DIAMETER ESTIMATION BASED ON LIDAR POINT CLOUDS AT STAND LEVEL OF LOBLOLLY PINE PLANTATIONS

Clinton Talmage, Yuhui Weng, Yanli Zhang, and Jason Grogan

ABSTRACT

Variables such as average tree diameter at breast height (d.b.h.) and total height are important stand characteristics for forest management. Conventionally, tree d.b.h. and height are measured; however, due to the difficulty of measuring tree height in the field, this variable is often estimated from tree d.b.h. Recent developments in remote sensing techniques, in particular the application of light detection and ranging (LiDAR) technology to forestry, provides a promising alternative. LiDAR point clouds can be used to accurately estimate tree total height and crown area, which can be used to calculate the number of trees per unit area. This study used LiDAR data-derived stand variables including total height, crown size, and the number of trees per acre collected from permanent growth and yield plots established in loblolly pine (*Pinus taeda*) plantations in east Texas to develop alternative models to estimate stand average d.b.h. The models were evaluated based on the mean difference between observed and predicted d.b.h., mean absolute difference, and root mean square error, respectively. Results indicate that stand-level d.b.h. could be accurately estimated using derived stand variables. The estimated stand d.b.h., paired with LiDAR-derived height and number of trees per acre, can then be employed to estimate stand-level volumes.

INTRODUCTION

The Southeastern United States produces more than 50 percent of the country's timber and as much as 25 percent of total world timber supply (Prestemon and Abt 2002). Covering more than 39.5 million acres in the Southeastern United States as of 2010 (Hartsell and Connor 2013), the commonly planted, rapidly growing, and disease resistant native loblolly pine (*Pinus taeda*) is the most important commercial forest species in the United States. Due to the significance of loblolly pine and timber production in the South, effective management of this resource is particularly important. Effective management requires knowledge of current stand volume and other variables, such as average diameter at breast height (d.b.h.), height, and basal area (Avery 1975).

Traditionally, stand volume estimations have relied on direct field-based measurements of average tree dimensions and stocking estimated from sample plots. However, measurements of some variables, such as tree height, can often be difficult, time consuming, or inefficient to obtain, especially in stands with complex canopy structure (Larjavaara and Muller-Landau 2013). To address this, height-diameter models often are utilized in conventional forest inventories to make inferences about stand height.

Recent advancements in remote sensing technology, such as light detection and ranging (LiDAR) have been gaining popularity due to their ability to quickly collect measurements of entire forest stands. For example, measurements of tree location, height, and crown size are more

Clinton Talmage, Graduate Research Assistant, Arthur Temple College of Forestry and Agriculture, Stephen F. Austin State University, Nacogdoches, TX 75962

Yuhui Weng, Associate Professor of Forestry, Arthur Temple College of Forestry and Agriculture, Stephen F. Austin State University, Nacogdoches, TX 75962

Yanli Zhang, Marilyn Cranford Distinguished Professor of Geospatial Sciences, Arthur Temple College of Forestry and Agriculture, Stephen F. Austin State University, Nacogdoches, TX 75962

Jason Grogan, Research Specialist, Arthur Temple College of Forestry and Agriculture, Stephen F. Austin State University, Nacogdoches, TX 75962

effective to gather from aerial LiDAR than from manual measurements (Aubry-Kientz and others 2019). Popescu and others (2003) were successful in estimating individual tree crown diameter derived from LiDAR point clouds. Koch and others (2006) found delineating individual crowns in conifer stands to be more consistent than mixed hardwoods.

While aerial LiDAR is particularly good at measuring tree height (Hyyppä and others 2000), direct measurement of tree d.b.h. using LiDAR is a challenge. However, LiDAR-derived metrics have been used to estimate individual tree or stand-level d.b.h., as strong evidence of allometric relationships between tree height and tree diameter exists for some species (Zhang and others 2020). For instance, models relating d.b.h. to height are well established (Curtis 1967, Hanus and others 1999, Temesgen and others 2006) and previous studies have shown that LiDAR-derived metrics could be used to estimate tree d.b.h. from height (Cao and Dean 2013, Hyyppä and others 2001, Parker and Evans 2004, Salas and others 2010). The objective of this paper was to determine if it was possible to develop a height-to-d.b.h. equation for intensively managed loblolly pine plantations throughout east Texas and western Louisiana using both field-collected data from permanent plots established by the East Texas Pine Plantation Research Project and corresponding LiDAR data.

MATERIALS AND METHODS

Field Data

The East Texas Pine Plantation Research Project (ETPPRP) initiated its Phase II growth and yield study in 2004 by installing 136 permanent plots in intensively managed loblolly pine plantations throughout east Texas and western Louisiana (fig. 1). All plots were treated similarly to the plantations they were within, i.e., if a plantation was thinned or fertilized, then the plot within also was thinned or fertilized. The dimensions for each plot were 100 feet by 100 feet. Within the plot, all trees were tagged and measurements of tree d.b.h., height, height to live crown, and defects were made every 3 years. Individual tree d.b.h. was measured using diameter calipers and heights were measured using a clinometer or Vertex hypsometer. The individual tree data were summarized to plot level. While the plots were repeatedly measured, only data which corresponded temporally with LiDAR data were used. Table 1 represents summary statistics of plot-level mean tree height, d.b.h., and age.

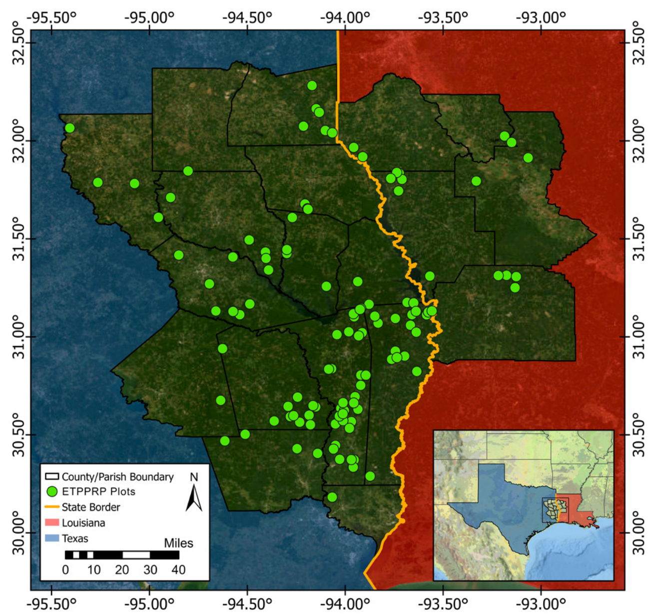

Figure 1—East Texas Pine Plantation Research Project (ETPPRP) plot locations in east Texas and western Louisiana.

Table 1—Summary statistics of field measured and LiDAR-derived stand attributes

Source	Attribute	Mean	Minimum	Maximum	Standard deviation
Field	Height (feet)	52.0	30.0	88.0	10.0
	d.b.h. (inches)	7.9	5.0	14.0	1.8
	Age (years)	16.0	7.0	23.0	3.4
LiDAR	Height (feet)	49.0	26.0	86.0	9.9
	Crown area (square feet)	18.0	1.5	48.0	7.5
	Crown diameter (feet)	4.6	1.2	7.6	1.0

Aerial LiDAR Data

The LiDAR-based 3D Elevation Program (3DEP) is managed by the U.S. Geological Survey to produce high-quality topographic data to update the National Elevation Dataset. LiDAR is an efficient method to collect elevation data within a desired location. When used in conjunction with global navigation satellite systems, LiDAR can produce highly accurate three-dimensional point measurements, also known as point clouds (White and others 2016). Using these point clouds, it is possible to produce accurate representations of a forest canopy. Canopy height models were derived using the difference between tree canopy returns and corresponding ground returns. Digital terrain models of ground returns and digital surface sampled models of first returns were derived from point clouds. LiDAR data was collected during the month of March between 2016 and 2018. ETPPRP field data was collected during the summer between May and August of corresponding years. In most cases, data from LiDAR and field-measured plots were matched year to year. However, in some instances an exact year-to-year match could not be made. In this situation, growth and yield modeling (Coble and others 2017) was utilized to prorate field data to the nearest appropriate year.

The grid size for each elevation model was 2.3 feet or the horizontal accuracy of the LiDAR dataset. This process allows for point clouds to be preprocessed into a product for further analysis. The resulting raster then was further processed by the ForestTools package developed for R (Plowright 2022). This package employs a variable window filter and watershed algorithm to delineate tree canopies, resulting in metrics of individual tree count, height, and canopy size (Plowright 2022). LiDAR parameters for this study include a nominal point density of 0.23 points per square foot, nominal point spacing of 2 feet, a scanning angle of 15 degrees, and vertical accuracy of 0.32 feet. Point cloud samples were collected from the same coordinates as each ETPPRP plot. With plots located throughout a large geographic area (eastern Texas and western Louisiana), it was not feasible to match the thousands of measured trees with their LiDAR data, especially given the low point density of the 3DEP LiDAR data. Instead, a plot-level average was used for both LiDAR-derived and field-measured variables.

Statistical Models

Both linear and nonlinear regression were used to predict plot d.b.h. using LiDAR-derived tree traits. Multiple height-to-d.b.h models were explored in preliminary analyses but only the top five models were selected to use in this study (table 2). SAS software was used in model development (Plowright 2022). Model assumptions (equality of variance, independence, and normality) were checked using residual plots.

Table 2—Candidate equations for estimating d.b.h. from height

Model number	Models	Equations	References
1	Parker and Evans	d.b.h. = a + bH^c * TPA^d	Parker and Evans 2004
2	Cao and Dean	d.b.h. = a + bH^c	Cao and Dean 2013
3	Hyyppä and others	d.b.h. = a + bH + cCD	Hyyppä and others 2001
4	Curtis	d.b.h. = H^2/(a + bH + cH^2)	Curtis 1967
5	Talmage	d.b.h. = a + bH + cTPA * dCA	–

H = tree height; *TPA* = trees per acre; *CD* = crown diameter; *CA* = crown area; a, b, c, d = parameter coefficients.

Model Accuracy Assessment

The data (field data and LiDAR data) were randomly split into two groups with 80 percent used for model training and 20 percent for model testing. Model performance was assessed using the test dataset by calculating coefficients of determination (R^2), root mean square error (RMSE) and relative root mean square error (rRMSE), and mean absolute difference (MAD) as follows:

$$R^2 = 1 - \left(\frac{SSR}{SST}\right) \quad (1)$$

$$\text{RMSE} = \sqrt{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (2)$$

$$\text{rRMSE} = \frac{\sqrt{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}}{(\sum_{i=1}^{N} y_i)/N} \times 100 \quad (3)$$

$$\text{MAD} = \frac{\sum_{i=1}^{N}|\hat{y}_i - y_i|}{N} \quad (4)$$

where

SSR = sum of square residuals

SST = sum of square total

y_i is the field-measured variable

$\hat{y}_i$ is the LiDAR-estimated variable

RESULTS

The correlation between field-measured and LiDAR-derived height was 0.895 before prorating height using the ETPPRP model to account for difference in year. After adjusting field data, the correlation between field-measured and LiDAR-derived height increased to 0.95. Table 3 provides the metrics for the top five assessed models based on testing datasets. Model 1 (Parker and Evans 2004) had an R^2 of 0.67 and a RMSE of 0.80, which was approximately 10.2 percent of the mean (table 3, fig. 2). The other four equations all resulted in an R^2 of 0.28 or less (table 3, fig. 2). Even though model 3 had an R^2 of 0.28, its rRMSE of 15.5 percent was considered acceptable.

Our resulting model, based on model 1 (Parker and Evans 2004), was as follows:

$$\text{d.b.h.} = 3.54 + 0.0603(H^{1.36483648})(TPA^{-0.2087}) \quad (5)$$

where

H = tree height (feet)

TPA = trees per acre

Table 3—The five models with metrics of analyses that were greater than others, processed using the validation dataset

Model number	Author(s)	Coefficient of determination (R^2)	Root mean square error (RMSE)	Relative RMSE	Mean absolute difference
			inches	*percent*	*inches*
1	Parker and Evans	0.67	0.80	10.24	0.75
2	Cao and Dean	0.23	1.20	15.07	1.65
3	Hyyppä and others	0.28	1.30	15.47	1.03
4	Curtis	0.24	1.23	14.96	1.26
5	Talmage	0.26	1.33	16.30	1.40

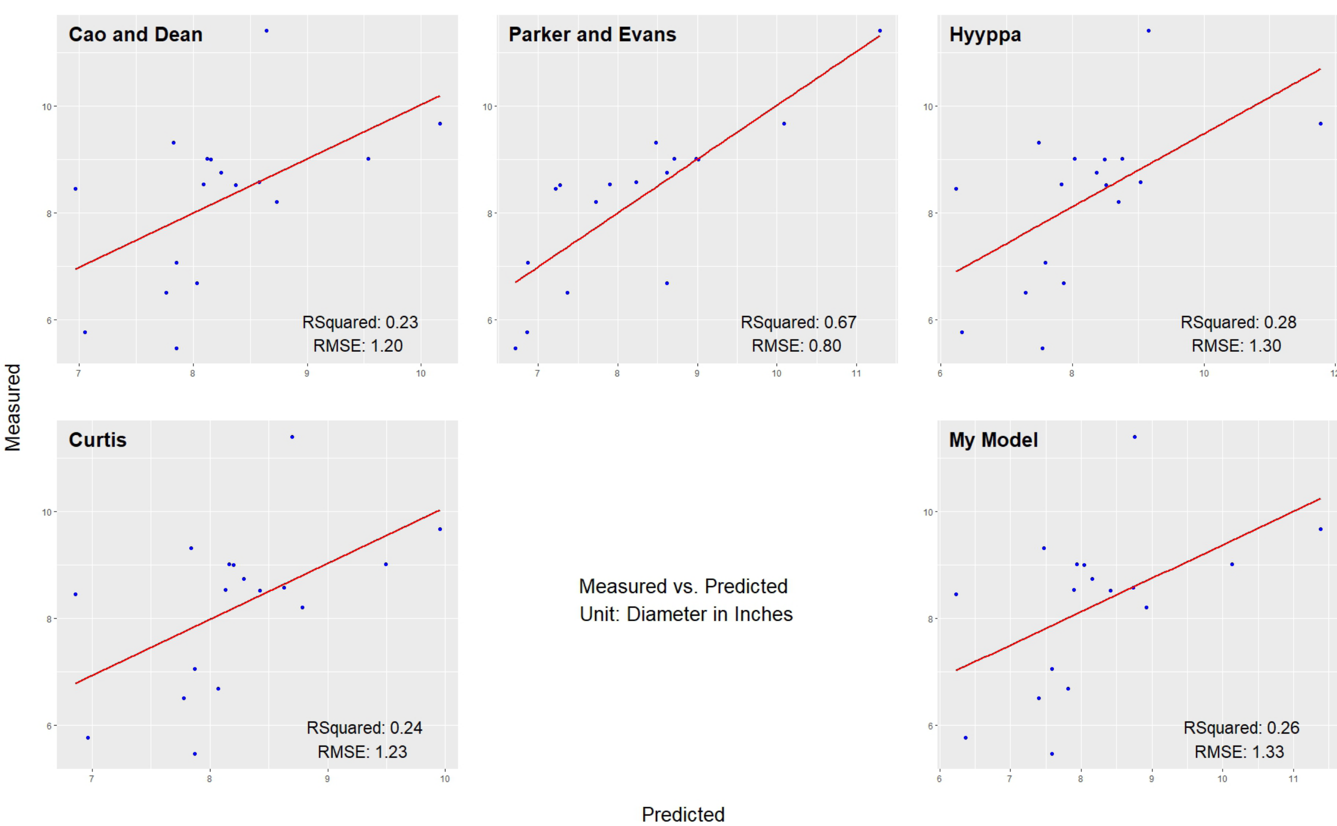

Figure 2—Field-measured versus LiDAR-predicted diameter at breast height of East Texas Pine Plantation Research Project plots.

The standard errors of the parameter coefficients were 1.12 (a = 3.54), 0.11 (b = 0.06), 0.38 (c = 1.36), and 0.13 (d = -0.21). Figure 3 shows the residual plots for the model.

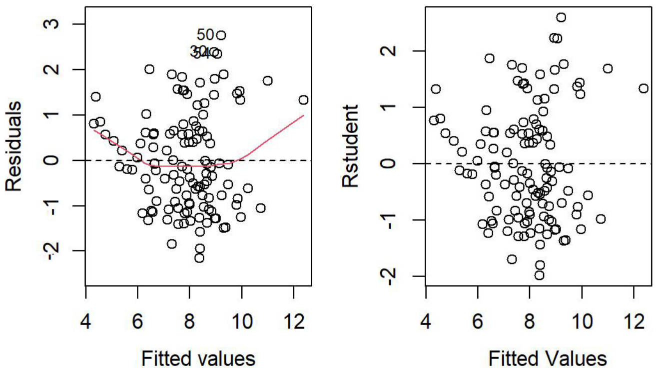

Figure 3—Residual plots of test data for Model 1.

DISCUSSION

We sought to determine if a conventional model could adequately estimate d.b.h. from LiDAR-derived variables. Toward this end, an R^2 of 0.67 and a rRMSE of 10 percent from model 1 is encouraging and seems to indicate that this methodology is viable (table 2). The model requires two predictors, LiDAR-based plot tree height and trees per acre. While LiDAR-based height correlated strongly with field measurement-based height, the association between LiDAR-based trees per acre (TPA) and field measurement-based TPA was weak (fig. 4). Thus, further study is needed to improve the LiDAR-based TPA estimates, which may improve the accuracy of model 1.

Models evaluated in this study did not include variables that could not be obtained by LiDAR data. For example, of the models we evaluated, many did not consider tree density, which plays an important role in predicting diameter. Similar results were experienced in other studies as well (e.g., Cao and Dean 2013). This likely could be attributed to the concept that tree diameter and height are also density dependent. However, because of computerized record-keeping in tools such as geographic information systems, other

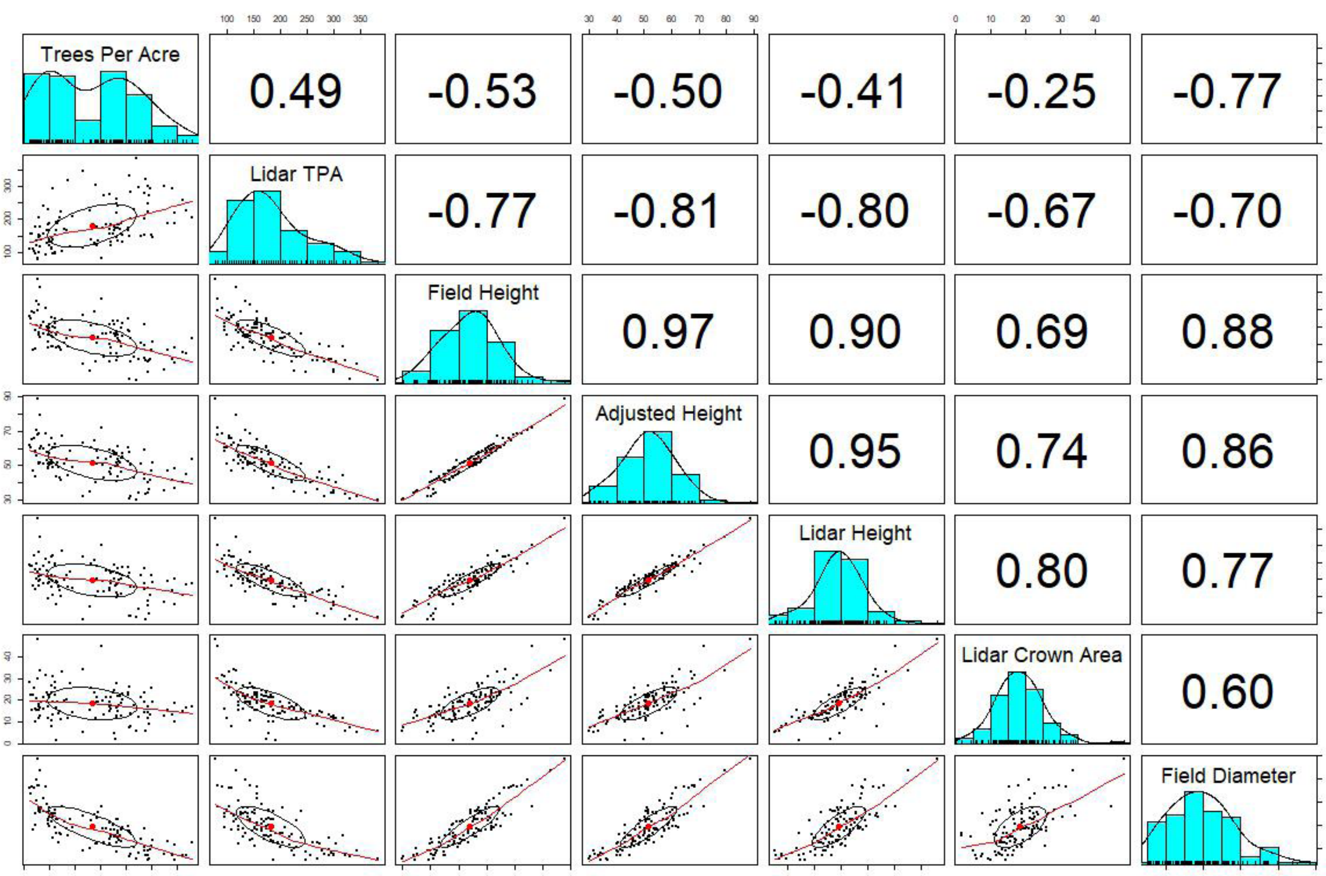

Figure 4—Correlation matrix of all analyzed variables including field-estimated trees per acre, height, height after adjustment, and field-measured diameter at breast height. LiDAR-derived variables included tree height, crown area, and field-measured diameter.

stand attributes are often known. For some, these additional attributes may include stand age, thinning density and timing, site index, and timing of other silvicultural treatments, and such attributes could contribute to further explaining the variation within a height-to-d.b.h. model.

Another factor that potentially could contribute to the accuracy and reliability within the assessed models is the LiDAR dataset itself. The 3DEP is a partnership program which utilizes a collection of elevation data sourced throughout the United States. Data collection for this program relies on assumptions that should imply consistency across the data collections. However, the LiDAR data is collected at different times and by various vendors. Given that this is a cooperative program, 3DEP has developed a baseline program-level set of specifications that are required for data to be considered for inclusion into the program. These specifications include a point spacing of 2.3 feet, an average of 0.19 points per square foot, and a vertical accuracy RMSE no greater than 0.3 feet as of the April 2022 revision (USGS 2022). With only 0.19 points per square foot, it is likely that the LiDAR scan failed to record the absolute top of every tree. Magnussen and others (1999) noted that laser scanners tend to miss treetops and underestimate the standard deviation of derived tree heights. Although the 3DEP LiDAR does not have a high pulse density (only 0.23 pulses per square foot) it did prove to be suitable for the purpose of this study. The relative success of this low-density LiDAR data favors the use of higher-density LiDAR to collect measurements of forest stands.

CONCLUSION

These results demonstrate the viability of estimating stand-level loblolly pine d.b.h. from LiDAR-derived variables and hence may contribute valuable information towards making forest management decisions. Our findings emphasize the value and potential of LiDAR as a tool to improve efficiency of forest inventories and support sustainable forestry practices. As technology continues to advance, utilizing LiDAR to assist in forest mensuration can help foresters meet the demands of the forestry industry.

LITERATURE CITED

Aubry-Kientz, M.; Dutrieux, R.; Ferraz, A. [and others]. 2019. A comparative assessment of the performance of individual tree crowns delineation algorithms from ALS data in tropical forests. Remote Sensing. 11(9): 1086. 21 p. https://doi.org/10.3390/rs11091086.

Avery, T.E. 1975. Natural resources measurements. 2nd ed. New York, NY: McGraw-Hill. 339 p.

Cao, Q.V.; Dean, T.J. 2013. Predicting diameter at breast height from total height and crown length. In: Guldin, J.M., ed. Proceedings of the 15th biennial southern silvicultural research conference. e-Gen. Tech. Rep. SRS-175. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station: 201–205.

Coble, D.W.; McTague, J.P.; Weng, Y. 2017. A whole-stand growth and yield model for intensively and extensively managed loblolly pine plantations in east Texas prior to first thin. Rep. 72. Nacogdoches, TX: Stephen F. Austin State University, Arthur Temple College of Forestry and Agriculture, East Texas Pine Plantation Research Project. 25 p.

Curtis, R.O. 1967. Height-diameter and height-diameter-age equations for second-growth Douglas-fir. Forest Science. 13(4): 365–375.

Hanus, M.L.; Marshall, D.D.; Hann, D.W. 1999. Height-diameter equations for six species in the coastal regions of the Pacific Northwest. Res. Contrib. 25. Corvallis, OR: Forest Research Laboratory, Oregon State University. 11 p.

Hartsell, A.J.; Conner, R.C. 2013. Forest area and conditions: a 2010 update of chapter 16 of the Southern Forest Resource Assessment. e-Gen. Tech. Rep. SRS-174. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 32 p. https://doi.org/10.2737/SRS-GTR-174.

Hyyppä, J.; Kelle, O.; Lehikoinen, M.; Inkinen, M. 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. IEEE Transactions on Geoscience and Remote Sensing. 39(5): 969–975. https://doi.org/10.1109/36.921414.

Hyyppä, J.; Pyysalo, U.; Hyyppä, H.; Samberg, A. 2000. Elevation accuracy of laser scanning-derived digital terrain and target models in forest environment. EARSeL eProceedings 1: 139–147. http://eproceedings.uni-oldenburg.de/website/vol01_1/01_1_hyyppae1.pdf.

Koch, B.; Heyder, U.; Weinacker, H. 2006. Detection of individual tree crowns in airborne lidar data. Photogrammetric Engineering and Remote Sensing. 72(4): 357–363. https://doi.org/10.14358/PERS.72.4.357.

Larjavaara, M.; Muller-Landau, H.C. 2013. Measuring tree height: a quantitative comparison of two common field methods in a moist tropical forest. Methods in Ecology and Evolution. 4(9): 793–801. https://doi.org/10.1111/2041-210X.12071.

Magnussen, S.; Eggermont, P.; LaRiccia, V.N. 1999. Recovering tree heights from airborne laser scanner data. Forest Science. 45(3): 407–422.

Parker, R.C.; Evans, D.L. 2004. An application of LiDAR in a double-sample forest inventory. Western Journal of Applied Forestry. 19(2): 95–101. https://doi.org/10.1093/wjaf/19.2.95.

Plowright, A. 2022. ForestTools: tools for analyzing remote sensing forest data. R package version 1.0.0, https://github.com/andrew-plowright/ForestTools. [Date accessed: August 12, 2022].

Popescu, S.C.; Wynne, R.H.; Nelson, R.F. 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. Canadian Journal of Remote Sensing. 29(5): 564–577. https://doi.org/10.5589/m03-027.

Prestemon, J.P.; Abt, R.C. 2002. The southern timber market to 2040. Journal of Forestry. 100(7): 16–22. https://doi.org/10.1093/jof/100.7.16.

Salas, C.; Ene, L.; Gregoire, T.G. [and others]. 2010. Modelling tree diameter from airborne laser scanning derived variables: a comparison of spatial statistical models. Remote Sensing of Environment. 114(6): 1277–1285. https://doi.org/10.1016/j.rse.2010.01.020.

Temesgen, H.; Hann, D.W.; Monleon, V.J. 2006 Regional height-diameter equations for major tree species of southwest Oregon. Western Journal of Applied Forestry. 22(3): 213–219. https://doi.org/10.1093/wjaf/22.3.213.

U.S. Geological Survey [USGS]. 2022. Lidar base specification: revision history. https://www.usgs.gov/ngp-standards-and-specifications/LiDAR-base-specification-revision-history. [Date accessed: September 15, 2022].

White, J.C.; Coops, N.C.; Wulder, M.A. [and others]. 2016. Remote sensing technologies for enhancing forest inventories: a review. Canadian Journal of Remote Sensing. 42(5): 619–641. https://doi.org/10.1080/07038992.2016.1207484.

Zhang, B.; Sajjad, S.; Chen, K. [and others]. 2020. Predicting tree height-diameter relationship from relative competition levels using quantile regression models for Chinese fir (*Cunninghamia lanceolata*) in Fujian Province, China. Forests. 11(2): 183. 15 p. https://doi.org/10.3390/f11020183.