

Article

Crown Structure Metrics to Generalize Aboveground Biomass Estimation Model Using Airborne Laser Scanning Data in National Park of Hainan Tropical Rainforest, China

Chenyun Li ¹, Zhexiu Yu ^{2,*}, Shaojie Wang ¹, Fayun Wu ³, Kunjian Wen ², Jianbo Qi ^{2,*}  and Huaguo Huang ² 

¹ Industry Development and Planning Institute, National Forestry and Grassland Administration, Beijing 100010, China; lcy913@163.com (C.L.); wsjcgzw@163.com (S.W.)

² Research Center of Forest Management Engineering of State Forestry and Grassland Administration, Beijing Forestry University, Beijing 100083, China; wenkunjian@bjfu.edu.cn (K.W.); huaguo_huang@bjfu.edu.cn (H.H.)

³ Academy of Inventory and Planning, National Forestry and Grassland Administration, Beijing 100714, China; wufayun@sina.com

* Correspondence: yzhexiu@bjfu.edu.cn (Z.Y.); jianboqi@bjfu.edu.cn (J.Q.)

Abstract: Forest aboveground biomass (AGB) is an important indicator for characterizing forest ecosystem structures and functions. Therefore, how to effectively investigate forest AGB is a vital mission. Airborne laser scanning (ALS) has been demonstrated as an effective way to support investigation and operational applications among a wide range of applications in the forest inventory. Moreover, three-dimensional structure information relating to AGB can be acquired by airborne laser scanning. Many studies estimated AGB from variables that were extracted from point cloud data, but few of them took full advantage of variables related to tree crowns to estimate the AGB. In this study, the main objective was to evaluate and compare the capabilities of different metrics derived from point clouds obtained from ALS. Particularly, individual tree-based alpha-shape, along with other traditional and commonly used plot-level height and intensity metrics, have been used from airborne laser scanning data. We took the random forest and multiple stepwise linear regression to estimate the AGB. By comparing AGB estimates with field measurements, our results showed that the best approach is mixed metrics, and the best estimation model is random forest ($R^2 = 0.713$, RMSE = 21.064 t/ha, MAE = 15.445 t/ha), which indicates that alpha-shape may be a good alternative method to improve AGB estimation accuracy. This method provides an effective solution for estimating aboveground biomass from airborne laser scanning.

Keywords: tree crown; alpha-shape; airborne laser scanning (ALS)



Citation: Li, C.; Yu, Z.; Wang, S.; Wu, F.; Wen, K.; Qi, J.; Huang, H. Crown Structure Metrics to Generalize Aboveground Biomass Estimation Model Using Airborne Laser Scanning Data in National Park of Hainan Tropical Rainforest, China. *Forests* **2022**, *13*, 1142. <https://doi.org/10.3390/f13071142>

Academic Editor: Dmitry Schepaschenko

Received: 30 May 2022

Accepted: 14 July 2022

Published: 20 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Forest biomass is a fundamental parameter for characterizing forest ecosystem structures and functions, which is the data of basic features for studying forest ecosystems [1]. Forest aboveground biomass (AGB) plays a vital role in the carbon cycle and the greenhouse effect reduction, as the energy base and material source for forest ecosystem works [2]. Especially, the rainforest is the most resistant and stable ecosystem on earth, with hot perennial climate, rich rainfall, rapid biological community succession, and abundant biodiversity [3,4]. Therefore, its biomass surveys are important for their water cycle, climate regulation, and organic matter conversion.

Although traditional surveys have a high precision on the ground, they are destructive to forest environments and there are some uncertainties in the selection of sample plots; there can also be residual variability, parameter estimation errors in the calculations, and a large labor force is required [5–7]. Remote sensing has the advantages of large scale, fast operation and low cost, and has become an important technique for investigating and monitoring forest resources [8].

LiDAR, in particular, is an active remote sensing technique that acquires three-dimensional (3-D) structure information of the vegetation directly and accurately, making up for the drawbacks of optical remote sensing applications in forestry [9]. Moreover, airborne laser scanning has the ability to accurately indicate the 3-D structure of the forest canopy [10,11]. It is also fast and efficient in the acquisition of spatial information of a large area and identifies abundant surface features in a short time, greatly reducing the workload and shortening the time required for field measurement [12].

Most studies of AGB modeling have used the correlation between the point cloud variables and measured AGB in situ. Laurin et al. studied tropical forest AGB estimation in Ghana, where LiDAR metrics were calculated using ALS data, and Multivariate Adaptive Regression Splines (MARS) was used in AGB estimation. The results showed that MARS provided a better accuracy for estimating AGB ($R^2 = 0.72$, RMSE = 47.1 t/ha). Moreover, this study indicated that the employed LiDAR metrics can be used for tree species richness assessment [13]. Manuri et al. assessed the effect of LiDAR return density on estimating AGB in tropical peat swamp forests in Kalimantan, Indonesia, where they found cumulative return proportion (CRP) variables in the power model has explanatory of 90.9% for the AGB variation [14]. Knapp et al. proposed a general estimation method of AGB across forest types from different continents. They studied five different sites including temperate and tropical forests in North America, Central America, South America, Europe, and Africa, respectively. The results indicated the importance of the employed structure descriptors in AGB estimation and the mechanisms and relationships (nRMSE = 12.4%, $R^2 = 0.74$) [15]. Jiang et al. suggested a stratification-based category for estimating AGB, which was named stratification based on forest stand structure (SBFSS). They used ALS to estimate the AGB of Gaofeng Forest Farm in China, and their model was able to improve estimation accuracy efficiently and corresponding to specific forest types in a subtropical forest ecosystem ($R^2 = 0.78$, RMSE = 0.78 t/ha, RMSEr = 23.65%) [16]. Oliveira et al. used ALS data to estimate total aboveground carbon (TAGC) and total aboveground biomass (TAGB) of dry tropical forests in Brazil, and established a model using multiple stepwise linear regression methods. The results showed that LiDAR can be used for estimating biomass and total carbon in dry tropical forests ($R^2 = 0.533$, RMSE = 14.76 t/ha) [17].

In conclusion, ALS data can observe the forest vertical structure and identify the features of tree height and crown, which have been frequently used as a proxy for estimating AGB. In order to estimate AGB with the point cloud, a widely-used general approach is using height, intensity, and density parameters derived from point cloud data, analyzing the relationship between these variables and AGB, and establishing estimation models. However, few studies have used point cloud data to extract individual tree crown morphological parameters to estimate AGB, such as crown surface area, crown volume, and crown width; these parameters can affect the AGB in the actual forest environment. The crown is one of the important parts of a tree and an important source of nutrients; the size reflects the competition between trees and the level of environmental effects. At the same time, the crown is the comprehensive result of tree growth and interaction with the environment [18]. Crowns are also related to tree growth and affect the quantity of AGB. Therefore, the research on volume, surface area, and crown width are important for the tree biomass. It is usually difficult to measure the volume and surface area of tree crowns in the field, but the 3-D characteristics of crowns can be highlighted by using airborne laser scanning.

Tree objects can be formed by the alpha-shape algorithm, which was used to reconstruct crown and stem attributes of trees from scattered point clouds. However, most research with the alpha-shape algorithm in forestry is related to the forestry scenes simulation or individual tree reconstruction; these parameters of crown construction based on alpha-shape had not been used to establish regression models of forestry parameters. Therefore, we employed the alpha-shape algorithm to obtain crown features and construct parameters for each tree in this paper. We took the National Park of Hainan Tropical Rainforest in southern China as the research area. In this forest, we established AGB

estimation models with field measurements and some variables extracted from the point cloud, and compared the advantages of crown metrics and other metrics for performing AGB estimation. Research from this study will provide insights into the rainforest situation at Hainan Island that improve the estimation accuracy in understanding forestry resources for scientific management and better development.

2. Materials and Methods

2.1. Materials

2.1.1. Study Area

The study area is located in the National Park of Hainan Tropical Rainforest in Hainan Island, China ($108^{\circ}36' \sim 109^{\circ}57'$ E, $18^{\circ}23' \sim 19^{\circ}11'$ N, Figure 1), which is located in the northwestern part of the South China Sea. The island has dense tropical rainforests, and the hydrothermal conditions are superior. Forest coverage exceeds 50% and the composition of vegetation is rich and diverse with a tropical monsoon maritime climate. Hainan Island is rich in rainfall, rivers, and hydrological resources. Therefore, monitoring and recording an inventory of aboveground biomass of the island is beneficial to sustaining biodiversity, forest management, forest resources development, and the tourism industry [19,20]. The study area covers about 4900 km², which is one-seventh of the area of Hainan Island. The dominant tree species in the plantation forest are mixed broad-leaved forest, *Hevea* spp., *Eucalyptus* spp., *Acacia* spp., *Cunninghamia lanceolata*, and Foreign Pines in the study area [21]. It is located on a hilly landform at an altitude of 100–1876 m. The annual rainfall is about 1759 mm with an average temperature of 21.6 °C, and the relative humidity is 80%. May to November are the rainy months and December to April are usually dry.

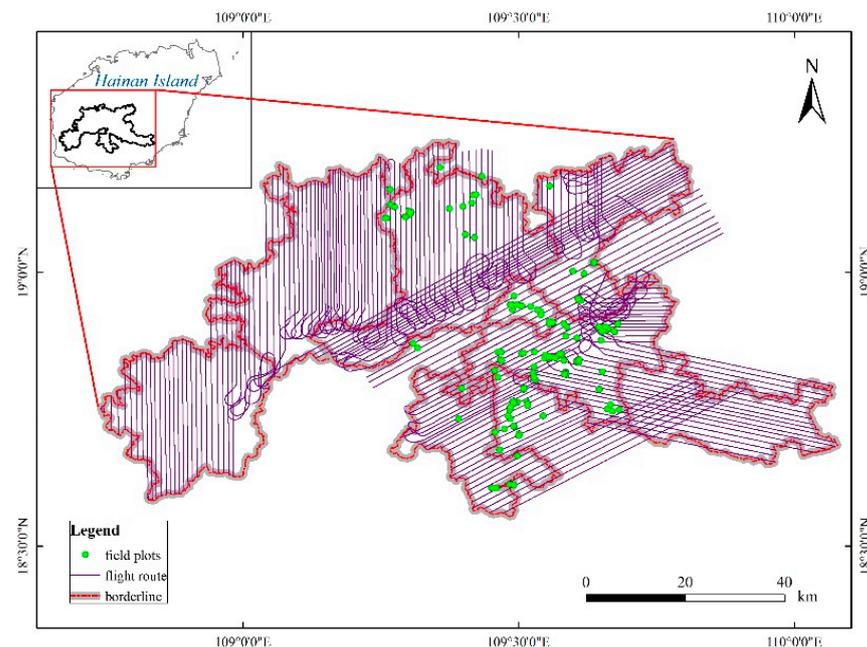


Figure 1. Location of the study area and the distribution of sample plots.

2.1.2. ALS Data Acquisition

The airborne laser scanning data were collected by National Forestry and Grassland Administration from March 2020 to February 2021, using a RIGEL VQ1560i-DW laser scanning system carried on a Cessna 208B aircraft on two sorties. The data format was LAS, and the average point cloud density of the plantation was 10 points/m². We collected ALS data for 217 routes with 71 h flight duration. The detailed scanning parameters of ALS are shown in Table 1.

Table 1. Airborne laser scanning system parameters.

Parameters	Value
Wavelength (nm)	1064
Divergence angle (mrad)	0.25
Pulse repetition rate (KHz)	2000
Scanning rate	2 × 666 kHz @ 60° scan angle
Width (m)	1980
Relative flight altitude (m)	1800
Flight speed (km/h)	260
Side overlap (%)	22
Average point spacing (m)	0.45

2.1.3. ALS Data Processing

The raw ALS data coordinates were calculated using position and orientation system (POS) and continuously operating reference stations (CORS). A POS was integrated on the LiDAR sensor. The POS model was Applanix AP60 (produced by Trimble Inc., Sunnyvale, CA, USA), which has an integrated global navigation satellite system (GNSS) receiver (that performs georeferencing of the data) and an inertial measurement unit (IMU) (which measures the drone's multi-directional movements and orientation and assists in increasing the accuracy of data georeferencing). The continuously operating reference stations (CORS) data were provided by Hainan Administration of Surveying Mapping and Geoinformation. Following this we used LiDAR360 (version 4.1 developed by GreenValley Co., Ltd., Beijing, China) to process within strip mosaic, point cloud classification, noise reduction, height normalization, individual tree detection, and individual tree segmentation. Among them, the individual tree segmentation algorithm used Li et al.'s [22] development algorithms, which were integrated into LiDAR360.

2.1.4. Inventory Data

The field data include sample-plot survey results collected during the ALS data acquisition period. On account of the topographic relief, accessibility of sample plots location, and the operability of setting sample plots, some plots were set outside the study area but the tree species were the same as within the study area. Meanwhile, these plots were recorded by airborne laser scanning and had corresponding point cloud data. As a result, we obtained 166 circular plots (Figure 1), each with a radius of 15 m, with real-time kinematic (RTK) used to measure the position of the center of each plot. The laser altimeter (Haglof Vertex Laser developed by Haglöf Sweden) was used to measure the height of each tree and the DBH of each tree was measured with a diameter tape. The crown width of the north–south and west–east directions of each tree were measured with tapes. The distribution of sample plots is shown in Figure 1. The specific sample plot information is shown in Table 2.

Table 2. Sample plots information.

Tree Species	Number of Sample Plots	Diameter at Breast Height (DBH/cm)	Tree Height (m)
Coniferous and broad-leaved mixed forest	6	21.4 ± 9.2	16.0 ± 7.6
broad-leaved mixed forest	18	19.1 ± 22.7	11.9 ± 11.4
Other coniferous trees	12	21.8 ± 20.6	15.6 ± 8.4
Chinese fir	2	18.1 ± 2.8	12.0 ± 2.2
<i>Eucalyptus robusta</i> Smith	49	17.2 ± 15.4	16.9 ± 10.7
<i>Acacia confusa</i> Merr.	30	21.7 ± 19.2	15.8 ± 11.2
<i>Hevea brasiliensis</i>	49	17.9 ± 20.2	14.7 ± 8.6

Note: $m \pm n$, m is the median of the tree parameters for each tree species, n is the maximum value by which this parameter fluctuates up or down.

2.2. Methods

2.2.1. AGB Calculation

The AGB of all tree species were calculated by using the AGB formula in [23]. Each of the AGB models is shown in Table 3.

Table 3. AGB models for each of the tree species.

Tree Species	AGB Models of Individual Tree	AGB of Plot ($t \cdot ha^{-1}$)
Coniferous and broad-leaved mixed forest	$AGB = 0.2253 \times D^{2.4213}$	124.6 ± 106.4
broad-leaved mixed forest	$AGB = 0.1131 \times (D^2H)^{0.8407}$	113.5 ± 260.2
Other coniferous trees	$AGB = 0.2309 \times (D^2H)^{0.6838}$	185.0 ± 174.6
Chinese fir	$AGB = 0.0182 \times (D^2H)^{0.9710}$	46.8 ± 7.3
<i>Eucalyptus robusta</i> Smith	$AGB = 0.0576 \times (D^2H)^{0.8587}$	89.2 ± 175.3
<i>Acacia confusa</i> Merr.	$AGB = 1.4240 \times (D^2H)^{0.5680}$	127.6 ± 170.0
<i>Hevea brasiliensis</i>	$AGB = 0.0524 \times D^{2.7451}$	46.5 ± 87.7

Note: AGB is aboveground biomass of individual trees, D is the diameter at breast height, and H is the tree height.

2.2.2. Crown Features Extraction

Alpha-shape is a classical algorithm of point cloud that outlines extraction, in which polyhedral generated precision was controlled from a parameter of α [24,25]. Selecting an appropriate α value to construct alpha-shape for a given point cloud data is able to return the original shapes approximately [26]. An alpha-shape polyhedral of a plot is shown in Figure 2.

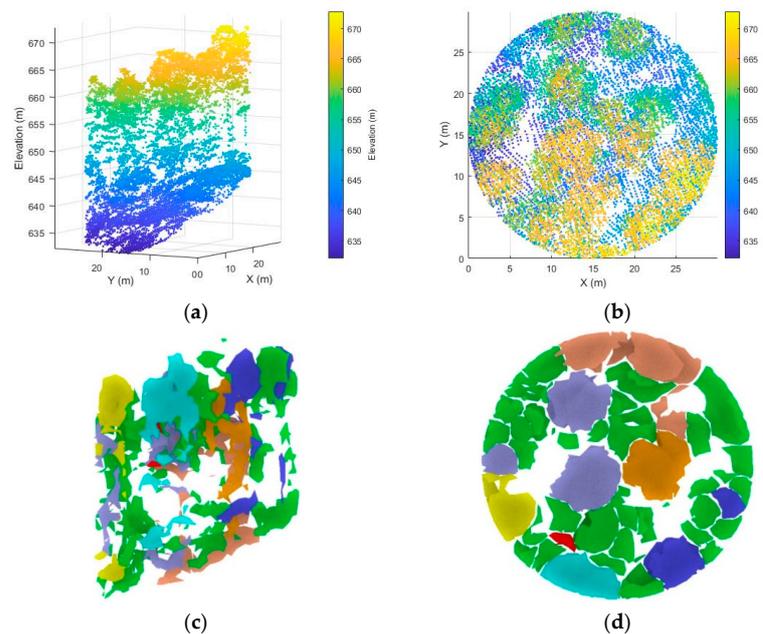


Figure 2. Original point cloud and alpha-shape in a plot: (a) Elevation view of the original point cloud; (b) Top view of the original point cloud; (c) Elevation view of alpha-shape for each individual wood; (d) Top view of alpha-shape.

2.2.3. Feature Variables Extraction

According to the structural features of point cloud data, considering the forest stand characteristics, canopy index, and ecological index, the four categories of feature parameters were extracted from LiDAR data for each field plot.

(a) Height metrics

Height metrics are statistical parameters related to the normalized height of the point cloud, which can be used in regression analysis with forest parameters of field plot

measurements. Otherwise, 10 density variables were also extracted at different elevations. The point cloud elevation is divided into 10 layers with the same height interval from lowest to highest within each plot. The proportion of echoes in each height slice is the corresponding density metric. As a result, we extracted 56 statistical parameters related to elevation. The depiction of elevation metrics is shown in Table 4.

(b) Intensity metrics

The intensity metrics are similar to the elevation metrics, but they are different from the elevation information of point cloud. A total of 42 intensity variables were calculated and the variable descriptions are shown in Table 4.

(c) Alpha-shape metrics

Alpha-shape algorithm was constructed in MATLAB, and we extracted 72 variables corresponding to tree crown with alpha-shape algorithm and statistical metrics derived like Table 4. These variables are shown in Table 5.

(d) Stand metrics

For the forest stand feature, four feature parameters were calculated, including canopy density, leaf area index, gap fraction and density of trees [27]. Canopy cover is the percentage of the vertical projection of forest canopy to forest land area, which plays an important role in forest ecology and resource management [28]. It is also an essential factor for estimating forest aboveground biomass. The leaf area index (*LAI*) is one of the most fundamental parameters for forest canopy structure and is a composite indicator of optical energy utilization and tree crown structures [29]. Additionally, the physiological and physical processes of the vegetation were reflected from the *LAI*, therefore it is closely related to forest biomass. Gap fraction can reflect illumination in forest stand and growth of the understory layer [1]. The competition of soil, water, and fertilizer among plants is also reflected indirectly in the gap fraction. The density of trees can indicate the utilization level of the space occupied by trees. As a result, the features of Tables 4–6 were calculated in LiDAR360. The four variables are shown in Table 6.

Table 4. ALS-derived tree height metrics, density metrics, and intensity metrics.

Variable Abbreviation	Description	Reference
$H_{max}(I_{max})$	Maximum tree height (intensity)	
$H_{min}(I_{min})$	Minimum tree height (intensity)	
$H_{mean}(I_{mean})$	Mean tree height (intensity)	
$H_{med}(I_{med})$	Median tree height (intensity)	
$H_{var}(I_{var})$	Variance of tree heights (intensity)	
$H_{std}(I_{std})$	Standard deviation of tree heights (intensity)	
$H_{aad}(I_{aad})$	Average absolute deviation of tree heights (intensity)	
H_{crr}	$H_{crr} = \frac{mean-min}{max-min}$	
$H_{cv}(I_{cv})$	Variation coefficient of tree heights (intensity)	
$H_{kurt}(I_{kurt})$	Kurtosis of tree heights (intensity)	[30]
$H_{skew}(I_{skew})$	Skewness of tree heights (intensity)	
$H_{mm}(I_{mm})$	Median of median absolute deviation of tree heights (intensity)	
$H_n(I_n)$	Tree height or intensity 1st, 5th, 95th, 99th percentile, 15 features were extracted in total	
$AIH_n(AIH_n)$	Accumulative interpercentile height (intensity)	
$H_{iq}(I_{iq})$	Interquartile range ($H_{75} - H_{25}$) or ($I_{75} - I_{25}$)	
AIH_{iq}	AIH Interquartile range ($AIH_{75} - AIH_{25}$)	
H_{sqr}	Generalized means for the 2nd power	
H_{cube}	Generalized means for the 3rd power	
D_n	Density metrics (0th, 1st, , 9th)	

Note: H_{xxx} is elevation, I_{xxx} is intensity.

Table 5. Tree crown metrics of alpha-shape complexity derived.

Abbreviation	Description
$CV_{max}, CV_{min}, CV_{mean}, CV_{std}, CV_{var}, CV_{med}, CV_{cv}, CV_{kurt}, CV_{skew}, CV_{iq}$	The volume as the alpha-shape complexity, unit: m^3
$CSA_{max}, CSA_{min}, CSA_{mean}, CSA_{std}, CSA_{var}, CSA_{med}, CSA_{cv}, CSA_{kurt}, CSA_{skew}, CSA_{iq}$	The surface area as the alpha-shape complexity, unit: m^2
$CX_{max}, CX_{min}, CX_{mean}, CX_{std}, CX_{var}, CX_{med}, CX_{cv}, CVX_{kurt}, CVX_{skew}, CVX_{iq}$	The width of the X-axis as the alpha-shape complexity, unit: m
$CY_{max}, CY_{min}, CY_{mean}, CY_{std}, CY_{var}, CY_{med}, CY_{cv}, CY_{kurt}, CY_{skew}, CY_{iq}$	The width of the Y-axis as the alpha-shape complexity, unit: m
$CZ_{max}, CZ_{min}, CZ_{mean}, CZ_{std}, CZ_{var}, CZ_{med}, CZ_{cv}, CZ_{kurt}, CZ_{skew}, CZ_{iq}$	The length of the Z-axis as the alpha-shape complexity, unit: m
$CW_{max}, CW_{min}, CW_{mean}, CW_{std}, CW_{var}, CW_{med}, CW_{cv}, CW_{kurt}, CW_{skew}, CW_{iq}$	The average crown width, $CW = \frac{1}{2}(C_X + C_Y)$
$CSR_{max}, CSR_{min}, CSR_{mean}, CSR_{std}, CSR_{var}, CSR_{med}, CSR_{cv}, CSR_{kurt}, CSR_{skew}, CSR_{iq}$	Crown shape ratio, $CSR = \frac{1}{2}(C_X + C_Y)/C_Z$
Vol	The volume of the 3-D alpha-shape
Surf	The surface area of the 3-D alpha-shape

Note: Alpha-shape metrics statistics for each plot as same as height and intensity metrics of point cloud data, maximum (max), minimum (min), average (mean), standard deviation (std), variance (var), median (med), coefficient variation (cv), kurtosis (kurt), skewness (skew), interquartile range (iq), respectively.

Table 6. Forest stand metrics.

Variable Abbreviation	Description	References
LAI	Leaf Area Index, half of the surface area of all leaves projected on the surface area of a plot	
CC	Canopy cover, the ratio of the first vegetation echoes to the total number of first echoes	[28,29]
GF	Gap fraction, the ratio of ground points to total points in a plot	
TD	Density of trees, the ratio of tree numbers of individual tree segmented from point cloud in each plot area	

2.2.4. Regression Modeling of AGB

In this paper, random forest and stepwise linear regression were used for establishing AGB models. The random forest and linear regression models were established by R programming.

Random forest (RF) is a robust machine learning algorithm that is constructed by combining the results of various decision trees and bagging the original dataset to select samples. Meanwhile, random forest is also a common feature variable selection method, and the features importance was sorted by increasing mean square error (*IncMSE*) [31,32], whose formula is shown in Equation (1):

$$IncMSE = \frac{1}{ntree} \sum (OOB_{error} - OOB'_{error}) \tag{1}$$

where: *ntree* is the number of random forest trees; OOB (out of bag) is a randomly selected sampling dataset; OOB_{error} is the error of OOB when the sampling dataset is not changed; and OOB'_{error} is the error of OOB when the sample set is changed.

Multiple stepwise linear regression (MSLR) considers the contribution of all independent variables to the dependent variable, step-by-step iterative establishment of a regression model, and finally the selection of independent variables to be built ultimate model [33]. We used the selected independent variables in random forest to populate the equation. The MSLR formula is shown in Equation (2):

$$y = a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_nx_n + \varepsilon \tag{2}$$

where: $a_1, a_2, a_3, \dots, a_n$ are constants, $x_1, x_2, x_3, \dots, x_n$ are independent variables, and ε is the error term.

2.2.5. Precision Assessment

The correspondence of the estimates with the reference data was evaluated by the coefficient of determination R^2 , root mean squared error (RMSE), and mean absolute error (MAE). However, R^2 value will always increase when models add independent variables in MSLR, so we use adjusted R^2 to assess the models' precision. The four assessment indicators were calculated according to:

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

$$R_{adj}^2 = 1 - \frac{n-1}{n-k-1} (1 - R^2) \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (5)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (6)$$

where: \hat{y}_i is predicted values; \bar{y} is the average of predicted values; y_i is observed values, n is the number of samples, k is the number of independent variables.

Where: \hat{y}_i is predicted values, \bar{y} is average of predicted values, y_i is observed values; n is the number of samples; and k is the number of independent variables.

3. Results

3.1. Feature Selection

We set the proportion of training samples and testing samples as 60%:40%. Among the 174 feature variables extracted from the point cloud data, the top 15 feature variables importance was sorted from height, intensity, alpha-shape, and mixed metrics, respectively, and four features importance were sorted by stand metrics. The importance rankings selected are shown in Figure 3.

Figure 3 shows the most important variables according to increasing the mean square error (*IncMSE*). The variable with the maximum *IncMSE* is tree density (*TD*), which is a suitable variable in forest stand features, followed by the surface area of 3-D alpha-shape (*Surf*), *I_{mm}*, *AIH_{60th}*, which are extracted from alpha-shape metrics, intensity metrics, and height metrics, respectively.

Furthermore, Figure 3e showed the top 15 features of all variables. Among them, the number of height metrics is the largest, followed by the alpha-shape, stand, and intensity metrics, respectively, of the whole variables. It indicates that the tree height and alpha-shape metrics are dominant, and they are more important to AGB.

As can be seen from the variables, the percentiles and accumulative interpercentiles are more important than the descriptive statistical variables in the height metrics and intensity metrics. There are significant *IncMSE* for in stand metrics, the four variables *IncMSE* > 1000. It is apparent that tree density, canopy cover, gap fraction, and *LAI* are important to AGB. The difference in the extremal value range of *IncMSE* between alpha-shape metrics and other metrics is very wide, besides, the descriptive statistical variables occupy the main part of the top 15 important features.

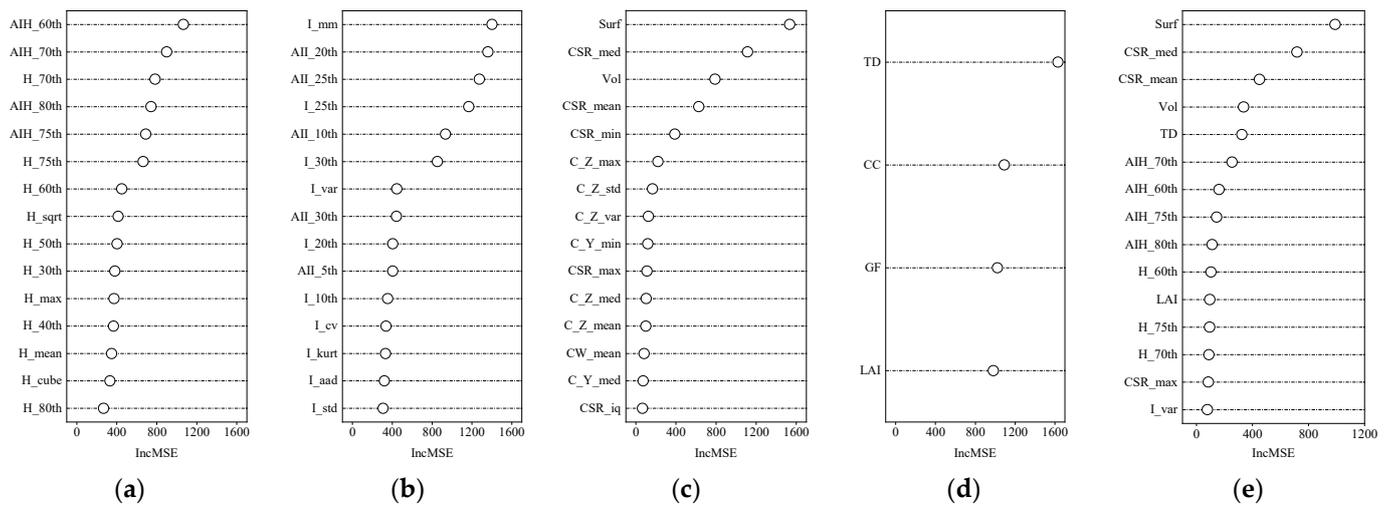


Figure 3. Feature importance rankings of top 15 in height, intensity, and alpha-shape metrics, respectively, and feature importance ranking of stand metrics. (a) Height metrics, (b) Intensity metrics, (c) Alpha-shape metrics, (d) Stand metrics, (e) All variables.

3.2. Correlation Analysis

The correlation coefficient analysis was carried out between AGB and the importance features. The bar charts in Figure 4 show the correlation between variables of each group and AGB.

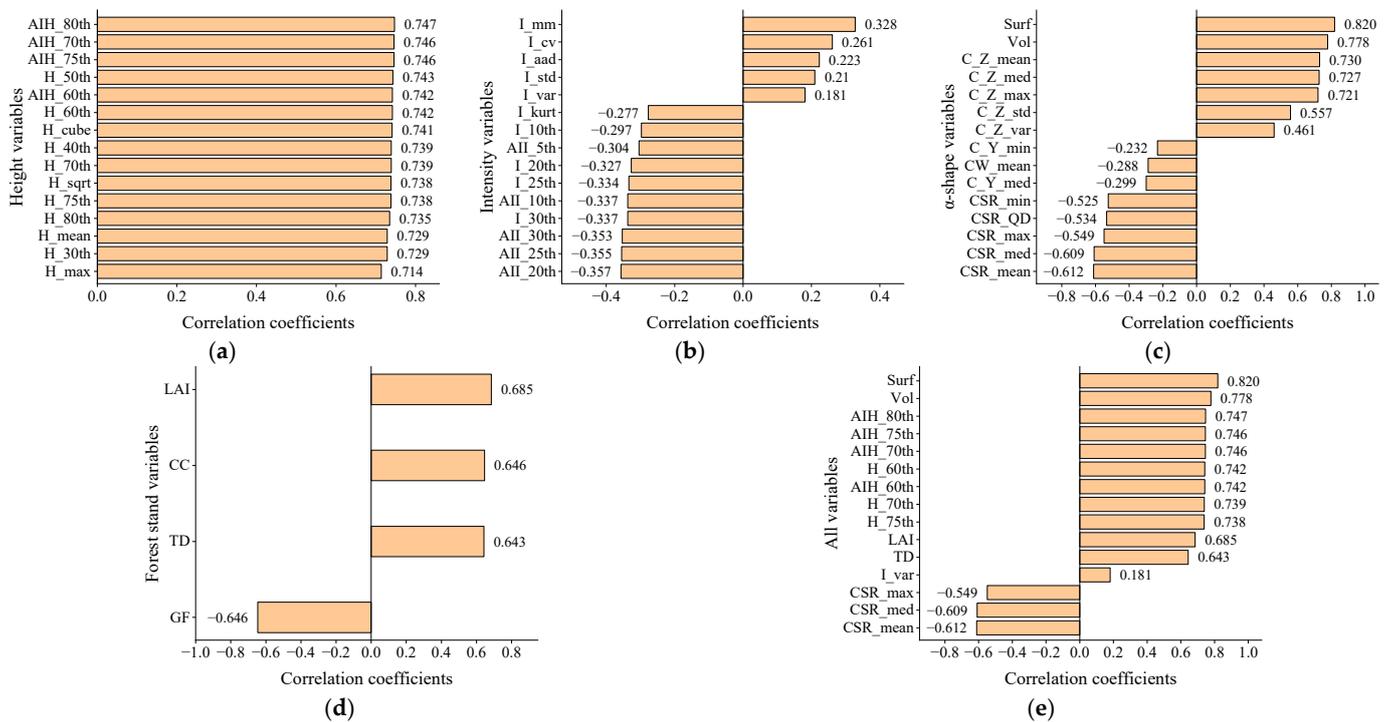


Figure 4. Features correlation of each group features: (a) Height metrics; (b) Intensity metrics; (c) Alpha-shape metrics; (d) Stand metrics; (e) Mixed metrics.

Among the 15 height features of the point cloud data, their correlation coefficient value was significantly and positively correlated with AGB ($r > 0.70, p < 0.01$). Overall, elevation percentiles features and AGB are the most correlated. However, the correlation between variables in intensity metrics, alpha-shape metrics, and stand metrics are negatively and positively correlated with AGB ($p < 0.01$), as shown in Figure 4b–d. Among all variables,

the surface area is the most correlative to AGB. From the bar charts, it can be seen that the absolute extreme difference of the correlation coefficient is large. In addition, the variables associated with the Z-axis of alpha-shape are positively correlated the same as height metrics. While the descriptive statistical variables of CSR and crown width are negatively correlated with AGB.

3.3. AGB Estimation Models

Figure 5 and Table 7 show the aboveground biomass models of random forest parameters. First, the data were normalized to remove the effect of dimension, then the selected variables were set to establish the AGB models.

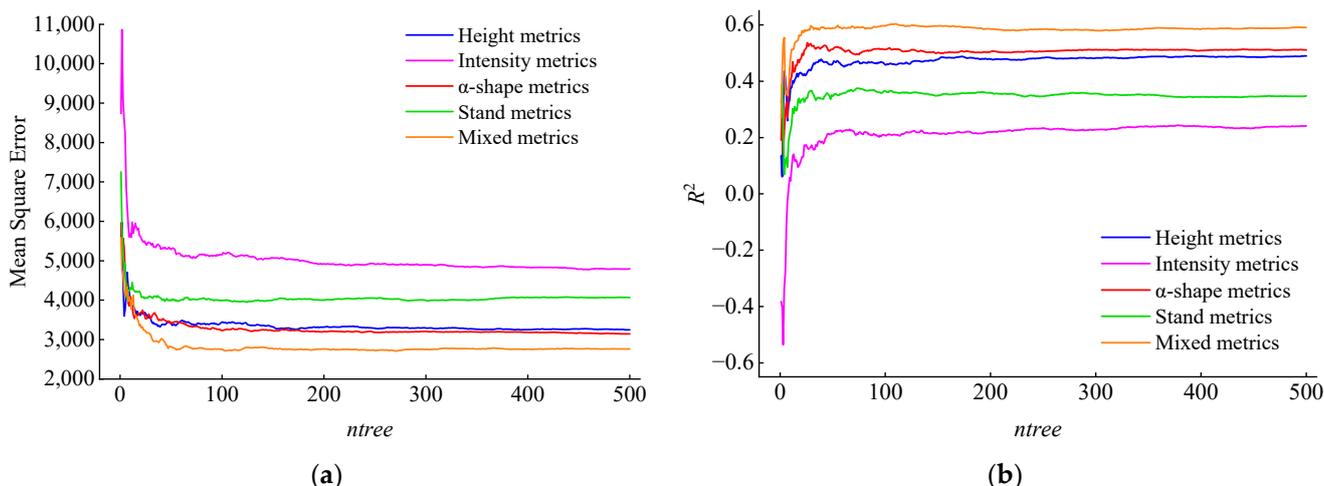


Figure 5. Hyperparameter tuning results for random forest: (a) The relationship between MSE and the number of trees; (b) The relationship between R^2 and the number of trees.

Table 7. The optimal parameters for each of the random forest model corresponding variables.

Feature Parameters	Corresponding the Minimum MSE			Corresponding the Maximum R^2		
	Min MSE	<i>ntree</i>	Corresponding R^2	Max R^2	<i>ntree</i>	Corresponding MSE
Height metrics	3259.615	173	0.487	0.487	173	3259.615
Intensity metrics	4910.851	198	0.219	0.229	66	5123.448
alpha-shape metrics	3194.515	197	0.502	0.537	26	3526.796
Stand metrics	3949.003	125	0.358	0.375	75	4012.294
Mixed metrics	2711.566	270	0.582	0.603	109	2721.430

The results of the random forest models are shown in Figure 5, where it can be seen that mean square error (MSE) decreases and R^2 increases gradually as the number of trees (*ntree* in random forest algorithm) increases. It is obvious that the trend is stable when the value of *ntree* reaches approximately 200. Combining the MSE and R^2 shows that after stabilization the error of intensity metrics is the largest, and R^2 is the smallest. There is a clear characteristic that $R^2 < -1$ when a few the *ntree* are applied, indicating that the fitting is poor. The alpha-shape metrics have the smallest MSE and the largest R^2 . This indicates that the accuracy of random forest regression based on variables of alpha-shape is the best, followed by height, stand, and intensity metrics.

In addition, the model started training with a small *ntree* and the MSE and R^2 performance were unstable, while the performance tends to be stable when the number of trees was increasing and taking into account the removal $R^2 < -1$ of intensity metrics model. Secondly, Excessive *ntree* will lead to overfitting. Therefore, the optimal parameters were found by counting the variation in MSE and R^2 between 8 to 200 trees. Table 7 lists the optimal parameters for random forest regression of each variables group. It can be seen from the *ntree* in Table 7 that the *ntree* with the best MSE and R^2 of height variables was the

same, while there is a great difference for the other three groups of variables. Overall, the most stable model is the height metric-based model, while the most accurate model is the alpha-shape based model. To compare the difference between min MSE and max R^2 , the optimal number of trees is 26 for alpha-shape variables.

Table 8 illustrates the aboveground biomass models of MSLR. The AGB model of each linear regression was established by using the selected important variables based on the four categories of feature parameters. The table below illustrates some of the main characteristics of the regression coefficients and important independent variables. The larger the absolute value of the regression coefficient, the stronger explanation of the variable for AGB, and the higher the number of independent variables, the better the relationship between the response of AGB and LiDAR signal. It is apparent from this table that the mixed metrics model is the same as the alpha-shape metrics model. We used the Durbin–Watson test, F statistics and R^2_{adj} to assess the MSLR models, which showed that the surface area of the crown was a better and important variable to estimate the AGB, followed by the height, stand, and intensity metrics.

Table 8. Different linear regression models.

Feature Parameters	Model	Durbin-Watson Test	R^2_{adj}	Sig.
Height metrics	$AGB = -36.035 + 9.028 \times AIH_{70}$	2.182	0.501	$p < 0.001$
Intensity metrics	$AGB = 437.465 - 0.047 \times AII_{20th} + 0.036 \times I_{30th}$	2.360	0.204	$p < 0.001$
alpha-shape metrics	$AGB = 2.332 + 0.015 \times Surf$	2.215	0.607	$p < 0.001$
Stand metrics	$AGB = 5.011 + 0.134 \times TD + 11.616 \times LAI$	2.253	0.450	$p < 0.05$
Mixed metrics	$AGB = 2.332 + 0.015 \times Surf$	2.215	0.607	$p < 0.001$

3.4. Regression Model Precision Assessment

Figures 6 and 7 summarize the accuracy of the evaluation results of AGB estimation models. Figure 6i,j show that the result using random forest based on mixed metrics is closer to the red line (the 1:1 line) ($R^2 = 0.713$, RMSE = 21.064 t/ha, MAE = 15.445 t/ha), compared to others results, followed by alpha-shape, height, stand, and intensity metrics. In the model of alpha-shape metrics based, the training $R^2 > 0.7$ and the testing R^2 are closed to 0.7. It shows that the performance is more stable than height, stand and intensity metrics.

As can be seen from Figure 7, there was an optimal relationship between measured AGB and alpha-shape metrics in stepwise linear regression, the same as random forest. The $R^2_{adj} > 0.6$ of the training model ($R^2_{adj} = 0.607$) and testing model ($R^2_{adj} = 0.711$) derived by alpha-shape variables, indicating that variables are more related to individual trees for the great estimation effect, followed by height, stand and intensity, but the conspicuous difference between RMSE and MAE in training and testing model. However, in the testing datasets, intensity metrics performed extremely badly.

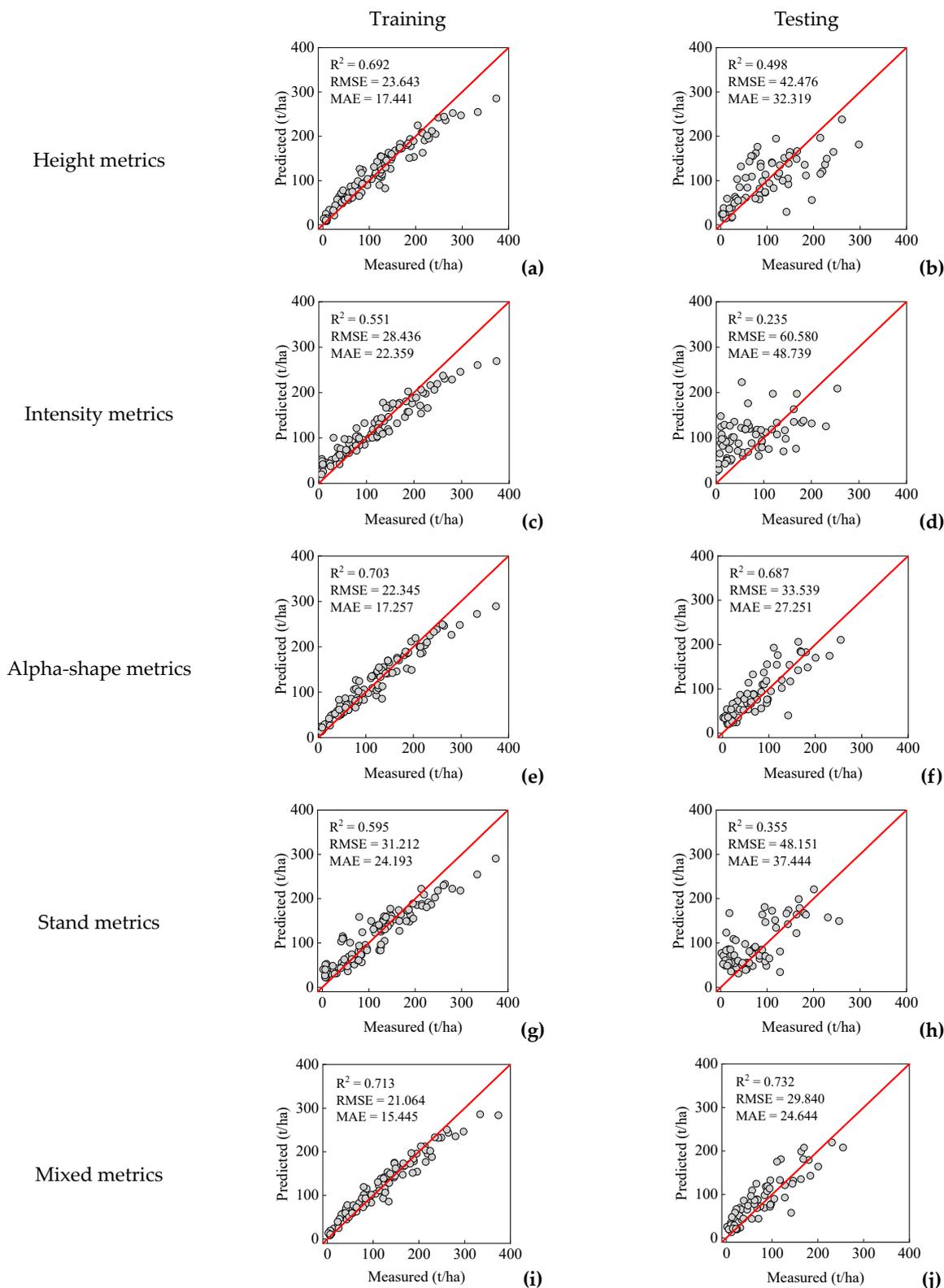


Figure 6. Scatterplot between measured and predicted AGB from the random forest model of each type of variables. (a) Training dataset of height metrics; (b) Testing dataset of height metrics; (c) Training dataset of intensity metrics; (d) Testing dataset of intensity metrics; (e) Training dataset of alpha-shape metrics; (f) Testing dataset of alpha-shape metrics; (g) Training dataset of stand metrics; (h) Testing dataset of stand metrics; (i) Training dataset of mixed metrics; (j) Testing dataset of mixed metrics.

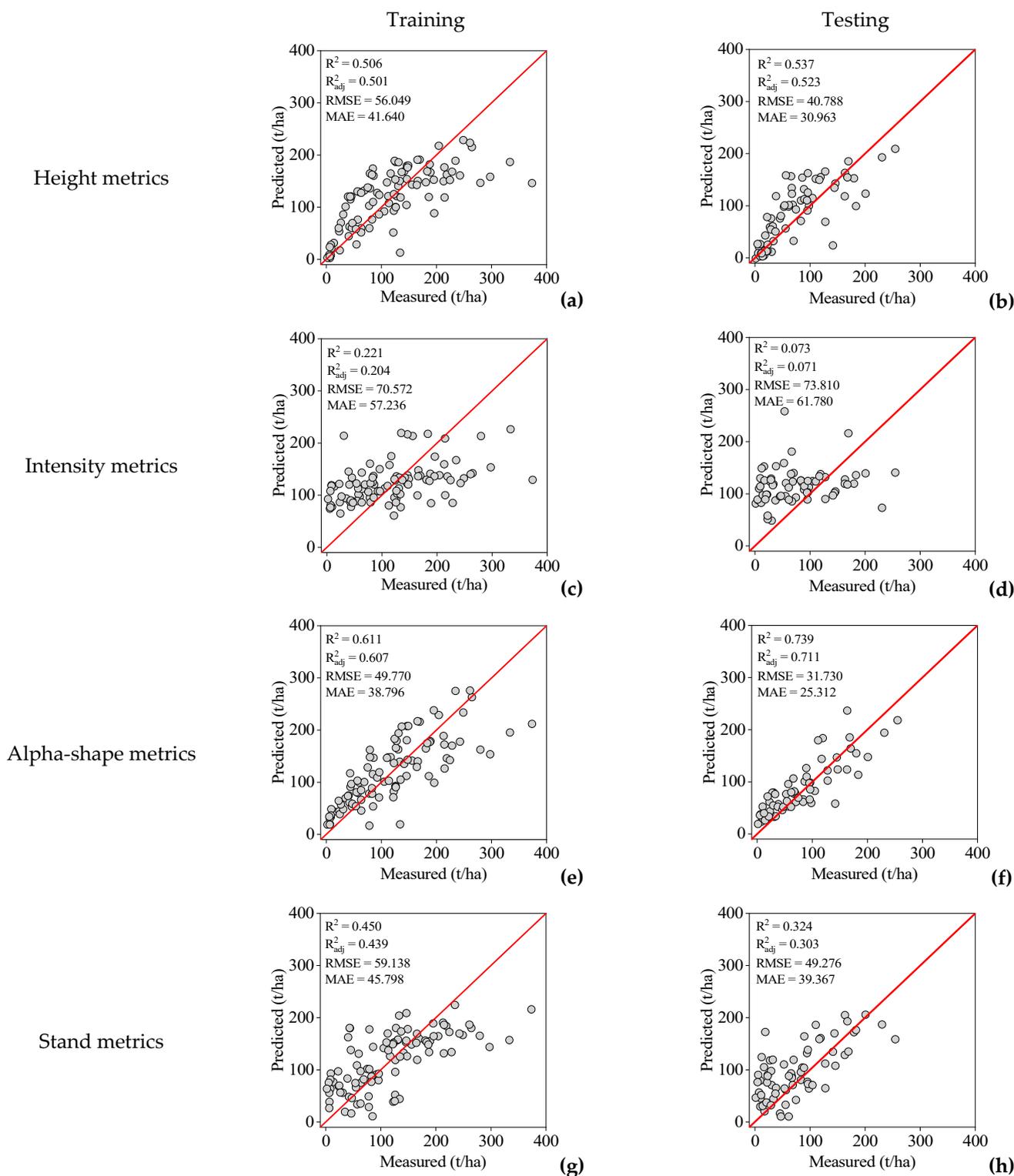


Figure 7. Scatterplot between measured and predicted AGB from the stepwise linear regression of each type of variables. (a) Training dataset of height metrics; (b) Testing dataset of height metrics; (c) Training dataset of intensity metrics; (d) Testing dataset of intensity metrics; (e) Training dataset of alpha-shape metrics; (f) Testing dataset of alpha-shape metrics; (g) Training dataset of stand metrics; (h) Testing dataset of stand metrics.

4. Discussion

4.1. The Important Roles of Variables from ALS for AGB Estimation at Stand Level

Figures 3 and 4 show the relationship between AGB and the variables extracted from ALS data. In stand metrics, *IncMSE* all variables are greater than 1000, indicating that tree density (*TD*), leaf area index (*LAI*), canopy cover (*CC*) and gap fraction (*GF*) are important to AGB. Tree density delineated not only the number of trees in the stand, but also an important indicator for structure of the ecological system and biological cycling [34,35]. Canopy cover is also an indicator of stand density the same as tree density. Furthermore, it is an important parameter in forest management and estimating forest volume [36–38]. Since the study area is in a tropical rainforest with dense trees, obvious vertical stratification and complex structure, the competition among tree crowns for light and living space is extremely strong, and the growth of the understory and forest floor are affected. Therefore, the biomass is also affected, and additionally, the leaf area index is intimately related to canopy and a feature of forest health; it can affect many physiological and physical processes in vegetation, such as photosynthesis, respiration, transpiration, and carbon cycling and precipitation retention [39,40]. The gap fraction can reflect the characteristics of canopy structure and spatial distribution of biomass [41,42].

Tree height as one of the most important quantitative forest observation parameters, especially in tropical rainforests, has a strong influence on tree growth, understory, and photosynthesis [11,43]. While it is important to summarize observations in terms of max, min, median, etc., it is more comprehensive when it is combined with percentiles [44]. The quantile is a measurement used to compute the position of the data, and the percentile provides information about the distribution of all data between the min and max, so the percentile step can be also used as an indicator to depict the degree of discrete distribution of data. The value of the percentile is fairly stable when there are many observations. Therefore, it is helpful to analyze stand height characteristics and aboveground biomass of the stand by combining elevation metrics of point cloud data and quantiles. Accumulative interpercentile height and height percentile performed better in AGB estimation.

The intensity reflects the intensity of laser pulse echoes, which are affected by the length of the laser path, scanning angle, distribution of branches and leaves, and the terrain in the forest [45–47]. Moreover, the intensity values are intimately related to tree species, and living and dead standing wood [48]. The median absolute deviation (*MAD*) is a robust measurement of sample differences in univariate datasets. *MAD* is a robust statistic, which is more flexible than standard deviation in dealing with outliers in data sets and can greatly reduce the impact of outliers on datasets [49,50]. Hence, the robust statistical metrics of intensity are benefit to vertical structure and AGB of forest stand.

4.2. Effects on Structure of Individual Trees Based on Alpha-Shape Analysis

The height, intensity and forest stand metrics delineated stand features at plot level; however, the principal part of aboveground biomass was made of individual trees on a plot and consequently, some characteristics were extracted by using the alpha-shape algorithm at tree-level, which is of benefit to describe the relationship between AGB and the variables extracted by the point cloud data; thereby we can make full use of point cloud data and take advantage of airborne laser scanning. We extracted surface area and volume of crown by using alpha-shape, which are important parameters describing the crown structure, responding the size and shape of crown directly and reflecting the photosynthetic ability and biomass accumulation indirectly [1].

The crown fullness ratio (*CFR*) is not only a quantitative index reflecting the crown size, but also indicating the growth of trees and the ability to occupy growth space [51,52]. Therefore, the size of the crown fullness ratio has a great influence on the growth of trees [53]. As a result of the laser scanning data acquisition and the high crown density in tropical rainforest, the segmented single tree point clouds cannot fully express details of the individual tree, and the alpha-shape polyhedron contains the tree height information. We use the ratio of average crown width to tree height as the evaluation factor of crown

fullness ratio. Furthermore, corresponding variables of the length of the Z axis (C_Z) also characterized tree height the same as height metrics. Therefore, the extraction of individual tree structure metrics based on alpha-shape after single tree point clouds segmentation is beneficial to estimating forest aboveground biomass.

4.3. Regression Models of AGB Estimation

We used two regression methods to estimate AGB. In both them, multiple stepwise linear regression is the parametric regression technique, which made clear the relationship between AGB and LiDAR derived metrics. Nevertheless, there were some assumptions for the data, such as sample independence, normal distribution, and equal variance [54,55]. On the other hand, the forest environment is an ecosystem that is too complex and non-linear to be depicted by linear regression. Besides, the individual wood is affected by site quality, stand competition, and other factor interactions, and it is difficult to express highly complex systems by using linear regression [56]. Moreover, there were possibilities of missing treetops during laser scanning in this study area, and the crown of each tree was less accurately depicted by the variables that extracted ALS data.

However, random forest is a common non-parametric regression method that makes no assumptions about the distribution of input data but also analyzes the relationship between AGB and LiDAR metrics well, deep mining valuable information of forest parameters and establishing a better regression model [57,58]. The RF model had a good inversion effect on the AGB in all kinds of metrics. Moreover, the random forest model was good at dealing with nonlinear regression and multicollinearity, such as parameters of the forestry environment in particular [56]. It is easy to improve performance and decrease overfitting. Additionally, it has some drawbacks, as the hyperparameters could not be adjusted well inside the RF model [59], therefore it may not perform well for other similar forests as it did in this study.

And the focus is not to discuss the regression method in this paper. In terms of regression strategies, more machine learning methods can be used and compared, for instance, support vector machine (SVM), artificial neural network (ANN), k-nearest neighbor (k-NN) etc., but the sample size and model generalization, which are also important problems to be considered. Figures 6 and 7 show that RF and MSLR models produced biases when $AGB > 200$ t/ha. The biases showed that the random effects of forestry environment for AGB. In addition, AGB has the characteristics of spatial autocorrelation and heterogeneous in forest [36,60,61]. The spatial effects will lead to the error of AGB estimation without considering the spatial autocorrelation and heterogeneity, but spatial regression model is a great method, the same as geographically weighted regression (GWR). In addition to considering the analyzed attributes and the weight of spatial distance, the spatial variance can be better characterized [62].

4.4. Other Factors for Characterizing AGB Estimation

Another advantage of ALS is the ability to generate a high spatial resolution digital elevation model (DEM). The stand characteristics were also affected by terrain, such as slope, aspect, relief, and curvature [50]. Moreover, topography has a significant impact on species' richness, hydrothermal conditions, and soil nutrient supply, especially in tropical areas [9,63,64]. In this paper, we did not consider topographic factors, or more importantly, the relationship between structure characteristics of stand or individual trees and aboveground biomass. The quality of generating DEM by ALS is affected by forest structure, off-nadir angle and interpolation algorithm [64,65]. Hence, high resolution DEM is used to extract topographic factors, which are used as independent variables for modeling in future research.

In terms of tree species, the shapes of conifer and broad-leaved trees species are quite different, which also produces different responses for laser scanning signals. Accuracy of the AGB estimation model can be improved by distinguishing tree species [66]. Due to the

limitation of conditions in situ, the number of sample plots for each dominant tree species is different, with merely a few sample plots for individual tree species.

5. Conclusions

In this research, the aim was to evaluate aboveground biomass of explanatory potential using different kinds of features extracted from airborne laser scanning data, such as the individual tree crown features that can play an important role in AGB estimation. The random forest model performed better than multiple stepwise linear regression. The multicollinearity and nonlinear were avoided, but also other variables related to AGB were chosen by random forest. The findings of this study suggest that the shape of the point cloud clusters representing tree crowns can be geometrically reconstructed by alpha-shape algorithm, which corresponds to some extracted features that are able to estimate AGB well. There was a fine relationship between AGB and the surface area, and the volume of crowns along with the crown shape ratio and tree height extracted from the alpha-shape polyhedron. In future work, the method may be extended to distinguish tree species mapping at stand level that could improve estimates of regional aboveground biomass.

Author Contributions: Conceptualization, Z.Y. and J.Q.; methodology, Z.Y., K.W. and J.Q.; software, F.W.; validation, S.W., J.Q. and H.H.; formal analysis, Z.Y., F.W., J.Q. and H.H.; investigation, S.W. and K.W.; resources, F.W.; data curation, C.L. and F.W.; writing—original draft preparation, C.L., Z.Y., S.W. and J.Q.; writing—review and editing, C.L., Z.Y. and J.Q.; visualization, Z.Y. and J.Q.; supervision, J.Q. and H.H.; project administration, C.L., J.Q. and H.H.; funding acquisition, Z.Y. and J.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This paper has been supported by National Natural Science Foundation of China Youth Fund, China (42001279).

Data Availability Statement: Data not available due to regulations restrictions.

Acknowledgments: This work was completed under the guidance of the Quantitative Remote Sensing of Forestry Team of Beijing Forestry University. We grateful acknowledge the experts who were involved in the aerial remote sensing data collection and field survey for this research project.

Conflicts of Interest: The authors declare no conflict of interest.

References

- West, P.W. *Tree and Forest Measurement*; Springer International Publishing: Cham, Switzerland, 2015; ISBN 978-3-319-14707-9.
- Wang, Q.; Pang, Y.; Chen, D.; Liang, X.; Lu, J. Lidar Biomass Index: A Novel Solution for Tree-Level Biomass Estimation Using 3D Crown Information. *For. Ecol. Manag.* **2021**, *499*, 119542. [[CrossRef](#)]
- Iida, S.; Shimizu, T.; Tamai, K.; Kabeya, N.; Shimizu, A.; Ito, E.; Ohnuki, Y.; Chann, S.; Levia, D.F. Evapotranspiration from the Understory of a Tropical Dry Deciduous Forest in Cambodia. *Agric. For. Meteorol.* **2020**, *295*, 108170. [[CrossRef](#)]
- Almeida, D.R.A.; Almeyda Zambrano, A.M.; Broadbent, E.N.; Wendt, A.L.; Foster, P.; Wilkinson, B.E.; Salk, C.; Papa, D.D.; Stark, S.C.; Valbuena, R.; et al. Detecting Successional Changes in Tropical Forest Structure Using GatorEye Drone-borne Lidar. *Biotropica* **2020**, *52*, 1155–1167. [[CrossRef](#)]
- Kaasalainen, S.; Holopainen, M.; Karjalainen, M.; Vastaranta, M.; Kankare, V.; Karila, K.; Osmanoglu, B. Combining Lidar and Synthetic Aperture Radar Data to Estimate Forest Biomass: Status and Prospects. *Forests* **2015**, *6*, 252–270. [[CrossRef](#)]
- Qin, L.; Liu, Q.; Zhang, M.; Saeed, S. Effect of Measurement Errors on the Estimation of Tree Biomass. *Can. J. For. Res.* **2019**, *49*, 1371–1378. [[CrossRef](#)]
- Qin, L.; Meng, S.; Zhou, G.; Liu, Q.; Xu, Z. Uncertainties in above Ground Tree Biomass Estimation. *J. For. Res.* **2021**, *32*, 1989–2000. [[CrossRef](#)]
- Gu, X.; Chen, B.; Yun, T.; Li, G.; Wu, Z.; Kou, W. Spatio-temporal Changes of Forest in Hainan Island from 2007 to 2018 Based on Multi-source Remote Sensing Data. *Chin. J. Trop. Crops* **2022**, *43*, 418–429.
- Maltamo, M.; Næsset, E.; Vauhkonen, J. *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*; Managing Forest Ecosystems; Springer: Dordrecht, The Netherlands, 2014; Volume 27, ISBN 978-94-017-8662-1.
- Jucker, T.; Hardwick, S.R.; Both, S.; Elias, D.M.O.; Ewers, R.M.; Milodowski, D.T.; Swinfield, T.; Coomes, D.A. Canopy Structure and Topography Jointly Constrain the Microclimate of Human-modified Tropical Landscapes. *Glob. Chang. Biol.* **2018**, *24*, 5243–5258. [[CrossRef](#)]

11. Reis, C.R.; Gorgens, E.B.; Almeida, D.R.A.D.; Celes, C.H.S.; Rosette, J.; Lima, A.; Higuchi, N.; Ometto, J.; Santana, R.C.; Rodriguez, L.C.E. Qualifying the Information Detected from Airborne Laser Scanning to Support Tropical Forest Management Operational Planning. *Forests* **2021**, *12*, 1724. [[CrossRef](#)]
12. Hilker, T.; van Leeuwen, M.; Coops, N.C.; Wulder, M.A.; Newnham, G.J.; Jupp, D.L.B.; Culvenor, D.S. Comparing Canopy Metrics Derived from Terrestrial and Airborne Laser Scanning in a Douglas-Fir Dominated Forest Stand. *Trees* **2010**, *24*, 819–832. [[CrossRef](#)]
13. Vaglio Laurin, G.; Puletti, N.; Chen, Q.; Corona, P.; Papale, D.; Valentini, R. Above Ground Biomass and Tree Species Richness Estimation with Airborne Lidar in Tropical Ghana Forests. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 371–379. [[CrossRef](#)]
14. Manuri, S.; Andersen, H.-E.; McGaughey, R.J.; Brack, C. 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. Geoinfor.* **2017**, *56*, 24–35. [[CrossRef](#)]
15. Knapp, N.; Fischer, R.; Cazcarra-Bes, V.; Huth, A. Structure Metrics to Generalize Biomass Estimation from Lidar across Forest Types from Different Continents. *Remote Sens. Environ.* **2020**, *237*, 111597. [[CrossRef](#)]
16. Jiang, X.; Li, G.; Lu, D.; Chen, E.; Wei, X. Stratification-Based Forest Aboveground Biomass Estimation in a Subtropical Region Using Airborne Lidar Data. *Remote Sens.* **2020**, *12*, 1101. [[CrossRef](#)]
17. de Oliveira, C.P.; Ferreira, R.L.C.; da Silva, J.A.A.; Lima, R.B.; Silva, E.A.; Silva, A.F.; Lucena, J.D.S.; dos Santos, N.A.T.; Lopes, I.J.C.; Pessoa, M.M.D.; et al. Modeling and Spatialization of Biomass and Carbon Stock Using LiDAR Metrics in Tropical Dry Forest, Brazil. *Forests* **2021**, *12*, 473. [[CrossRef](#)]
18. Zou, W.; Zeng, W.; Zhang, L.; Zeng, M. Modeling Crown Biomass for Four Pine Species in China. *Forests* **2015**, *6*, 433–449. [[CrossRef](#)]
19. Lin, M.; Ling, Q.; Pei, H.; Song, Y.; Qiu, Z.; Wang, C.; Liu, T.; Gong, W. Remote Sensing of Tropical Rainforest Biomass Changes in Hainan Island, China from 2003 to 2018. *Remote Sens.* **2021**, *13*, 1696. [[CrossRef](#)]
20. Zhu, Z.-X.; Harris, A.J.; Nizamani, M.M.; Thornhill, A.H.; Scherson, R.A.; Wang, H.-F. Spatial Phylogenetics of the Native Woody Plant Species in Hainan, China. *Ecol. Evol.* **2021**, *11*, 2100–2109. [[CrossRef](#)]
21. Liu, Y.; Wu, F.; Sun, Z.; Fu, A.; Gao, J.; Gao, X.; Gao, J.; Cui, C.; Chao, Z. Comprehensive Experiment Substitute for Multi-Payload Data of Terrestrial Ecosystem Carbon Inventory Satellite in Hainan. *For. Resour. Manag.* **2021**, *4*, 138–148. [[CrossRef](#)]
22. Li, W.; Guo, Q.; Jakubowski, M.K.; Kelly, M. A New Method for Segmenting Individual Trees from the Lidar Point Cloud. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 75–84. [[CrossRef](#)]
23. Luo, Y.; Wang, X.; Lu, F. *Comprehensive Database of Biomass Regressions for China's Tree Species*; China Forestry Publishing House: Beijing, China, 2015; ISBN 978-7-5038-8306-4.
24. Edelsbrunner, H.; Kirkpatrick, D.; Seidel, R. On the Shape of a Set of Points in the Plane. *IEEE Trans. Inf. Theory* **1983**, *29*, 551–559. [[CrossRef](#)]
25. Gardiner, J.D.; Behnsen, J.; Brassey, C.A. Alpha Shapes: Determining 3D Shape Complexity across Morphologically Diverse Structures. *BMC Evol. Biol.* **2018**, *18*, 184. [[CrossRef](#)]
26. Vauhkonen, J. Geometrically Explicit Description of Forest Canopy Based on 3D Triangulations of Airborne Laser Scanning Data. *Remote Sens. Environ.* **2016**, *173*, 248–257. [[CrossRef](#)]
27. Caselli, M.; Loguercio, G.A.; Urretavizcaya, M.F.; Defosse, G.E. Stand Level Volume Increment in Relation to Leaf Area Index of *Austrocedrus Chilensis* and *Nothofagus Dombeyi* Mixed Forests of Patagonia, Argentina. *For. Ecol. Manag.* **2021**, *494*, 119337. [[CrossRef](#)]
28. Ma, Q.; Su, Y.; Guo, Q. Comparison of Canopy Cover Estimations From Airborne LiDAR, Aerial Imagery, and Satellite Imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 4225–4236. [[CrossRef](#)]
29. Richardson, J.J.; Moskal, L.M.; Kim, S.-H. Modeling Approaches to Estimate Effective Leaf Area Index from Aerial Discrete-Return LiDAR. *Agric. For. Meteorol.* **2009**, *149*, 1152–1160. [[CrossRef](#)]
30. Korpela, I.; Ørka, H.; Maltamo, M.; Tokola, T.; Hyyppä, J. Tree Species Classification Using Airborne LiDAR—Effects of Stand and Tree Parameters, Downsizing of Training Set, Intensity Normalization, and Sensor Type. *Silva Fenn.* **2010**, *44*, 319–339. [[CrossRef](#)]
31. Li, X.; Lin, H.; Long, J.; Xu, X. Mapping the Growing Stem Volume of the Coniferous Plantations in North China Using Multispectral Data from Integrated GF-2 and Sentinel-2 Images and an Optimized Feature Variable Selection Method. *Remote Sens.* **2021**, *13*, 2740. [[CrossRef](#)]
32. Jiang, F.; Kutia, M.; Sarkissian, A.J.; Lin, H.; Long, J.; Sun, H.; Wang, G. Estimating the Growing Stem Volume of Coniferous Plantations Based on Random Forest Using an Optimized Variable Selection Method. *Sensors* **2020**, *20*, 7248. [[CrossRef](#)]
33. Han, H.; Wan, R.; Li, B. Estimating Forest Aboveground Biomass Using Gaofen-1 Images, Sentinel-1 Images, and Machine Learning Algorithms: A Case Study of the Dabie Mountain Region, China. *Remote Sens.* **2022**, *14*, 176. [[CrossRef](#)]
34. Sun, Y.; Li, Z.; He, H.; Guo, L.; Zhang, X.; Xin, Q. Counting Trees in a Subtropical Mega City Using the Instance Segmentation Method. *Int. J. Appl. Earth Obs. Geoinfor.* **2022**, *106*, 102662. [[CrossRef](#)]
35. Liu, D.; Zhou, C.; He, X.; Zhang, X.; Feng, L.; Zhang, H. The Effect of Stand Density, Biodiversity, and Spatial Structure on Stand Basal Area Increment in Natural Spruce-Fir-Broadleaf Mixed Forests. *Forests* **2022**, *13*, 162. [[CrossRef](#)]
36. Zhao, J.; Zhao, L.; Chen, E.; Li, Z.; Xu, K.; Ding, X. An Improved Generalized Hierarchical Estimation Framework with Geostatistics for Mapping Forest Parameters and Its Uncertainty: A Case Study of Forest Canopy Height. *Remote Sens.* **2022**, *14*, 568. [[CrossRef](#)]

37. Meng, P.; Wang, H.; Qin, S.; Li, X.; Song, Z.; Wang, Y.; Yang, Y.; Gao, J. Health Assessment of Plantations Based on LiDAR Canopy Spatial Structure Parameters. *Int. J. Digit. Earth* **2022**, *15*, 712–729. [[CrossRef](#)]
38. Li, S.; Hou, Z.; Ge, J.; Wang, T. Assessing the Effects of Large Herbivores on the Three-Dimensional Structure of Temperate Forests Using Terrestrial Laser Scanning. *For. Ecol. Manag.* **2022**, *507*, 119985. [[CrossRef](#)]
39. Yang, X.; Wang, C.; Pan, F.; Nie, S.; Xi, X.; Luo, S. Retrieving Leaf Area Index in Discontinuous Forest Using ICESat/GLAS Full-Waveform Data Based on Gap Fraction Model. *ISPRS J. Photogramm. Remote Sens.* **2019**, *148*, 54–62. [[CrossRef](#)]
40. de Almeida, D.R.A.; Stark, S.C.; Shao, G.; Schiatti, J.; Nelson, B.W.; Silva, C.A.; Gorgens, E.B.; Valbuena, R.; Papa, D.D.; Brancalion, P.H.S. Optimizing the Remote Detection of Tropical Rainforest Structure with Airborne Lidar: Leaf Area Profile Sensitivity to Pulse Density and Spatial Sampling. *Remote Sens.* **2019**, *11*, 92. [[CrossRef](#)]
41. Pascu, I.-S.; Dobre, A.-C.; Badea, O.; Andrei Tanase, M. Estimating Forest Stand Structure Attributes from Terrestrial Laser Scans. *Sci. Total Environ.* **2019**, *691*, 205–215. [[CrossRef](#)]
42. Fisher, A.; Armston, J.; Goodwin, N.; Scarth, P. Modelling Canopy Gap Probability, Foliage Projective Cover and Crown Projective Cover from Airborne Lidar Metrics in Australian Forests and Woodlands. *Remote Sens. Environ.* **2020**, *237*, 111520. [[CrossRef](#)]
43. Gadov, K.V.; Hui, G. Modelling Forest Development. *For. Sci.* **1999**, *57*, 1146–1158. [[CrossRef](#)]
44. Robinson, A.P.; Hamann, J.D. *Forest Analytics with R: An Introduction*, 2011th ed.; Springer: New York, NY, USA, 2010; ISBN 978-1-4419-7761-8.
45. You, H. Inversion Study of Forest Structural Parameters Based on Footprint LiDAR Data. Master's Thesis, Northeast Forestry University, Harbin, China, 2014.
46. Laslier, M.; Hubert-Moy, L.; Dufour, S. Mapping Riparian Vegetation Functions Using 3D Bispectral LiDAR Data. *Water* **2019**, *11*, 483. [[CrossRef](#)]
47. Shang, C.; Treitz, P.; Caspersen, J.; Jones, T. Estimation of Forest Structural and Compositional Variables Using ALS Data and Multi-Seasonal Satellite Imagery. *Int. J. Appl. Earth Obs. Geoinform.* **2019**, *78*, 360–371. [[CrossRef](#)]
48. Kim, Y.; Yang, Z.; Cohen, W.B.; Pflugmacher, D.; Lauver, C.L.; Vankat, J.L. 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.* **2009**, *113*, 2499–2510. [[CrossRef](#)]
49. Legendre, P.; Legendre, L. *Numerical Ecology*, 3rd ed.; Developments in Environmental Modeling 24; Elsevier: Amsterdam, The Netherlands, 2012; ISBN 978-0-444-53868-0.
50. Dong, P.; Chen, Q. *LiDAR Remote Sensing and Applications*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2017.
51. Comfort, E.J.; Roberts, S.D.; Harrington, C.A. Midcanopy Growth Following Thinning in Young-Growth Conifer Forests on the Olympic Peninsula Western Washington. *For. Ecol. Manag.* **2010**, *259*, 1606–1614. [[CrossRef](#)]
52. Bahadur, R.B.K.; Sharma, R.P.; Bhandari, S.K. A generalized aboveground biomass model for juvenile individuals of rhododendron *Arboreum* (SM.) in Nepal. *Cerne* **2019**, *25*, 119–130. [[CrossRef](#)]
53. Von Gadov, K.; Hui, G. *Modelling Forest Development*, 1st ed.; Forestry Sciences 57; Springer: Dordrecht, The Netherlands, 1999; ISBN 978-1-4020-0276-2.
54. Wu, C.; Shen, H.; Shen, A.; Deng, J.; Gan, M.; Zhu, J.; Xu, H.; Wang, K. Comparison of Machine-Learning Methods for above-Ground Biomass Estimation Based on Landsat Imagery. *J. Appl. Remote Sens.* **2016**, *10*, 035010. [[CrossRef](#)]
55. Pham, T.D.; Yokoya, N.; Xia, J.; Ha, N.T.; Le, N.N.; Nguyen, T.T.T.; Dao, T.H.; Vu, T.T.P.; Pham, T.D.; Takeuchi, W. Comparison of Machine Learning Methods for Estimating Mangrove Above-Ground Biomass Using Multiple Source Remote Sensing Data in the Red River Delta Biosphere Reserve, Vietnam. *Remote Sens.* **2020**, *12*, 1334. [[CrossRef](#)]
56. Lei, X. Applications of machine learning algorithms in forest growth and yield prediction. *J. Beijing For. Univ.* **2019**, *41*, 23–36.
57. García-Gutiérrez, J.; Martínez-Álvarez, F.; Troncoso, A.; Riquelme, J.C. A Comparison of Machine Learning Regression Techniques for LiDAR-Derived Estimation of Forest Variables. *Neurocomputing* **2015**, *167*, 24–31. [[CrossRef](#)]
58. Yu, X.; Hyypä, J.; Vastaranta, M.; Holopainen, M.; Viitala, R. Predicting Individual Tree Attributes from Airborne Laser Point Clouds Based on the Random Forests Technique. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 28–37. [[CrossRef](#)]
59. Sun, Z.; Gao, J.; Wu, F.; Gao, X.; Hu, Y.; Gao, J. Estimating Forest Stock Volume via Small-Footprint LiDAR Point Cloud Data and Random Forest Algorithm. *Sci. Silvae Sin.* **2021**, *57*, 68–81.
60. Pringle, M.J.; Schmidt, M.; Tindall, D.R. Multi-Decade, Multi-Sensor Time-Series Modelling—Based on Geostatistical Concepts—to Predict Broad Groups of Crops. *Remote Sens. Environ.* **2018**, *216*, 183–200. [[CrossRef](#)]
61. Zhang, T.; Lin, H.; Long, J.; Zhang, M.; Liu, Z. Analyzing the Saturation of Growing Stem Volume Based on ZY-3 Stereo and Multispectral Images in Planted Coniferous Forest. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 50–61. [[CrossRef](#)]
62. Ou, G.; Lv, Y.; Xu, H.; Wang, G. Improving Forest Aboveground Biomass Estimation of *Pinus densata* Forest in Yunnan of Southwest China by Spatial Regression Using Landsat 8 Images. *Remote Sens.* **2019**, *11*, 2750. [[CrossRef](#)]
63. Banasiak, P.Z.; Berezowski, P.L.; Zaplata, R.; Mielcarek, M.; Duraj, K.; Sterenczak, K. Semantic Segmentation (U-Net) of Archaeological Features in Airborne Laser Scanning—Example of the Bialowieza Forest. *Remote Sens.* **2022**, *14*, 995. [[CrossRef](#)]
64. Cao, L.; Zhang, Z.; Yun, T.; Wang, G.; Ruan, H.; She, G. Estimating Tree Volume Distributions in Subtropical Forests Using Airborne LiDAR Data. *Remote Sens.* **2019**, *11*, 97. [[CrossRef](#)]

-
65. Balenovic, I.; Gasparovic, M.; Milas, A.S.; Berta, A.; Seletkovic, A. Accuracy Assessment of Digital Terrain Models of Lowland Pedunculate Oak Forests Derived from Airborne Laser Scanning and Photogrammetry. *Croat. J. For. Eng.* **2018**, *39*, 117–128.
 66. Gao, L.; Zhang, X. Above-Ground Biomass Estimation of Plantation with Complex Forest Stand Structure Using Multiple Features from Airborne Laser Scanning Point Cloud Data. *Forests* **2021**, *12*, 1713. [[CrossRef](#)]