using small-footprint Lidar-distributional parameters on a per-segment basis. *Forest Science*, *52*(6), 636–649.

Woodcock, C. E., & Strahler, A. H. (1987). The factor of scale in remote sensing. *Remote Sensing of Environment*, *21*, 311–332.

Yu, X., Hyypp, J., Kaartinen, H., & Maltamo, M. (2004). Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. *Remote Sensing of Environment*, *90*, 451–462.

Zhao, K., & Popescu, S. C. (2007). Hierarchical watershed segmentation of canopy height model for multi-scale forest inventory. In P. Rönnholm, H. Hyyppä, & J. Hyyppä (Eds.), *Proceedings of the ISPRS working group "Laser Scanning 2007 and SilviLaser 2007", ISPRS Volume XXXVI, Part3/W52. Espoo, September 12–14, 2007, Finland* (pp. 436–442).

Zhao, K., Popescu, S. C., & Nelson, R. F. (2008a). Quantifying the uncertainty for the line-intercept sampling estimators of canopy cover. *Journal of Forest Planning (Japanese Society of Forest Planning), 13*, 195–205.

Zhao, K., Popescu, S. C., & Zhang, X. (2008b). Bayesian learning with Gaussian processes for supervised classification of hyperspectral data. *Photogrammetric Engineering & Remote Sensing*, *74*(10), 1223–1234.

Zavitkovski, J. (1976). Ground vegetation biomass, production, and efficiency of energy utilization in some northern Wisconsin forest ecosystems. *Ecology*, *57*, 694–706.