

Estimating forest structural attributes using UAV-LiDAR data in Ginkgo plantations

Kun Liu¹, Xin Shen¹, Lin Cao*, Guibin Wang, Fuliang Cao

Co-Innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, 159 Longpan Road, Nanjing, Jiangsu 210037, China



ARTICLE INFO

Keywords:

LiDAR
UAV
Forest structural attributes
Ginkgo
Planted forest
Point cloud density

ABSTRACT

Estimating forest structural attributes in planted forests is crucial for sustainably management of forests and helps to understand the contributions of forests to global carbon storage. The Unmanned Aerial Vehicle-Light Detecting and Ranging (UAV-LiDAR) has become a promising technology and attempts to be used for forest management, due to its capacity to provide highly accurate estimations of three-dimensional (3D) forest structural information with a lower cost, higher flexibility and finer resolution than airborne LiDAR. In this study, the effectiveness of plot-level metrics (i.e., distributional, canopy volume and Weibull-fitted metrics) and individual-tree-summarized metrics (i.e., maximum, minimum and mean height of trees and the number of trees from the individual tree detection (ITD) results) derived from UAV-LiDAR point clouds were assessed, then these metrics were used to fit estimation models of six forest structural attributes by parametric (i.e., partial least squares (PLS)) and non-parametric (i.e., k-Nearest Neighbors (k-NN) and Random Forest (RF)) approaches, within a Ginkgo plantation in east China. In addition, we assessed the effects of UAV-LiDAR point cloud density on the derived metrics and individual tree segmentation results, and evaluated the correlations of these metrics with aboveground biomass (AGB) by a sensitivity analysis. The results showed that, in general, models based on both plot-level and individual-tree-summarized metrics ($CV-R^2 = 0.66\text{--}0.97$, $rRMSE = 2.83\text{--}23.35\%$) performed better than models based on the plot-level metrics only ($CV-R^2 = 0.62\text{--}0.97$, $rRMSE = 3.81\text{--}27.64\%$). PLS had a relatively high prediction accuracy for Lorey's mean height ($CV-R^2 = 0.97$, $rRMSE = 2.83\%$), whereas k-NN performed well for predicting volume ($CV-R^2 = 0.94$, $rRMSE = 8.95\%$) and AGB ($CV-R^2 = 0.95$, $rRMSE = 8.81\%$). For the point cloud density sensitivity analysis, the canopy volume metrics showed a higher dependence on point cloud density than other metrics. ITD results showed a relatively high accuracy (F_1 -score > 74.93%) when the point cloud density was higher than 10% (16 pts m^{-2}). The correlations between AGB and the metrics of height percentiles, lower height level of canopy return densities and canopy cover appeared stable across different point cloud densities when the point cloud density was reduced from 50% (80 pts m^{-2}) to 5% (8 pts m^{-2}).

1. Introduction

Planted forests account for 7.3% (290 million ha) of the world's forest cover (3999 million ha) (FAO, 2015), which plays a key role in meeting demands for wood, fuel material, and other forest products, and in providing ecological and environmental services such as soil and water conservation and carbon sequestration (Carnus et al., 2006). China has the largest area of planted forests all over the world (CSFA, 2014) and accounts for approximately 36% of the total national forest coverage. As one of the dominant planted tree species in China, Ginkgo (*Ginkgo biloba* L.) is widely distributed over 20 provinces, and the middle and northern regions of Jiangsu Province are the major ginkgo

producing region (Cao, 2007; Schmid and Balz, 2005). The structural attributes of planted forests such as diameter at breast height (DBH), Lorey's mean height, stem density, basal area, and volume, that provide detailed spatial information are critical for studies on forest productivity and exchange between mass and energy (Couteron et al., 2010; Huang et al., 2009). Therefore, accurate and reliable estimations of planted forest structural attributes are significantly important for forest managers to make decisions on long-term sustainable forest management (Dash et al., 2015; Ozdemir and Karnieli, 2011).

Compared with traditional remote-sensing technologies, Light Detection and Ranging (LiDAR) can provide detailed characteristics of forest canopy structures in three-dimensions (Harding et al., 2001;

* Corresponding author.

E-mail address: lincao@njfu.edu.cn (L. Cao).

¹ Kun Liu and Xin Shen are co-first authors.

Lefsky et al., 2002; Næsset and Gobakken, 2008). In the past few years, studies have focused on using airborne LiDAR data to estimate forest structural attributes such as DBH (Clark et al., 2004), crown height (Maltamo et al., 2009), stem density (Latifi et al., 2012), basal area (Hudak et al., 2006; Means et al., 2000), volume (Maltamo et al., 2004; Næsset, 2002), and aboveground biomass (AGB) and other biomass components (Cao et al., 2016; García et al., 2010). However, the expenses for acquiring airborne LiDAR data are usually high and limited by mission safety and flight conditions etc., which constrain the application of airborne LiDAR in forestry (Guo et al., 2017; Wallace et al., 2012).

Unmanned Aerial Vehicle (UAV) systems, benefit from low material and operational costs and data measurement flexibility and repeatability (Puliti et al., 2017; Tang and Shao, 2015), making UAV systems an alternative platform to aircrafts and satellites for remote sensing of forest resources (Torresan et al., 2016). In the early development stage, due to the limited payload of UAV, these systems were only equipped with lightweight optical sensors to acquire digital imageries (Laliberte et al., 2011). Point clouds can be generated from these digital photographs using image matching and photogrammetric methods (Zarco-Tejada et al., 2014). Lisein et al. (2013) used a UAV platform to acquire ordinary photographs to produce high spatial resolution three-dimensional measurements of vegetation structure, and dominant height models were generated with a relatively high accuracy ($R^2 = 0.82$, rRMSE = 8.4%) at the plot level. However, the point clouds generated from using image matching are difficult to be used to extract the complete vertical information of forest canopy as well as derive terrain information due to their limited penetration capability (Guo et al., 2017). Recently, owing to the increased UAV payload and the availability of lighter LiDAR sensors, several studies have attempted to integrate LiDAR sensors with UAV platforms for characterizing forest structural attributes. Chisholm et al. (2013) developed a UAV-LiDAR system to estimate the DBH of 12 trees. The results showed that 73% of trees with a DBH greater than 20 cm were detected, and DBH estimations by UAV-LiDAR were positively correlated with field measures ($R^2 = 0.45$, $p = 0.02$). Wang et al. (2017) investigated a UAV-LiDAR system for estimating the aboveground biomass in the Hulunbe grassland ecosystem. They found that the mean canopy height derived from UAV-LiDAR was the most reasonable metric for predicting AGB (rRMSE = 14.1%). Nonetheless, previous studies based on UAV-LiDAR data mostly focused on the estimation of forest structural attributes such as tree height, DBH and crown area, and showed little interest in the estimation of basal area, volume and AGB. Furthermore, due to the limited payload, short flight time and limited capability of data processing of early UAV-LiDAR, most of the previous studies had been conducted on sample trees in small areas.

The distributional metrics extracted from LiDAR data, including height percentile metrics, distribution moment metrics and canopy return density metrics (Næsset, 2004; Véga et al., 2016), have been commonly used to estimate forest structural attributes (Lim et al., 2003a; Means et al., 2000; Zhao et al., 2011). Hall et al. (2005) used discrete-return LiDAR data and derived height-related metrics to predict the height, basal area, stem density and AGB in ponderosa pine forests, and found that these attributes were well predicted ($R^2 = 0.57\text{--}0.87$) by the 75th height percentile and coefficient of variation of height metrics. Montealegre et al. (2016) used airborne LiDAR data derived height related and canopy return density metrics to estimate height, DBH, basal area, stem density, and volume in an Aleppo pine forest. The results showed that the most selected metrics were the 95th height percentile, 50th height percentile and 50th canopy return density. However, there are still some limitations in distributional metric-only models. First, the distributional metrics derived from airborne LiDAR are usually strongly inter-correlated (Chen, 2013). Second, the distributional metrics represent the vertical distribution of forests, whereas canopy horizontal arrangement information is insufficiently considered (Bouvier et al., 2015; Véga et al., 2016). Lefsky

et al. (1999) suggested an approach, i.e., the canopy volume profile (CVP) method, to simulate the vertical and horizontal structure and the volume of forest canopies in voxels and used this approach to estimate DBH, stem density, basal area and LAI with a relatively high accuracy ($R^2 = 0.51\text{--}0.91$). Coops et al. (2007) refined the CVP approach and conducted a Weibull-fitted model to describe foliage profiles. They derived relevant LiDAR metrics and used these metrics to predict height and basal area ($R^2 = 0.65\text{--}0.85$, $p < 0.05$) in a Douglas-fir dominated forest. However, most previous studies used only one type of metrics, and the performance of prediction models based on both distributional and canopy volume metrics derived from UAV-LiDAR data has not been fully assessed.

Manuri et al. (2017) conducted studies to examine the relationships between airborne LiDAR point density and biomass estimations in tropical forests, and the results showed that reducing the point data to a relatively low density (1/49 returns per m^2) provided a relatively stable estimation accuracy (standard error = 13.7%) of AGB by using the distributional metrics. In a mixed conifer forest, Jakubowski et al. (2013) reduced the point density from 9 to $0.01 \text{ pts}\cdot m^{-2}$, and found that the correlations between LiDAR derived metrics and attributes (i.e., height, DBH and basal area) were rarely affected until the point density was below $1 \text{ pts}\cdot m^{-2}$. However, studies assessing the effects of point cloud densities on the estimation of forest structural attributes have rarely focused on planted forests using UAV-LiDAR data with high point density (Hansen et al., 2015; Næsset, 2009; Treitz et al., 2012). In this study, UAV-LiDAR data derived plot-level and individual-tree-summarized metrics were used separately and in combination to estimate forest structural attributes (i.e., DBH, Lorey's mean height, basal area, stem density, volume and AGB) using different approaches (i.e., PLS, k-NN model and RF) over a planted forest in east China. The main objectives are: (1) to select the optimal UAV-LiDAR-derived plot-level and individual tree- summarized metrics and evaluate their capability for estimating forest structural attributes; (2) to assess the capability of different modeling approaches for predicting these attributes; and (3) to evaluate the effects of point cloud densities on individual tree detection results and the UAV-LiDAR data derived metrics (for predicting forest structural attributes).

2. Materials and methods

2.1. Study area

This study was conducted in Pizhou Ginkgo Plantation Base, a planted forest managed by the local government and people in the northern plains of Jiangsu Province ($118^\circ 4'10''\text{E}$, $34^\circ 32'19''\text{N}$) (Fig. 1). The study area has a semi-humid continental climate and an annual temperature of 14.0°C , an annual mean precipitation of 867 mm, and 211 annual mean frost-free days. The study area is 2190 ha, with an elevation ranging from 29 to 32 m above sea level, and is covered with ginkgo plantation forests under different silvicultural treatments.

Field works were conducted in October 2016. Forty-five circular plots (radius = 15 m) were established within 5 of $1 \text{ km} \times 1 \text{ km}$ UAV-LiDAR data acquisition square sites. The coordinates of each plot center were located using a GPS handheld (Trimble, Sunnyvale, CA, USA) receiving real-time differential signals from Jiangsu Continuously Operating Reference Stations (JSCORS) (Song et al., 2009). For trees (with DBH $> 5 \text{ cm}$) in all plots, the DBH were measured using a diameter tape, and the width of the tree crown, tree top height, and height to the crown base were measured using the Vertex V[®] hypsometer (Langsele, Sweden). The coordinates (x, y) of tree locations were measured within 23 of the total 45 plots using an ultrasound-based Haglöf PosTex positioning instrument (Långsele, Sweden). These plots were used for the ITD analysis. In this study, DBH, stem density, basal area, Lorey's mean height (Lim et al., 2003b), volume and AGB were calculated by summing to the individual tree data in plot-level. The individual tree based volume and AGB values were calculated

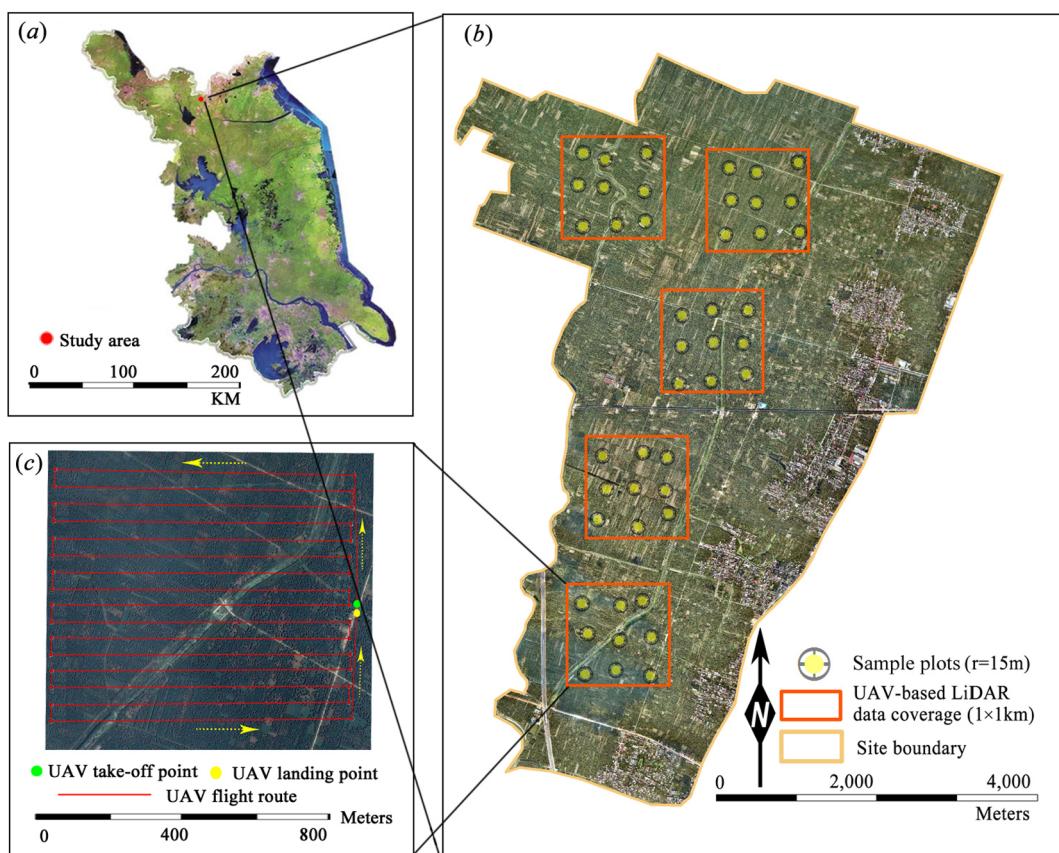


Fig. 1. The location of the study area: (a) the study area in Jiangsu Province; (b) the Pizhou ginkgo plantation study site, the location of sample plots and UAV- LiDAR plots, and (c) the UAV flight route.

Table 1
Summary of the mean and range values of the derived forest structural attributes within the plots.

Variables	Group 1 (<i>n</i> = 15)		Group 2 (<i>n</i> = 15)		Group 3 (<i>n</i> = 15)	
	Range	Mean	Range	Mean	Range	Mean
DBH (cm)	10.15–17.6	13.87	17.28–19.74	18.39	18.93–23.39	21.25
$H_{lorey's}(m)$	6.64–11.05	8.98	11.25–12.26	11.84	12.39–14.6	13.1
Stem Density (N)	594–1329	796	481–594	540	311–481	423
Basal Area (m^2)	13.89–43.26	31.67	45.11–50.83	48.26	50.96–74.82	57.96
Volume ($m^3 ha^{-1}$)	16.48–87.22	47.02	72.53–84.15	79.98	84.24–128.33	99.6
$W_{AGB}(Mg ha^{-1})$	6.73–40.28	26.26	40.84–48.63	45.82	49.7–76.01	53.11

according to the DBH and tree height, based on the volume equations and biomass allometric models developed by Guo et al. (2016) and Liu et al. (2017) respectively. The measured forest structural attributes for the sample plots are summarized in Table 1. Overall, the sample plots were divided into three categories according to the derived UAV-LiDAR metric of canopy cover above 1.3 m.

2.2. UAV platform

The GV1300 multi-rotor UAV (GreenValley International, USA) was used in this study and the UAV platform consisted of two main systems: a multi-rotor UAV and a ground control system (Fig. 2 a-1 and a-2). The technical specifications of the UAV platform are listed in Table 2. This UAV has eight brushless motors, which can provide a better stability and decreased vibration in comparison with other platforms (Wallace et al., 2012). A Novatel IMU (IMU-IGM-S1) and a dual-frequency GPS (Novatel) were used to provide the real-time trajectory information and parameters of flying speed, attitude etc. The ground control system can track the aircraft, control the UAV-LiDAR system, and continuously

monitor the UAV flying parameters. In addition, the system included a Novatel GNSS ground base station. Once the UAV took off manually, it needed to circle the base station to ensure that the GPS accuracy reached a stable state. Then, the UAV could automatically execute a predefined flight mission with an autopilot system to make effective use of flying time and accomplish the surveys. The real-time UAV-LiDAR observations were transferred to the ground data terminal through a long-range Wi-Fi system connected to the UAV.

2.3. UAV-LiDAR data

The raw UAV-LiDAR point clouds were acquired from 13 to 17th, October 2016 using a lightweight Velodyne Puck VLP-16 sensor, with a flight altitude of 60 m above ground level, flight speed of $4.8 m s^{-1}$ and swath width of 50 m. The beam divergence was 3 mrad, and the footprint size was 18 cm in diameter. The LiDAR operated at 21.7 kHz of pulse repetition frequency with 16 scanning channels. The average ground point distance of the dataset was 0.3 m (flying direction) and 0.05 m (scanning direction) for each strip, with an average point

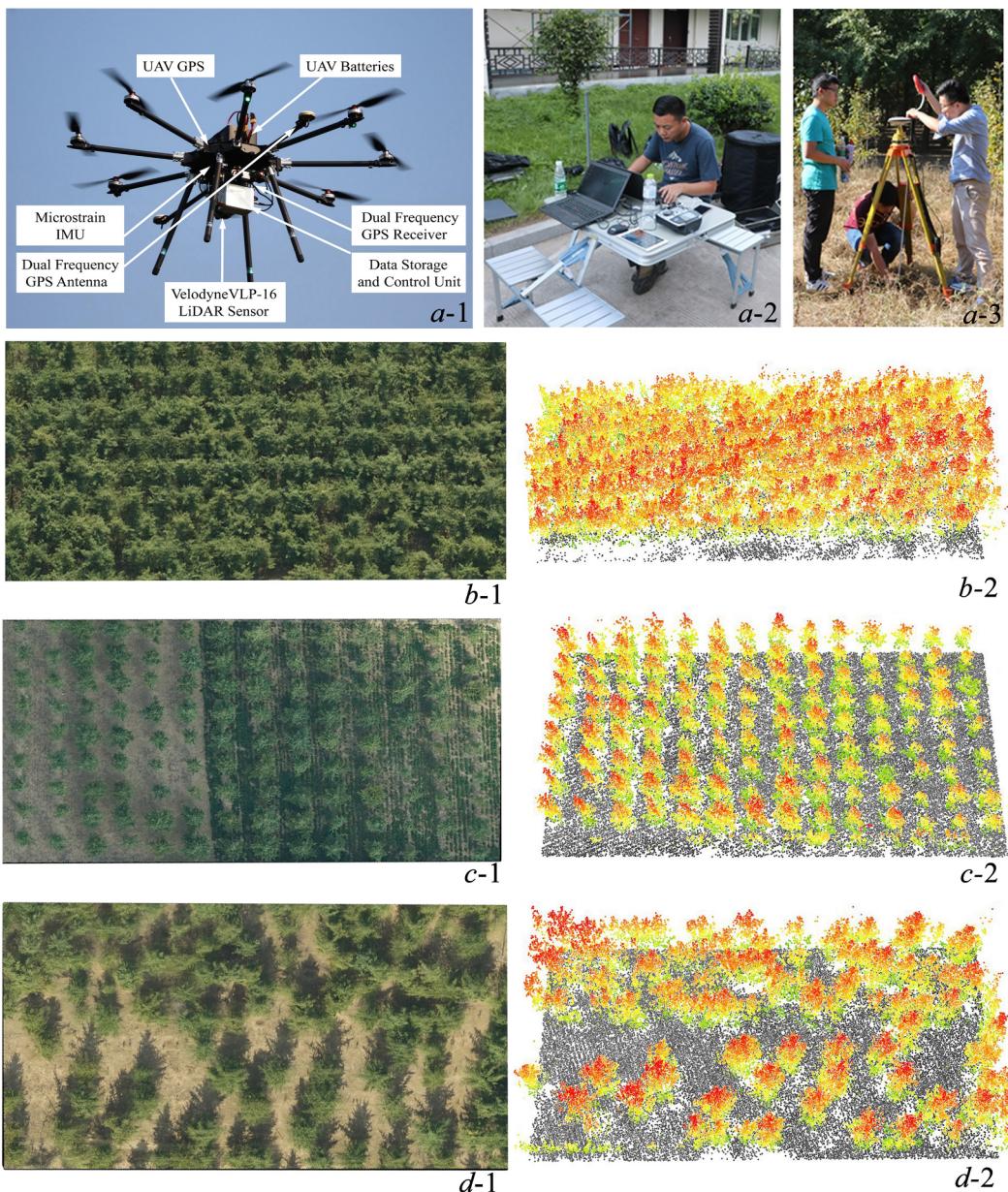


Fig. 2. The main components of the multi-rotor UAV platform and the ortho-photos and LiDAR point clouds acquired using this platform. *a-1*: UAV platform with the Velodyne VLP-16 LiDAR sensor system; *a-2*: UAV-LiDAR ground control system; *a-3*: NovAtel GPS system; *b-1* to *d-2*: the aerial photos (*b-1*, *c-1*, and *d-1*) and UAV-LiDAR point clouds (*b-2*, *c-2* and *d-2*) of three different planted stem densities within the gingko plantation.

Table 2
Technical specifications of the UAV platform used in the study.

Parameters	Value
Weight of loadings (kg)	7
Take-off-weight (kg)	17
Diagonal wheelbase (mm)	1280
Maximum flying time (min)	32
Maximum flying distance (km)	10
Flying height (m)	60 m in this study (0–1000 m)
Flying speed (m/s)	4.8 m/s in this study (0–10 m/s)

density of approximately 160 pts·m⁻².

The main steps of UAV-LiDAR data processing included UAV-LiDAR point cloud coordinate computation, strip adjustment, and point cloud denoising (Fig. 3). To generate a spatially accurate UAV-LiDAR point cloud, first, base station GPS data and UAV platform GPS data were

integrated using post-processing technology. Second, the processed GPS data and IMU data were used to produce higher accuracy flight trajectories by a Sigma Point Kalman Filter (Crassidis, 2006). A nonlinear dynamic state-space model of the system under control was used to predict the state of the IMU/GPS integrated navigation based on the observations from multiple sensors. Then, these positioning data were used for converting the UAV-LiDAR data coordinates. The georeferencing UAV-LiDAR equation was used to determine the variables during the trajectory determination and point cloud generation stages as follows (Guo et al., 2017; Wallace et al., 2012):

$$[x, y, z]^T = P_t + R_0(R_b r^s + L_a) \quad (1)$$

where $[x, y, z]^T$ is the location of a point in the given time, $[x, y, z]$ in the mapping frame represents (North, East, Up), P_t is the UAV trajectory position at time T, R_0 is the orientation matrix, R_b is the bore-sight matrix and L_a is the lever arm, which is used for system calibration. r^s is the laser scanner observation matrix, which can be modified

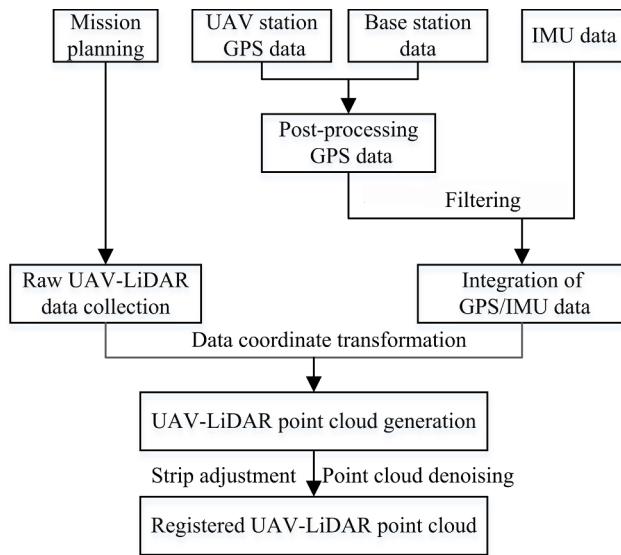


Fig. 3. Overview of the UAV-LiDAR point cloud processing from the UAV system.

as follows:

$$\mathbf{r}^s = \begin{bmatrix} \cos \theta_L & -\sin \theta_L \sin \theta_E & -\sin \theta_L \cos \theta_E \\ 0 & \cos \theta_E & -\sin \theta_E \\ \sin \theta_L & -\cos \theta_L \sin \theta_E & \cos \theta_L \cos \theta_E \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ r \end{bmatrix} \quad (2)$$

where r is the LiDAR-measured range, θ_L is the addition of a layer angle, and θ_E is the encoder angle. To ensure a gapless coverage of the study area, the side-lap of the UAV-LiDAR strips was 70%. To reduce the redundant information within the overlap strips and improve the matching performance, a surface matching method using variants of the iterative closest point (ICP) algorithm was used for strip adjustment (Glira et al., 2015). In this approach, a point selection method of maximum leverage sampling was proposed, and Euclidean distance was used to minimize the sum of the squared distance between the

established correspondences with the following equation:

$$E = \sum_i \Delta S_i^2 = \|T(p_i) - q_i\|^2 \quad (3)$$

where E is the value of the objective function, ΔS_i is the error metric, T is a transformation matrix, p_i and q_i are the corresponding points selected by maximum leverage sampling. The fixed point cloud is formed by q_i , and the loose point cloud is formed by p_i . With this method, the selected correspondences were established on the basis of the point cloud of overlapping strips, so the full resolution of the UAV-LiDAR data could be effectively used to increase the matching accuracy and avoid interpolation of point clouds. In addition, a distance-based deterministic minimum covariance determinant (DetMCD) method was used for point cloud denoising (MiaHubert et al., 2012):

$$D(z_i, \hat{\mu}, \hat{\Sigma}) = \sqrt{\frac{(z_i - \hat{\mu})^T (z_i - \hat{\mu})}{\hat{\Sigma}}} \quad (4)$$

$$d_{ik} = D(z_i, \hat{\mu}_k(Z), \hat{\Sigma}_k(Z)) \quad (5)$$

where D is the statistical distance based on minimum covariance method, z_i is the i -th observation of point, $\hat{\mu}$ is the minimum covariance determinant estimate of location, $\hat{\Sigma}$ is a multiple of the covariance matrix of the scatter estimation, d_{ik} is the statistical distance based on deterministic minimum covariance method ($k = 1, \dots, 6$), and Z represents the standardized data using the $n \times p$ matrix with rows z_i^T ($i = 1, \dots, n$) and columns Z_j ($j = 1, \dots, p$). If $d_{ik} \leq \sqrt{\chi_{p,a}^2}$, the point would be reserved, otherwise, the point would be removed. Where $\chi_{p,a}^2$ is the a -quantile of the χ_p^2 distribution.

Before obtaining forest structural information from the UAV-LiDAR point clouds, some pre-processing was required to create the digital terrain model (DTM) in the study site. First, we classified the point clouds as ground and non-ground points using the algorithm that adapted from Zhao et al. (2016); second, the 1 m DTM was calculated using the ground points by the Kriging interpolation (Guo et al., 2010). Finally, the height of all non-ground points was normalized by the DTM (Fig. 4).

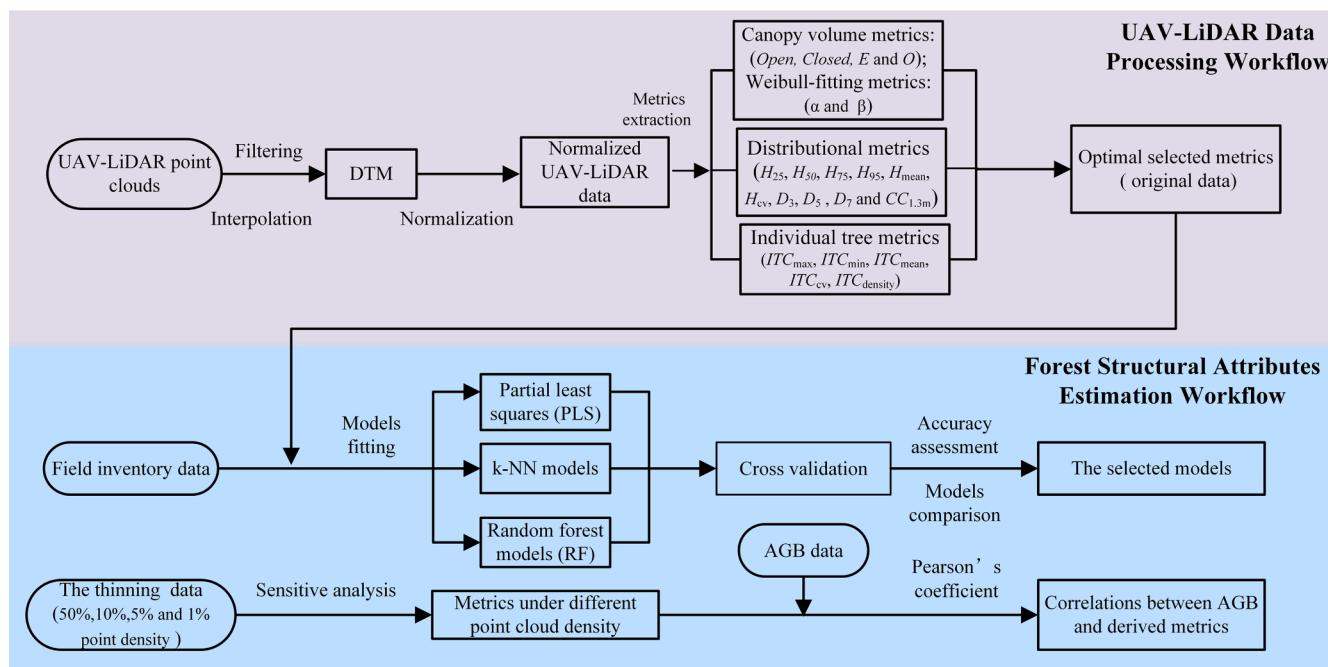


Fig. 4. An overview of the workflow for forest structural attribute estimation using UAV-LiDAR data. The upper panel shows the UAV-LiDAR data processing, and the lower panel shows the steps of forest structural attribute estimations. See Table 3 for the description of the LiDAR-derived metrics.

2.4. Sensitivity analysis of UAV-LiDAR data decimation

For sensitivity analysis, the point cloud data acquired from the UAV-LiDAR system were decimated from the original point density ($160 \text{ pts} \cdot \text{m}^{-2}$) to lower densities of 50% ($80 \text{ pts} \cdot \text{m}^{-2}$), 10% ($16 \text{ pts} \cdot \text{m}^{-2}$), 5% ($8 \text{ pts} \cdot \text{m}^{-2}$) and 1% ($1.6 \text{ pts} \cdot \text{m}^{-2}$) of the original density using the ‘Decimate LAS File(s)’ utility in BACL LiDAR Tools (BACL LiDAR Tools, 2016). Considering the inconsistent point spacing and point density of UAV-LiDAR data over the sample plots, we used a reduction algorithm based on the height of the UAV-LiDAR points. In contrast with other point density reduction approaches such as based on point density or point spacing (Anderson et al., 2006; Magnusson et al., 2007), in this approach, all point clouds were stratified by height for each return, and the reduced point clouds were selected by the percentage of points with the same height (Singh et al., 2015). This approach not only ensured the consistency of sampling across the plots but also maintained a spatial distribution in the decimated data that was similar to that of the original UAV-LiDAR data.

The individual tree detection was based on a point cloud segmentation (PCS) algorithm method (Li et al., 2012). In this method, all the points were classified gradually from the top to the bottom, and the highest point was first identified and then considered as the top of the tree. A specified threshold was used to classify the points, and the threshold value was set similar to the crown radius. Points with spacing smaller than a specified threshold were included into the target tree. Then, a target tree was “grown” by including neighboring points and excluding points of other trees based on relative spacing. We used three statistical parameters, namely, “recall” (r , the detection rate of trees), “precision” (p , the detection accuracy of detected trees) and F_1 -score (F_1 , the overall accuracy including the omission and commission trees), to evaluate the performance of individual tree segmentation. Three measures were calculated as follows (Goutte and Gaussier, 2005):

$$r = \frac{N_t}{N_t + N_o} \quad (6)$$

$$p = \frac{N_t}{N_t + N_c} \quad (7)$$

$$F_1 = 2 \times \frac{r \times p}{r + p} \quad (8)$$

where N_t is the number of detected trees in the plot, N_o is the number of trees omitted by PCS, and N_c is the number of trees that do not exist in the plot but were falsely detected.

2.5. UAV-LiDAR metrics

The first returns were significantly related to the target height, and according to previous studies, have proved to be effective for determining forest structural attributes (Kim et al., 2009; Ritchie, 1993). In our study, the UAV-LiDAR point cloud metrics were calculated from the first return, and three sets of metrics were computed: (1) distributional metrics, i.e., percentile heights (H_{25} , H_{50} , H_{75} and H_{95}), mean height (H_{mean}), coefficient of variation of height (H_{cv}), kurtosis of heights (H_{kurtosis}), Interquartile distance of height, Variance of height (H_{variance}), canopy return density (D_3 , D_5 and D_7) and canopy cover above 1.3 m ($CC_{1.3 \text{ m}}$) (Lim et al., 2003b; Næsset and Gobakken, 2008; Thomas et al., 2006); (2) canopy volume-related metrics, (i.e., Open gap (Open), Closed gap (Closed), Euphotic (E) and Oligophotic (O)) (Lefsky et al., 1999); (3) Weibull model -fitted related metrics, i.e., α (scale parameter) and β (shape parameter) (Coops et al., 2007); (4) The individual tree metrics, (i.e., maximum height (ITC_{max}), minimum height (ITC_{min}), mean height (ITC_{mean}), variance in height (ITC_{cv}) and stem density (ITC_{density})) (Table 3).

The Canopy Volume Model (CVM), based on the three-dimensional geometry of forest canopies, was adopted to calculate the additional UAV-LiDAR data point cloud metrics. The crown space was simulated

as the sum of laser pulses in the $5 \times 5 \text{ m}$ array at 0.5 m height intervals, and according to the presence or absence of return energy in the waveform, each element of the voxels was classified as “filled” or “empty”. The “filled” elements were classified as “euhohotic” (if they occupied the uppermost (65%) of all filled elements) and “oligopfotic” (which consisted of the balance of the filled element). In addition to these two classifications, the elements were classified as “closed gap” (empty volume within the canopy) and “open gap” (the empty space between the top of each of the voxel and the maximum height in the array) (Lefsky et al., 1999).

The foliage profile (FP) was defined as the vertical distribution of the phytoelement density above the ground (Lovell et al., 2003). Due to the flexibility of Weibull distributions in characterizing the foliage distribution, these distributions are the most commonly used to examine the distribution of FPs (Harding et al., 2001; Zhang et al., 2017). In this study, the Weibull cumulative density function was calculated as follows,

$$L(z) = 1 - \left[e^{-\left(\frac{1-\frac{z}{H}}{\alpha}\right)^{\beta}} \right] \quad (9)$$

where α and β are fitted parameters, α determines the basic shape of a vertical distribution, and β represents the increase or decrease in the breadth of the distribution. z is a function of height, and H is the maximum canopy height (i.e., the highest return) (Bailey and Dell, 1973; Coops et al., 2007).

The individual tree metrics were summarized from the descriptive statistics results of individual trees, which were segmented using the PCS algorithm (Li et al., 2012). These metrics were maximum height (ITC_{max}), minimum height (ITC_{min}), mean height (ITC_{mean}), variance in height (ITC_{cv}) and stem density (ITC_{density}) (Ferster et al., 2009).

2.6. Development of forest structural attribute estimations model

In this study, three modeling approaches, i.e., partial least squares (PLS), k-NN model and Random Forest (RF), were used to develop ginkgo plantation forest structural attributes estimation models based on UAV-LiDAR metrics.

PLS regression is a multivariate statistical modeling approach that generalizes and combines multivariate line regression (MLR), principal component analysis (PCA) and canonical correlation analysis (Costa et al., 2015; Wold et al., 2001). Unlike MLR, PLS is suitable for dealing with the variables with multicollinearity and high dimensionality (Dormann et al., 2013; Schmidlein et al., 2012), and it assumes that variance in the dependent variable can be largely characterized by a few number of principle components. The limited number of components produced from large quantities of potentially inter-correlated variables can be used to reduce the potential of model overfitting and improve the accuracy. The PLS has been employed for forest structural attribute estimation in the previous studies (Laurin et al., 2014). In this study, we applied the “pls” package within the R environment to determine the predictive abilities of plot-level and individual tree metrics for estimating forest structural attributes (Wehrens and Mevik, 2007). Before the analysis, all the variables had to be transformed to a fairly symmetrical distribution, and the data were centered and scaled before fitting models. Following the model estimation, the optimal number of components in each model was selected according to the minimizing Root-Mean-Squared-Error (RMSE) of the cross-validation. Then, the predictive models were validated using the ten-fold cross-validated coefficient of determination (R^2) and relative RMSE (rRMSE) (Wold et al., 2001). In addition, the models’ regression coefficients and the variable importance for projection (VIP) were computed and used for further selecting variables. Finally, the final model was developed after removing the predictor variables with small regression coefficients and low VIP values (< 0.8) (Goodbody et al., 2018).

Table 3

Description of metrics derived from UAV-LiDAR data.

LiDAR metrics	Description
Plot-level metrics	
Height-related metrics	Percentile heights (H_5 , H_{25} , H_{50} , H_{75} , H_{95} and H_{99})
	Mean height (H_{mean})
	The coefficient of variation of height (H_{cv})
	Kurtosis of heights (H_{kurtosis})
	Interquartile distance of height (H_{IQ})
	Variance of height (H_{variance})
Density-related metrics	Canopy return density (D_1 , D_3 , D_5 , D_7 and D_9)
	Canopy cover above 1.3 m ($CC_{1.3m}$)
Canopy volume metrics	Open and Closed gap zones of CVM (i.e., Open gap (Open) and Closed gap (Closed))
	Euphotic and Oligophotic zones of CVM (i.e., Euphotic (E) and Oligophotic (O))
Weibull-fitted metrics	α and β parameter of Weibull distribution
Individual tree metrics	
Individual tree segmentation summarized metrics	Maximum height (ITC_{max})
	Minimum height (ITC_{min})
	Mean height (ITC_{mean})
	Coefficient of Variance in height (ITC_{cv})
	Stem density (ITC_{density})
	Maximum height above ground of all UAV-LiDAR first return for individual tree
	Minimum height above ground of all UAV-LiDAR first return for individual tree
	Mean height above ground of all UAV-LiDAR first return for individual tree
	Coefficient of variation of heights above ground of all UAV-LiDAR first return for individual tree
	The numbers of trees in plots measured from individual tree segmentation

The k-NN method used a weighting method for imputation, and the estimation of each location was calculated as weighted average values. As a result, the estimation accuracy was influenced by the selection of the k value and type of distance (Chirici et al., 2008). An optimal value for k was a trade-off between the prediction accuracy and the observation of the variance structure (Chirici et al., 2016; White et al., 2017). In this study, forest structural attributes measured in the field were used as target observations, distributional and canopy metrics derived from UAV-LiDAR data were used as predictors, and a RF distance metric was used for the k-NN approach (Väistaranta et al., 2013). We applied the “yalimpute” package which offered a nearest neighbor distance metric in R for the k-NN method with the RF distance metric (Crookston and Finley, 2008; Liaw and Wiener, 2002). A total of 1000 regression trees were generated by replacement from two-thirds of the data for training and one-third of the data for testing for each tree, considering the randomness, RF was run 50 times, and the final results were the average of these runs. The number of neighbors for imputation was set to five.

For RF, there were two parameters that had to be set: (1) the total number of trees in the model run ($n\text{tree}$) and (2) the number of predictor variables in the random subset at each node ($m\text{try}$). With the advantages of using discrete or continuous datasets, the RF was insensitive to noisy data and hardly overfit (Ismail et al., 2010; Vincenzi et al., 2011); the value of $n\text{tree}$ could be as large as needed and was usually set to several hundred (Hornung, 2010). In the RF approach, rather than choosing the best split from all predictors, $m\text{try}$ was randomly selected in classic regression trees. We applied the “random forest” package within the R environment to model the relationship between the UAV-LiDAR metrics and forest structural attributes of the ginkgo plantation. In this study, the value of $n\text{tree}$ was set to 1000, the error distribution could be shown by the error trend map produced by the program, and the optimal $n\text{tree}$ was selected when the error trend became stable. All the metrics listed in Table 3 were used for $m\text{try}$, and the value of $m\text{try}$ was set to 8. The relative importance of each metric was obtained by quantifying the increase in the mean squared error of the model after the removal of a variable. Finally, leave-one-out cross-validation was used for RF model assessment.

The model performances of PLS, k-NN, and RF were compared based on differences in the adjusted R-square ($Adj-R^2$), RMSE, and

relative root squared error (rRMSE) (Zolkos et al., 2013). The k-NN and RF models’ prediction accuracies were evaluated via leave-one-out cross validation (Bengio and Grandvalet, 2003; Silva et al., 2017a).

$$Adj - R^2 = 1 - \frac{n - 1}{n - p - 1} (1 - R^2) \quad (10)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_i)^2} \quad (11)$$

$$rRMSE = \frac{RMSE}{\bar{x}} \times 100\% \quad (12)$$

where x_i represents observed forest structural attributes for the i plot, \bar{x} is the observed mean value, \hat{x}_i is the estimated forest structural attributes for the i plot, n is the number of plots, and p is the number of variables.

3. Results

3.1. PLS method for forest structural attribute estimations

The forest structural attribute estimation models based on plot-level and combo (plot-level metrics and individual tree metrics) metric had the lowest Root-Mean-Squared-Error (RMSE) with two components (Fig. 5). Overall, all the attributes were generally predicted well ($R^2 = 0.77\text{--}0.97$, rRMSE = 2.83–11.33%) using PLS, except stem density ($R^2 = 0.62$ and 0.67, rRMSE = 23.35% and 26.64%) (Table. 4). The combo models based on plot-level and individual tree metrics generally had a higher performance than the models fitted by the plot-level metrics only for DBH and stem density. For Lorey’s mean height, basal area, volume and aboveground biomass, the separate models and the combo models had relatively close prediction accuracies ($R^2 = 0.90\text{--}0.97$, rRMSE = 2.83–11.33%). For DBH and stem density, the R^2 (rRMSE) of the model based on plot-level metrics were 0.74 (13.75%) and 0.62 (27.64%), respectively, while after adding the individual tree metrics, the $Adj-R^2$ values increased to 0.77 and 0.66, and the rRMSE decreased to 10.16% and 23.35%, respectively. Scatterplots comparing the forest structural attributes of the field-estimated and model-fitted results were calculated for cross-validation (Fig. 6). The

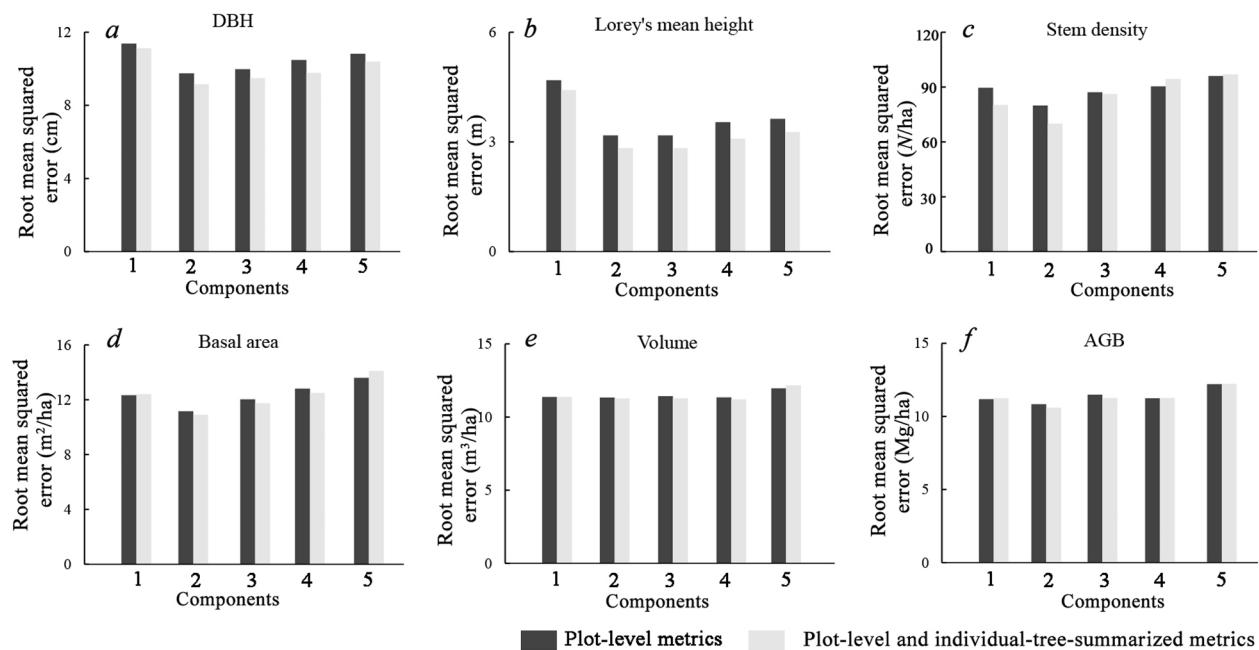


Fig. 5. Changes in the value of Root-Mean-Square-Error (RMSE) across different components (in combination) for six forest structural attributes, i.e., *a*: DBH; *b*: Lorey's mean height; *c*: stem density; *d*: basal area; *e*: volume; *f*: aboveground biomass.

scatterplot relationships for the Lorey's mean height, AGB, volume, basal area, and DBH predictions were close to the 1:1 line, while the stem density relationship slightly deviated from the 1:1 line, as shown in Fig. 6c. For the stem density estimation, the higher stem density had a slightly higher deviation.

A discriminant analysis using the first two components of the separated models and combo models showed that, in general, two PCs accounted for 84.84% and 88.71% of the total variance contained in the selected sets of 19 (Fig. 7a) and 22 (Fig. 7b) UAV-LiDAR metrics, respectively. PC1 and PC2 accounted for 70.15%, 13.69% and 74.37%, 14.34% of the total variance, respectively. The metrics selected by the two models were predominantly the same, with only slight differences. For the plot-level metrics, most height-related, density-related and canopy metrics were included in the combo models. The combo models also included several individual tree metrics (ITC_{max} , ITC_{mean} , and $ITC_{density}$). The plot-level metric of H_{mean} and the individual tree metric of ITC_{mean} showed similar performances and indicated relatively high multiple squared correlations with forest structural attributes.

3.2. k-NN Method for forest structural attribute estimations

The k-NN method for the prediction of plot-level forest attributes based on plot-level metrics and combined plot-level and individual tree are summarized in Table 5. In general, the k-NN method for predicting forest structural attributes based solely on plot-level metrics had a similar performance as the combo metric model for Lorey's mean height ($CV-R^2 = 0.95$, $rRMSE = 5.05\%$; $CV-R^2 = 0.96$, $rRMSE = 4.96\%$), aboveground biomass ($CV-R^2 = 0.94$, $rRMSE = 9.01\%$; $CV-R^2 = 0.95$, $rRMSE = 8.81\%$) and volume ($CV-R^2 = 0.94$, $rRMSE = 8.95\%$; $CV-R^2 = 0.94$, $rRMSE = 9.07\%$). When the individual tree metrics were added to the k-NN models, the prediction accuracy was improved for DBH ($CV-R^2 = 0.85$, $rRMSE = 10.13\%$; $CV-R^2 = 0.89$, $rRMSE = 9.77\%$), stem density ($CV-R^2 = 0.69$, $rRMSE = 18.57\%$; $CV-R^2 = 0.75$, $rRMSE = 17.91\%$) and basal area ($CV-R^2 = 0.86$, $rRMSE = 10.02\%$; $CV-R^2 = 0.87$, $rRMSE = 9.95\%$). The better fitting results of the two models is shown in Fig. 8. The DBH (Fig. 8a), Lorey's mean height (Fig. 8b), basal area (Fig. 8d), volume (Fig. 8e) and AGB (Fig. 8f) models exhibited a good performance, with $CV-R^2 > 0.85$ and $rRMSE < 10\%$. The stem density model performed slightly poorer ($CV-R^2 = 0.75$, $rRMSE = 17.91\%$; $CV-R^2 = 0.69$, $rRMSE = 18.57\%$) than

Table 4

Summary of the best iteration of the PLS model with the cumulative explained variability (%), and cross-validation coefficient of determination ($CV-R^2$) and relative root mean squared error ($rRMSE$) for DBH, Lorey's mean height (H_L), stem density (N), basal area (G), volume (V) and aboveground biomass (W_{AGB}) estimation.

Variable	Models	Total components	Cumulative explained variability (%)	$CV-R^2$	$rRMSE$ (%)
DBH	Plot-level metrics	2	80.47	0.74	13.75
	Plot-level and individual tree metrics	2	81.21	0.77	10.16
H_L	Plot-level metrics	2	82.29	0.97	3.18
	Plot-level and individual tree metrics	2	80.29	0.97	2.83
N	Plot-level metrics	2	82.27	0.62	27.64
	Plot-level and individual tree metrics	2	80.25	0.66	23.35
G	Plot-level metrics	2	82.14	0.84	11.16
	Plot-level and individual tree metrics	2	79.98	0.84	10.90
V	Plot-level metrics	2	78.50	0.90	11.33
	Plot-level and individual tree metrics	2	82.16	0.90	11.26
W_{AGB}	Plot-level metrics	2	79.13	0.91	10.83
	Plot-level and individual tree metrics	2	80.60	0.92	10.59

$CV-R^2$: R-square of cross-validation; $rRMSE$: relative Root-Mean-Square-Error.

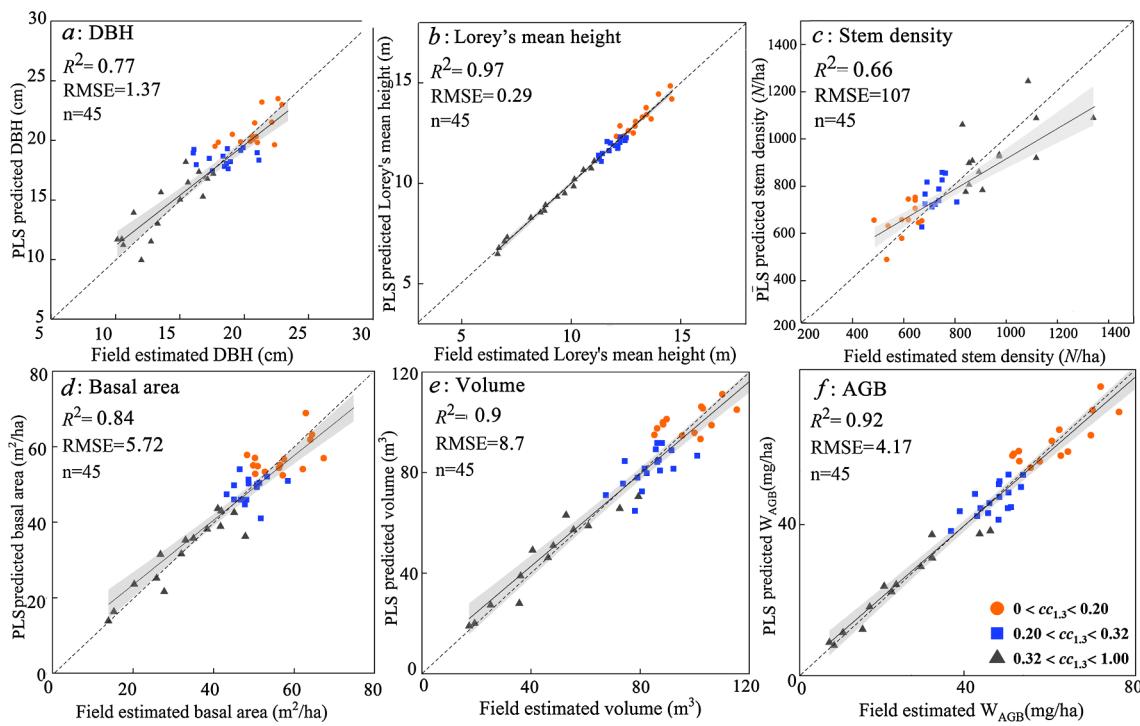


Fig. 6. Scatterplots for the PLS regression model of forest structural attributes i.e., *a*: DBH; *b*: Lorey's mean height; *c*: stem density; *d*: basal area; *e*: volume; *f*: aboveground biomass. Note: CC_{1,3} represents the canopy cover above 1.3 m.

the other forest attributes but still had a better accuracy than the PLS.

3.3. RF method for forest structural attributes estimation

The importance of the variables determined for the RF model of each of the forest structural attributes is shown in Fig. 9. The results of the RF approach also indicated that the three most important predictor variables for modeling DBH, Lorey's mean height, and AGB were height related metrics (H_{25} , H_{95} and H_{mean}), each of which accounted for over 10% of the importance of model estimations. For basal area and volume estimation, the most important variable was the canopy metric *open*, which contributed to more than 12% of the importance for estimation.

For the stem density model, the importance of all variables were similar and less than 5%, except for $ITC_{density}$.

The RF method for the prediction of forest structural attributes using plot-level metrics and combined plot-level and individual tree metrics are summarized in Table 6. The results of the cross-validation of combo metrics are shown in Fig. 10. Overall, the fitted models based on the RF method performed well, with all CV-R² > 0.7 and rRMSE < 22%. The combo models based on plot-level and individual tree metrics had a higher prediction accuracy than the models based on plot-level metrics for DBH (CV-R² = 0.89, rRMSE = 8.89%; CV-R² = 0.87, rRMSE = 9.46%) and stem density (CV-R² = 0.77, rRMSE = 15.54%; CV-R² = 0.71, rRMSE = 15.54%). The Lorey's mean height model

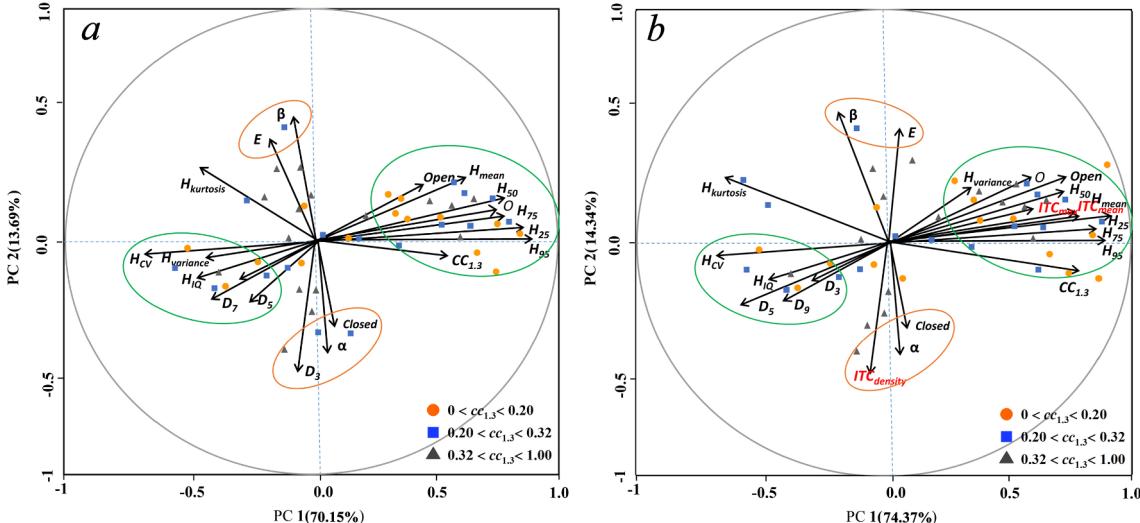


Fig. 7. Projection of the first two PC scores from the selected UAV-LiDAR metrics for the model based on plot-level metrics (a) and plot-level and individual tree metrics (b). The metrics included in the model with associated vectors indicating the direction and magnitude of the coefficient weights on the final model (the black vectors are plot-level metrics, and the red vectors are individual tree metrics). The field-measured samples are classified into three color-coded classes. See Table 3 for a description of the LiDAR-derived metrics. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

Table 5

Summary of the cross-validation results of k-NN for CV-R², relative RMSE (rRMSE) and differences among cross-validation models for DBH, Lorey's mean height (H_L), stem density (N), basal area (G), volume (V) and aboveground biomass (W_{AGB}) estimation.

Variable	Models	CV-R ²	rRMSE (%)	Differences in cross-validation		
				Mean	Std-dev	Range
DBH	Plot-level metrics	0.85	10.13	-0.61	1.74	-5.53 to 1.96
	Plot-level and individual tree metrics	0.89	9.77	-0.53	1.59	-4.98 to 1.64
H_L	Plot-level metrics	0.95	5.05	0.07	0.51	-1.49 to 0.44
	Plot-level and individual tree metrics	0.96	4.96	0.06	0.55	-1.42 to 0.38
N	Plot-level metrics	0.69	18.57	-40.29	106.94	-326 to 220
	Plot-level and individual tree metrics	0.75	17.91	-38.15	98.84	-301 to 195
G	Plot-level metrics	0.86	10.02	-0.05	3.68	-9.2 to 7.55
	Plot-level and individual tree metrics	0.87	9.95	-0.04	3.88	-8.8 to 7.22
V	Plot-level metrics	0.94	8.95	0.41	5.62	-10.8 to 9.47
	Plot-level and individual tree metrics	0.94	9.07	0.45	5.52	-11.3 to 9.33
AGB	Plot-level metrics	0.94	9.01	0.15	2.94	-4.27 to 5.14
	Plot-level and individual tree metrics	0.95	8.81	0.19	2.89	-4.12 to 5.06

CV-R²: R-square of cross-validation; rRMSE: relative Root-Mean-Square-Error; Std-dev: Standard deviation.

showed the highest prediction accuracy ($R^2 = 0.95$, rRMSE = 3.69%; CV-R² = 0.95, rRMSE = 3.74%), followed by the AGB (CV-R² = 0.94, rRMSE = 8.36%; CV-R² = 0.94, rRMSE = 8.72%), volume (CV-R² = 0.93, rRMSE = 9.94%; CV-R² = 0.93, rRMSE = 9.94%), basal area (CV-R² = 0.89, rRMSE = 9.71%; CV-R² = 0.89, rRMSE = 9.84%), and DBH (CV-R² = 0.89, rRMSE = 8.89%; CV-R² = 0.87, rRMSE = 9.46%) models, and stem density had the lowest fitting accuracy (CV-R² = 0.77, rRMSE = 15.54%; CV-R² = 0.71, rRMSE = 16.16%).

Stratification has been regarded as an effective approach to improve forest structural attributes estimation accuracy (Gobakken et al., 2013; William et al., 2015). To evaluate the performance of stratification on models, we assessed the accuracies by fitting models in different groups according to canopy cover for estimating forest structural attributes

(Table 7). The results indicated that overall, the stratified models showed a slight gain in prediction accuracy compared to the use of non-stratified model in forest structural attributes. Specifically, for the stratified models, the Group 2 model had the highest prediction accuracy among three stratified models in DBH (CV-R² = 0.92, rRMSE = 7.94%), Lorey's mean height (CV-R² = 0.97, rRMSE = 2.68%), stem density (CV-R² = 0.87, rRMSE = 12.86%), basal area (CV-R² = 0.94, rRMSE = 8.34%), volume (CV-R² = 0.95, rRMSE = 8.94%) and AGB (CV-R² = 0.96, rRMSE = 7.62%) estimations. For different forest structural attributes, the stratified model of stem density had highest gain in predication accuracy ($\Delta rRMSE = 3.68\%$) compared to non-stratified model.

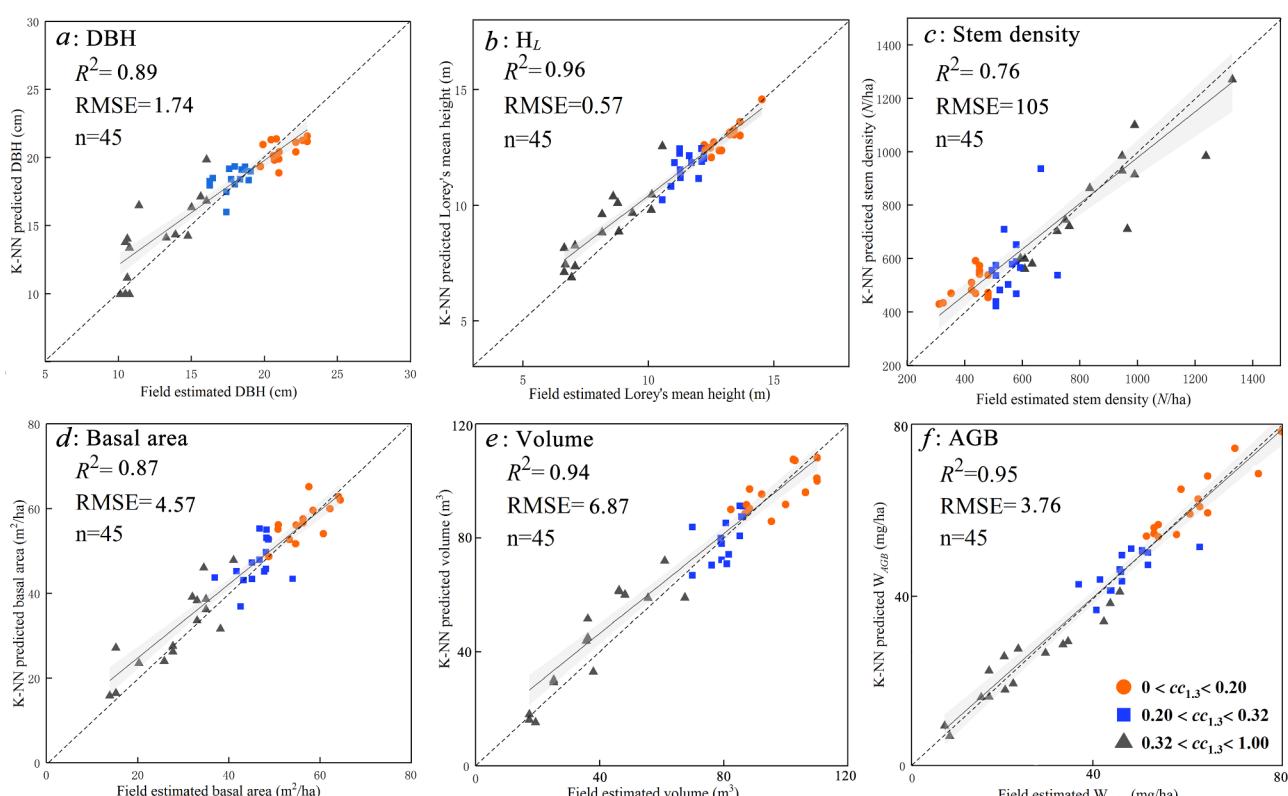


Fig. 8. Scatterplots of the k-NN model of forest structural attributes, i.e., a: DBH; b: Lorey's mean height; c: stem density; d: basal area; e: volume; f: aboveground biomass. CC1.3 represents the canopy cover above 1.3 m.

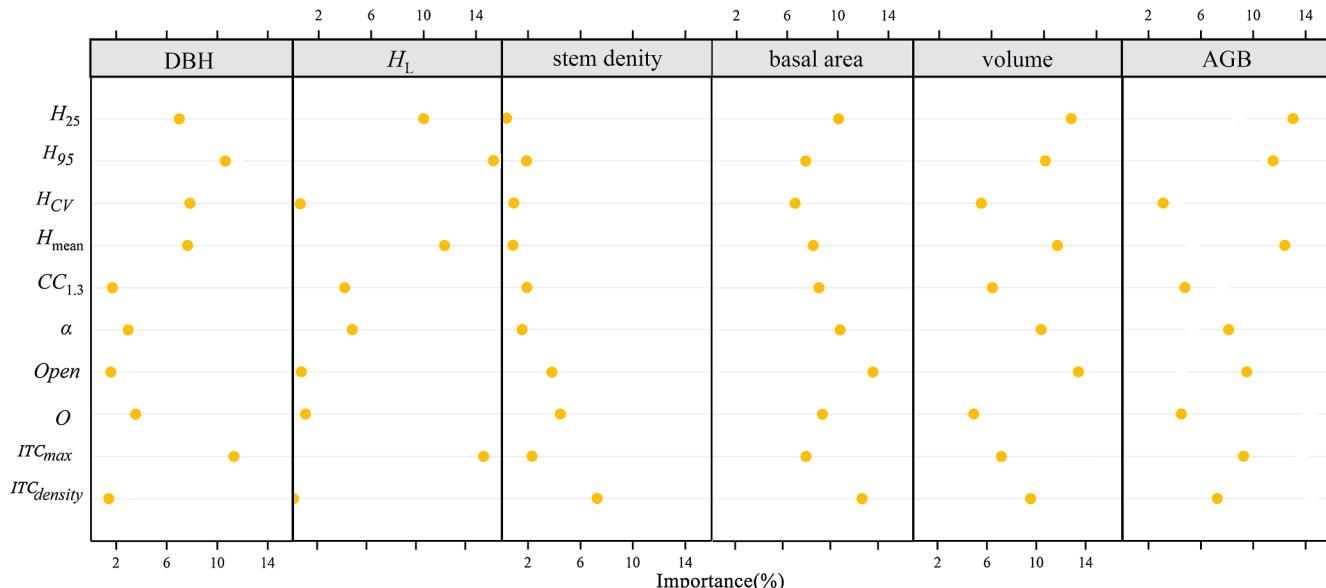


Fig. 9. Variable importance determined by the RF of the most important ten metrics for forest structural attribute estimation. See Table 3 for a description of the LiDAR-derived metrics.

3.4. UAV-LiDAR point density effects on performance of the derived metrics

To investigate the effects of UAV-LiDAR point density on UAV-LiDAR-derived metrics, we calculated 16 plot-level metrics (Fig. 11) with different sampling densities (50% ($80 \text{ pts}\cdot\text{m}^{-2}$), 10% ($16 \text{ pts}\cdot\text{m}^{-2}$), 5% ($8 \text{ pts}\cdot\text{m}^{-2}$), and 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$)) and the Pearson's correlation between AGB and the reduced UAV-LiDAR metrics. The resulting box-plots (Fig. 11) showed the change in metric values by UAV-LiDAR point density reduction. From a-1 to a-4, with the decrease in point density, there were no significant differences in the metric values of H_{25} , H_{50} , H_{75} and H_{95} . Meanwhile, the Pearson's correlation (r) values between AGB and these four metrics were relatively stable (0.94–0.95). For the metrics H_{cv} and $CC_{1,3 \text{ m}}$, the r values ranged from 0.60 to 0.64 (H_{cv}) and 0.79 to 0.82 ($CC_{1,3 \text{ m}}$). As shown in b-1 to b-4, the metric values of D_3 , D_5 , D_7 and D_9 changed across different UAV-LiDAR point densities. For D_3 , D_5 and D_7 , as the point density decreased from 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) to 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$), the metric values showed little variation, and the similar tendency was shown in Pearson's correlation when the point density decreased from 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) to 5% ($8 \text{ pts}\cdot\text{m}^{-2}$). However, when the point density decreased to 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$), there was a marked decrease in the r values, and the effect of point density was more pronounced. For D_9 , as the point density dropped from 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) to 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$), the mean value of the metric rapidly increased from 0.01 to 0.03. The change in the r values of D_9 had a similar tendency as that of other point density metrics. c-3 and c-4 in Fig. 11 and a-3 to e-3 in Fig. 12 show the effect of UAV-LiDAR density on the Weibull-fitted metrics. The metric values had a slight decreasing tendency as the density dropped from 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) to 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$). In addition, for the Pearson's correlation values between AGB and these two metrics, as the point density thinned from 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) to 50% ($80 \text{ pts}\cdot\text{m}^{-2}$), the r values decreased slightly, and as the density decreased from 50% ($80 \text{ pts}\cdot\text{m}^{-2}$) to 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$), there was a marked decrease in the r values. As shown in d-1 (Fig. 11), the $open$ values monotonically increased with decreasing UAV-LiDAR point density, and the associated r value first increased from 0.64 (100% ($160 \text{ pts}\cdot\text{m}^{-2}$)) to 0.81 (5% ($8 \text{ pts}\cdot\text{m}^{-2}$)) and then slightly decreased from 0.81 (5% ($8 \text{ pts}\cdot\text{m}^{-2}$)) to 0.8 (1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$)). d-2 and d-3 (Fig. 12) show a rapid decline in the E and O metric values as the UAV-LiDAR point cloud densities decrease from 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) to 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$). For the $closed$ metric, the mean value slightly increased (from 0.27 to 0.34) as the point density dropped from 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) to 5% ($8 \text{ pts}\cdot\text{m}^{-2}$) and markedly decreased as the point density dropped from 5% ($8 \text{ pts}\cdot\text{m}^{-2}$) to 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$). The r values of these metrics were only slightly affected by a decrease in the point density.

$\text{pts}\cdot\text{m}^{-2}$) to 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$), the mean value of the metric rapidly increased from 0.01 to 0.03. The change in the r values of D_9 had a similar tendency as that of other point density metrics. c-3 and c-4 in Fig. 11 and a-3 to e-3 in Fig. 12 show the effect of UAV-LiDAR density on the Weibull-fitted metrics. The metric values had a slight decreasing tendency as the density dropped from 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) to 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$). In addition, for the Pearson's correlation values between AGB and these two metrics, as the point density thinned from 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) to 50% ($80 \text{ pts}\cdot\text{m}^{-2}$), the r values decreased slightly, and as the density decreased from 50% ($80 \text{ pts}\cdot\text{m}^{-2}$) to 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$), there was a marked decrease in the r values. As shown in d-1 (Fig. 11), the $open$ values monotonically increased with decreasing UAV-LiDAR point density, and the associated r value first increased from 0.64 (100% ($160 \text{ pts}\cdot\text{m}^{-2}$)) to 0.81 (5% ($8 \text{ pts}\cdot\text{m}^{-2}$)) and then slightly decreased from 0.81 (5% ($8 \text{ pts}\cdot\text{m}^{-2}$)) to 0.8 (1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$)). d-2 and d-3 (Fig. 12) show a rapid decline in the E and O metric values as the UAV-LiDAR point cloud densities decrease from 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) to 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$). For the $closed$ metric, the mean value slightly increased (from 0.27 to 0.34) as the point density dropped from 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) to 5% ($8 \text{ pts}\cdot\text{m}^{-2}$) and markedly decreased as the point density dropped from 5% ($8 \text{ pts}\cdot\text{m}^{-2}$) to 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$). The r values of these metrics were only slightly affected by a decrease in the point density.

Table 6

Summary of the RF cross-validation CV-R² and relative RMSE (rRMSE) results and differences among cross-validation models for DBH, Lorey's mean height (H_L), stem density (N), basal area (G), volume (V) and aboveground biomass (W_{AGB}) estimation.

Variable	Models	CV-R ²	rRMSE (%)	Differences in cross-validation		
				Mean	Std-dev	Range
DBH	Plot-level metrics	0.87	9.46	-0.13	1.38	-3.57 to 2.26
	Plot-level and individual tree metrics	0.89	8.89	-0.1	1.26	-3.34 to 2.37
H_L	Plot-level metrics	0.95	3.74	0.01	0.39	-1.04 to 0.95
	Plot-level and individual tree metrics	0.95	3.69	0.01	0.45	-1.12 to 0.87
N	Plot-level metrics	0.71	16.16	4.31	86.4	-116 to 225
	Plot-level and individual tree metrics	0.77	15.54	3.89	79.4	-105 to 204
G	Plot-level metrics	0.89	9.84	-1.24	4.59	-9.45 to 10.51
	Plot-level and individual tree metrics	0.89	9.71	-1.06	4.48	-9.29 to 10.34
V	Plot-level metrics	0.93	9.87	-0.54	5.2	-12.02 to 11.87
	Plot-level and individual tree metrics	0.93	9.94	-0.59	5.5	-11.49 to 12.63
W_{AGB}	Plot-level metrics	0.94	8.72	-0.33	3.61	-8.89 to 8.43
	Plot-level and individual tree metrics	0.94	8.36	-0.28	3.22	-7.81 to 8.65

CV-R²: R-square of cross-validation; rRMSE: relative Root-Mean-Square-Error; Std-dev: Standard deviation.

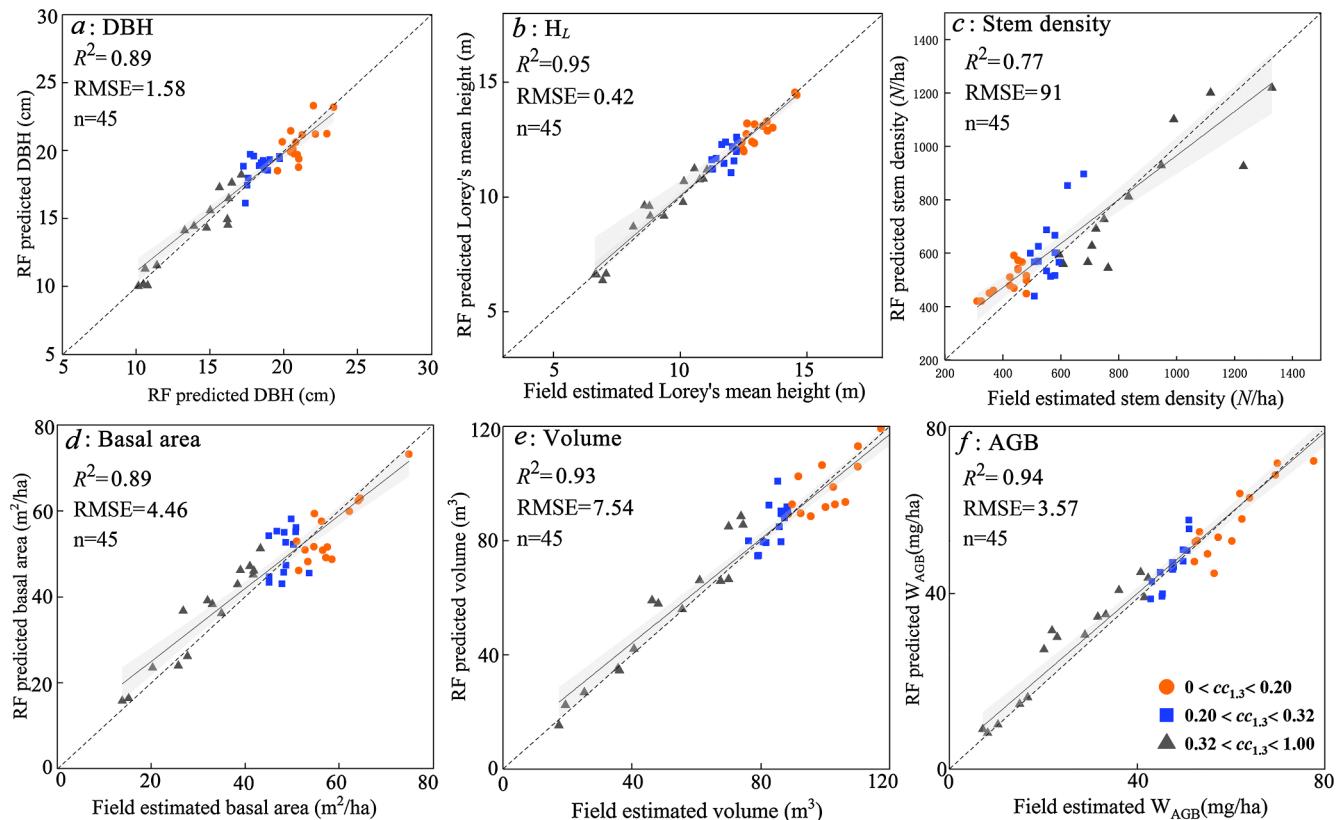


Fig. 10. Scatterplots for RF model of forest structural attributes, i.e., a: DBH; b: Lorey's mean height; c: stem density; d: basal area; e: volume; f: aboveground biomass. CC1.3 represents the canopy cover above 1.3 m.

density from 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) to 5% ($8 \text{ pts}\cdot\text{m}^{-2}$); for E and *closed*, the r values increased, while the r value of O decreased. However, when the point density decreased to 1% ($1.6 \text{ pts}\cdot\text{m}^{-2}$), there was a marked decrease in the r values of all three metrics.

3.5. Effects of UAV-LiDAR point density on individual tree detection

To test the sensitivity of individual tree detection to point cloud density, the LiDAR point cloud were decimated to 50% ($80 \text{ pts}\cdot\text{m}^{-2}$), 10% ($16 \text{ pts}\cdot\text{m}^{-2}$), 5% ($8 \text{ pts}\cdot\text{m}^{-2}$) and 3% ($4.8 \text{ pts}\cdot\text{m}^{-2}$). Fig. 13 shows the point cloud of one plot decimated to 50%, 10%, 5% and 3% and the detection results within a sample plot with five point cloud densities by using the PCS algorithm. These results indicated that the PCS algorithm succeeded in detecting individual trees. Individual tree detection had a relatively high accuracy (F_1 -score $> 74.93\%$) when the point cloud density was higher than 10%. The segmentation using point clouds with 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) density had the highest number of correct trees (161), the lowest number of omission trees (26) and the best performance of tree detection (F_1 -score = 80.90%), and as the point density

decreased from 100% ($160 \text{ pts}\cdot\text{m}^{-2}$) to 3% ($4.8 \text{ pts}\cdot\text{m}^{-2}$) the number of omission trees increased (from 26 to 132), and the number of correctly detected trees and the overall accuracy decreased (from 161 to 55, from 80.9% to 44.9%, respectively) within the sample plot. The accuracy assessment of individual tree detection results for the plots under different point cloud densities were shown in Table 8.

4. Discussion

In this study, we used a multi-rotor UAV, which could offer a stable platform for the LiDAR sensor by reducing vibration during the flight mission (Wallace et al., 2012). In addition, large-capacity batteries (Table 2) and lightweight LiDAR sensor (Velodyne Puck VLP-16) were mounted on the platform to ensure a high endurance. The results demonstrated that UAV-LiDAR data are suitable for estimating planted forest structural attributes at the plot level. When the PLS was used, the DBH, Lorey's mean height, basal area, volume and AGB estimations had rRMSE values of 8.59%, 2.29%, 11.94%, 11.48% and 9.76%, respectively; thus, this method had a better performance than previous

Table 7

Performance of RF model on different groups of forest structural attributes CV-R²: R-square of cross-validation; rRMSE: relative Root-Mean-Square-Error. DBH, Lorey's mean height (H_L), stem density (N), basal area (G), volume (V) and aboveground biomass (W_{AGB}).

	Group 1 (n = 15)		Group 2 (n = 15)		Group 3 (n = 15)		Total (n = 45)	
	CV-R ²	rRMSE (%)						
DBH	0.92	7.76	0.92	7.94	0.90	8.25	0.89	8.89
H_L	0.96	3.05	0.97	2.68	0.95	3.80	0.95	3.69
N	0.85	12.29	0.87	11.86	0.80	12.67	0.77	15.54
G	0.93	9.37	0.94	8.34	0.91	11.04	0.93	9.71
V	0.95	8.71	0.95	8.94	0.92	10.10	0.93	9.87
AGB	0.95	8.17	0.96	7.62	0.94	8.18	0.94	8.36

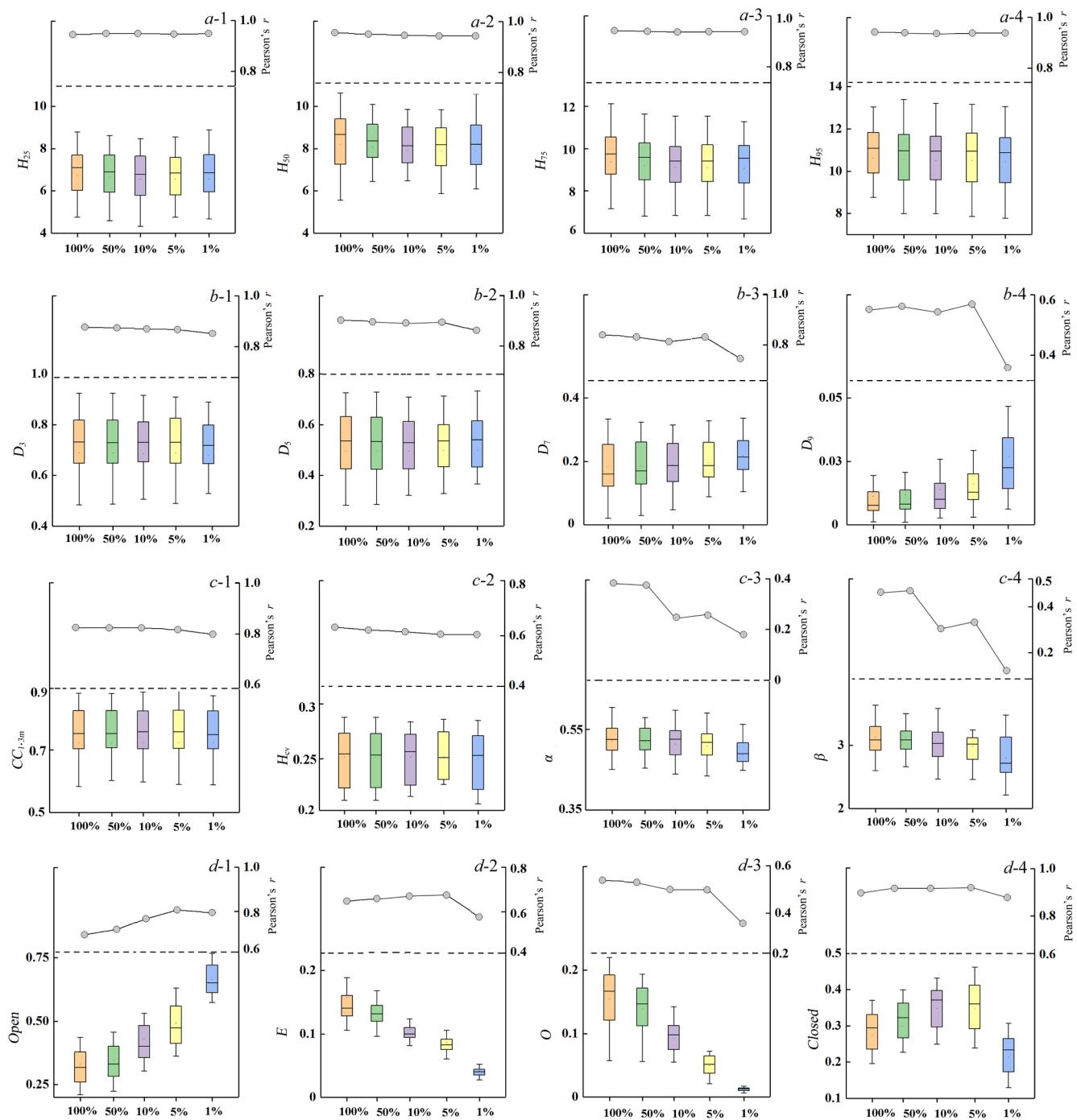


Fig. 11. Change in the values of the UAV-LiDAR metrics and the correlations between AGB and the metrics across reduced UAV-LiDAR point densities. The boxplots show the change in each metric value with a reduction in the UAV-LiDAR point density (a-1 to a-4, c-1 to c-2: height-related metrics, b-1 to b-4: density-related metrics, c-3 to c-4: Weibull-fitted metrics, d-1 to d-4: canopy volume metrics). The line chart in the top of each boxplot shows the Pearson's correlation between AGB and the reduced UAV-LiDAR metrics. See Table 3 for the description of the LiDAR-derived metrics.

airborne scanning LiDAR estimation methods (the rRMSE of DBH, Lorey's mean height, basal area, volume and AGB estimation were 11.7%, 4.6%, 14.8%, 15.9% and 14.3%, respectively) (Yu et al., 2015). This improved performance probably resulted from the different point cloud densities. In this study, the UAV-LiDAR had a lower flight altitude above ground level (60 m) and a slower flight speed (4.8 ms^{-1}) than the ALS system, and the average UAV-LiDAR point density of the plots was approximately $160 \text{ pts}\cdot\text{m}^{-2}$, which exceeds the typical point density of manned airborne LiDAR data (usually $< 25 \text{ pts}\cdot\text{m}^{-2}$) (Kato et al., 2009). The higher the point density is, the better the generated terrain quality will be because tree height measurements depend on the

DTM quality, and tree tops are detected with a higher accuracy as point density increases (Kato et al., 2009). In addition, higher point densities are more appropriate for capturing canopy structure and enable the three-dimensional structure of forests to be reconstructed at a finer scale with a higher precision (Jakubowski et al., 2013; Thomas et al., 2006).

Stem density estimation is crucial for understanding the overall structure of a forest stand; however, stem density is the most difficult forest structural attribute estimated from LiDAR metrics and usually has the poorest estimation accuracy when compared with the estimation accuracy of other attributes (Kandel et al., 2011). In our study, RF had

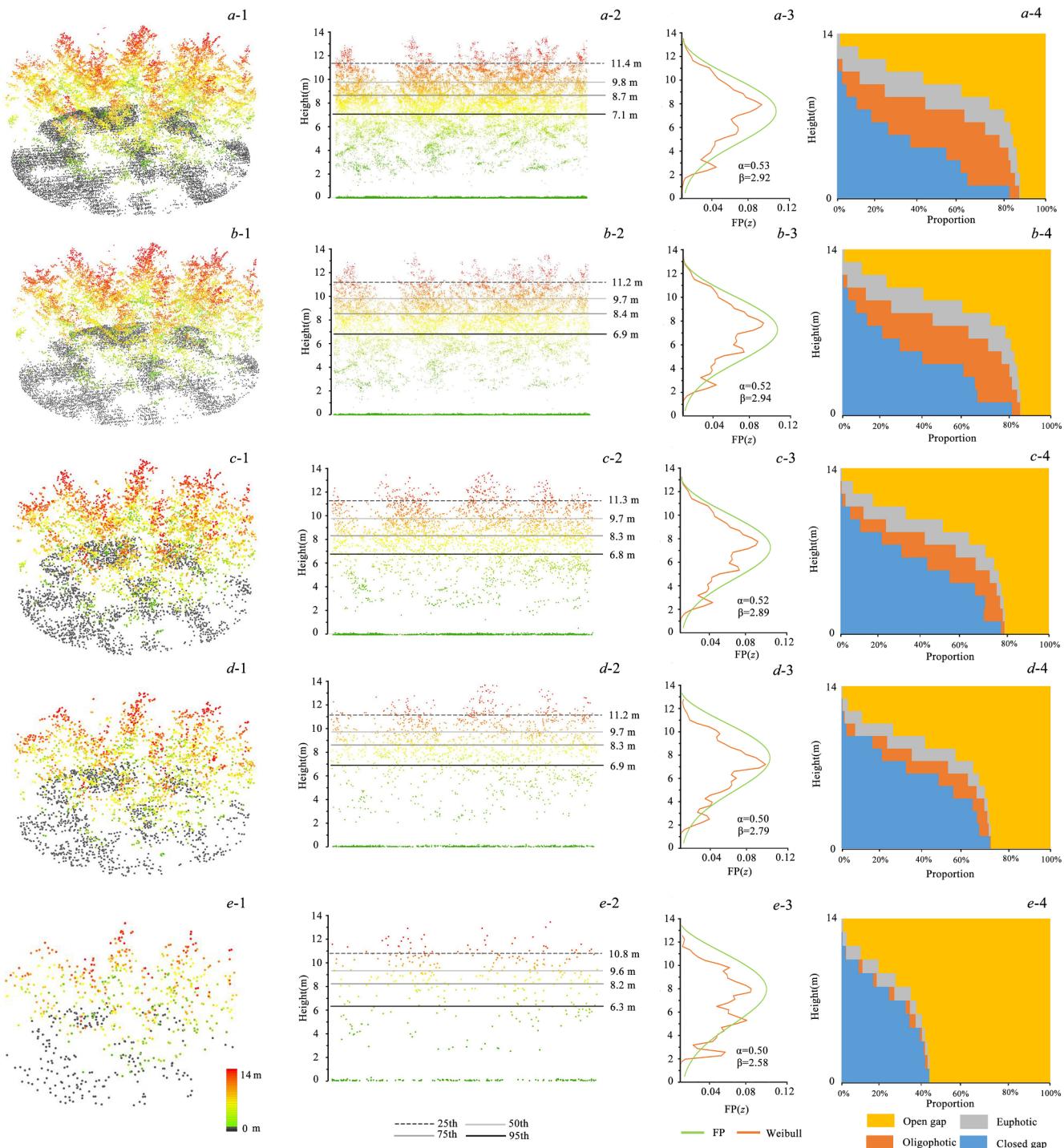


Fig. 12. An illustration of the percentage-based UAV-LiDAR data density reduction at the plot level. a-1, b-1, c-1, d-1 and e-1: 100%, 50%, 10%, 5% and 1% UAV-LiDAR point cloud density, respectively. a-2, b-2, c-2, d-2 and e-2: UAV-LiDAR point cloud profile at 100%, 50%, 10%, 5% and 1% density, respectively. a-3, b-3, c-3, d-3 and e-3: UAV-LiDAR-derived Weibull-fitted metrics at 100%, 50%, 10%, 5% and 1% density, respectively. a-4, b-4, c-4, d-4 and e-4: UAV-LiDAR-derived canopy volume metrics at 100%, 50%, 10%, 5% and 1% density, respectively.

the highest stem density prediction accuracy, with a CV- R^2 of 0.76 and rRMSE of 13.67%. Dash et al. (2015) used the k-NN method to estimate stem density in the managed forest of *Pinus radiata* in New Zealand and reported an R^2 of 0.70 and a relative mean deviation of -1.8%, which is a higher prediction accuracy than that based on the k-NN method in our study (CV- R^2 = 0.69, rRMSE = 18.57%). The study was conducted in coniferous forest, and stem density estimation in coniferous forests are more accurate than those in broad leaved forests (Heurich and Thoma, 2008). Lim et al. (2003a) developed a linear fitting model for

stem density with airborne LiDAR data for a hardwood forest in Canada with an R^2 of 0.86, which is a significantly higher accuracy than the result (R^2 = 0.66) of the linear model in our study. It was expected that the estimation of stem density with UAV-LiDAR metrics for a ginkgo plantation forest would be less accurate than that with airborne LiDAR data in other planted forests, due to the following reasons: the LiDAR sensor (Velodyne Puck VLP-16) used in this study is a lightweight sensor and, compared with airborne LiDAR sensors, has a weaker penetration, which makes it more difficult to detect the bottoms of trees;

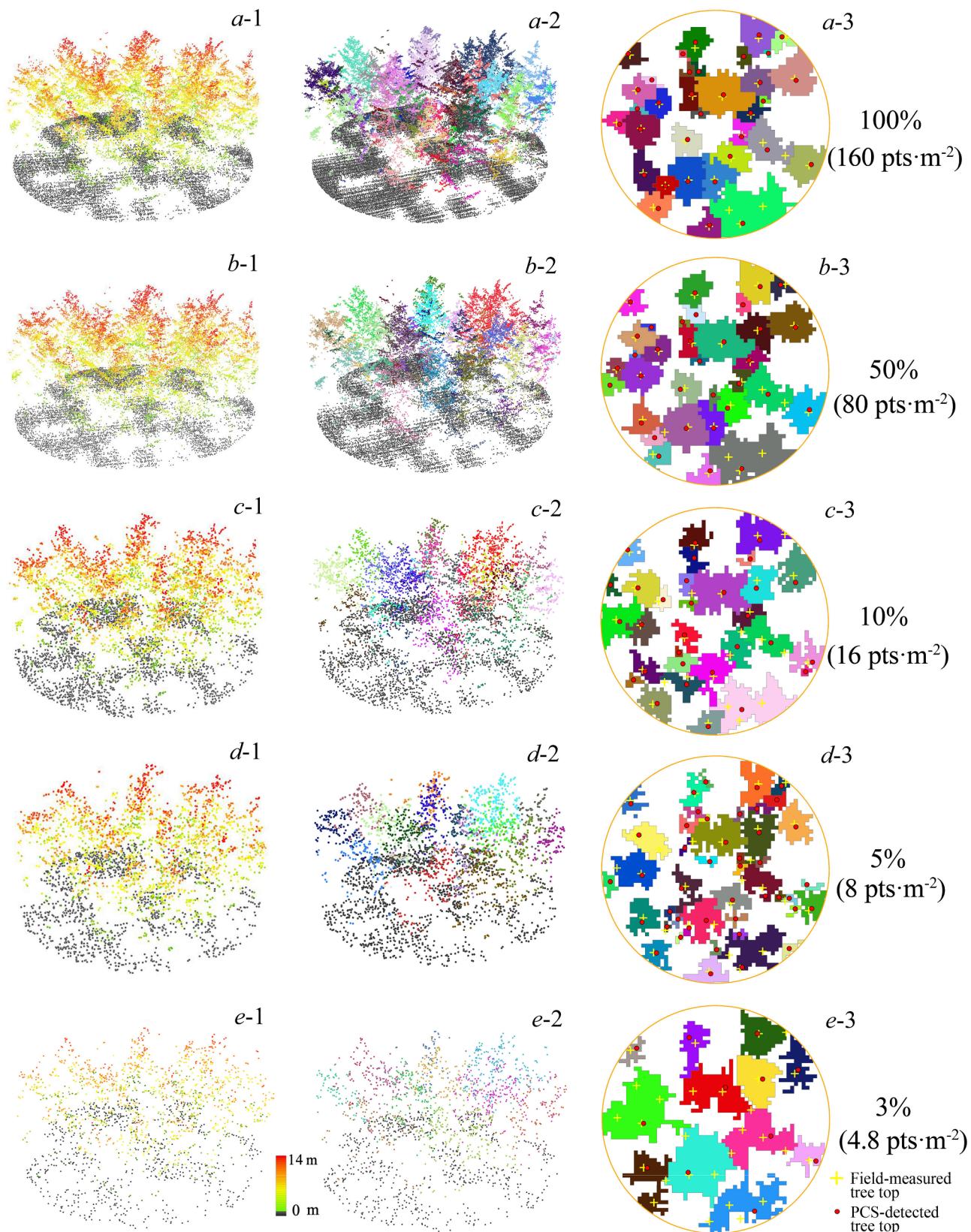


Fig. 13. An illustration of the percentage-based UAV-LiDAR data density reductions at the plot level and the segmentation results derived by using the PCS algorithm for this plot with five point cloud density grades (each tree corresponds to a color).

the beam divergence and flight height of LiDAR sensor also affect the estimate performance, as the low flying height (50 m) of the UAV and large angles ($\pm 30^\circ$) of incidence of the LiDAR sensor produce a

shadowing effect (Wallace et al., 2012); in addition, the unique properties of the ginkgo trees increased the difficulties of individual tree detection in sample plots (Wallace et al., 2014).

Table 8

The accuracy assessment of individual tree detection results using the PCS algorithm for plots with 100%, 50%, 10%, 5% and 3% point cloud density.

Density	Correct/Nt	Omission/No	Commission/Nc	r (%)	p (%)	F ₁ (%)
100% (160 pts·m ⁻²)	161	26	50	86.10	76.30	80.90
50% (80 pts·m ⁻²)	147	40	44	78.61	76.96	77.78
10% (16 pts·m ⁻²)	121	66	15	64.71	88.97	74.93
5% (8 pts·m ⁻²)	96	91	10	51.34	90.57	65.53
3% (4.8 pts·m ⁻²)	55	132	3	29.41	94.83	44.90

Nt is the number of detected trees in the plot, No is the number of trees omitted by PCS, and Nc is the number of trees that do not exist in the plot but were falsely detected. See the calculations of r, p, F₁ in Eqs. (5)–(7).

Table 9

The results of reducing sample size by random subsampling. The R² and rRMSE are the average cross-validation results calculated using RF with 100 iterations.

Plot number	DBH		H _{lorey's}		Stem Density		Basal Area		Volume		W _{AGB}	
	CV-R ²	rRMSE (%)	CV-R ²	rRMSE (%)	CV-R ²	rRMSE (%)	CV-R ²	rRMSE (%)	CV-R ²	rRMSE (%)	CV-R ²	rRMSE (%)
all	0.92	6.77	0.95	3.80	0.76	13.67	0.90	9.70	0.92	10.10	0.93	8.18
36	0.92	6.79	0.95	4.02	0.75	15.48	0.90	9.65	0.91	10.26	0.93	8.43
27	0.90	7.33	0.95	4.24	0.79	11.86	0.89	10.68	0.92	11.08	0.93	9.62
18	0.92	8.45	0.94	5.49	0.75	18.71	0.88	11.46	0.90	14.06	0.91	10.83
9	0.91	8.83	0.92	9.13	0.81	7.40	0.84	14.54	0.91	11.24	0.93	15.74

CV-R²: R-square of cross-validation; rRMSE: relative Root-Mean-Square-Error.

Previous studies have demonstrated the effects of sample size on the estimation of forest structural attributes (Gobakken et al., 2013; Gobakken and Næsset, 2009; Shin et al., 2016). Gobakken et al. (2013) found that the optimized sample size and stratified sample design can maintain the area-based forest structural attribute estimation accuracy. Fassnacht et al. (2014) reported that the sample sizes chosen in previous studies ranged from 10 to 40. Thus, considering the trade-offs, 45 plots were used in this study, and the plots were stratified into three groups according to canopy cover. In addition, the predictive models indicated that the stratified models are more accurate than the general models. In this study, we performed a sensitivity analysis to reduce the sample size for analyzing the effects of sample size on forest structural attribute estimation. The sample size was reduced using simple random sampling to 80%, 60%, 40% and 20% of the original plot number (i.e., 36, 27, 18 and 9 plots), and the R² and rRMSE values of the cross-validation results (of Random Forest) were calculated at each iteration after refitting the predictive model (of the forest structural parameters). The result of reducing the sample size by random subsampling is shown in Table 9, which indicates that the cross-validation accuracies were decreased by reducing the sample size. When the sample size was decreased to 9, the predictive models had the lowest accuracies (rRMSE = 7.40, 15.74%). However, when the sample size was decreased from 45 to 27, the increase in the rRMSE was less than 2%. Therefore, 27 to 45 could be the suitable sample size range in this study for forest structural attribute estimations.

Planted forests are composed by trees established through planting or seeding (Petersen et al., 2016). In general, for the plantations, the trees in a stand usually with homogeneous forest attributes (e.g., age, stem density and tree species, etc.) and the species were commonly keep in monoculture (Carle and Holmgren, 2003; Carnus et al., 2006; Shi et al., 2016). Moreover, the ginkgo planted forest is a typical plantation in China (Cao et al., 2010), and the plots used in this study were established across different stem density, tree height and DBH of ginkgo. So this study will provide important meanings for the precise silvicultural treatment and sustainable management of ginkgo plantation. However, this study was only focused on ginkgo plantations in relatively homogeneous forest conditions, making it difficult to transfer these results to other species. Additional studies assessing the approaches are needed to be undertaken in a number of different and highly variable forest conditions before more general conclusions can be drawn. Moreover, the sample size used in this study may not be

suitable for other studies. In other forests with complex forest structure, such as natural or secondary forests, more plots should be established and the plots should be stratified according to forest type or development stages of silviculture treatments. More sophisticated modeling approaches such as dummy variables or mixed-effects models could also be implemented to maintain the predictive accuracy with relatively small number of sample plots of each group after stratification.

This study examined how UAV-LiDAR point cloud density affects the derived metrics and the correlations between these metrics and AGB. The results showed that the metrics had a different performance with reduced point cloud densities. The UAV-LiDAR point density was reduced from 100% (160 pts·m⁻²) to 50% (79.5 pts·m⁻²), 10% (15.9 pts·m⁻²), 5% (8 pts·m⁻²) and 1% (1.6 pts·m⁻²). In our study, the point cloud density reduction did not significantly affect the variabilities of most distributional metrics (H_{25} , H_{50} , H_{75} , H_{95} , H_{cv} , $CC_{1.3}$, D_3 , D_5 , and D_7). This result agreed with previous studies (Hansen et al., 2015; Silva et al., 2017b). Unlike the distribution metrics, our results showed that the reduction in UAV-LiDAR point density had a significant effect on the canopy volume model metrics (*Open*, *E*, *O* and *Closed*). This can be explained by the fact that distributional metrics are computed directly from the entire return point cloud (Garcia et al., 2017), whereas the canopy volume metrics are computed based on voxels, and a matrix composed of voxels organizes the canopy spaces (Coops et al., 2007). Pearson's correlation showed that the correlations between AGB and the distributional metrics were reliable until UAV-LiDAR point cloud densities down to 5%, which is in agreement with various studies that large LiDAR point cloud density reductions do not significantly affect AGB estimations (Gonzálezferreiro et al., 2012; Treitz et al., 2012).

5. Conclusions

In this study, the effectiveness of plot-level metrics (i.e., distributional, canopy volume and Weibull-fitted metrics) and individual-tree-summarized metrics (i.e., maximum, minimum and mean height of trees and the number of trees from the individual tree detection (ITD) results) derived from UAV-LiDAR point clouds were assessed within a Ginkgo plantation in east China. Estimation models of six forest structural attributes were fitted by parametric (i.e., partial least squares (PLS)) and non-parametric (i.e., k-Nearest Neighbors (k-NN) and Random Forest (RF)) approaches. The results showed that, in general, models based on both plot-level and individual-tree-summarized

metrics ($CV-R^2 = 0.66\text{--}0.97$, $rRMSE = 2.83\text{--}23.35\%$) performed better than models based on the plot-level metrics only ($CV-R^2 = 0.62\text{--}0.97$, $rRMSE = 3.81\text{--}27.64\%$). PLS had a relatively high prediction accuracy for Lorey's mean height ($CV-R^2 = 0.97$, $rRMSE = 2.83\%$), whereas k-NN performed well for predicting volume ($CV-R^2 = 0.94$, $rRMSE = 8.95\%$) and AGB ($CV-R^2 = 0.95$, $rRMSE = 8.81\%$). For the point cloud density sensitivity analysis, the canopy volume metrics showed a higher dependence on point cloud density than other metrics. ITD results showed a relatively high accuracy ($F_1\text{-score} > 74.93\%$) when the point cloud density was higher than 10% ($16 \text{ pts}\cdot\text{m}^{-2}$). The correlations between AGB and the metrics of height percentiles, lower height level of canopy return densities and canopy cover appeared stable across different point cloud densities when the point cloud density was reduced from 50% ($80 \text{ pts}\cdot\text{m}^{-2}$) to 5% ($8 \text{ pts}\cdot\text{m}^{-2}$).

Acknowledgments

We acknowledge grants from the National Key R&D Program of China (2017YFD0600904) and the National Natural Science Foundation of China (31770590). We also acknowledge grants from the Natural Science Foundation of Jiangsu Province (BK20151515) and the Doctorate Fellowship Foundation of Nanjing Forestry University. This research was also supported by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD). The authors gratefully acknowledge the foresters in Pizhou forests for their assistances with data collection and sharing their rich knowledge and working experiences of the local forest ecosystems.

References

- Anderson, E.S., Thompson, J.A., Crouse, D.A., Austin, R.E., 2006. Horizontal resolution and data density effects on remotely sensed LiDAR-based DEM. *Geoderma* 132, 406–415.
- Bailey, R.L., Dell, T.R., 1973. Quantifying Diameter Distributions with the Weibull Function. *Forest Science* 19, 97–104.
- BCAL LiDAR Tools, 2016. Version bcal_lidar_tools_envi5X for ENVI 5.x, ed. Idaho State University, Department of Geosciences, Boise Center Aerospace Laboratory (BCAL), Boise, Idaho.
- Bengio, Y., Grandvalet, 2003. No unbiased estimator of the variance of K-fold cross-validation. *J. Mach. Learn. Res.* 5, 1089–1105.
- Bouvier, M., Durrieu, S., Fournier, R.A., Renaud, J.P., 2015. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote Sens. Environ.* 156, 322–334.
- Cao, F.-L., Kimmins, J., Jolliffe, P., Wang, J., 2010. Relative competitive abilities and productivity in Ginkgo and broad bean and wheat mixtures in southern China. *Agrofor. Syst.* 79, 369–380.
- Cao, F.L., 2007. Chinese Notes of Ginkgo Biloba. Chinese Forestry Press, Beijing.
- Cao, L., Gao, S., Li, P., Yun, T., Shen, X., Ruan, H., 2016. Aboveground biomass estimation of individual trees in a coastal planted forest using full-waveform airborne laser scanning data. *Remote Sens.* 8, 729.
- Carle, J., Holmgren, P., 2003. Definitions related to planted forests, UNFF Intercessional Expert Meeting on the Role of Planted Forests in Sustainable Forest Management. Zealand, Wellington, New, pp. 329–343.
- Carnus, J.-M., Parrotta, J., Brockerhoff, E., Arbez, M., Jactel, H., Kremer, A., Lamb, D., O'Hara, K., Walters, B., 2006. Planted forests and biodiversity. *J. Forest* 104, 65–77.
- Chen, Q., 2013. LiDAR Remote Sensing of Vegetation Biomass. In: Taylor & Francis: Remote Sensing of Natural Resources CRC Press, pp. 399–420.
- China's State Forestry Administration, 2014. The eighth (2009–2013) national forest resources inventory of china. *For. Resour. Manage.* 1–2.
- Chirici, G., Barbat, A., Corona, P., Marchetti, M., Travaglini, D., Maselli, F., Bertini, R., 2008. Non-parametric and parametric methods using satellite images for estimating growing stock volume in alpine and Mediterranean forest ecosystems. *Remote Sens. Environ.* 112, 2686–2700.
- Chirici, G., Mura, M., McInerney, D., Py, N., Tomppo, E.O., Waser, L.T., Travaglini, D., McRoberts, R.E., 2016. A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data. *Remote Sens. Environ.* 176, 282–294.
- Chisholm, R.A., Cui, J., Lum, S.K., Chen, B.M., 2013. UAV LiDAR for below-canopy forest surveys. *J. Unmanned Veh. Syst.* 1, 61–68.
- Clark, M.L., Clark, D.B., Roberts, D.A., 2004. Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. *Remote Sens. Environ.* 91, 68–89.
- Coops, N.C., Thomas, H., Michaela, W., Benoit, S.O., Glenn, N., Anders, S., Trofymow, J.A., 2007. Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees* 21, 295.
- Costa, C., Menesatti, P., Spinelli, R., 2015. Performance modelling in forest operations through partial least square regression. *Silva Fennica* 46, 241–252.
- Couteron, P., Pelissier, R., Nicolini, E.A., Paget, D., 2010. Predicting tropical forest stand structure parameters from Fourier transform of very high-resolution remotely sensed canopy images. *J. Appl. Ecol.* 42, 1121–1128.
- Crassidis, J.L., 2006. Sigma-point Kalman filtering for integrated GPS and inertial navigation. *IEEE Trans. Aerosp. Electron. Syst.* 42, 750–756.
- Crookston, N.L., Finley, A.O., 2008. yalmpute: An R Package for kNN Imputation. *J. Stat. Softw.* 23.
- Dash, J.P., Marshall, H.M., Rawley, B., 2015. Methods for estimating multivariate stand yields and errors using k-NN and aerial laser scanning. *Forestry* 88, 237–247.
- Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36, 27–46.
- Fassnacht, F., Hartig, F., Latifi, H., Berger, C., Hernández, J., Corvalán, P., Koch, B., 2014. Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sens. Environ.* 154, 102–114.
- Ferster, C.J., Coops, N.C., Trofymow, J.A., 2009. Aboveground large tree mass estimation in a coastal forest in British Columbia using plot-level metrics and individual tree detection from lidar. *Can. J. Remote Sens.* 35, 270–275.
- Food and Agriculture Organization of the United Nations (FAO), 2015. Global Forest Resources Assessment 2015: how are the world's forests changing? Rome, p. 9.
- García, M., Riaño, D., Chuvieco, E., Danson, F.M., 2010. Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sens. Environ.* 114, 816–830.
- García, M., Saatchi, S., Ferraz, A., Silva, C.A., Ustin, S., Koltunov, A., Balzter, H., 2017. Impact of data model and point density on aboveground forest biomass estimation from airborne LiDAR. *Carbon Bal. Manage.* 12, 4.
- Gobakken, T., Korhonen, L., Næsset, E., 2013. Laser-assisted selection of field plots for an area-based forest inventory. *Silva Fennica* 47, 943.
- Gobakken, T., Næsset, E., 2009. Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. *Can. J. For. Res.* 39, 1036–1052.
- Gonzálezferreiro, E., Diéguezaranda, U., Miranda, D., 2012. Estimation of stand variables in Pinus radiata D. Don plantations using different LiDAR pulse densities. *Forestry* 85, 281–292.
- Goodbody, T.R.H., Coops, N.C., Hermosilla, T., Tompalski, P., McCartney, G., Maclean, D.A., 2018. Digital aerial photogrammetry for assessing cumulative spruce budworm defoliation and enhancing forest inventories at a landscape-level. *ISPRS J. Photogramm. Remote Sens.* 142, 1–11.
- Goutte, C., Gaussier, E., 2005. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In: European Conference on Information Retrieval. Springer, pp. 345–359.
- Guo, J., Wu, Y., Wang, B., Lu, Y., Cao, F., Wang, G., 2016. The effects of fertilization on the growth and physiological characteristics of Ginkgo biloba L. *Forests* 7, 293.
- Guo, Q., Li, W., Yu, H., Alvarez, O., 2010. Effects of topographic variability and lidar sampling density on several DEM interpolation methods. *Photogramm. Eng. Remote Sens.* 76, 701–712.
- Guo, Q., Su, Y., Hu, T., Zhao, X., Wu, F., Li, Y., Liu, J., Chen, L., Xu, G., Lin, G., 2017. An integrated UAV-borne lidar system for 3D habitat mapping in three forest ecosystems across China. *Int. J. Remote Sens.* 38, 2954–2972.
- Hall, S.A., Burke, I.C., Box, D.O., Kaufmann, M.R., Stoker, J.M., 2005. Estimating stand structure using discrete-return lidar: an example from low density, fire-prone ponderosa pine forests. *For. Ecol. Manage.* 208, 189–209.
- Hansen, E., Gobakken, T., Næsset, E., 2015. Effects of pulse density on digital terrain models and canopy metrics using airborne laser scanning in a tropical rainforest. *Remote Sens.* 7, 8453.
- Harding, D.J., Lefsky, M.A., Parker, G.G., Blair, J.B., 2001. Laser altimeter canopy height profiles: methods and validation for closed-canopy, broadleaf forests. *Remote Sens. Environ.* 76, 283–297.
- Heurich, M., Thoma, F., 2008. Estimation of forestry stand parameters using laser scanning data in temperate, structurally rich natural European beech (*Fagus sylvatica*) and Norway spruce (*Picea abies*) forests. *Forestry* 81, 645–661.
- Horning, N., 2010. Random Forests: An algorithm for image classification and generation of continuous fields data sets. Proceedings of the International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences, Osaka, Japan.
- Huang, H.B., Gong, P., Cheng, X., Clinton, N., Li, Z.Y., 2009. Improving measurement of forest structural parameters by co-registering of high resolution aerial imagery and low density LiDAR data. *Sensors* 9, 1541–1558.
- Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M.S., Gessler, P.E., Morgan, P., 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. *Can. J. Remote Sens.* 32, 126–138.
- Ismail, R., Mutanga, O., Kumar, L., 2010. Modeling the potential distribution of pine forests susceptible to Sirex noctilio infestations in Mpumalanga, South Africa. *Trans. Gis* 14, 709–726.
- Jakubowski, M.K., Guo, Q., Kelly, M., 2013. Tradeoffs between lidar pulse density and forest measurement accuracy. *Remote Sens. Environ.* 130, 245–253.
- Kandel, Y.P., Fox, J.C., Arndt, S.K., Livesley, S.J., 2011. LiDAR estimation of mean dominant height, quadratic mean canopy height and stem density in native sclerophyll forests. Proceedings of SilviLaser. Australia, Hobart, Tasmania, pp. 1–12.
- Kato, A., Moskal, L.M., Schiess, P., Swanson, M.E., Calhoun, D., Stuetzel, W., 2009. Capturing tree crown formation through implicit surface reconstruction using airborne lidar data. *Remote Sens. Environ.* 113, 1148–1162.
- Kim, Y., Yang, Z., Cohen, W.B., Pflugmacher, D., Lauver, C.L., Vankat, J.L., 2009.

- Distinguishing between live and dead standing tree biomass on the North Rim of Grand Canyon National Park, USA using small-footprint lidar data. *Remote Sens. Environ.* 113, 2499–2510.
- Caliberte, Andrea S., Goforth, Marl A., Steele, Caitriana M., Rango, Albert, 2011. Multispectral remote sensing from unmanned aircraft: Image processing workflows and applications for rangeland environments. *Remote Sens. Environ.* 3, 2529–2551.
- Latifi, H., Fassnacht, F., Koch, B., 2012. Forest structure modeling with combined airborne hyperspectral and LiDAR data. *Remote Sens. Environ.* 121, 10–25.
- Laurin, G.V., Chen, Q., Lindsell, J.A., Coomes, D.A., Frate, F.D., Guerriero, L., Pirotti, F., Valentini, R., 2014. Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS J. Photogramm. Remote Sens.* 89, 49–58.
- 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 Sens. Environ.* 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, W., Guo, Q., Jakubowski, M.K., Kelly, M., 2012. A new method for segmenting individual trees from the lidar point cloud. *Photogramm. Eng. Remote Sens.* 78, 75–84.
- Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. *R News*, pp. 23.
- Lim, K., Treitz, P., Baldwin, K., Morrison, I., Green, J., 2003a. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. *Can. J. Remote Sens.* 29, 658–678.
- Lim, K., Treitz, P., Wulder, M., Stonge, B., Flood, M., 2003b. LiDAR remote sensing of forest structure. *Prog. Phys. Geogr.* 27, 88–106.
- Lisein, J., PierrotdeSeilligny, M., Bonnet, S., Lejeune, P., 2013. A photogrammetric workflow for the creation of a forest canopy height model from small unmanned aerial system imagery. *Forests* 4, 922–944.
- Liu, K., Cao, L., Wang, G., Cao, f., 2017. Biomass allocation patterns and allometric models of Ginkgo biloba L. *J. Beijing For. Univ.* 12–20.
- Lovell, J.L., Jupp, D.L.B., Culvenor, D.S., Coops, N.C., 2003. Using airborne and ground-based ranging lidar to measure canopy structure in Australian forests. *Can. J. Remote Sens.* 29, 607–622.
- Magnusson, M., Jes, F., Holmgren, J., 2007. Effects of estimation accuracy of forest variables using different pulse density of laser data. *For. Sci.* 53, 619626 (618).
- Maltamo, M., Eerikäinen, K., Pitkänen, J., Hyypä, J., Vehmas, M., 2004. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. *Remote Sens. Environ.* 90, 319–330.
- Maltamo, M., Peuhkurinen, J., Malinen, J., Vauhkonen, J., Packalén, P., Tokola, T., 2009. Predicting tree attributes and quality characteristics of Scots pine using airborne laser scanning data. *Silva Fennica* 43, 507–521.
- Manuri, S., Andersen, H.-E., McCaughey, R.J., Brack, C., 2017. Assessing the influence of return density on estimation of lidar-based aboveground biomass in tropical peat swamp forests of Kalimantan, Indonesia. *Int. J. Appl. Earth Obs. Geoinf.* 56, 24–35.
- Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L., Hendrix, C.J., 2000. Predicting forest stand characteristics with airborne scanning lidar. *Photogramm. Eng. Remote Sens.* 66, 1367–1371.
- Hubert, M., Rousseeuw, P.J., Verdonck, T., 2012. A deterministic algorithm for robust location and scatter. *J. Comput. Graph. Statist.* 21, 618–637.
- Montealegre, A.L., Lamelas, M.T., Riva, J.D.L., García-Martín, A., Escribano, F., 2016. Use of low point density ALS data to estimate stand-level structural variables in Mediterranean Aleppo pine forest. *Forestry* 89, 373–382.
- Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* 80, 88–99.
- Næsset, E., 2004. Effects of different flying altitudes on biophysical stand properties estimated from canopy height and density measured with a small-footprint airborne scanning laser. *Remote Sens. Environ.* 91, 243–255.
- Næsset, E., 2009. Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. *Remote Sens. Environ.* 113, 148–159.
- Næsset, E., Gobakken, T., 2008. Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sens. Environ.* 112, 3079–3090.
- Ozdemir, I., Karnieli, A., 2011. Predicting forest structural parameters using the image texture derived from WorldView-2 multispectral imagery in a dryland forest, Israel. *Int. J. Appl. Earth Obs. Geoinf.* 13, 701–710.
- Petersen, R., Goldman, E., Harris, N., Sargent, S., Aksenen, D., Manisha, A., Espipova, E., Shevade, V., Loboda, T., Kuksina, N., 2016. Mapping tree plantations with multispectral imagery: preliminary results for seven tropical countries. *World Resources Institute*, Washington, DC.
- Puliti, S., Ene, L.T., Gobakken, T., Næsset, E., 2017. Use of partial-coverage UAV data in sampling for large scale forest inventories. *Remote Sens. Environ.* 194, 115–126.
- Ritchie, J.C., 1993. Measuring canopy structure with an airborne laser altimeter. *Trans. Asae* 36, 1235–1238.
- Schmid, W., Balz, J., 2005. Cultivation of Ginkgo biloba L. on three continents. *Acta Hortic.*
- Schmidlein, S., Feilhauer, H., Bruelheide, H., 2012. Mapping plant strategy types using remote sensing. *J. Veg. Sci.* 23, 395–405.
- Shi, H., Wen, Z., Paull, D., Jiao, F., 2016. Distribution of natural and planted forests in the yanhe river catchment: have we planted trees on the right sites? *Forests* 7, 258.
- Shin, J., Temesgen, H., Strunk, J.L., Hilker, T., 2016. Comparing modeling methods for predicting forest attributes using LiDAR metrics and ground measurements. *Can. J. Remote Sens.* 42, 739–765.
- Silva, C.A., Hudak, A.T., Klauber, C., Vierling, L.A., Gonzalez-Benecke, C., Carvalho, S.d.P.C., Rodriguez, L.C.E., Cardil, A., 2017a. Combined effect of pulse density and grid cell size on predicting and mapping aboveground carbon in fast-growing Eucalyptus forest plantation using airborne LiDAR data. *Carbon Bal. Manage.* 12, 13.
- Silva, C.A., Hudak, A.T., Vierling, L.A., Klauber, C., Garcia, M., Ferraz, A., Keller, M., Eitel, J., Saatchi, S., 2017b. Impacts of airborne lidar pulse density on estimating biomass stocks and changes in a selectively logged tropical forest. *Remote Sens.* 9, 1068.
- Singh, K.K., Chen, G., McCarter, J.B., Meentemeyer, R.K., 2015. Effects of LiDAR point density and landscape context on estimates of urban forest biomass. *ISPRS J. Photogramm. Remote Sens.* 101, 310–322.
- Song, Y.B., Ding, Y.P., Shen, F., 2009. The development and latest progress of JSCORS. *Bull. Surv. Mapp.* 2, 73–74.
- Tang, L., Shao, G., 2015. Drone remote sensing for forestry research and practices. *J. For. Res.* 26, 791–797.
- Thomas, V., Treitz, P., McCaughey, J., Morrison, I., 2006. Mapping stand-level forest biophysical variables for a mixedwood boreal forest using lidar: an examination of scanning density. *Can. J. For. Res.* 36, 34–47.
- Torresan, C., Berton, A., Carotenuto, F., Gennaro, S.F.D., Gioli, B., Matese, A., Miglietta, F., Vagnoli, C., Zaldei, A., Wallace, L., 2016. Forestry applications of UAVs in Europe: a review. *Int. J. Remote Sens.* 38, 2427–2447.
- Treitz, P., Lim, K., Woods, M., Pitt, D., Nesbitt, D., Etheridge, D., 2012. LiDAR sampling density for forest resource inventories in Ontario, Canada. *Remote Sens.* 4, 830–848.
- Vastaranta, M., Wulder, M.A., White, J.C., Pekkarinen, A., Tuominen, S., Ginzler, C., Kankare, V., Holopainen, M., Hyypä, J., Hyypä, H., 2013. Airborne laser scanning and digital stereo imagery measures of forest structure: comparative results and implications to forest mapping and inventory update. *Can. J. Remote Sens.* 39, 382–395.
- Véga, C., Renaud, J.P., Durrieu, S., Bouvier, M., 2016. On the interest of penetration depth, canopy area and volume metrics to improve Lidar-based models of forest parameters. *Remote Sens. Environ.* 175, 32–42.
- Vincenzi, S., Zucchetta, M., Franzoi, P., Pellizzato, M., Pranovi, F., Leo, G.A.D., Torricelli, P., 2011. Application of a Random Forest algorithm to predict spatial distribution of the potential yield of Ruditapes philippinarum in the Venice lagoon, Italy. *Ecol. Model.* 222, 1471–1478.
- Wallace, L., Lucieer, A., Watson, C., Turner, D., 2012. Development of a UAV-LiDAR System with Application to Forest Inventory. *Remote Sens.* 4, 1519–1543.
- Wallace, L., Lucieer, A., Watson, C.S., 2014. Evaluating tree detection and segmentation routines on very high resolution UAV LiDAR data. *IEEE Trans. Geosci. Remote Sens.* 52, 7619–7628.
- Wang, D., Xin, X., Shao, Q., Brolly, M., Zhu, Z., Chen, J., 2017. Modeling aboveground biomass in hulunbeir grassland ecosystem by using unmanned aerial vehicle discrete lidar. *Sensors* 17, 180.
- Wehrens, R., Mevik, B.-H., 2007. The pls package: principal component and partial least squares regression in R.
- White, J.C., Tompalski, P., Vastaranta, M., Wulder, M.A., Saarinen, N., Stepper, C., Coops, N.C., 2017. A model development and application guide for generating an enhanced forest inventory using airborne laser scanning data and an area-based approach, Victoria, British Columbia, Canada.
- William, M.E., Theodor, E.L., Martin, B.O., Terje, G., Erik, N., Ernest, M.R., Eliakimu, Z., 2015. Modelling aboveground forest biomass using airborne laser scanner data in the miombo woodlands of Tanzania. *Carbon Bal. Manage.* 10, 28.
- Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. *Chemomet. Intell. Lab.* 58, 109–130.
- Yu, X., Hyypä, J., Karjalainen, M., Nurminen, K., Karila, K., Vastaranta, M., Kankare, V., Kaartinen, H., Holopainen, M., Honkavaara, E., 2015. Comparison of laser and stereo optical, SAR and InSAR point clouds from air- and space-borne sources in the retrieval of forest inventory attributes. *Remote Sens.* 7, 15933–15954.
- Zarco-Tejada, P.J., Diaz-Varela, R., Angileri, V., Loudjani, P., 2014. Tree height quantification using very high resolution imagery acquired from an unmanned aerial vehicle (UAV) and automatic 3D photo-reconstruction methods. *Eur. J. Agron.* 55, 89–99.
- Zhang, Z., Cao, L., She, G., 2017. Estimating forest structural parameters using canopy metrics derived from airborne LiDAR data in subtropical forests. *Remote Sens.* 9, 940.
- Zhao, K., Popescu, S., Meng, X., Pang, Y., Agca, M., 2011. Characterizing forest canopy structure with lidar composite metrics and machine learning. *Remote Sens. Environ.* 115, 1978–1996.
- Zhao, X., Guo, Q., Su, Y., Xue, B., 2016. Improved progressive TIN densification filtering algorithm for airborne LiDAR data in forested areas. *ISPRS J. Photogramm. Remote Sens.* 117, 79–91.
- Zolkos, S.G., Goetz, S.J., Dubayah, R., 2013. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sens. Environ.* 128, 289–298.