Supplemental materials for:

Individual Tree Diameter Estimation in Small-Scale Forest Inventory Using UAV Laser Scanning

Yuanshuo Hao, Faris Rafi Almay Widagdo, Xin Liu, Ying Quan, Lihu Dong * and Fengri Li

Key Laboratory of Sustainable Forest Ecosystem Management-Ministry of Education, School of Forestry, Northeast Forestry University, Harbin 150040, Heilongjiang, China; haoyuanshuo@nefu.edu.cn (Y.H.); farisalmay26@nefu.edu.cn (F.R.A.W.); xin_liu@nefu.edu.cn (X.L.); quanying@nefu.edu.cn (Y.Q.)

* Correspondence: lihudong@nefu.edu.cn (L.D.); fengrili@nefu.edu.cn (F.L.); Tel.: +86-451-82191751 (L.D.); +86-451-82190609 (F.L.)

Appendix S1. LiDAR metrics description

Table S1. Summary of the tree- and plot-level metrics derived from unmanned aerial vehicle light detection and ranging (UAVLS) point clouds that were used for the DBH estimation.

Metrics	Definition	Reference
Tree-level		
H[1]	The height of the highest point within individual tree point clouds	–
CA	The area of the 2D convex hull of projected individual tree points on the x-y plane	[1]
CD	Crown diameter calculated by $2 \times \sqrt{CA/\pi}$	[2]
CR_{max}, CR_{medium}	The maximum and medium values of crown radius	–
Cw_V, Cw_S	Crown volume and surface area calculated from the 3D convex hull of crown points	[3]
$H_5^T, H_{10}^T, H_{20}^T, H_{25}^T, H_{30}^T, H_{40}^T, H_{50}^T, H_{60}^T, H_{70}^T, H_{75}^T, H_{80}^T, H_{90}^T, H_{95}^T, H_{99}^T$	The percentiles of the individual tree point height distributions (5th, 10th, …, 95th and 99th).	[4]
H_{mean}^T	The mean height of the individual tree points	[4]
$H_{var}^T, H_{std}^T, H_{cv}^T, H_{ske}^T, H_{kur}^T$	The variance, standard deviations, coefficient of variation, skewness, and kurtosis values of heights	[4]
$D_5^T, D_{10}^T, D_{20}^T, D_{25}^T, D_{30}^T, D_{40}^T, D_{50}^T, D_{60}^T, D_{70}^T, D_{75}^T, D_{80}^T, D_{90}^T, D_{95}^T, D_{99}^T$	The proportion of points above the percentiles ($H_5^T, H_{10}^T, ... H_{99}^T$) to total number of points within a subject tree	[4]
D_{mean}^T	The proportion of points above the mean height to the total number of points	[4]
RH_{max}, RH_{mean}	The ratios of a target tree's height to the maximum and mean tree height	[5]
$RCA_{max}, RCA_{mean}, RCA_{total}$	The ratios of a target tree's crown area to the maximum and mean crown area and total crown area	[5]
$CC_{p25}, CC_{p50}, CC_{p75}, CC_{p100}$	The ratio between crown areas computed at a reference height equal to p% of the height of the subject tree (h_p) and the total crown areas [2]	[5]
Plot-level		
$H_5^P, H_{10}^P, H_{20}^P, H_{25}^P, H_{30}^P, H_{40}^P, H_{50}^P, H_{60}^P, H_{70}^P, H_{75}^P, H_{80}^P, H_{90}^P, H_{95}^P, H_{99}^P$	The percentiles of the points' height distributions (1th, 5th, …, 95th and 99th) of all returns within a plot	[6]
H_{max}^P, H_{mean}^P	The maximum and mean values of points' height within a plot	[6]

$H^P_{var}, H^P_{std}, H^P_{cv}, H^P_{ske}, H^P_{kur}$	The variance, standard deviations, coefficient of variation, skewness, and kurtosis values of heights	[6]
$D^P_5, D^P_{10}, D^P_{20}, D^P_{25}, D^P_{30}, D^P_{40}, D^P_{50}, D^P_{60}, D^P_{70}, D^P_{75}, D^P_{80}, D^P_{90}, D^P_{95}, D^P_{99}$	The proportion of points above the percentiles ($H^P_5, H^P_{10}, \ldots H^P_{99}$) to total number of points within a plot	[7]
D^P_{mean}	The proportion of points above the mean height (H^P_{mean}) to total number of points within a plot	[7]
D^P_{above3}	The proportion of points above 3m to the total number of points representing the canopy fraction in a plot	[7]
Cnp_H, Cnp_C, Cnp_R	The canopy height, canopy cover, and canopy rugosity [3]	[8]
$Aspect$	Aspect	[9]
$Cur_{plan}, Cur_{profile}, Cur_{mean}$	Plan, profile and mean curvature	[9]
$Alt_{mean}, Alt_{mean}, Alt_{std}$	The mean, range, and standard deviations of altitude	[9]
$Slope$	Slope	[9]
$Wetness$	Wetness index	[9]

[1]In our study, LiDAR-derived tree height (H) were applied as a sole predictor for DBH modeling. The relationship between field-measured and LiDAR-derived tree height was shown in Figure S1. The coefficient of determination (R^2) of linear regression was 0.9570 and the slope of regression line is 0.9466, which showed the relatively high goodness-of-fit. The relative root mean square error (RMSE%) between LiDAR-derived and field-measured tree height was was about 7%, which calculated by $RMSE\% = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(H_{field} - H_{LiDAR})^2} / \overline{H_{field}}$ where n, H_{field}, H_{LiDAR}, and $\overline{H_{field}}$ represent the observations, LiDAR-derived, field-measured and the mean value of field-measured tree height, respectively.

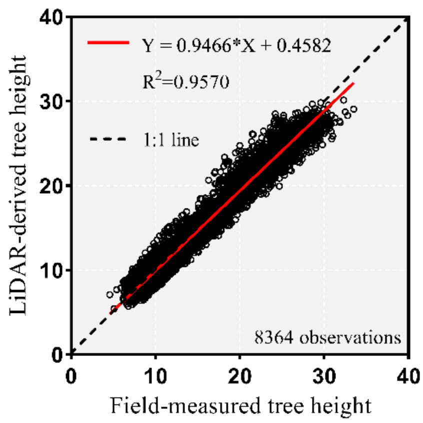

Figure S1. A linear fit between field-measured and LiDAR-derived tree height.

[2]The measures based on crown projection area evaluated at a certain percentage of the height of the subject tree could be calculated as Eq. (S1) and Figure S2.

$$CC_p = \frac{1}{S_a}\sum_{i=1}^{n} CA_{p_i} \tag{S1}$$

where CC_p is crown cover computed at h_p, S_a is plot area, CA_{p_i} is the crown cross-sectional area of one tree at the same height. As shown in Figure S2, if the base of the crown of a competitor is above this height, the full crown area is used instead ($Tree_1$); whereas if the tree height is below the reference height the tree is not considered ($Tree_3$). In practice, we easily calculated the area based on LiDAR-driven CHMs.

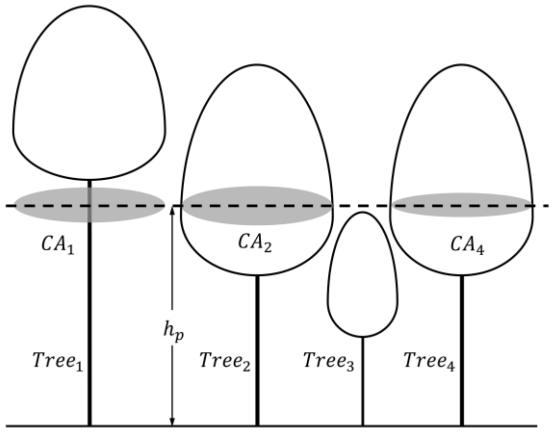

Figure S2. Computation of the competition index based on crown cross-sectional areas calculated at a reference height equal to a certain percentage of the height of the subject tree.

[3]The canopy height, canopy rugosity, and canopy cover were calculated by the mean, standard deviation values, and proportion of pixels above 3m on pit-free CHM in our study, which is a little different from the calculation of Almeida et al. (2019) that used the local maximum points within 2m-intervals to represent the canopy surface attributes.

Appendix S2. Base model selection

In this study, we applied LiDAR-derived tree height (H) as the single predictor for developing the base model. Over the last 60 years, more than 30 H-DBH equation forms have been developed for various species across several types of forest. The inverse functions of these equations can be applied for the LiDAR-derived DBH-H estimations [10]. According to the recommendation from Bi et al. (2012) and the characteristics of our data, six models were selected as candidates and listed in Table S2. Model 1-5 are the inverse functions of commonly used height-DBH models, the Original H-DBH function forms are also presented. A logical constraint was incorporated to ensure a zero DBH when tree height equals breast height (1.3 m). Besides, the parameter of the asymptotic tree height needs to be previously set to ensure the transferability of the models. Therefore, parameter a in both Model (4) and (5) was set as 33.7 m (maximum total tree height of 35 m), referring to the historical data observation (i.e., tree height curve, etc.) in the larch plantation sites. The functions were fitted independently to all observations using weighted nonlinear least square regressions in R software (www.r-project.org). As a result, Model (6) was selected as the base model for DBH estimation due to lower the mean difference (BIAS), root mean square error (RMSE), Akaike information criterion (AIC), and higher coefficient of determination (R_a^2) than others.

Table S2. The basic equation forms considered for the base model selection.

Model	Equation form	Original H-DBH functions	Bias	RMSE	R_a^2	AIC
(1)	$D = \left(a^{-1}(H-1.3)\right)^{1/b}$	$H = 1.3 + aD^b$	0.0032	2.7399	0.7959	40598.39
(2)	$D = a(H-1.3)^{1/2}/(1-b(H-1.3)^{1/2})$	$H = 1.3 + D^2/(a + bD)^2$	-0.0109	2.6468	0.8095	40020.05
(3)	$D = a^{-1}(H-1.3)^{1/b}/(1 - (a^{-1}(H-1.3))^{1/b})$	$H = 1.3 + a(D/(1 + D))^b$	0.0220	2.6498	0.8091	40038.87
(4)	$D = -b^{-1}ln\,(1-(a^{-1}(H-1.3))^{1/c})$	$H = 1.3 + a(1-e^{-bD})^c$	-0.0053	2.6479	0.8094	40026.90
(5)	$D = (-b^{-1}ln\,(1-(a^{-1}(H-1.3))))^{1/c}$	$H = 1.3 + a(1-e^{-bD^c})$	-0.0075	2.6600	0.8076	40103.67
(6)	$D = a(H-1.3)^b e^{c(H-1.3)}$	$-$	-0.0005	2.6400	0.8105	39976.27

Note: D and H in the equations represent DBH (in cm) and tree total height (in m), respectively; a, b, and c are model parameters, $a = 33.7$ in Model (4) and (5).

Appendix S3. Random forests variable selection

Out-of-bag error (mean square error of the out-of-bag samples) in a backward stepwise variable selection of random forests was presented in Figure S3. When 85 predictors were removed, the continued variable reduction will cause significantly increased OOB error. So the remained 15 variables (H_{95}^{T}, H, H_{99}^{T}, RCA_{total}, H_{90}^{T}, Cw_{V}, H_{80}^{T}, CD, H_{75}^{T}, CA, H_{70}^{T}, Cw_{S}, CC_{p75}, H_{75}^{P} and Cnp_{R}) were selected by the variable selection method.

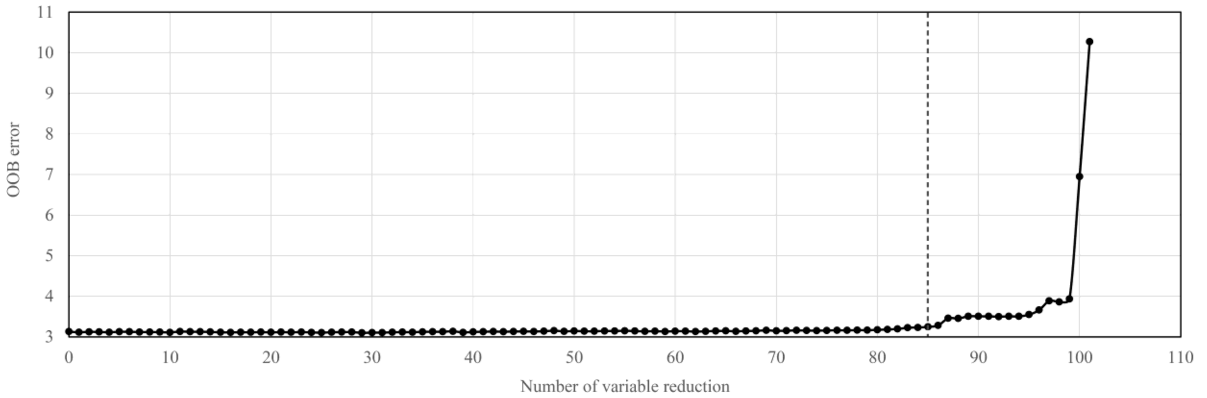

Figure S3. Out-of-bag (OOB) error with variables being removed by a backward stepwise variable selection of random forests. The dash line represents the number of variables equal to 85.

References

1. Dalponte, M.; Coomes, D.A. Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. *Methods Ecol. Evol.* **2016**, *7*, 1236–1245, doi:10.1111/2041-210X.12575.
2. Coomes, D.A.; Dalponte, M.; Jucker, T.; Asner, G.P.; Banin, L.F.; Burslem, D.F.R.P.; Lewis, S.L.; Nilus, R.; Phillips, O.L.; Phua, M.H.; et al. Area-based vs tree-centric approaches to mapping forest carbon in Southeast Asian forests from airborne laser scanning data. *Remote Sens. Environ.* **2017**, *194*, 77–88, doi:10.1016/j.rse.2017.03.017.
3. Tao, S.; Guo, Q.; Li, L.; Xue, B.; Kelly, M.; Li, W.; Xu, G.; Su, Y. Airborne Lidar-derived volume metrics for aboveground biomass estimation: A comparative assessment for conifer stands. *Agric. For. Meteorol.* **2014**, *198–199*, 24–32, doi:10.1016/j.agrformet.2014.07.008.
4. Yu, X.; Hyyppä, J.; Vastaranta, M.; Holopainen, M.; Viitala, R. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 28–37, doi:10.1016/j.isprsjprs.2010.08.003.
5. Biging, G.S.; Dobbertin, M. Evaluation of competition indexes in individual tree growth models. *For. Sci.* **1995**, *41*, 360–377, doi:10.1093/forestscience/41.2.360.
6. McGaughey, R.J. FUSION/LDV: Software for LIDAR Data Analysis and Visualization Available online: http://forsys.sefs.uw.edu/Software/FUSION/FUSION_manual.pdf. (accessed on October 7, 2020)
7. Cao, L.; Liu, H.; Fu, X.; Zhang, Z.; Shen, X.; Ruan, H. Comparison of UAV LiDAR and digital aerial photogrammetry point clouds for estimating forest structural attributes in subtropical planted forests. *Forests* **2019**, *10*, 1–26, doi:10.3390/f10020145.
8. Almeida, D.R.A.; Stark, S.C.; Chazdon, R.; Nelson, B.W.; Cesar, R.G.; Meli, P.; Gorgens, E.B.; Duarte, M.M.; Valbuena, R.; Moreno, V.S.; et al. The effectiveness of lidar remote sensing for monitoring forest cover attributes and landscape restoration. *For. Ecol. Manage.* **2019**, *438*, 34–43, doi:10.1016/j.foreco.2019.02.002.
9. Paris, C.; Bruzzone, L. A Growth-Model-Driven Technique for Tree Stem Diameter Estimation by

Using Airborne LiDAR Data. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 76–92, doi:10.1109/TGRS.2018.2852364.
10. Bi, H.; Fox, J.C.; Li, Y.; Lei, Y.; Pang, Y. Evaluation of nonlinear equations for predicting diameter from tree height. *Can. J. For. Res.* **2012**, *42*, 789–806, doi:10.1139/X2012-019.