Estimating Basal Area and Stem Volume for Individual Trees from Lidar Data

Qi Chen, Peng Gong, Dennis Baldocchi, and Yong Q. Tian

Abstract

This study proposes a new metric called canopy geometric volume G, which is derived from small-footprint lidar data, for estimating individual-tree basal area and stem volume. Based on the plant allometry relationship, we found that basal area B is exponentially related to G ($B = \beta_1 G^{3/4}$, where β_1 is a constant) and stem volume V is proportional to G ($V = \beta_2 G$, where β_2 is a constant). The models based on these relationships were compared with a number of models based on tree height and/or crown diameter. The models were tested over individual trees in a deciduous oak woodland in California in the case that individual tree crowns are either correctly or incorrectly segmented. When trees are incorrectly segmented, the theoretical model $B = \beta_1 G^{3/4}$ has the best performance (adjusted R^2, $R_a^2 = 0.78$) and the model $V = \beta_2 G$ has the second to the best performance ($R_a^2 = 0.78$). When trees are correctly segmented, the theoretical models are among the top three models for estimating basal area ($R_a^2 = 0.77$) and stem volume ($R_a^2 = 0.79$). Overall, these theoretical models are the best when considering a number of factors such as the performance, the model parsimony, and the sensitivity to errors in tree crown segmentation. Further research is needed to test these models over sites with multiple species.

Introduction

Accurate forest structural information is crucial to a number of applications including forest management (Maltamo *et al.*, 2004a), fire behavior analysis (Riaño *et al.*, 2004), and global warming and carbon management (Birdsey and Lewis, 2002). Compared to optical remotely sensed imagery, lidar (Light Detection and Ranging) can directly measure the vertical canopy information and is gaining popularity in forest and ecological studies (Lefsky *et al.*, 2002a; Hall *et al.*, 2005). With the high pulse-density of small-footprint lidar data, nowadays it is possible to isolate individual trees and directly extract individual-tree locational and dimensional parameters including treetop locations, tree heights, crown sizes, and even crown boundaries (e.g., Popescu *et al.*, 2003; Hyyppä *et al.*, 2001; Persson *et al.*, 2002). We developed a new method to delineate individual tree crowns in a savanna woodland in California using small-footprint lidar data (Chen *et al.*, 2006). This companion study is to further extract individual-tree structural parameters including basal area and stem volume, which cannot be directly measured by lidar, but can be potentially related to tree dimensional information such as tree height and crown size. This study is part of a large project, in which we parameterize an individual-tree based biogeochemical model called "MAESTRA" (Medlyn, 2003) for spatially-explicit ecological modeling.

Extracting individual-tree structural parameters from small-footprint lidar data is useful not only for our detailed ecological modeling but also for the large-scale forest inventory. There are at least three kinds of methods to perform the large-scale forest inventory with small-footprint lidar data: the first is called *stand-level regression method*. This method is to create regression models between canopy structural parameters and the laser pulses statistics at the stand or plot levels (Means *et al.*, 2000; Holmgren *et al.*, 2003; Næsset *et al.*, 2004; Næsset, 2004; Riaño *et al.*, 2004; Hall *et al.*, 2005). The statistics used in the regression models typically include the height mean, minimum, maximum, variance, coefficient of variance, percentiles, and the percentage of canopy returns.

The second method is called *individual-tree integration method*. If the tree dimensional information such as individual tree height and crown size can be accurately extracted, the individual tree canopy structure parameters can be derived based on allometric equations or regression models (Hyyppä, *et al.*, 2001; Persson *et al.*, 2002; Næsset and Økland, 2002; Riaño *et al.*, 2004; Roberts *et al.*, 2005). Then, the canopy structure information over a large area can be obtained by simply integrating individual tree values. Compared to the *stand-level regression method*, the ground truth canopy structure measurements for developing the regression models are needed only for a sample of trees instead of many of plots or stands, which can significantly reduce the fieldwork. However, not many studies have used

Qi Chen is with the Department of Geography, University of Hawaii at Manoa, Honolulu, HI 96822 (qichen@hawaii.edu).

Dennis Baldocchi is with the Department of Environmental Science, Policy, and Management, 137 Mulford Hall, University of California at Berkeley, Berkeley, CA 94720 (qch@nature.berkeley.edu).

Peng Gong is with the Center for the Assessment and Monitoring of Forest and Environmental Resources (CAM-FER), 137 Mulford Hall, University of California at Berkeley, Berkeley, CA 94720, and the State Key Lab of Remote Sensing Science (jointly sponsored by Institute of Remote Sensing Applications, Chinese Academy of Sciences, and Beijing Normal University), Postal Box 9718, Beijing, P.R. China 100101.

Yong Q. Tian is with the Department of Environmental, Earth and Ocean Sciences, University of Massachusetts-Boston, Boston, MA 02125.

the *individual-tree integration method* for the large-scale forest inventory because previous research has shown that the accuracy of isolating individual trees is typically low due to the complexity of canopy surface (Persson *et al.*, 2002; Brandt-berg *et al.*, 2003; Leckie *et al.*, 2003; Maltamo *et al.*, 2004a; Morsdorf *et al.*, 2004). We applied a method by Popescu and Wynne (2004) to our study site, a woodland of deciduous oak trees, and obtained an accuracy of 37 percent; when we used an improved method that can reduce both the commission errors and omission errors, the accuracy increased to 64.1 percent, which is still not very high (Chen *et al.*, 2006).

Although the low accuracy of the individual tree analysis methods makes it difficult to integrate individual tree values to obtain the forest canopy structure information over a large area (Riaño *et al.*, 2004), especially in deciduous forests, some researchers found that the individual tree analysis results are useful for predicting stand-level information with a method called *hybrid regression method*. In this method, a series of plot or stand-level statistics are first derived based on the individual tree isolation results, such as the total number of trees, the maximum and mean tree height, and the mean crown size. Then, these statistics are used in the regression models for predicting canopy structure parameters at the plot or stand-level (Holmgren *et al.*, 2003; Popescu *et al.*, 2003; Popescu *et al.*, 2004; Maltamo *et al.*, 2004b). Such a method might improve the accuracy of forest plot/stand structural information prediction (Popescu *et al.*, 2003) but cannot reduce the fieldwork since the ground-truth measurements still need be collected at the plot or stand level for developing the models.

Since the *individual-tree integration* method can minimize the fieldwork, it is expected to have great potential in the large-scale forest inventory, especially if we can develop some individual-tree level models that are not much affected by the errors in tree isolation and crown delineation. To search for such models, let us assume that the structure parameter Y_i for a tree i is related to a lidar metric X_i by:

$$Y_i = f(X_i). \tag{1}$$

Suppose this tree is over-segmented into n parts and the corresponding metrics are: $X_{i,1}, X_{i,2}, \ldots, X_{i,n}$; then, the predicted structure parameter $Y_{p,i}$ is:

$$Y_{p,i} = f(X_{i,1}) + f(X_{i,2}) + \cdots + f(X_{i,n}). \tag{2}$$

If assume that (a) the structural parameter Y is proportional to the lidar metric X ($Y = \alpha X$, where α is a constant) even at the scale of a part of a tree, which means:

$$f(X_{i,1}) + f(X_{i,2}) = f(X_{i,1} + X_{i,2}) \tag{3}$$

and (b) the sum of the lidar metrics for each part is equal to the lidar metric for the tree:

$$X_{i,1} + X_{i,2} + \cdots + X_{i,n} = X_i; \tag{4}$$

then, $Y_{p,i} = f(X_{i,1}) + f(X_{i,2}) + \cdots + f(X_{i,n}) = f(X_{i,1} + X_{i,2} + \cdots + X_{i,n}) = f(X_i) = Y_i$. This means that if Equations 3 and 4 are satisfied, the predicted canopy structure parameter will not be affected by the over-segmentation of tree crowns. It is easy to verify that this is also true in the case of the under-segmentation of tree crowns. Our hypothesis is that if either or both equations are satisfied the prediction of canopy structural parameters is not much affected by the errors of tree isolation and crown segmentation. So, now the question is to find some metrics that satisfy Equation 4 and test whether the canopy structural parameter is proportional to the metrics (Equation 3).
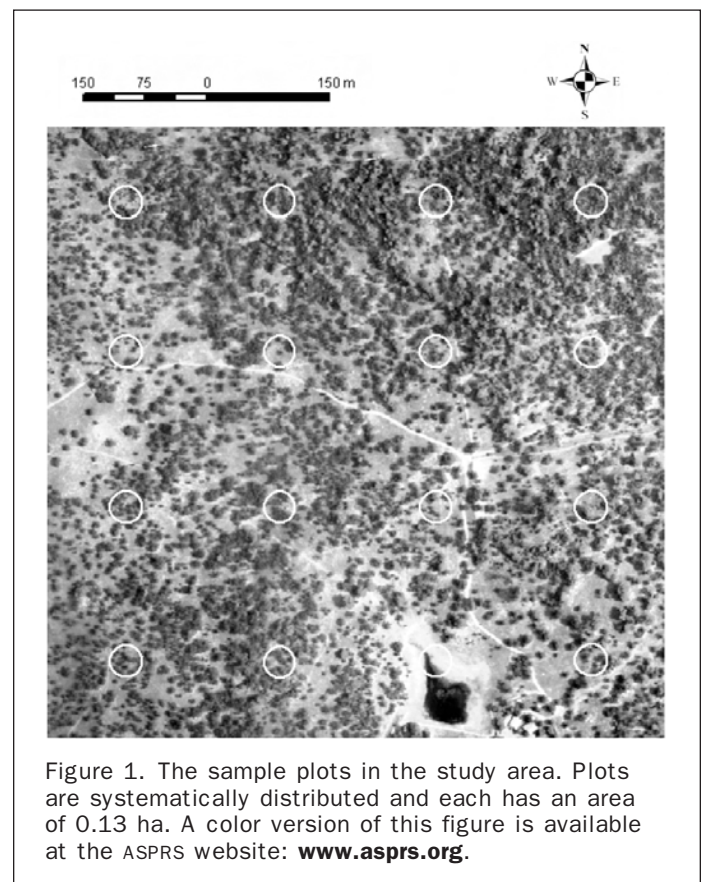
There are at least two lidar metrics that satisfy Equation 4: crown area and canopy geometric volume. Crown area has been used in a number of studies for estimating canopy structure parameters (e.g., Holmgren *et al.*, 2003; Maltamo *et al.*,

2004b). However, canopy geometric volume (denoted as G) is a new metric first proposed in this study. Therefore, the main objective of this study is to examine (a) whether basal area and stem volume are proportional to canopy geometric volume, and if not, what the functional relationships are, and (b) how the different models based on canopy geometric volume perform when tree crowns are incorrectly segmented. Since trees are growing in a three-dimensional space, it is hypothesized that their canopy structural characteristics can be better predicted with three-dimensional metrics such as canopy geometric volume. Thus, the second objective of our study is to compare the models based on canopy geometric volume with those based on either canopy height or crown size metrics. In the companion study on tree isolation (Chen *et al.*, 2006), we found that some tree crowns were over-segmented (a tree on the ground is segmented into several parts) or under-segmented (several trees on the ground correspond to only one segment) due to the irregular canopy shape and the coexistence of dominant and co-dominant trees in deciduous forests. Correspondingly, we will address the above research questions by testing the models over correctly and incorrectly segmented trees, respectively.

Materials and Methods

Study Area
The study site is an open oak savanna woodland, located near Ione, California (latitude: 38.26°N, longitude: 120.57°W) (Figure 1). The site is on a private ranch and is part of the AmeriFlux network of eddy covariance field sites (Baldocchi *et al.*, 2004). The landscape is characterized by flat terrain (with a maximum slope of less than 15 percent) with a

Figure 1. The sample plots in the study area. Plots are systematically distributed and each has an area of 0.13 ha. A color version of this figure is available at the ASPRS website: **www.asprs.org**.

scattered, clumped distribution of blue oaks (*Quercus douglasii*) and a minority of grey pines (*Pinus sabiniana*) over a continuous layer of Mediterranean annual grasses.

Lidar Data

On 24 August 2003, laser altimetry data were acquired with an Optech ALTM 2025, which recorded both first and last returns for each laser pulse. The scanning pattern was z-shaped. The scanning angle was 17°, and the flying altitude was about 500 m, corresponding to a swath of about 300 m. The data provider claimed that the vertical accuracy and horizontal accuracy were 18 cm and 17 cm, respectively, with 95 percent confidence. The footprint size was about 18 cm. The average posting density was 9.5 points per square meter, resulting in an average spot spacing of about 32 cm. To obtain such a high pulse density, the site was flown over twice. The data covering 800 by 800 m around the eddy covariance tower were used to segment individual tree crowns (Chen *et al.*, 2006).

Field Data

Ground truth data were collected for 16 circular plots systematically distributed over the study area, with a spacing of 200 m in two perpendicular directions (Figure 1). Each plot has an area of about 0.13 ha and a radius of 20 m. The centers of the plots were located with a Trimble AgGPS 132 receiver. The claimed "pass-to-pass" accuracy measured over a 15-minute time period is 10 to 30 cm. Under forest canopy, GPS systems tend to produce from 1.5 to 3 times less accurate solutions (Popescu *et al.*, 2003). Since the trees in this woodland are more sparsely distributed than those in the Popescu *et al.* study site, it is expected that the accuracy of plot center measurements was at the sub-meter level.

The tree height and diameter at breast height (DBH) were measured for all trees larger than 12.5 cm DBH. A total of 313 oak trees and eight grey pines were located in these plots. Since the number of grey pines is too small to perform reliable regression analysis, this study only focused on the blue oak trees. The tree height was measured to the nearest 0.1 m with a laser rangefinder (Opti-LOGIC, Tullahoma, Tennessee) and the DBH was measured to the nearest 0.01 m. The basal area of individual trees

was calculated from the measurements of DBH. Based on the measurements of tree height and DBH, the stem volume for individual trees was calculated using the blue oak allometric equation developed by Pillsbury and Kirkley (1984).

Lidar Metrics

The tree crown segments automatically generated from lidar data (Chen *et al.*, 2006) were used to extract relevant the lidar metrics for basal area and stem volume estimation. There were a total of 291 segments corresponding to the blue oaks in these plots. For each tree segment, a total of 44 metrics (42 height metrics, crown area, and canopy geometric volume G) were extracted (Table 1).

All of these metrics except the canopy geometric volume G have been applied in previous studies (Means *et al.*, 2000; Næsset and Økland, 2002; Popescu *et al.*, 2003). The canopy height metrics H_x were calculated based on the laser pulses falling in each segment. The height statistics, including percentile from 0 to 100 by 10, mean, standard deviation, and coefficient of variance, were calculated for first, last, and all canopy returns, respectively. Canopy returns were defined as the returns that are 0.5 m higher than the ground. The crown area C_d is the area of each crown segment. The canopy geometric volume G of each segment is the volume under the canopy height model (CHM). To produce a CHM, a DEM was first created with a method developed by Chen *et al.* (in press) and then the relative canopy height of laser pulses was interpolated into a CHM by kriging. The details of generating the CHM were described in Chen *et al.* (2006).

Regression Analysis

The regression models were developed only based on the correctly segmented trees, which are those that have one-to-one relationship with the lidar-derived tree segments. A one-to-one relationship means that the overlaying area S_o between a crown polygon validated in the field and its corresponding segment is within the range of $S_r \pm 10\% * S_r$, where S_r is the area of the validated crown polygon (Chen *et al.*, 2006). The regression models for predicting basal area and stem volume are summarized in Tables 2 and 3, respectively.

TABLE 1. THE LIDAR METRICS USED IN THE REGRESSION MODELS

Category	Metrics	Dependent Variables
Canopy height (H_x, m)	For first canopy returns, $H_{i,f}$, where i=0,10, …,100%, percentile height $H_{mean,f}$, mean of height $H_{std,f}$, stand deviation of height $H_{cv,f}$, coefficient of variance of height For last canopy returns, $H_{i,l}$, where i=0,10, …,100%, percentile height $H_{mean,l}$, mean of height $H_{std,l}$, stand deviation of height $H_{cv,l}$, coefficient of variance of height For all canopy returns, $H_{i,a}$, where i=0,10, …,100%, percentile height $H_{mean,a}$, mean of height $H_{std,a}$, stand deviation of height $H_{cv,a}$, coefficient of variance of height	Basal Area: (B, m²) Stem Volume: (V, m³)
Crown area (C_a, m)	The area of individual crown segment	
Canopy geometric volume (G, m³)	The geometric volume for a segment under the CHM	

TABLE 2. THE MODELS FOR PREDICTING BASAL AREA

No.	General Form of the Models	Number of Specific Models
B.1	$B = \beta_1 H_x^c$	42
B.2	$B = \beta_1 C_a^c$	1
B.3	$B = \beta_1 G^c$	1
B.4	$B = \beta_1 G$	1
B.5	$B = \beta_1 G^{3/4}$	1
B.6	DBH $= \beta_0 + \beta_1 H_x + \beta_2 C_a$	42
	$B = \pi(\text{DBH}/2)^2$	
B.7	B=stepwise($\{H_x\}$, C_a, G)	1
B.8	B=stepwise($\{H_x\}$, C_a)	1
B.9	B=stepwise($\{H_x\}$)	1

The univariate power models were developed with tree height, crown area, or geometric canopy volume as independent variables, respectively (Models B.1–3 and V.1–3). These power functions are justified by the theories in plant science: many structure and functional variables of organisms (Y) scale as power functions of measures of sizes (S) such as body mass, length, diameter, area, and volume (Norberg, 1988; West *et al.*, 1999; Enquist, 2002):

$$Y = \alpha S^\beta, \tag{5}$$

where α is a constant that varies with the type of variables and the kind of plants, and β is the allometric exponent. The power models, especially with height, have been used in a number of studies (e.g., Lim *et al.*, 2003; Hall *et al.*, 2005).

With canopy geometric volume as the independent variable, we test two kinds of models. First, a simple linear regression model with no intercept was tested for basal area and stem volume, respectively, since such a model satisfies Equations 3 and 4 and therefore has an advantage of not being affected by the errors of tree isolation (Models B.4 and V.4). Second, we developed theoretical models based on the recent progresses in the plant allometry; Enquist (2002) found such allometric relationships: DBH $\propto M^{3/8}$ and $H \propto M^{1/4}$, where M is the plant mass and H is the tree height. If assuming that the plant is filled with a tissue density that is approximately constant across sizes, the plant mass M is

TABLE 3. THE MODELS FOR PREDICTING STEM VOLUME

No.	General Form of the Model	Number of Specific Models
V.1	$V = \beta_1 H_x^c$	42
V.2	$V = \beta_1 C_a^c$	1
V.3	$V = \beta_1 G^c$	1
V.4	$V = \beta_1 G$	1
V.5	DBH $= \beta_0 + \beta_1 H_x + \beta_2 C_a$	42
	V=allometric(DBH, $H_{100,a}$)	
V.6	V=stepwise($\{H_x\}$, C_a, G)	1
V.7	V=stepwise($\{H_x\}$, C_a)	1
V.8	V=stepwise($\{H_x\}$)	1

Note: $\{H_x\}$ means the set of all height metrics.

proportional to canopy geometric volume G (West *et al.*, 1999; Enquist, 2002). Therefore, DBH $\propto G^{3/8}$ and $H \propto G^{1/4}$. Since basal area $B \propto \text{DBH}^2$ and stem volume $V \propto \text{DBH}^2*H$,

$$B \propto \text{DBH}^2 \propto G^{3/4}, \text{ and} \tag{6}$$

$$V \propto \text{DBH}^2*H \propto V*V^{1/4} \propto G \tag{7}$$

Equation 6 indicates that theoretically the basal area is not proportional to the canopy geometric volume G, based on which Model B.5 is developed. However, Equation 7 indeed reveals that the stem volume V is proportional to canopy geometric volume G. This equation has significant implications since it can provide theoretical support for Model V.4.

Although these allometric equations are developed at the individual tree level, there is a biological basis for them to be applied over a part of canopy: most terrestrial plants have a transport system that moves water, minerals, and nutrients through the plant body by plant tissues such as xylem and phloem. To support the maintenance and growth of leaves and branches within any certain portion of the canopy, there are corresponding conducting tubes in the stem for transporting water and nutrients (West *et al.*, 1999), which correspond to a portion of basal area or stem volume. Therefore, canopy geometric volume can be related to basal area or stem volume even at the scale of a portion of an individual tree.

Besides the above univariate models, we also tested a multiple linear regression model that depend on both a height metric and crown area (Models B.6 and V.5). The hypothesis for these models is that the combination of the height and the crown area can characterize the canopy structure in a three-dimensional space (Hyyppä *et al.*, 2001; Holmgren *et al.*, 2003; Hall *et al.*, 2005). Finally, considering their popularity in many studies (Means *et al.*, 2000; Næsset, 2002; Popescu *et al.*, 2003), we tested stepwise regression models that depend on all metrics (Models B.7 and V.6), all height metrics plus crown area (Models B.8 and V.7), or all height metrics (Models B.9 and V.8), respectively.

Model Evaluation

The adjusted coefficient of determination and root mean square error were used for model evaluation. Although different formulas of R^2 exist, Kvålseth (1985) recommended the following equation to calculate the adjusted R^2:

$$R^2 = 1 - \frac{n-1}{n-p}\frac{\Sigma(y-\hat{y})^2}{\Sigma(y-\bar{y})^2} \tag{8}$$

where $\hat{y}$ is the fitted value of basal area or stem volume y, $\bar{y}$ is the mean of all y's, n is the number of observations, and p is the number of parameters in the regression model, not including the residual variance. The coefficient of determination (R^2) is perhaps the single most extensively used measure of *goodness of fit* for regression models (Kvålseth, 1985). However, there are problems with R^2 for the no-intercept model and for transformed variables (Anderson-Sprecher, 1994). Therefore, in addition to the above two statistics, the Akaike's information criterion (AIC) was used to compare different models.

Akaike's information criterion was developed from the Kullback-Leibler information, which can be used to quantify the distance between a regression model and reality. Based on the philosophy that reality cannot be modeled, AIC is to calculate the "relative" distance between the regression model and reality. In practice, the second-order bias correction version called AIC$_c$ is used, especially when the sample size is small (Burnham and Anderson, 2002):

$$\text{AIC}_c = n\log\left(\frac{\Sigma(y-\hat{y})^2}{n}\right) + 2K + \frac{2K(K+1)}{n-K-1}. \tag{9}$$

In this case, K is the total number of parameters in the model, including the residual variance. Therefore, K equals p plus 1. The smaller the AIC_c, the more closely a model approaches reality.

The individual AIC_c values are not interpretable since they contain arbitrary constants (Burnham and Anderson, 2002). When comparing different models, the common practice is to rescale these values with the minimum value of these models:

$$\Delta_i = \text{AIC}_i - \text{AIC}_{\min}, \qquad (10)$$

where $\text{AIC}_{\min}$ is the minimum of the different AIC_i values. This transformation forces the best model to have $\Delta = 0$, while the rest of the models have positive values. Some simple rules are often useful in assessing the relative merits of models: Models with $\Delta_i \leq 2$ have substantial support (evidence), those in which $4 \leq \Delta_i \leq 7$ have considerably less support, and models with $\Delta_i > 10$ have essentially no support (Burnham and Anderson, 2002). In the following analysis, AIC_c is preferred for evaluating different models. However, adjusted R^2 and RMSE are also analyzed since they have been widely used and are likely to be in continued use in the future (Anderson-Sprecher, 1994).

It is noteworthy that although the regression models were developed only based on correctly-segmented trees, they were evaluated for correctly-segmented and mis-segmented trees, respectively. Due to the errors of the tree isolation algorithm, only 181 trees of all 313 trees in the field have a one-to-one relationship. The remaining trees could be over-segmented (1-to-m) (Figure 2a) or under-segmented (n-to-1) (Figure 2b). In some cases, although the outer boundaries of a group of trees were correctly delineated, the internal boundaries to divide these crown segments were incorrect, leading to an n-to-m relationship (Figure 2c). To test the model performance when mis-segmentation occurs, a segment (in the case of n-to-1) or a group of segments (in the case of 1-to-m and m-to-n) were linked to the corresponding tree(s) observed on the ground so that the overlaying area between this or these segmented and the

ground located tree(s) was within the range of $S_{r,g} \pm 10\% * S_{r,g}$, where $S_{r,g}$ is the area of the corresponding tree(s) identified on the ground. In doing so, a total of 67 pairs were formed between 110 segments and 132 trees. After the pairs of lidar-derived segments and trees identified on the ground had been formed, the total basal area and stem volume for segments and the corresponding trees on the ground were used to calculate the adjusted R^2, root mean square error, and AIC_c. The test of the models over the mis-segmented trees can help evaluate how the models are sensitive to the errors in tree isolation and crown delineation.

Results

There are eight and nine general models for estimating basal area and stem volume, respectively. Note that for the general models that include height as independent variables, the number of the corresponding specific models should be multiplied by 42, i.e., the number of the height metrics. Therefore, there are a total of 89 and 131 specific models for basal area and stem volume estimation, respectively. For the models depending on any individual height metric, the maximum height of all laser pulses always leads to the best performance in this study, which will be discussed later in more detail. Therefore, only the fitting statistics for the models that depend on the maximum height are listed in Tables 4 and 5.

Among the models of predicting basal area, the power Model B.1, which depends on the maximum height, is the worst. The best model for the correctly segmented trees is the Model B.6, which depends on tree height and crown area. However, for the mis-segmented trees, the best model goes to the power Model B.5, which predict basal area with the $\frac{3}{4}$ power of the canopy geometric volume G.

Among the models of predicting stem volume, neither the height metric H_x nor the crown area C_a is a good predictor when a power model is used. Like basal area, the power Model V.1, which depends on the height metric H_x, is the worst for predicting stem volume. The best model for both correctly and incorrectly segmented trees is stepwise regression Model V.6, which has all metrics as input variables initially and depends on the canopy geometric volume G and a height metric in the final model.

Discussion

Height Metrics

Although the height metrics are the major predictors for most studies for estimating canopy structure parameters from lidar (e.g., Means *et al.*, 2000; Hyyppä *et al.*, 2001; Næsset 2002; Persson *et al.*, 2002; Lim *et al.*, 2003; Popescu *et al.*, 2003; Maltamo *et al.*, 2004b; Hall *et al.*, 2005), this study shows that when a power model is used the height metric H_x is worse than the crown area C_a or the canopy geometric volume G for predicting either basal area or stem volume. When all the height metrics are combined to create stepwise regression models (Models B.9 and V.8), the models improve but are still worse than the power models depending on C_a or G (Models B.2–3 and V.2–3) for predicting basal area. For predicting stem volume, the stepwise regression model with all height metrics as input variables can achieve a slightly better performance than the power model depending on crown area C_a while the performance is still much worse than the power model depending on G.

Despite that the height metrics are not powerful for predicting basal area and stem volume at the individual tree level, they have been most widely used for canopy structure
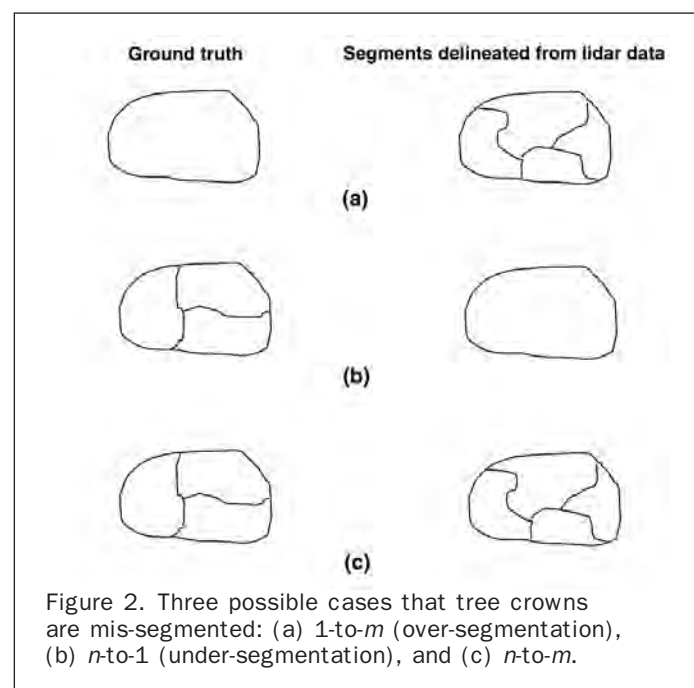
Figure 2. Three possible cases that tree crowns are mis-segmented: (a) 1-to-m (over-segmentation), (b) n-to-1 (under-segmentation), and (c) n-to-m.

TABLE 4. THE FITTING STATISTICS FOR THE MODELS THAT PREDICT BASAL AREA

Model	Correctly Segmented Trees				Mis-Segmented Trees				Equation
	R_a^2	RMSE (10^{-3}m^2)	AIC_c	Δ	R_a^2	RMSE (10^{-3}m^2)	AIC_c	Δ	
B.1	0.51	31.3	−1236.5	146.6	0.48	70.2	−314.4	51.9	$B = 0.0012 H_{100,a}^{1.8}$
B.2	0.70	24.7	−1320.7	62.4	0.68	55.1	−343.5	22.8	$B = 0.0011 C_a^{1.1}$
B.3	0.76	21.8	−1365.7	17.3	0.77	47.0	−362.6	3.8	$B = 0.0015 G^{0.73}$
B.4	0.69	24.7	−1321.4	61.7	0.64	58.3	−337.8	28.5	$B = 0.0003 G$
B.5	0.77	21.3	−1374.7	8.4	0.78	46.0	−366.3	0.0	$B = 0.0014 G^{3/4}$
B.6	0.79	20.7	−1383.1	0.0	0.62	60.2	−331.6	34.7	$\text{DBH} = 2.29 + 1.6 H_{100,a} + 0.26 C_a$; $B = \pi(\text{DBH}/2)^2$
B.7	0.79	20.7	−1380.2	2.9	0.74	50.2	−350.7	15.6	$B = -0.008 + 0.0001 G + 0.005 H_{100,f} - 0.003 H_{80,l} + 0.0006 C_a$
B.8	0.78	20.9	−1374.6	8.4	0.72	51.6	−344.4	21.9	$B = -0.03 + 0.012 H_{mean,a} + 0.005 H_{100,f} + 0.008 H_{0,l} - 0.007 H_{10,l} - 0.004 H_{80,l} + 0.001 C_a$
B.9	0.62	27.4	−1279.8	103.3	0.61	61.1	−327.1	39.2	$B = -0.03 + 0.02 H_{mean,a} - 0.006 H_{mean,l} - 0.03 H_{90,f} + 0.04 H_{100,l}$

Notes:
(1) For models B.1 and B.6, only the fitting statistics of the best model are listed.
(2) The equations whose Δ values are the lowest three are underlined.

TABLE 5. THE FITTING STATISTICS FOR THE MODELS THAT PREDICT STEM VOLUME

Model	Correctly Segmented Trees				Mis-Segmented Trees				Equation
	R_a^2	RMSE (m^3)	AIC_c	Δ	R_a^2	RMSE (m^3)	AIC_c	Δ	
V.1	0.54	327.6	−395.4	146.0	0.58	730.8	−33.2	42.9	$V = 0.0026 H_{100,a}^{2.4}$
V.2	0.64	289.4	−439.8	101.5	0.58	725.8	−34.1	42.0	$V = 0.0027 C_a^{1.4}$
V.3	0.78	227.5	−525.9	15.4	0.77	538.1	−70.0	6.1	$V = 0.0028 G^{1.008}$
V.4	0.79	219.1	−540.5	0.8	0.78	528.4	−73.3	2.8	$V = 0.0029 G$
V.5	0.79	218.8	−540.6	0.7	0.57	738.1	−30.8	45.3	$\text{DBH} = -7.7 + 1.82 H_{100,a} + 0.78 C_a$; $V = \text{allometric} (\text{DBH}, H_{100,a})$
V.6	0.80	217.3	−541.3	0.0	0.80	506.0	−76.1	0.0	$V = -0.06 + 0.003 G + 0.08 H_{std,f}$
V.7	0.76	234.0	−507.8	33.5	0.77	541.5	−58.9	17.2	$V = -0.07 + 0.1 H_{mean,a} - 1.6 H_{cv,l} + 0.08 H_{0,l} - 0.16 H_{10,l} - 0.06 H_{30,l} - 0.05 H_{80,l} + 0.16 H_{100,l} + 0.01 C_a$
V.8	0.65	283.4	−441.7	99.6	0.68	635.3	−43.1	33.0	$V = 0.04 + 0.14 H_{mean,a} - 1.82 H_{cv,a} - 0.12 H_{70,f} - 0.14 H_{10,l} - 0.1 H_{30,l} + 0.29 H_{100,l}$

Notes:
(1) For models V.1 and V.5, only the fitting statistics of the best model are listed.
(2) The equations whose Δ values are the lowest three are underlined.

parameter estimation since they can directly be derived from the lidar measurements. To explore more about these metrics, we compared the performance of the power models that depend on any individual height metric (Tables 6 and 7). Note that the adjusted R^2 values for some models are negative, which highlighted the preference of using AIC_c for model evaluation. The patterns of AIC_c values for different height metrics are similar for basal area and stem volume: among the four height statistics (including maximum, mean, standard deviation, and coefficient of variance), the order of overall performance from the best to the worst was maximum, standard deviation, mean, and coefficient of variance (see Tables 6 and 7). Among all percentile height metrics, the maximum height (100% percentile height) has the lowest AIC_c values and therefore the best performance (Figure 3). For the first canopy returns, there is a trend of improving performance when the height percentile increases from 0 to 100. Such a trend also exists for the last canopy returns except when the percentile is 0, where its AIC_c value is low than those at the immediate neighboring percentiles.

The low AIC_c values (better performance) at the 0th height percentile (the minimum height) of the last canopy returns can be explained by such a fact: for a larger tree it is more possible that the lowest last canopy return is from the bottom of the tree. If examining the trees with basal area greater than 0.15 m, we can find that the minimum height of last returns for a tree is very low and around 1 m (Figure 4b). This pattern leads to a negative relationship. This phenomenon was not observed for the first canopy returns (Figure 4a). In dense forest, it is possible that all first canopy returns are from the upper portion of a tree, and therefore the minimum of their percentile heights is relatively large. It is also possible that the minimum of the percentile heights is small if a pulse hits the lower-portion of a tree, which often happens if a pulse hits the side of a large isolated tree. Therefore, there is no much association between the basal area and the minimum height of the first canopy returns. Overall, the minimum heights of both first and last canopy returns were poor predictors for basal area. At most percentiles, the percentile height of the first canopy returns has better performance than that of last canopy

TABLE 6. THE FITTING STATISTICS FOR THE POWER MODELS THAT PREDICT BASAL AREA FROM HEIGHT METRICS

| | First Returns | | | | | | Last Returns | | | | | | All Returns | | | | | |
| | Correctly Segmented | | | Mis-Segmented | | | Correctly Segmented | | | Mis-Segmented | | | Correctly Segmented | | | Mis-Segmented | | |
Metric	R_a^2	RMSE (10⁻³m²)	AIC_c	R_a^2	RMSE (10⁻³m²)	AIC_c	R_a^2	RMSE (10⁻³m²)	AIC_c	R_a^2	RMSE (10⁻³m²)	AIC_c	R_a^2	RMSE (10⁻³m²)	AIC_c	R_a^2	RMSE (10⁻³m²)	AIC_c
H_0	-0.14	47.7	-1085.1	-0.19	106.5	-264.4	0.07	43.2	-1120.4	-0.07	101.0	-270.8	0.07	43.2	-1120.5	-0.07	100.8	-271.0
H_{10}	0.25	38.8	-1158.9	0.13	90.8	-283.5	-0.12	47.4	-1087.5	-0.23	108.1	-262.6	0.03	44.2	-1112.8	-0.03	99.2	-272.9
H_{20}	0.33	36.6	-1179.9	0.18	88.1	-287.1	-0.14	47.7	-1085.0	-0.19	106.4	-264.5	0.22	39.6	-1152.3	-0.02	98.5	-273.8
H_{30}	0.37	35.5	-1190.9	0.21	86.5	-289.3	-0.09	46.6	-1093.2	-0.16	105.0	-266.0	0.26	38.4	-1163.0	0.10	92.6	-281.2
H_{40}	0.36	35.8	-1187.7	0.22	86.0	-290.0	0.02	44.2	-1112.2	-0.11	102.6	-268.9	0.30	37.4	-1172.7	0.16	89.3	-285.5
H_{50}	0.39	35.1	-1195.3	0.26	83.8	-293.2	0.08	42.9	-1139.6	0.02	96.6	-276.0	0.34	36.2	-1183.5	0.20	87.1	-288.5
H_{60}	0.41	34.5	-1201.6	0.29	82.2	-295.4	0.16	41.0	-1150.5	0.12	91.6	-282.4	0.38	35.3	-1192.6	0.28	82.7	-294.7
H_{70}	0.42	34.2	-1204.5	0.34	79.4	-299.6	0.21	39.7	-1173.0	0.23	85.9	-290.2	0.40	34.5	-1201.1	0.32	80.7	-297.7
H_{80}	0.42	34.0	-1206.0	0.37	77.4	-302.7	0.30	37.3	-1207.7	0.35	78.9	-300.4	0.41	34.2	-1204.1	0.37	77.3	-302.8
H_{90}	0.44	33.6	-1210.4	0.40	75.6	-305.5	0.43	33.9	-1207.7	0.41	74.9	-306.6	0.44	33.6	-1210.5	0.39	75.9	-305.0
H_{100}	0.51	31.3	-1235.8	0.43	73.4	-309.1	0.51	31.3	-1236.5	0.48	70.2	-314.4	0.51	31.3	-1235.8	0.43	73.4	-309.1
H_{mean}	0.39	34.9	-1197.5	0.26	83.9	-293.1	0.10	42.4	-1127.4	-0.22	107.6	-263.1	0.34	36.3	-1183.3	0.16	89.4	-285.4
H_{std}	0.36	35.9	-1187.1	0.40	75.6	-305.5	0.44	33.5	-1212.1	0.46	71.5	-312.2	0.43	33.8	-1208.9	0.45	72.2	-310.9
H_{cv}	-0.10	46.9	-1091.1	-0.16	105.3	-265.8	0.25	38.7	-1160.1	0.23	85.7	-290.4	0.10	42.5	-1126.4	0.05	94.9	-278.2

Note: The maximum adjusted R^2, and the minimum RMSE and AIC_c were underlined for each column.

TABLE 7. THE FITTING STATISTICS FOR THE POWER MODELS THAT PREDICT STEM VOLUME FROM HEIGHT METRICS

| | First Canopy Returns | | | | | | Last Canopy Returns | | | | | | All Canopy Returns | | | | | |
| | Correctly Segmented | | | Mis-Segmented | | | Correctly Segmented | | | Mis-Segmented | | | Correctly Segmented | | | Mis-Segmented | | |
Metric	R_a^2	RMSE (m³)	AIC_c	R_a^2	RMSE (m³)	AIC_c	R_a^2	RMSE (m³)	AIC_c	R_a^2	RMSE (m³)	AIC_c	R_a^2	RMSE (m³)	AIC_c	R_a^2	RMSE (m³)	AIC_c
H_0	-0.15	0.52	-232.8	-0.19	1.23	28.8	0.01	0.48	-259.6	-0.13	1.19	25.6	0.01	0.48	-259.7	-0.13	1.19	25.5
H_{10}	0.27	0.41	-314.3	0.25	0.97	1.2	-0.15	0.51	-233.6	-0.24	1.25	31.3	0.03	0.47	-263.9	0.04	1.10	15.8
H_{20}	0.34	0.39	-331.2	0.31	0.93	-4.0	-0.15	0.52	-232.6	-0.17	1.22	27.9	0.25	0.42	-308.8	0.10	1.06	11.8
H_{30}	0.37	0.38	-341.7	0.37	0.89	-9.1	-0.10	0.51	-240.5	-0.08	1.17	23.3	0.26	0.41	-311.1	0.25	0.97	1.0
H_{40}	0.35	0.39	-336.0	0.37	0.89	-9.2	0.04	0.47	-265.9	-0.03	1.14	20.0	0.28	0.41	-317.5	0.32	0.93	-4.4
H_{50}	0.38	0.38	-343.5	0.41	0.87	-12.8	0.13	0.45	-282.6	0.09	1.07	12.8	0.33	0.39	-328.9	0.35	0.91	-7.4
H_{60}	0.40	0.37	-349.8	0.44	0.84	-16.3	0.20	0.43	-297.0	0.20	1.01	5.4	0.37	0.38	-340.2	0.42	0.86	-14.1
H_{70}	0.41	0.37	-353.2	0.49	0.81	-21.5	0.21	0.43	-300.8	0.32	0.92	-5.1	0.39	0.37	-347.5	0.46	0.82	-18.9
H_{80}	0.41	0.37	-352.6	0.52	0.78	-25.4	0.28	0.41	-317.7	0.46	0.83	-18.3	0.40	0.37	-349.6	0.51	0.78	-25.0
H_{90}	0.42	0.37	-354.1	0.55	0.76	-29.2	0.41	0.37	-352.4	0.53	0.77	-26.5	0.42	0.37	-353.9	0.54	0.76	-28.2
H_{100}	0.54	0.33	-395.4	0.58	0.73	-33.2	0.52	0.33	-390.0	0.60	0.71	-36.1	0.54	0.33	-395.4	0.58	0.73	-33.2
H_{mean}	0.38	0.38	-343.9	0.42	0.86	-14.0	0.06	0.47	-268.6	-0.13	1.20	25.8	0.30	0.40	-321.5	0.33	0.92	-5.2
H_{std}	0.36	0.38	-337.7	0.42	0.86	-14.1	0.44	0.36	-361.3	0.43	0.85	-15.4	0.43	0.36	-358.6	0.45	0.83	-17.7
H_{cv}	-0.12	0.51	-237.4	-0.18	1.22	28.4	0.19	0.43	-295.6	0.17	1.03	7.6	0.05	0.47	-267.7	0.02	1.11	17.3

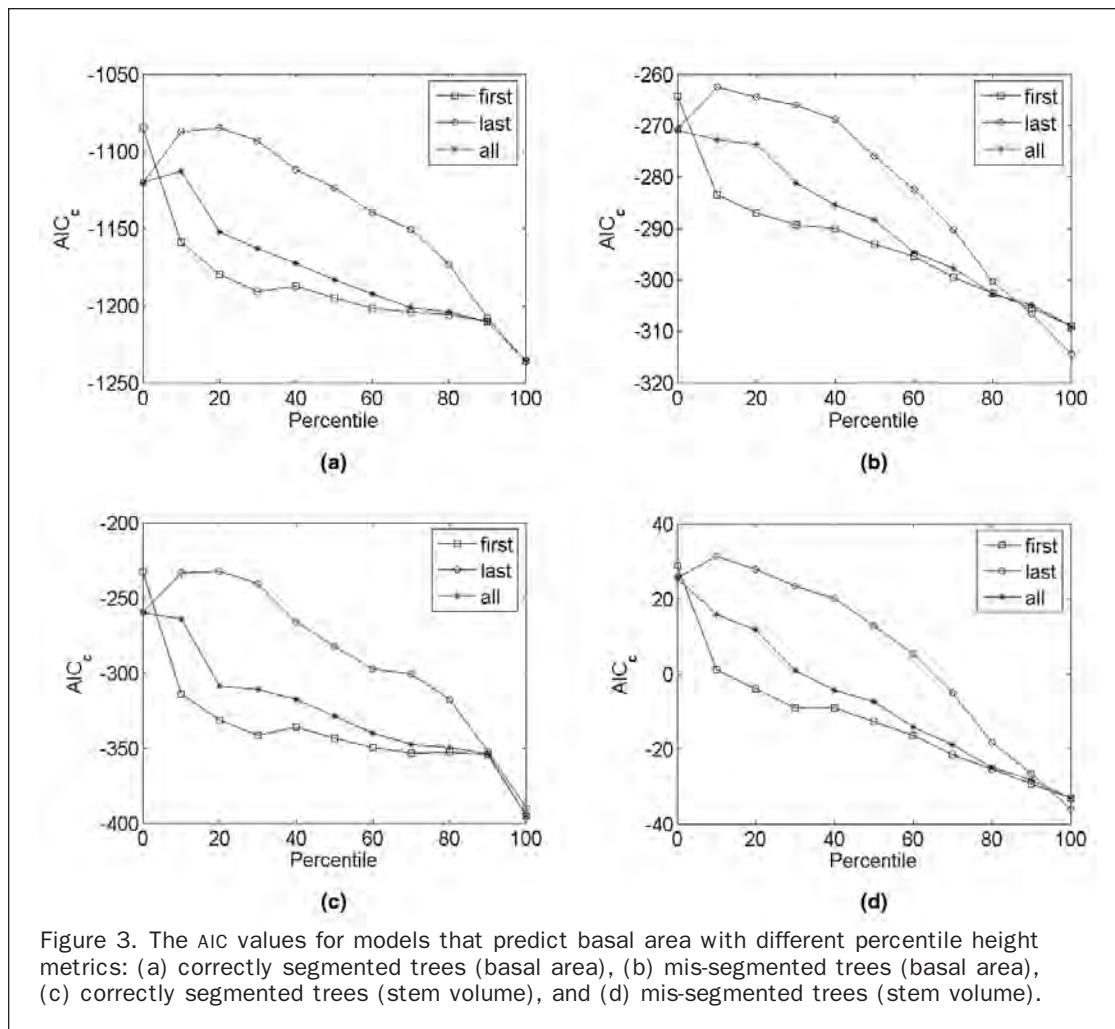Note: The maximum adjusted R^2, and the minimum RMSE and AIC_c were underlined for each column.

Figure 3. The AIC values for models that predict basal area with different percentile height metrics: (a) correctly segmented trees (basal area), (b) mis-segmented trees (basal area), (c) correctly segmented trees (stem volume), and (d) mis-segmented trees (stem volume).
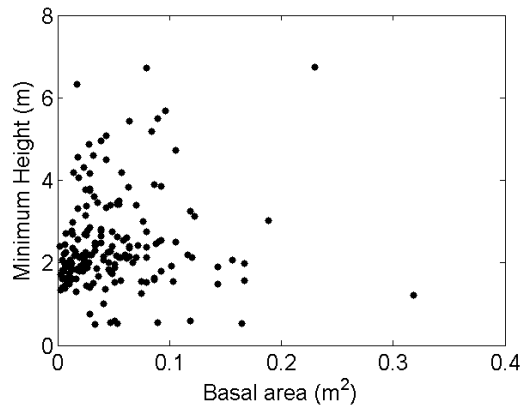
returns. However, when the percentile increases to 100, their performance becomes close or identical. Since the maximum height is the best percentile height metric, it seems that dividing the canopy returns into first and last returns has little effects on improving model fitting.

Crown Area
Overall, the crown area is a better predictor than the height metrics for both basal area and stem volume in the power models. For example, the adjusted R^2 for a power model of the crown area C_d is about 0.2 higher than that for a power model of the maximum height H_{max} for predicting basal area. Although crown area is a powerful metric, only a few studies have used it in the estimation of the canopy structure parameters (Hyyppä et al., 2001; Persson et al., 2002; Popescu et al., 2004; Maltamo et al., 2004b; Roberts et al., 2005), mainly due to the difficulty in extracting this metric.

As in Models B.6 and V.5, many studies used crown area or diameter and tree height to predict the diameter at breast height (DBH), based on which other canopy structure parameters were predicted. Hyyppä et al. (2001) is the first to extract crown diameter from lidar data for forest studies. They used a segmentation-based method to extract crown diameter and applied it to the stand-level stem volume estimation of a boreal forest in Finland. They first fitted a linear regression model that depends on crown diameter and

tree height to predict the diameter at breast height (DBH), based on which stem volume was calculated. When evaluating the accuracy at the stand level, they found that the stand errors for mean height, basal area, and stem volume are 1.8 m, 2.0 m²/ha, and 18.5 m³/ha, respectively. The precision is better than that in conventional standwise forest inventory. Maltamo et al. (2004b) fitted such a linear regression model at the logarithmic scale and found that the RMSE for the estimates of timber volume are 22.5 percent in the same study site. Persson et al. (2002) applied the product of tree height and crown diameter derived from lidar to a simple linear regression model for predicting DBH. With the predicted DBH, the stem volume was estimated for a boreal forest with dominant species of Norway spruce, Scots pine, and birch. They found that 91 percent of the total stem volume was detected. Popescu et al. (2004) tested stepwise regression models that depended on the statistics of individual tree height and crown diameter to predict plot-level biomass. They fou.nd that the crown diameter parameters are significant variables in the stepwise regression models. Roberts et al. (2005) is the first to extract individual tree leaf area from lidar data. At a 16-year-old loblolly pine spacing trial in Mississippi, they found that lidar-derived estimates of leaf area based on height and crown diameter were on average within 0.1 m² of ground-based estimates for trees on plots initially planned at 1.5 m × 1.5 m spacing.

Figure 4. The relationship between basal area and minimum height for first and last canopy returns: (a) first canopy returns, and (b) last canopy returns.

aspects: first, it has a theoretical basis in plant allometry; second, it has only one parameter; and third, the parameter is easy to interpret. In particular, for the stem volume model, the coefficient β_1 can be interpreted as the stem volume density, similar as the leaf area volume density for predicting leaf area. Note that Model B.5 will overestimate the basal area in the case of over-segmentation of tree crowns and underestimate it in the case of under-segmentation; however, the estimation of stem volume with Model V.4 is not affected by over-segmentation or under-segmentation theoretically. The $\frac{3}{4}$ and 1 exponents in Models B.5 and V.4 can also be confirmed to some extent by analyzing the confidence intervals of the exponents in Model B.3 and V.3. The 95 percent confidence interval of the exponent in Model B.3 is (0.67, 0.80), which includes the exponent $\frac{3}{4}$. Likewise, the 95 percent confidence interval of the exponent in Model V.3 is (0.93, 1.09), which includes the exponent 1. As expected, the simple linear model with no intercept for predicting basal area (Model B.4) has much worse performance than the Model B.5 since the analysis based on the plant allometry show that basal area is proportional to $G^{3}\!/_{4}$ instead of G.

Lefsky *et al.* (1999) proposed a canopy volume method (CVM) to describe the three-dimensional canopy structure and applied it to predict the total forest biomass for Douglas-fir/western hemlock forests in western Oregon. Their method is the first to take advantage of the ability of a waveform-recording sensor (SLICER) to directly measure the three-dimensional distribution of canopy structure (Lefsky *et al.*, 2002b). However, the canopy volume derived from their method is different from the one in this study. Lefsky *et al.* (1999) first derived the canopy height profile from the returned waveform within a footprint, and then divided the canopy within that footprint into four parts (closes gap, oligophotic zone, euphotic zone, and open space) and summed up the height of the euphotic and oligophotic zones to get the total volume of "filled" canopy (Figure 5). The problem with their method is that it cannot accurately describe the canopy structure variability in the horizontal plane within the footprint. Figure 5 shows the

When testing the above methods over our study site, we found that Models B.6 and V.5, which are based on the Hyyppä *et al.* (2001) method, achieved the best results. For clarity, only the statistics for the Hyyppä *et al.* (2001) method are shown in Tables 4 and 5. Compared to the power models of either tree height or crown area, the models depending on both tree height and crown area (Models B.6 and V.5) can achieve much better performance for correctly segmented trees. This implies that crown area and tree height are complementary to each other for estimating basal area and stem volume since they characterize the canopy structure information in the horizontal and vertical dimensions, respectively (Hyyppä *et al.*, 2001, Holmgren *et al.*, 2003; Hall *et al.*, 2005).

Canopy Geometric Volume
In the power models (Models B.1–3 and V.1–3), the canopy geometric volume G is a much better predictor than the tree height or crown area for predicting both the basal area and stem volume. The result confirms our hypothesis that canopy geometric volume can obtain better performance since it can characterize canopy structure in a three-dimensional space. Especially, the theoretical models based on the plant allometry (Models B.5 and V.4) have the lowest AIC$_c$ values among all univariate models. These models are attractive in several
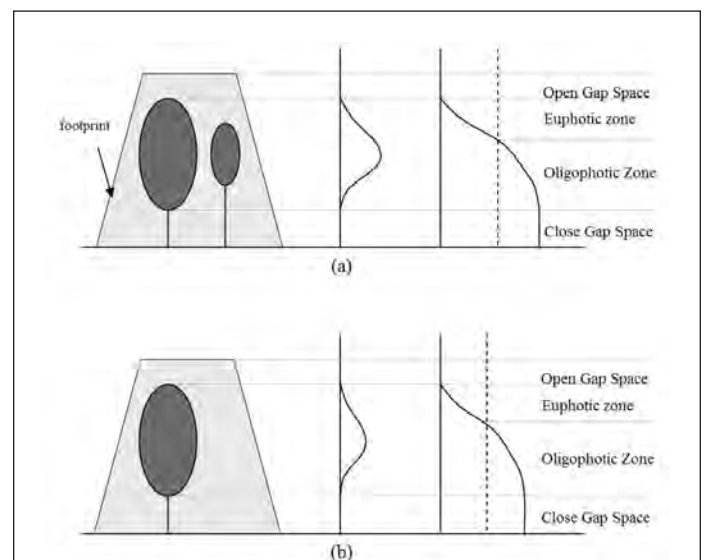

Figure 5. Illustration of the canopy volume method by Lefsky *et al.* (1999) (adapted from Figure 2 in Lefsky *et al.*, 1999).

waveforms and their cumulative power profiles within two footprints; there are two trees within footprint (a) and only one tree within (b). However, the total volume of "fill" canopy, which is the sum of the euphotic zone and oligophotic zone height, is the same for the two waveforms even though the total geometric canopy volumes of trees within these two footprints are different. Therefore, their method can indicate whether there is canopy along the vertical direction (Lefsky *et al.*, 1999) but cannot completely describe the canopy structure information in a three-dimensional space.

Comparison

The research on applying the *individual-tree integration* method in forest studies is still in its infancy. Most studies that extracted the individual tree dimensional information have been using the *hybrid-regression* method (e.g., Hyyppä *et al.*, 2001; Popescu *et al.*, 2003; Maltamo *et al.*, 2004b), and therefore their accuracy assessment is typically performed at the plot level (Riaño *et al.*, 2004). There are only a few studies (Persson *et al.*, 2002; Næsset and Økland, 2002; Riaño *et al.*, 2004; Roberts *et al.*, 2005) that evaluated the accuracy of canopy structure estimation at the individual tree level. Persson *et al.* (2002) used the product of tree height and crown diameter to predict DBH, and then calculate stem volume using allometric equations. They obtained a R^2 value of 0.83 and 0.88 for DBH and stem volume estimation, respectively, for a sample of 135 trees. Næsset and Økland (2002) used a stepwise regression model for predicting tree height and crown length properties, but not basal area and stem volume. Riaño *et al.* (2004) evaluated the accuracy of estimating crown bulk density for *Pinus sylvestris* in a naturally regenerated and planted Scots pine forest in Spain and obtained a coefficient of determination of 0.l4. Roberts *et al.* (2005) predicted the leaf area index with tree height and crown size for individual trees in loblolly pine plantations. For trees on plots originally planted at a square spacing of 2.4 m and 3.0 m, they reported an underestimate of 5.8 m^2 and 14.5 m^2, respectively, and attributed the errors to the inability of generating accurate lidar-based estimates of crown dimensions. The low accuracy reported in Riaño *et al.* (2004) and Roberts *et al.* (2005) has revealed the difficulty of estimating individual-tree canopy structural information with the previous methods. The accuracy in the Persson *et al.* (2002) study was relatively high. However, when their method was applied to our study site, the adjusted R^2 are 0.70 and 0.62 for basal area and stem volume estimation, respectively. Their study site is in a conifer forest, which could be a factor for their better model fit. Our theoretical models based on canopy geometric volume G can obtain a much higher adjusted R^2 of 0.77 and 0.79 for basal area and stem volume estimation, respectively. This proves to some extent the usefulness of the canopy geometric volume G and the associated theoretical models for estimating basal area and stem volume.

Among all models for basal area estimation, overall the theoretical model B.5 should be recommended for use because (a) for the correctly-segmented trees it is among the top three models, and (b) for mis-segmented trees it is the best model and much better than others. For stem volume estimation, the stepwise regression model V.6 is the best. However, the theoretical model V.4 should be recommended for use due to a number of considerations: (a) its close performance to the best models, (b) the model parsimony (Burnham and Anderson, 2002), and (c) the model produced from a stepwise regression is usually site-specific.

Conclusions

This study proposed a new lidar metric called canopy geometric volume G for estimating basal area and stem volume of individual trees. Based on the plant allometry, it was found that basal area is proportional to the $\frac{3}{4}$ power of G and stem volume is directly proportional to G. When tested over trees that were mis-segmented, the theoretical models based on these relationships have the best performance for basal area estimation and the second to best performance for stem volume estimation. Overall, we think these theoretical models have the best performance and should be recommended for basal area and stem volume estimation.

The major limitation of this research is that only one species was considered. If the linear model with no intercept (Model V.4) is used for stem volume estimation, the coefficient β_1 in the model, which is interpreted as the stem volume density, could differ for different species. It is unknown whether such a model has a better performance than other models over sites with multiple species. It is hypothesized that Model V.4 can still be used since the coefficient could be understood as the effective stem volume density for all species in an area. For example, the coefficient could be defined for boreal, temperate, tropic forests, etc. if an approximate estimation of stem volume is required at a large scale. Such a hypothesis needs to be tested in future studies. It is also possible to fuse lidar data with species information derived from optical imagery to estimate canopy structural parameters for each species separately.

References

Anderson-Sprecher, R. 1994. Model comparisons and R^2, *The American Statistician*, 48:113–117.

Baldocchi, D.D., L. Xu, and N. Kiang, 2004. How plant functional-type, weather, seasonal drought, and soil physical properties alter water and energy fluxes of an oak-grass savanna and an annual grassland, *Agricultural and Forest Meteorology*, 123:13–39.

Birdsey, R.A., and G.M. Lewis, 2002. *Carbon in United States Forests and Wood Products, 1987–1997*, Newtown Square, Pennsylvania, USDA Forest Service.

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

Burnham, K.P., and D.R. Anderson, 2002. *Model Selection and Multimodel Inference: A Practical Information-theoretic Approach*, Second edition, New York, Springer-Verlag Press.

Chen, Q., P. Gong, D.D. Baldocchi, and G. Xie, in press. Filtering airborne laser scanning data with morphological methods, *Photogrammetric Engineering & Remote Sensing*.

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

Enquist, B.J., 2002. Universal scaling in tree and vascular plant allometry: Toward a general quantitative theory linking plant form and function from cells to ecosystems, *Tree Physiology*, 22:1045–1064.

Hall, S.A., I.C. Burke, D.O. Box, M.R. Kaufmann, and J.M. Stoker, 2005. Estimating stand structure using discrete-return lidar: An example from low density, fire prone ponderosa pine forests, *Forest Ecology and Management*, 208:189–209.

Holmgren, J., M. Nilsson, and H. Olsson, 2003. Estimation of tree height and stem volume on plots using airborne laser scanning, *Forest Science*, 49:419–428.

Hyyppä, J., O. Kelle, M. Lehikoinen, and M. Inkinen, 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners, *IEEE Transactions on Geoscience and Remote Sensing*, 39:969–975.

Kvålseth, T.O. 1985. Cautionary note about R^2, *American Statistician*, 39:279–285.

Leckie, D., F. Gougeon, D. Hill, R. Quinn, L. Armstrong, and R. Shreenan, 2003. Combined high-density lidar and multi-spectral imagery for individual tree crown analysis, *Canadian Journal of Remote Sensing*, 29:1–17.

Lefsky, M., W. Cohen, A. Acker, G. Parker, T. Spies, and D. Harding, 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., W.B. Cohen, G.G. Parker, and D.J. Harding, 2002a. Lidar remote sensing of above-ground biomass in three biomes, *Global Ecology and Biogeography*, 11:393–399.

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

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

Maltamo, M., K., Mustonen, J. Hyyppä, J. Pitkänen, and X. Yu, 2004a. The accuracy of estimating individual tree variables with airborne laser scanning in a boreal nature reserve, *Canadian Journal of Forest Research*, 34:1791–1801.

Maltamo, M., K., Eerikäinen, J. Pitkänen J. Hyyppä, and M. Vehmas, 2004b. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions, *Remote Sensing of Environment*, 90:319–330.

Means, J., S. Acker, B. Fitt, M. Renslow, L. Emerson, and C. Hendrix, 2000. Predicting forest stand characteristics with airborne scanning lidar, *Photogrammetric Engineering & Remote Sensing*, 66(11):1367–1371.

Medlyn, B.E., 2003. A MAESTRO Retrospective, *Forests at the Land-Atmosphere Interface* (M. Mencuccini, J. Moncrieff, K. McNaughton, and J. Grace, editors), CABI Publishing.

Morsdorf F., E. Meier, B. Kötz, K. I. Itten, M. Dobbertin, and B. Allgöwer, 2004. LIDAR-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management, *Remote Sensing of Environment*, 3(92):353–362.

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., T. Gobakken, J. Holmgren, H. Hyyppä, J. Hyyppä, M. Maltamo, M. Nilsson, H. Olsson, Å. Persson, and U. Söderman, 2004. Laser scanning of forest resources: The Nordic experience, *Scandinavian Journal of Forest Research*, 19:482–499.

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

Norberg, A., 1988. Theory of growth geometry of plants and self-thinning of plant populations: Geometry similarity, elastic similarity, and different growth modes of plants parts, *American Naturalist*, 131:220–256.

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

Pillsbury, N.H., and M.L. Kirkley, 1984. *Equations for Total, Wood, and Saw-Log Volume for Thirteen California Hardwoods, Research Note PNW-414*, Portland, Oregon, U.S. Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station.

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

Popescu, S.C., and R.H. Wynne, 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(5):589–604.

Popescu, S.C., R.H. Wynne, and J.A. Scrivani, 2004. Fusion of small-footprint lidar and multispectral data to estimate plot-level volume and biomass in deciduous and pine forests in Virginia, U.S.A., *Forest Science*, 50:551–565

Riaño, D., E. Chuvieco, S. Condis, J. Gonzalez-Matesanz, and S.L. Ustin, 2004. Generation of crown bulk density for *Pinus sylvestris L.* from lidar, *Remote Sensing of Environment*, 923:345–352.

Roberts, S.D., T.J. Dean, D.L. Evans, J.W. McCombs, R.L. Harrington, P.A. Glass, 2005. Estimating individual tree leaf area in loblolly pine plantations using LiDAR-derived measurements of height and crown dimensions, *Forest Ecology and Management*, 213:54–70.

West, G.B., J.H. Brown, and B.J. Enquist, 1999. A general model for the structure and allometry of plant vascular systems, *Nature*, 400:664–667.