

Integration of UAV-LiDAR and Sentinel-2A Data for Modeling Robinia pseudoacacia Aboveground Biomass via Artificial Neural Networks

Jiaqi Cheng

School of Soil and Water Conservation, Beijing Forestry University

Xuexia Zhang

zhangxuexia@bjfu.edu.cn

School of Soil and Water Conservation, Beijing Forestry University

Jianjun Zhang

School of Soil and Water Conservation, Beijing Forestry University

Yawei Hu

School of Soil and Water Conservation, Beijing Forestry University

Jiongchang Zhao

School of Soil and Water Conservation, Beijing Forestry University

Yang Li

School of Soil and Water Conservation, Beijing Forestry University

Research Article

Keywords: Robinia pseudoacacia, Sentinel-2A image, UAV-LiDAR, artificial neural networks, aboveground biomass, remote sensing

Posted Date: October 22nd, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-7432796/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

1 Integration of UAV-LiDAR and Sentinel-2A 2 Data for Modeling Robinia pseudoacacia 3 Aboveground Biomass via Artificial Neural 4 Networks

5 Jiaqi Cheng¹, Xuexia Zhang^{1,2,*}, Jianjun Zhang^{1,2,3}, Yawei Hu¹, Jiongchang Zhao¹, Yang
6 Li¹

7 (¹School of Soil and Water Conservation, Beijing Forestry University, Beijing 100083,
8 China; ²Shanxi Jixian County Forest Ecosystem National Field Scientific Observation
9 and Research Station, Linfen 041000, Shanxi, China; ³State Forestry Administration
10 Key Laboratory of Soil and Water Conservation, Beijing 100083, China)

11 Corresponding author: Xuexia Zhang (zhangxuexia@bjfu.edu.cn)

12 Abstract

13 Aboveground forest biomass (AGB), a critical ecosystem component, plays
14 essential roles in maintaining ecological balance, conserving biodiversity, mitigating
15 climate change, and supporting sustainable development. Optical remote sensing
16 characterizes horizontal forest structures through multispectral and textural data,
17 whereas UAV-LiDAR captures vertical structural parameters. Integrating these
18 complementary technologies improves biomass estimation by enhancing model fit and
19 accuracy. This multimodal approach significantly advances aboveground biomass
20 quantification.

21 This study investigated aboveground biomass (AGB) estimation in artificial
22 Robinia pseudoacacia forests within the Caijiachuan watershed of Jixian County,
23 Shanxi Province, utilizing multisource remote sensing data. Field survey data served as
24 ground truths, whereas Sentinel-2A multispectral imagery and UAV-LiDAR point
25 clouds provided optical and radar remote sensing data, respectively. After extracting

feature variables from both datasets, we developed and compared AGB estimation models via four machine learning approaches—partial least squares regression (PLS), random forest regression (RF), backpropagation neural network (BPNN), and convolutional neural network (CNN)—to evaluate each method's performance with single-source and combined data inputs. The optimal model was subsequently applied to generate a spatial distribution map of *Robinia pseudoacacia* AGB density across the watershed, followed by analysis of biomass distribution patterns relative to topographic factors (elevation, slope, and aspect).

Our results provide important insights as follows:(1) Compared with the single-variable model, the model incorporating both variables demonstrates superior performance (MAE = 2.53, MBE = -0.31, RMSE = 3.96, NMSE = 0.06) (MAE = 4.57, MBE = -2.92, RMSE = 5.61, NMSE = 0.11).(2) Among the four modeling approaches evaluated—partial least squares (PLS) regression, random forest (RF) regression, backpropagation neural network (BPNN), and convolutional neural network (CNN)—the CNN algorithm demonstrated superior performance in reducing estimation errors for *Robinia pseudoacacia* aboveground biomass.(3) The mean aboveground biomass (AGB) density of *Robinia pseudoacacia* in the Caijiachuan plantation watershed was 61.49 Mg·ha⁻¹. Approximately 87% of the total biomass occurred within the 1,050--1,200 m elevation range. Slope gradient analysis revealed maximum AGB accumulation on moderately steep and steep slopes, whereas minimum values were observed on gentle slopes. The aspect-dependent distribution revealed peak biomass on

semii-shaded northeast-facing slopes, whereas the lowest biomass levels occurred on north-facing shaded slopes.

These findings provide an effective method for assessing the aboveground biomass of *Robinia pseudoacacia* forests in the region, as well as valuable data for evaluating the carbon sequestration capacity of soil and water conservation forests.

Keywords: *Robinia pseudoacacia*; Sentinel-2A image; UAV-LiDAR; artificial neural networks; aboveground biomass; remote sensing

Introduction

Forest aboveground biomass (AGB) serves as a critical component of global carbon cycling and climate regulation while simultaneously supporting ecosystem stability and biodiversity conservation^[1,2,3]. The development of accurate and efficient AGB estimation methodologies has significant implications for carbon cycle quantification, ecosystem health assessment, and sustainable forest management^[4,5]. In addition to accelerating global forest depletion and environmental change, precise AGB monitoring has become a critical research priority for ecological conservation^[6,7].

Traditional AGB estimation relies on field plot surveys, where sample plots are established to measure tree parameters (e.g., DBH, height, and species), and biomass is calculated via species-specific allometric equations. A common approach for developing these equations is the standard tree method, where representative sample trees are selected and felled, and their component biomass is measured through weighing^[10,11,12]. While this method provides relatively accurate estimates, it is labor

intensive and time consuming for large-scale applications and has limitations when applied to complex ecosystems with diverse vegetation types¹³.

The rapid advancement of remote sensing technologies, particularly high-resolution sensors (e.g., satellites, UAVs, and LiDAR), has enabled researchers to develop more sophisticated monitoring approaches^[14,15]. Optical remote sensing imagery, including high-resolution satellite and UAV-based images, provides critical vegetation spatial distribution and structural information^[16,17]. Spectral vegetation indices (e.g., the normalized difference vegetation index (NDVI)), spectral features, and texture metrics extracted from optical data can effectively support biomass estimation modeling^[18,19,20]. Regression models leveraging the strong correlation between vegetation indices and forest aboveground biomass (AGB) facilitate upscaling from plot-level to regional-scale assessments, thereby enabling comprehensive ecosystem monitoring^[21,22,23]. However, these methods face vegetation index saturation in dense canopies or high leaf area index (LAI) conditions, which weakens the AGB–vegetation index relationship and consequently reduces estimation accuracy^[24,25,26,27].

LiDAR technology, as an active remote sensing method, generates high-precision 3D point cloud data that provide detailed information about surface and vegetation structures^[28,29]. The point clouds enable accurate measurements of ground elevation, vegetation height distribution, and canopy volume parameters^[30]. LiDAR vertical detection enables superior AGB estimation in multilayered vegetation^[31]. However, this technology faces vegetation occlusion challenges in dense canopies where ground

surface detection becomes obstructed, limiting its effectiveness in complex forest environments^[32].

Multisource data fusion approaches, particularly the integration of optical imagery with LiDAR data, can significantly increase the accuracy and reliability of forest aboveground biomass (AGB) estimation by leveraging their complementary advantages^[33]. This synergistic combination overcomes individual limitations by jointly considering the optical characteristics and 3D structural features of vegetation. The integrated approach has distinct advantages, where optical data provide rich spectral information while LiDAR delivers precise vertical structural parameters, collectively improving estimation comprehensiveness and precision, especially in topographically complex and structurally heterogeneous forest environments^[34,35,36]. However, challenges, including complex data processing, high equipment costs, and demanding technical requirements, remain obstacles for widespread implementation. The development of solutions to effectively harness both datasets' complementary information represents a critical pathway for advancing AGB estimation methodologies^[37,38].

The Loess Plateau, a critical component of China's "Three Zones and Four Belts" ecological security barrier system, plays a vital role in national ecological security. Located in the western Shanxi loess region, the Caijiachuan watershed in Jixian County has undergone 30 years of afforestation and natural restoration, significantly improving local ecological conditions. This study provides an effective methodology for

estimating *Robinia pseudoacacia* aboveground biomass in the Caijiachuan watershed,
offering crucial data support for assessing carbon sequestration capacity in soil–water
conservation forests.

Materials and methods

Study area

The study area (Fig. 1) encompasses the Caijiachuan River Basin (110° 39' 45" – 110° 47' 45" E, 36° 14' 27" – 36° 18' 23" N) in Shanxi Province, China, which hosts the National Field Scientific Observation and Research Station for Forest Ecosystems. This 39.33 km² watershed features an east-high, west-low topography with elevations ranging from 900–1513 m. Characterized by a warm temperate continental climate (mean annual temperature: 10.2°C; precipitation: 575.9 mm, with 59.5% occurring from July -September), the basin contains slightly alkaline loess-derived brown soils. Since the 1970s, afforestation efforts have resulted in forest cover exceeding 80%, with vegetation varying longitudinally: natural secondary forests (*Populus davidiana*, *Quercus mongolica*) dominate upstream, artificial plantations (*Robinia pseudoacacia*, *Pinus tabulaeformis*) occupy mid-reaches, and lower areas comprise farmland/wasteland. The focal 2.06 km² research plot, established in the 1990s with *Robinia pseudoacacia* plantations, lies within this ecologically diverse basin and supports 194 woody species, 180 herbs, and 141 medicinal plants.

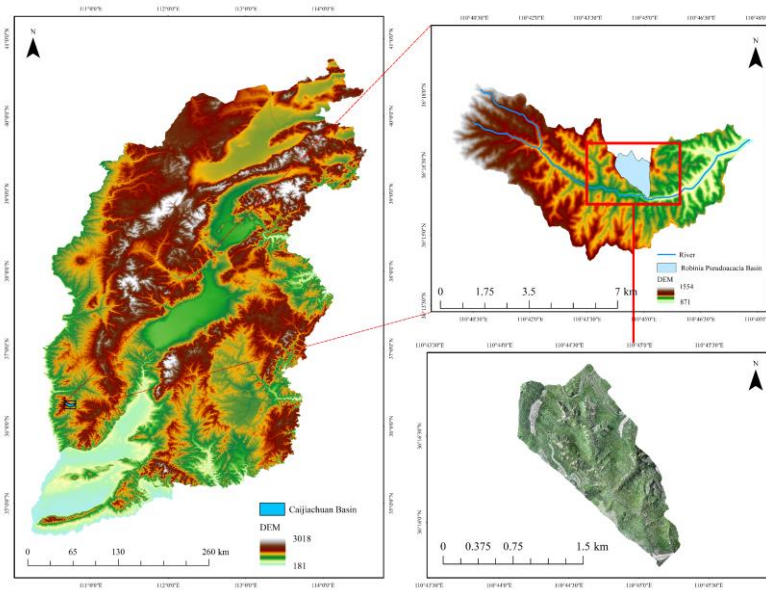


Figure 1 Study area

Field investigation

A field survey was conducted in July 2022 across the Caijiachuan artificial Robinia pseudoacacia forest watershed, involving the establishment of 45 representative 20 m×20 m sample plots. The sampling design systematically covered various stand density gradients, with each plot carefully selected to represent the local Robinia pseudoacacia distribution characteristics. The plot coordinates were precisely georeferenced via Qianxun Starfire RTK technology. The comprehensive dendrometric measurements included diameter at breast height (DBH; measured to ± 0.1 cm precision via calipers), tree height (± 0.1 m precision via telescopic poles), and crown dimensions for all Robinia pseudoacacia individuals within each plot. On the basis of complete tree census data, plot-level mean DBH and height values were calculated to identify representative sample trees for destructive sampling, which involved complete tree harvesting followed by component separation and weighing for biomass determination.

146

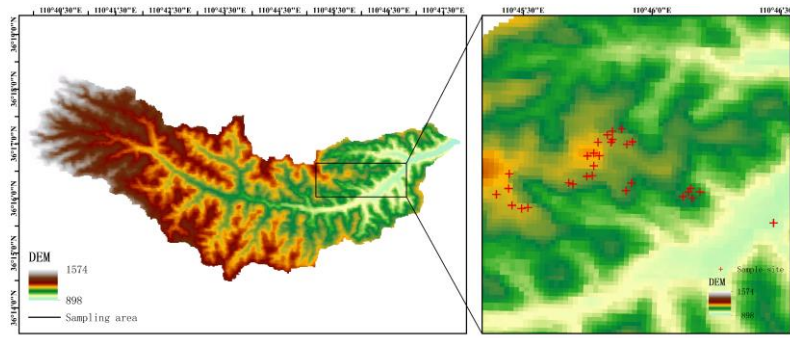


Figure 2 Sample site

147

148

149

150 UAV-LiDAR data and processing

151 UAV-LiDAR data acquisition was conducted synchronously with field surveys via

152 a DJI M300 RTK drone equipped with a Zenmuse L1 sensor. The flight parameters

153 were configured as follows: mission time window 10:00–14:00 local time, terrain-

154 following flight mode at 80 m altitude with 10 m/s velocity, 80% side overlap, and

155 average point density of 147 pts/m². The collected dataset comprised 56.14 GB of raw

156 LiDAR data (including aerial imagery), with 43.4 GB of UAV-LiDAR point clouds and

157 2,046 orthorectified images covering the entire study area.

158 The raw UAV-LiDAR data were preprocessed in DJI Terra and exported as .las

159 files. Subsequent processing in MATLAB involves the following steps: point cloud

160 denoising to remove outliers and enhance data quality; ground point classification via

161 an improved progressive TIN densification filtering algorithm; generation of 0.5 m-

162 resolution digital elevation models (DEMs) and digital surface models (DSMs) via

163 ordinary kriging interpolation of the ground and first returns, respectively; and height

164 normalization via subtraction of DEM values from each point's elevation^[17,39,40].

165 LiDAR metrics were extracted from the normalized point cloud (canopy height model),

as presented in Table 1.

Table 1 UAV-LiDAR Metrics Summary

Metrics	Description
Height variables	Percentile heights(h ₂₅ , h ₅₀ , h ₇₅ , h ₉₅) Percentiles of the canopy height distribution from the first return above 2 m (25th, 50th, 75th, 95th) The maximum/average/median/minimum height above ground of all points in a specific cell size
Density variables	Percentile densities(den ₀ , den ₁ , den ₂ , den ₃ , den ₄ , den ₅ , den ₆ , den ₇ , den ₈ , den ₉) Percentiles of the canopy height distribution from the first return above 2 m
Intensity variables	Percentile intersities(int ₂₅ , int ₅₀ , int ₇₅ , int ₉₅) Percentiles of the canopy height distribution from the first return above 2 m (int ₂₅ , int ₅₀ , int ₇₅ , int ₉₅) The maximum/average/median/minimum height above ground of all points in a specific cell size
Canopy cover(CC)	$CC = \frac{n_{veg}}{n_{total}}$
Gap fraction(GF)	$GF = \frac{n_{ground}}{n}$
Leaf area index(LAI)	$LAI = -\frac{\cos(ang) \times \ln(GF)}{k}$ $ang = \frac{\sum_{i=1}^n angle_i}{n}$

Sentinel-2A data and processing

This study utilized Sentinel-2A imagery (acquired from <https://earthexplorer.usgs.gov/>) with minimal cloud cover (0–5%) captured on July 24,

2022 (Tile: T49SDA), which coincided with the field survey period. The specific dataset (S2A_MSIL2A_20220724T031531. ZIP) was processed through the following workflow: initial downloading of multispectral bands at native resolutions, followed by resampling to uniform 20 m×20 m pixel matching plot dimensions; format conversion from native. SAFE to ENVI format using SNAP Desktop 8.0, including projection restoration to address coordinate system loss during conversion; spectral band extraction and composition (Bands 2, 3, 4, 6, 7, and 8A) for subsequent analysis; and final clipping using the study area boundary vector. The metrics extracted from the Sentinel-2A data are presented in Table 2.

Table 2 Sentinel-2A Metrics Summary

Metrics	Description	
Band information	Band2, Band3, Band4, Band6, Band7, Band8A	Extracted values of each band
Vegetation index	NDVI	$NDVI = \frac{NIR - R}{NIR + R}$
	SAVI	$SAVI = \frac{NIR - R}{(NIR + R + L)} (1 + L), L = 0.5$
	DVI	$DVI = NIR - R$
	RVI	$RVI = \frac{NIR}{R}$
Texture information	Mean	$\sum_{i,j=0}^{N-1} iP_{ij}$
	Variance	$\sum_{i,j=0}^{N-1} P_{ij}(i - ME)^2$
	Correlation	$\sum_{i,j=0}^{N-1} iP_{ij} \left \frac{(i - ME)(j - ME)}{\sqrt{VA_i VA_j}} \right $

Contrast	$\sum_{i,j=0}^{N-1} P_{ij}(i-j)^2$
Homogeneity	$\sum_{i,j=0}^{N-1} i \frac{P_{ij}}{1+(i-j)^2}$
Dissimilarity	$\sum_{i,j=0}^{N-1} i P_{ij} i-j $
Entropy	$\sum_{i,j=0}^{N-1} i P_{ij} (-\ln P_{ij})$
Second_Moment	$\sum_{i,j=0}^{N-1} i P_{ij}^2$

Assistance data

The study incorporated multiple geospatial datasets: orthomosaic imagery generated from UAV photogrammetric stitching; a 0.5 m-resolution DEM derived through kriging interpolation of classified UAV-LiDAR ground points, categorized into six elevation classes (≤ 1000 m, 1000–1050 m, 1050–1100 m, 1100–1150 m, 1150–1200 m, and ≥ 1200 m); slope data calculated from the DEM and classified into six gradients (0–5° flat, 6–15° gentle, 16–25° moderate, 26–35° steep, 36–45° abrupt, and $\geq 46^\circ$ extreme); aspect data classified into four sectors (shady, semishady, sunny, and semisunny); and field-collected plot data with watershed boundaries provided by the research team.

Methods

This study focused on *Robinia pseudoacacia* plantations in the Caijiachuan watershed of Shanxi Province and integrated multisource remote sensing data for

aboveground biomass (AGB) estimation. Field survey data served as ground references, whereas Sentinel-2A multispectral imagery and LiDAR point clouds provided optical and structural datasets, respectively. The feature variables extracted from both sources were used to develop AGB estimation models through four machine learning approaches: partial least squares regression (PLSR), random forest regression (RF), a backpropagation neural network (BPNN), and a convolutional neural network (CNN). Model performance was systematically evaluated across single-source and fused-data scenarios. The optimal model was applied to generate a spatial distribution map of AGB density, with subsequent analysis of biomass patterns across elevation gradients, slope classes, and aspect categories. Figure 3 is a technical roadmap.

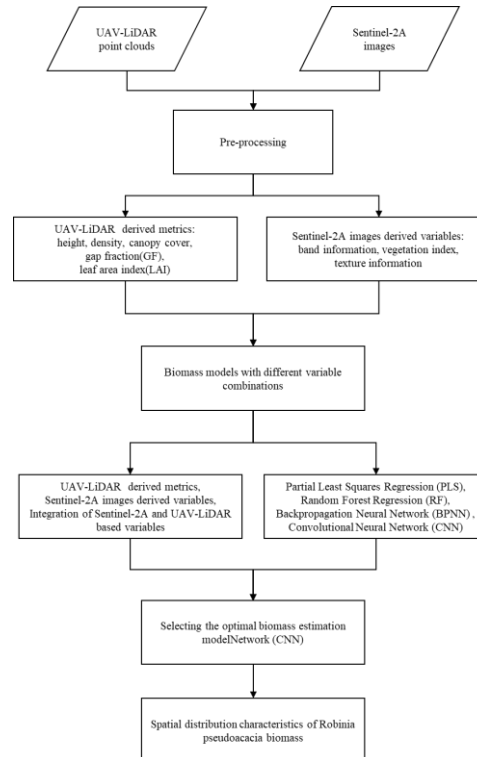


Figure 3 Technical roadmap

212 Biomass estimation modeling and accuracy assessment

213 1. Coefficient of determination (R^2)

214 The coefficient of determination (R^2) is a statistical measure for evaluating the
215 goodness-of-fit in regression models and is calculated as follows:

$$216 \quad R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

217 where n =sample size, y_i =observed value, $\hat{y}_i$ =predicted value, and $\bar{y}$ =mean of
218 observed values.

219 2. Pearson correlation coefficient (r)

220 The Pearson correlation coefficient is a statistical measure used to quantify the
221 strength and direction of the linear relationship between two continuous variables. It
222 evaluates the degree of linear association between variables and is calculated as follows:

$$r = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$

223 where $Cov(X, Y)$ = covariance between variables X and Y and where $\sigma_X \sigma_Y$ =
224 standard deviations of X and Y , respectively.

225 3. Mean absolute error (MAE)

226 The MAE is a metric used to measure the average absolute difference between the
227 predicted and actual values. Its calculation formula is as follows:

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

228 where MAE = mean absolute error, n = sample size, y_i = observed value, and $\hat{y}_i$
229 =predicted value.

230 4. Mean bias error (MBE)

231 The MBE is a metric used to quantify systematic bias in model predictions and
232 represents the average difference between the predicted and observed values. The
233 formula is as follows:

$$MBE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)$$

234 where MBE = mean bias error, n = sample size, y_i = observed value, and $\hat{y}_i$ =
235 model-predicted value.

236 5. Root mean square error (RMSE)

237 The RMSE is a metric used to evaluate model prediction accuracy by measuring
238 the square root of the average squared differences between the predicted and observed
239 values. The formula is as follows:

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

240 where RMSE = root mean square error, n = sample size, y_i = observed value, and
241 $\hat{y}_i$ = model-predicted value.

242 6. Normalized mean square error (NMSE)

243 The NMSE is a standardized metric that quantifies the model prediction error
244 relative to the variance of the target variable. The formula is as follows:

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

245 where NMSE = normalized mean square error, n = sample size, y_i = observed
246 value, and $\hat{y}_i$ = model-predicted value.

Results

Modeling biomass with UAV-LiDAR data

This study extracted 29 UAV-LiDAR-derived feature variables, which were categorized into 8 height metrics, 10 density measures, 8 intensity parameters, canopy cover, gap fraction, and leaf area index (LAI). Pearson correlation analysis (Fig. 4) was used to quantify the relationships between these features and the field-measured AGB, establishing critical predictors for the estimation models.

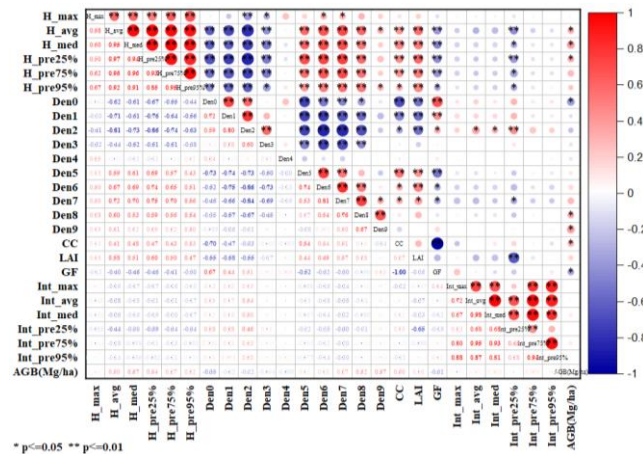


Figure 4 Correlations between UAV-LiDAR metrics and Robinia pseudoacacia AGB

Figure 4 shows significant positive correlations between AGB and six UAV-LiDAR-derived features: mean height (H_avg), median height (H_med), 25th percentile height (H_pre25%), density metrics (Den8, Den9), and canopy cover (CC). Conversely, negative correlations were observed with the ground return density (Den0) and gap fraction (GF). Notably, substantial multicollinearity existed among multiple feature variables.

Feature importance ranking was performed via random forest analysis, with the

results visualized in Figure 5. Figure 5 identifies Den0 as the most significant predictor for Robinia pseudoacacia aboveground biomass (AGB) estimation, followed by H_med, Den9, and Den2. To address variable multicollinearity, we integrated Pearson correlation and random forest importance results to select optimal features: height metrics (H_avg, H_med, H_25%), density measures (Den0, Den8, Den9), and canopy parameters (GF, CC) for the final AGB estimation model.

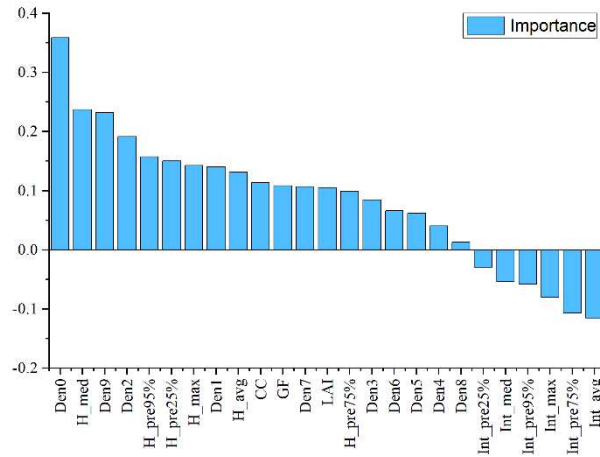


Figure 5 Importance ranking of LiDAR variables

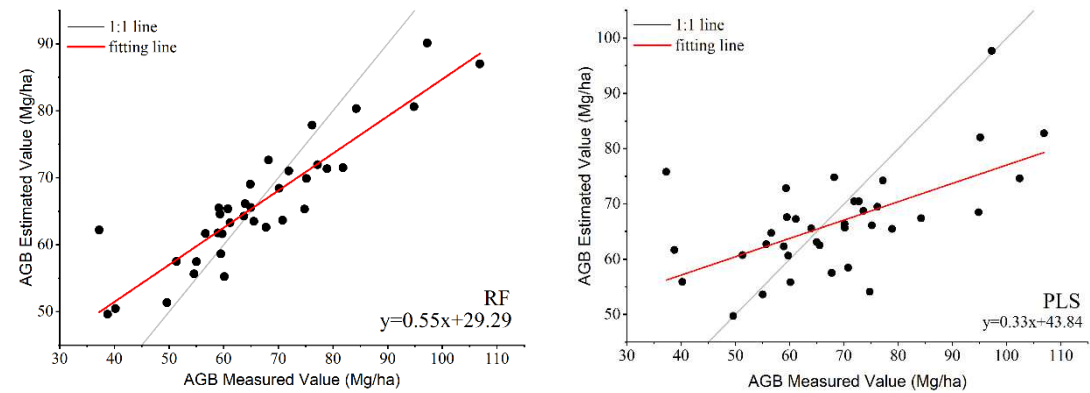
The study employed eight selected feature variables demonstrating the strongest influence on Robinia pseudoacacia aboveground biomass (AGB) to develop estimation models via four machine learning approaches: partial least squares (PLS) regression, random forest (RF) regression, a backpropagation neural network (BPNN), and a convolutional neural network (CNN). The dataset was randomly partitioned into 80% training and 20% testing subsets, with model performance metrics detailed in Table 3. The model training results demonstrated that both the RF regression and the CNN achieved reliable performance, with the explained variance exceeding 80%. The CNN

model achieved the highest accuracy ($R^2=0.85$; $r=0.92$; $MAE=4.59$; $MBE=-0.38$; $RMSE=12.02$; $NMSE=0.17$), outperforming PLS regression, which performed the poorest. While BPNN slightly surpassed PLS in accuracy, RF regression ranked second only to CNN in predictive performance.

Table 3 Model Accuracy Analysis

AGB Model	R^2	Pearson's r	MAE	MBE	RMSE	NMSE
RF	0.81	0.91	5.85	-0.20	7.84	0.27
PLS	0.36	0.60	10.10	-1.98	13.53	0.65
BP	0.49	0.70	9.54	-0.78	6.32	0.51
CNN	0.85	0.92	4.59	-0.38	12.02	0.17

Figure 6 presents scatterplots comparing the measured versus estimated Robinia pseudoacacia aboveground biomass (AGB) across all the models. The plots demonstrate the superior performance of the RF and CNN models, which have the strongest linear relationships ($R^2>0.80$), and systematic estimation biases, i.e., overestimation for $AGB<63 \text{ Mg}\cdot\text{ha}^{-1}$ and underestimation for $AGB>63 \text{ Mg}\cdot\text{ha}^{-1}$, indicating potential nonlinearity in high-biomass stands.



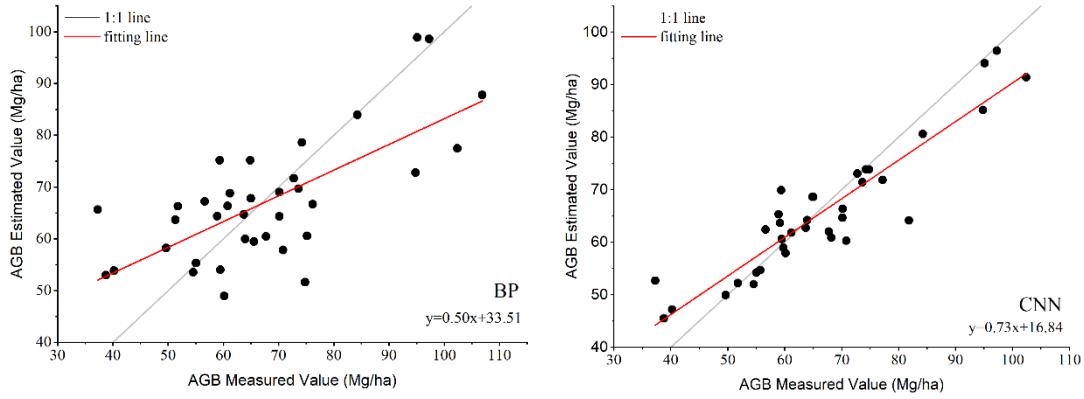


Figure 6 Measured vs. Estimated AGB of *Robinia pseudoacacia*

Analysis of Table 3 and Figure 6 reveals that while different estimation models introduce minor variations in *Robinia pseudoacacia* aboveground biomass (AGB) predictions, LiDAR-derived feature variables consistently demonstrate robust 1:1 linear relationships across all modeling approaches. These findings confirm the reliability of UAV-LiDAR data for accurate AGB estimation in plantation ecosystems.

The trained models ($R^2 > 0.8$) were applied to the 20% test set, with the results shown in Figure 7 and Table 4. The RF regression achieves MAE=5.87, MBE=-1.72, RMSE=8.38, and NMSE=0.23, whereas the CNN model achieves superior performance (MAE=4.57, MBE=-2.92, RMSE=5.61, NMSE=0.11). These results establish the CNN as the optimal approach for *Robinia pseudoacacia* aboveground biomass estimation via UAV-LiDAR data in the Caijiachuan watershed.

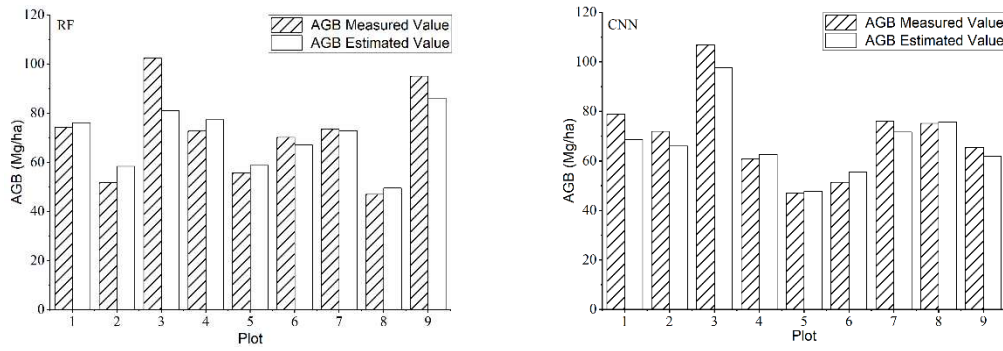


Figure 7 Prediction result

Table 4 Prediction accuracy

AGB Model	MAE	MBE	RMSE	NMSE
RF	5.87	-1.72	8.38	0.23
CNN	4.57	-2.92	5.61	0.11

Modeling biomass with Sentinel-2A images

This study extracted 18 Sentinel-2A-derived variables categorized into 6 spectral bands, 4 vegetation indices, and 8 texture features. Pearson correlation analysis (Fig. 8) was used to quantify their relationships with the field-measured Robinia pseudoacacia aboveground biomass (AGB), identifying optimal optical predictors for the estimation model.

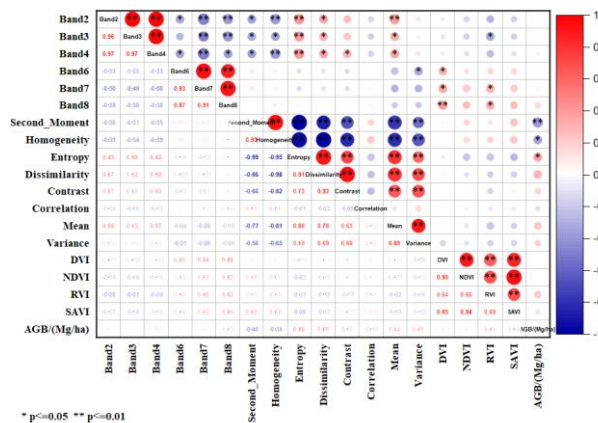


Figure 8 Correlation between Sentinel-2A variables and Robinia pseudoacacia AGB

Figure 8 shows that dissimilarity is significantly positively correlated with *Robinia pseudoacacia* aboveground biomass (AGB), whereas second_moment and homogeneity are negatively correlated. Substantial multicollinearity is observed among multiple feature variables, particularly between spectral bands and texture metrics.

Feature importance ranking was performed via random forest analysis, with the results visualized in Figure. Among the spectral features analyzed, Band 2's Second_Moment, Homogeneity, and Mean demonstrated the highest predictive importance and strongest correlations with *Robinia pseudoacacia* aboveground biomass (AGB). Among the six original spectral bands, Band 3 and Band 7 presented superior explanatory power, whereas SAVI emerged as the most significant vegetation index. The optimal feature combination selected for final modeling comprised Band2 (Second_Moment, Homogeneity, Mean), Band3, Band7, and SAVI.

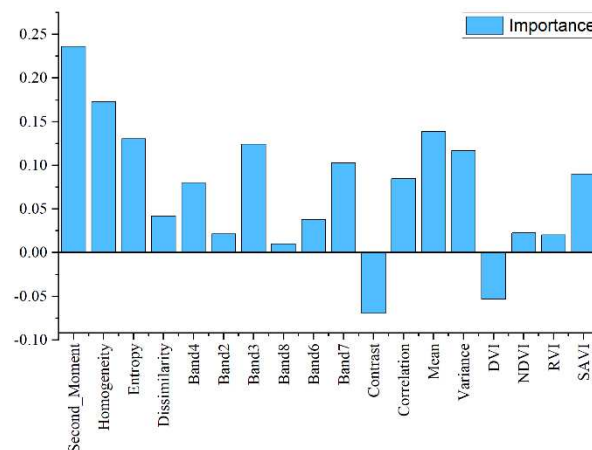


Figure 9 Sentinel-2A Variable Importance

This study employed Pearson correlation analysis and random forest feature importance ranking to identify six optimal predictors for *Robinia pseudoacacia*

aboveground biomass (AGB) estimation. Four modeling approaches—partial least squares (PLS) regression, random forest (RF) regression, backpropagation neural network (BPNN), and convolutional neural network (CNN)—were implemented via an 80:20 random split of plot data for training and testing, with the performance metrics detailed in Table 5. Among the four models evaluated, both RF and CNN demonstrated reliable performance, with explained variance exceeding 70%. The CNN achieved optimal accuracy ($R^2=0.93$; Pearson's $r=0.82$; MAE=4.89; MBE=3.11; RMSE=6.31; NMSE=0.16), whereas the BPNN yielded the poorest results. PLS regression exhibited intermediate performance, surpassing BPNN but remaining secondary to CNN in predictive capability.

Table 5 Model Accuracy Analysis

AGB Model	R^2	Pearson's r	MAE	MBE	RMSE	NMSE
RF	0.79	0.89	7.02	-0.36	9.15	0.33
PLS	0.18	0.42	11.43	-1.06	14.71	0.83
BP	0.07	0.005	13.09	1.05	16.99	1.02
CNN	0.93	0.87	4.89	3.11	6.31	0.16

Figure 10 displays scatterplots comparing the measured versus estimated Robinia pseudoacacia aboveground biomass (AGB) across all the models. The results demonstrate the superior performance of the RF and CNN models, with the CNN model exhibiting the strongest linear relationship ($R^2=0.93$) and systematic overestimation for $AGB < 90 \text{ Mg} \cdot \text{ha}^{-1}$ in the CNN model, which contrasts with more balanced UAV-LiDAR-based estimates.

361

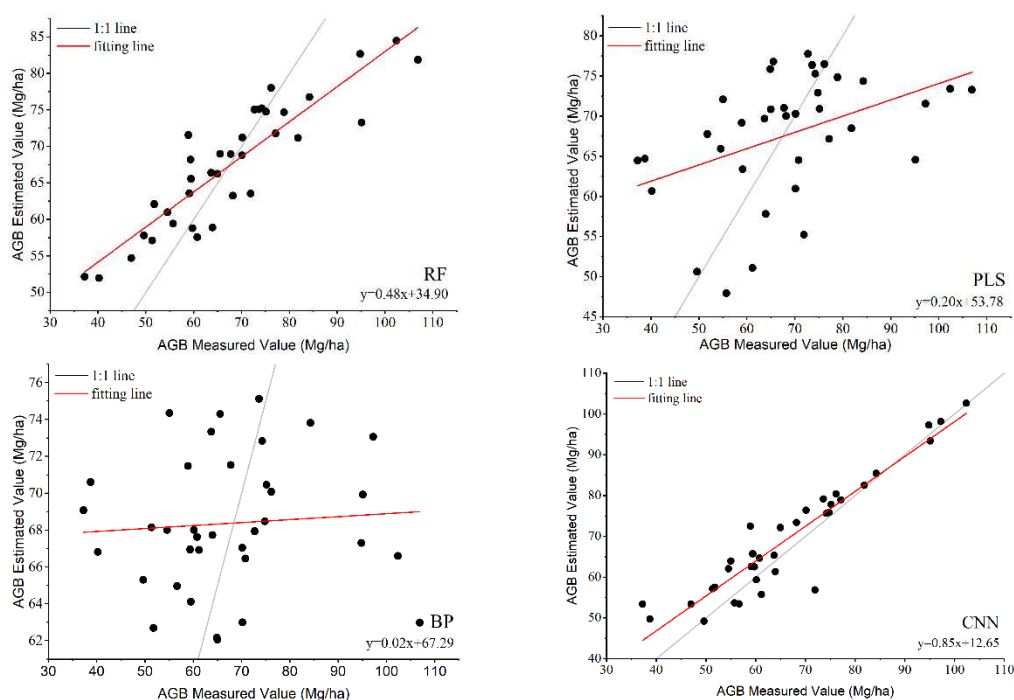


Figure 10 Measured vs. Estimated Biomass

362

363

364 An analysis of the data in Table 5 and Figure 10 reveals that among the Sentinel-
 365 2A-based estimation models, only the CNN neural network regression yields a
 366 consistent 1:1 linear relationship for the estimation of Robinia pseudoacacia
 367 aboveground biomass (AGB). The CNN model achieved superior performance
 368 ($R^2=0.93$, slope=0.98) with minimal bias (MBE=3.11), establishing it as the optimal
 369 choice for optical remote sensing-based AGB estimation in the Caijiachuan plantation
 370 watershed.

371 The CNN regression model was applied to the reserved 20% test set, with the
 372 evaluation results presented in Figure 11 and Table 6. The model demonstrated strong
 373 performance in estimating Robinia pseudoacacia aboveground biomass (AGB), with an
 374 MAE of $3.72 \text{ Mg} \cdot \text{ha}^{-1}$, MBE of $2.96 \text{ Mg} \cdot \text{ha}^{-1}$, RMSE of $4.71 \text{ Mg} \cdot \text{ha}^{-1}$, and NMSE of
 375 0.08. These results establish the CNN as the optimal approach for Robinia pseudoacacia

aboveground biomass estimation via UAV-LiDAR data in the Caijiachuan watershed.

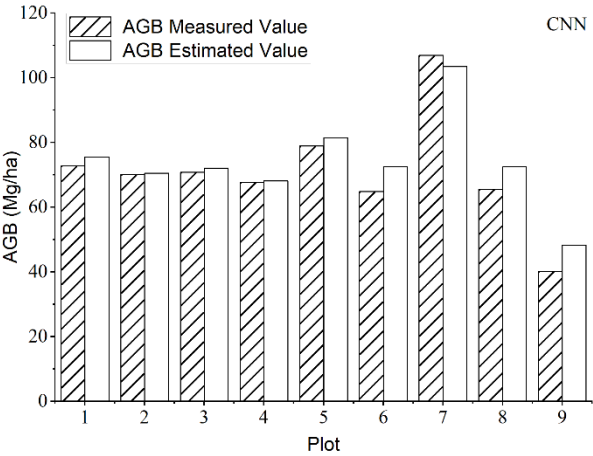


Figure 11 Prediction result

Table 6 Prediction accuracy

AGB Model	MAE	MBE	RMSE	NMSE
CNN	3.72	2.96	4.71	0.08

Modeling biomass by intergrating UAV-LiDAR and Sentinel-2A images

Pearson correlation analysis of 47 multisensor features revealed distinct relationships with Robinia pseudoacacia aboveground biomass (AGB) (Fig. 12). Structural metrics (H_avg, H_med, H_pre25%), canopy density parameters (Den8, Den9, CC), and texture features (entropy, dissimilarity) presented significant positive correlations, whereas ground returns (Den0), gap fraction (GF), and homogeneity indices (Second_Moment, homogeneity) presented negative correlations. Substantial multicollinearity was observed among multiple feature pairs, particularly between LiDAR-derived structural measurements and optical texture variables.

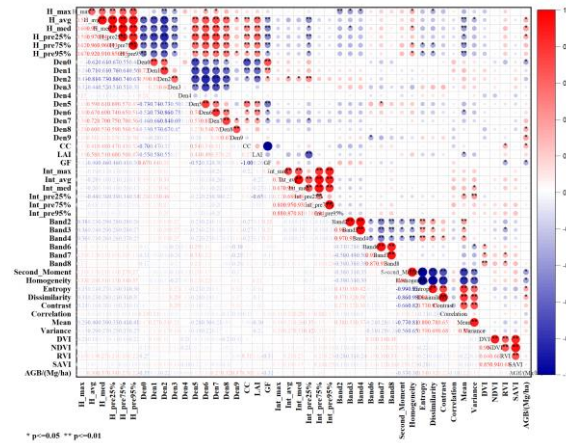


Figure 12 Correlations between remote sensing variables and Robinia pseudoacacia AGB

Random forest feature importance analysis (Fig. 13) identified Den0 as the most significant predictor for *Robinia pseudoacacia* aboveground biomass estimation, followed by H_med, Den9 and Den2. To address variable multicollinearity, we integrated both feature selection methods to retain 12 optimal predictors: eight LiDAR-derived variables (H_avg, H_med, H_25%, Den0, Den8, Den9, GF, CC) and four optical texture features (Band 2's Homogeneity, Dissimilarity, Entropy, Second_Moment).

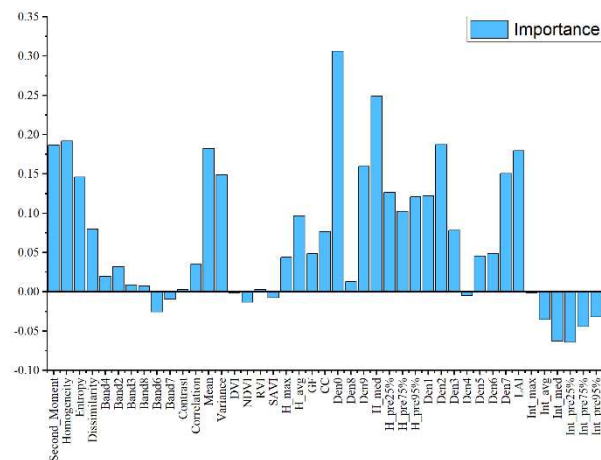


Figure 13 Importance Ranking

The 12 most significant feature variables were selected through Pearson

correlation analysis and random forest importance ranking, and biomass estimation models were subsequently developed via four methods: PLS regression, RF regression, BPNN, and CNN. As shown in Table 7, the RF and CNN achieved reliable performance, with explained variance >80%. The CNN model demonstrated optimal accuracy ($R^2=0.94$; Pearson's $r=0.97$; MAE=2.97; MBE=0.49; RMSE=3.91; NMSE=0.06), whereas the BPNN model yielded the poorest results. PLS regression slightly outperformed BPNN, with RF ranking second only to CNN in terms of predictive performance.

Table 7 Model Accuracy

AGB Model	R^2	Pearson's r	MAE	MBE	RMSE	NMSE
RF	0.81	0.90	5.53	-0.26	7.86	0.27
PLS	0.52	0.72	8.46	-1.13	10.99	0.50
BP	0.37	0.61	8.11	-1.26	14.25	0.72
CNN	0.94	0.97	2.97	0.49	3.91	0.06

Figure 14 presents scatterplots comparing the measured versus estimated aboveground biomass (AGB) across all the models. Both the RF and CNN models demonstrated superior performance, exhibiting strong linear relationships with field measurements. Systematic estimation biases were observed: AGB underestimation for high values ($>68 \text{ Mg} \cdot \text{ha}^{-1}$ in PLS/RF/BPNN; $>72 \text{ Mg} \cdot \text{ha}^{-1}$ in CNN) and overestimation for low values, indicating persistent model saturation effects at biomass extremes.

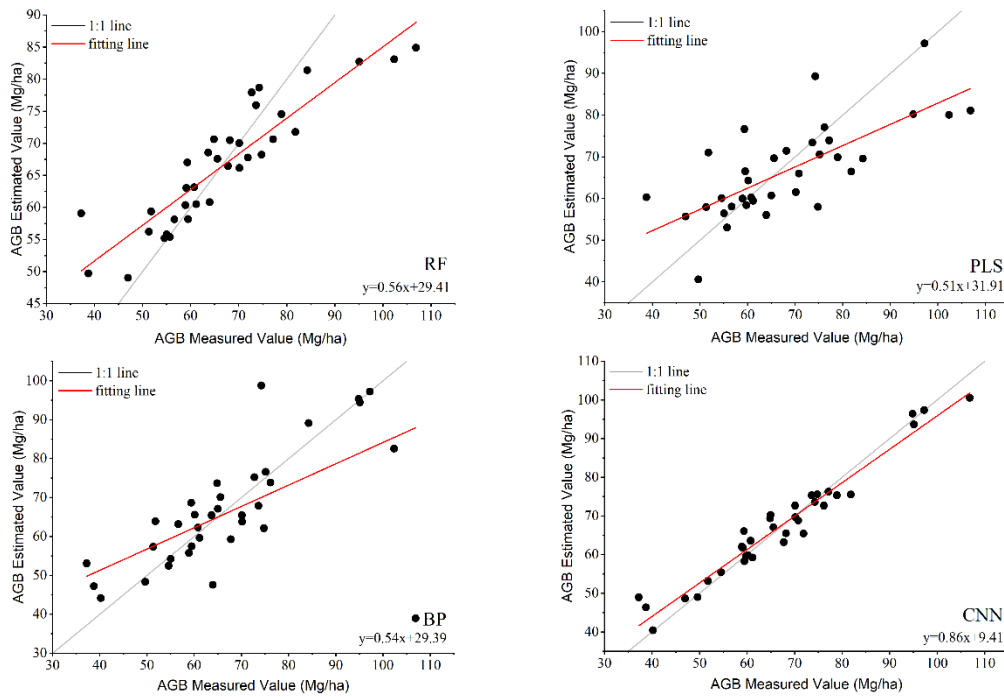


Figure 14 Measured vs. Estimated Biomass

An analysis of the data in Table 7 and Figure 14 reveals that while the different estimation methods introduce minor variations, the Sentinel-2A variables consistently demonstrate robust 1:1 linear relationships across all the models for the estimation of the aboveground biomass of *Robinia pseudoacacia*. This confirms the reliability of combining UAV-LiDAR and Sentinel-2A data for accurate biomass assessment.

The reserved 20% test set was evaluated via pretrained models ($R^2 > 0.80$), and the results are presented in Figure 15 and Table 8. Both the RF and CNN models demonstrated robust performance in estimating *Robinia pseudoacacia* aboveground biomass, with the CNN model exhibiting superior accuracy ($MAE = 2.53 \text{ Mg} \cdot \text{ha}^{-1}$, $MBE = -0.31 \text{ Mg} \cdot \text{ha}^{-1}$, $RMSE = 3.96 \text{ Mg} \cdot \text{ha}^{-1}$, $NMSE = 0.06$).

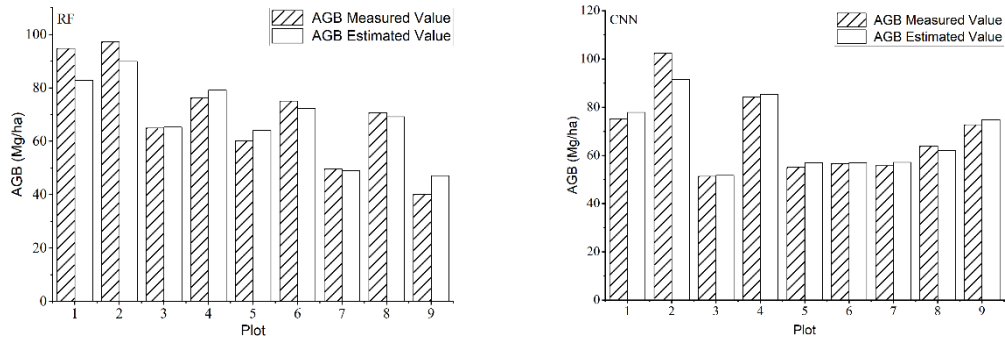


Figure 15 Prediction result

Table 8 Prediction accuracy

AGB Model	MAE	MBE	RMSE	NMSE
RF	4.27	-1.10	5.58	0.10
CNN	2.53	-0.31	3.96	0.06

Spatial distribution characteristics of aboveground biomass in Robinia pseudoacacia forest

Figure 16 presents the spatial distribution of Robinia pseudoacacia aboveground biomass (AGB) in the Caijiachuan watershed for 2022, estimated via a CNN model that integrates UAV-LiDAR and Sentinel-2A data. The results indicate a total watershed AGB of 3,727 Mg and an AGB density ranging from 23.49--148.68 $\text{Mg}\cdot\text{ha}^{-1}$, with most areas concentrated in the 50--85 $\text{Mg}\cdot\text{ha}^{-1}$ range, and a mean AGB density of 61.49 $\text{Mg}\cdot\text{ha}^{-1}$, which is consistent with field measurements (Appendix Table 1).

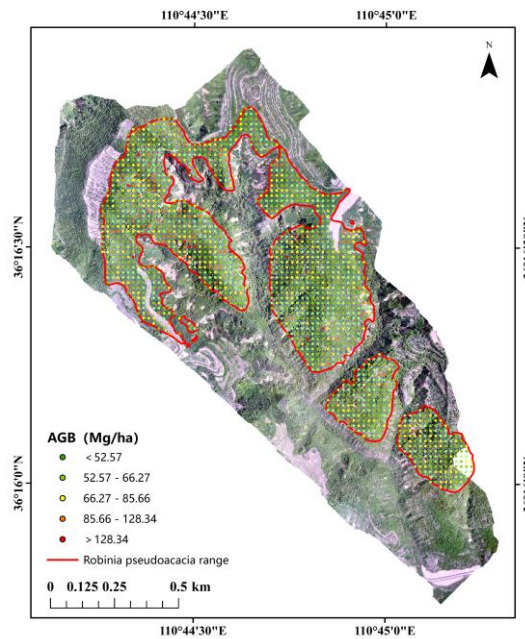


Figure 16 Spatial Distribution of Robinia pseudoacacia Biomass

On the basis of the DEM data analysis, the elevation of the artificial Robinia pseudoacacia forest watershed in Caijiachuan ranges from 951.73 m to 1234.06 m. The elevation data in the study area were divided into six elevation classes: below 1000 m, 1000–1050 m, 1050–1100 m, 1100–1150 m, 1150–1200 m, and above 1200 m. A DEM distribution map was generated (Figure 17). The distribution of the aboveground biomass of Robinia pseudoacacia across these elevation classes was then statistically analyzed, and the results are presented in Table 9 and Figure 18.

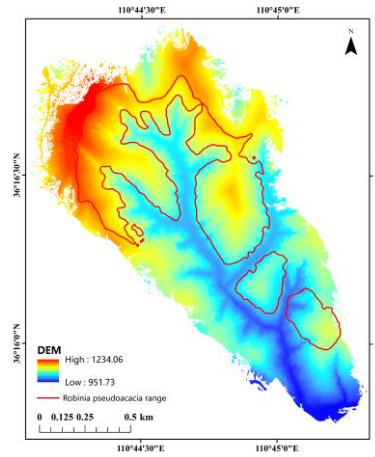


Figure 17 Digital elevation model

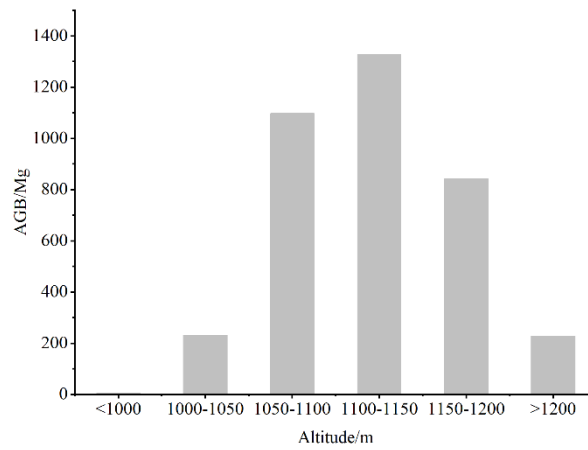


Figure 18 AGB Distribution by Altitude

Table 9 Biomass Distribution by Altitude

Altitude classification	Area proportion (%)	Biomass proportion (%)	Biomass density (Mg/ha)
<1000	0.03%	0.11%	199.60
1000-1050	6.15%	6.18%	60.75
1050-1100	30.26%	29.43%	58.79
1100-1150	36.04%	35.61%	59.72
1150-1200	22.19%	22.58%	61.51
>1200	5.32%	6.09%	69.16

As shown in Table 9 and Figure 18, the aboveground biomass of *Robinia pseudoacacia* in the Caijiachuan watershed is distributed primarily at elevations between 1050 and 1200 m, accounting for 87.62% of the total biomass, which is

consistent with actual afforestation practices. Areas above 1200 m have limited coverage, resulting in a lower biomass distribution. Conversely, elevations below 1050 m consist mainly of gullies or cliffs, making afforestation difficult and thus contributing to minimal biomass accumulation. In terms of biomass density, except for elevations below 1000 m, all the other elevation zones exhibited a density range of 60–70 Mg·ha⁻¹. The anomalously high biomass density below 1000 m may be attributed to the extremely small area of this zone.

On the basis of the DEM data, slope data were generated and classified into six grades: 0°–5° as flat slope, 6°–15° as gentle slope, 16°–25° as moderate slope, 26°–35° as steep slope, 36°–45° as abrupt slope, and >46° as extreme slope. A slope classification map was then created (Figure 19). Additionally, the distribution of the aboveground biomass of *Robinia pseudoacacia* across different slope grades was statistically analyzed, and the results are presented in Table 10 and Figure 20.

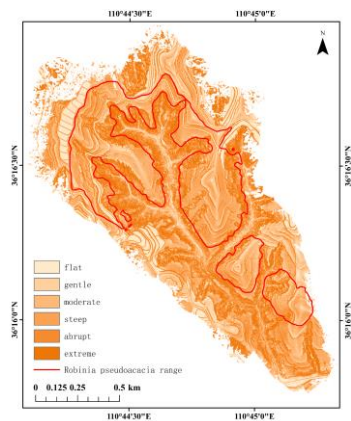


Figure 19 Slope grade map

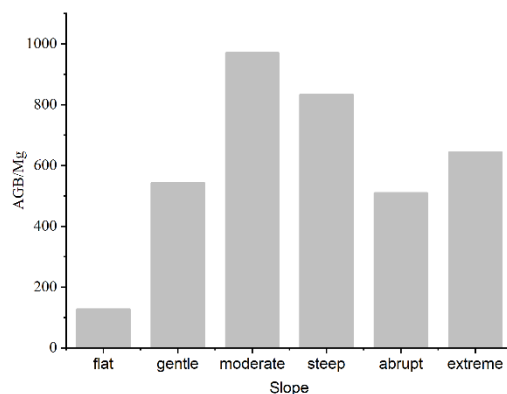


Figure 20 AGB Distribution by Slope

Table 10 Biomass Distribution by Slope

Slope classification	Area proportion (%)	Biomass proportion (%)	Biomass density (Mg/ha)
Flat slope	4.22%	3.48%	48.43
Gentle slope	15.44%	14.97%	56.97
Moderate slope	25.68%	26.76%	61.21
Steep slope	23.49%	23.00%	57.51
Abrupt slope	13.24%	14.05%	62.36
Extreme slope	17.94%	17.74%	58.05

According to Table 10 and Figure 20, moderate slopes (16–25°) cover the largest area and contribute the most biomass in the artificial *Robinia pseudoacacia* forest watershed of Caijiachuan, followed by steep slopes (26–35°). In contrast, flat slopes (0–5°) occupy the smallest area and yield the lowest biomass. In terms of biomass density, all slope grades except flat slopes maintain approximately 60 Mg·ha⁻¹, whereas flat slopes present significantly lower values, likely because of the local cultivation of cash crops (e.g., fruit and pear trees), as confirmed by field surveys.

On the basis of the DEM data, aspect data were generated and classified into four

categories: 0°–45° as a shady slope, 45°–135° as a semishady slope, 135°–225° as a sunny slope, 225°–315° as a semisunny slope, and 315°–360° as a shady slope. An aspect classification map was then created (see Figure 21). The distribution of the aboveground biomass of black Robinia pseudoacacia across the different aspect categories was statistically analyzed, and the results are presented in Table 11 and Figure 22.

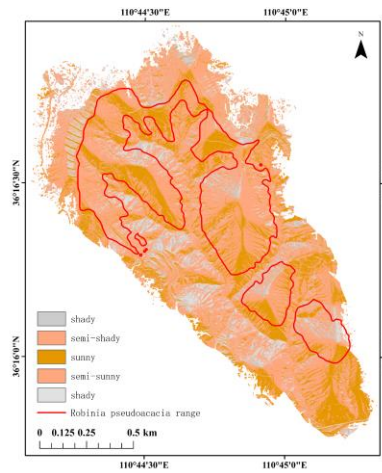


Figure 21 Slope classification diagram

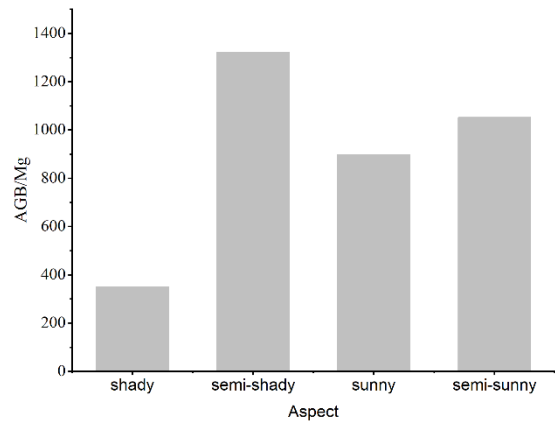


Figure 22 AGB Distribution by Aspect

Table 11 Biomass Distribution by Aspect

slope direction classification	Area proportion (%)	Biomass proportion (%)	Biomass density (Mg/ha)
--------------------------------------	------------------------	---------------------------	-------------------------

Shady slopes	9.75%	9.67%	58.27
Semishady slopes	37.73%	36.48%	56.81
Sunny slopes	27.16%	24.80%	53.66
Semisunny slopes	25.36%	29.05%	67.31

As shown in Table 11 and Figure 22, the semishady slope (45° – 135°) occupied the largest area and contributed the most biomass in the artificial black Robinia pseudoacacia forest watershed of Caijiachuan, whereas the shady slopes (0° – 45° and 315° – 360°) covered the smallest area with the lowest biomass. This pattern likely reflects black locust growth preference, as the species thrives under sufficient sunlight and moderate moisture conditions—conditions typically found on semishady and semisunny slopes. In contrast, the limited light availability on shady slopes may inhibit biomass accumulation. Furthermore, biomass density is notably lower on shady slopes, which is consistent with black locust's high light demand, whereas sunny, semisunny, and semishady slopes present higher biomass densities because of favorable light exposure.

Discussion

Estimating aboveground biomass (AGB), particularly for Robinia pseudoacacia, in the Caijiachuan watershed presents methodological challenges. While AGB distribution is influenced by multiple factors, including topography, soil properties, microclimate, and hydrology, this study specifically focuses on remote sensing-based estimation via field survey data and multisensor imagery (Sentinel-2A/UAV-LiDAR),

excluding nonremotely sensed environmental variables.

The height–DBH and height–DBH–biomass relationships were derived from field surveys and standard tree measurements. However, selecting standard trees on the basis of plot-level mean height and DBH may introduce bias for outliers (e.g., exceptionally tall trees or saplings). Comparative analysis revealed that, compared with the field measurements, the height-DBH-biomass model underestimated the total AGB by 88.76 kg (6.34 kg/tree), whereas the allometric equations had larger deviations (254.3 kg total, 18.16 kg/tree). These results demonstrate that ground-based model calibration critically impacts the accuracy of *Robinia pseudoacacia* AGB estimation.

This study utilized Sentinel-2A imagery (acquired July 24, 2022, with <5% cloud cover) for *Robinia pseudoacacia* aboveground biomass (AGB) estimation, prioritizing temporal alignment with field surveys and minimal cloud obstruction. However, no universal standards exist for remote sensing image selection in biomass studies: Yang et al. employed annual Sentinel-2A composites with <20% cloud cover for China's 30 m-resolution forest AGB mapping, whereas Zeng et al. used May–September MODIS temporal composites for Tibetan grassland AGB estimation, demonstrating significant variability in optimal image selection criteria across ecosystems and research objectives^[42,43].

The UAV-LiDAR data acquisition followed China's national standard "Technical Specification for Multisensor Consistency Testing of UAV Low-altitude Remote Sensing Monitoring" (GB/T 41450-2022) while adapting the parameters to the

Caijiachuan watershed conditions⁴⁴. This standard, which is applicable to both fixed-wing and multirotor UAV platforms, specifies fundamental requirements for radiometric and geometric consistency in low-altitude remote sensing, providing critical guidelines for *Robinia pseudoacacia* aboveground biomass estimation. Key operational parameters (flight planning, altitude, and side overlap) were customized on the basis of watershed characteristics (80 m altitude and 80% overlap), establishing a replicable protocol for similar Loess Plateau ecosystems.

This study evaluated four commonly used aboveground biomass (AGB) estimation models by integrating optical remote sensing (Sentinel-2A) and LiDAR point cloud data, ultimately identifying the optimal model for *Robinia pseudoacacia* plantations in the Caijiachuan watershed. The proposed approach provides not only an efficient methodology for local AGB estimation but also valuable references for biomass studies in other regions. Notably, all datasets (field surveys, Sentinel-2A imagery, and UAV-LiDAR acquisitions) were collected during July 2022, minimizing temporal discrepancies and thereby enhancing estimation accuracy by reducing phenological variations in biomass measurements.

Conclusions

This study developed aboveground biomass (AGB) estimation models for *Robinia pseudoacacia* plantations in the Caijiachuan watershed by integrating field survey data with multisensor remote sensing (Sentinel-2A optical imagery and UAV-LiDAR point clouds). Four modeling approaches—PLS regression, RF regression, BPNN, and

CNN—were implemented using both individual and combined data sources. The optimal model was selected through accuracy comparison and subsequently applied to generate a spatial AGB density distribution map, with topographic analysis conducted across elevation gradients, slope classes, and aspect categories.

Key conclusions:

(1) The integrated Sentinel-2A and UAV-LiDAR model outperformed single-source approaches ($R^2=0.94$, $RMSE=3.91 \text{ Mg}\cdot\text{ha}^{-1}$), highlighting the advantage of multisensor data fusion.

(2) Among the four modeling approaches evaluated—partial least squares (PLS) regression, random forest (RF) regression, backpropagation neural network (BPNN), and convolutional neural network (CNN)—the CNN algorithm demonstrated superior performance in reducing estimation errors for *Robinia pseudoacacia* aboveground biomass.

(3) The *Robinia pseudoacacia* aboveground biomass (AGB) in the Caijiachuan plantation watershed ranged from 50--85 $\text{Mg}\cdot\text{ha}^{-1}$, with a mean value of 61.49 $\text{Mg}\cdot\text{ha}^{-1}$. Spatial distribution analysis revealed concentrated biomass accumulation (87.2% of the total) within the 1,050--200 m elevation zone, predominantly on northeast-facing (semishaded) moderate (15--25°) and steep slopes (>25°).

References

1. Tsitsi, B. (2016). Remote sensing of aboveground forest biomass: A review. *Trop. Ecol*, 57(2), 125-132.
2. Návar, J. (2010). Measurement and assessment methods of forest aboveground biomass: A literature review and the challenges ahead. *Biomass*, 27-64.

- 595 3. Feng, Q., Yang, H., Liu, Y., Liu, Z., Xia, S., Wu, Z., & Zhang, Y. (2024).
596 Interdisciplinary perspectives on forest ecosystems and climate interplay: a review.
597 *Environmental Reviews*, 33, 1-21.
- 598 4. Lu, D. (2006). The potential and challenge of remote sensing-based biomass
599 estimation. *International journal of remote sensing*, 27(7), 1297-1328.
- 600 5. Rodríguez-Veiga, P., Wheeler, J., Louis, V., Tansey, K., & Balzter, H. (2017).
601 Quantifying forest biomass carbon stocks from space. *Current Forestry Reports*,
602 3(1), 1-18.
- 603 6. Tian, L., Wu, X., Tao, Y., Li, M., Qian, C., Liao, L., & Fu, W. (2023). Review of
604 remote sensing-based methods for forest aboveground biomass estimation:
605 Progress, challenges, and prospects. *Forests*, 14(6), 1086.
- 606 7. Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., & Moran, E. (2016). A survey of remote
607 sensing-based aboveground biomass estimation methods in forest ecosystems.
608 *International Journal of Digital Earth*, 9(1), 63-105.
- 609 8. Bond-Lamberty, B., Wang, C., & Gower, S. T. (2002). Aboveground and
610 belowground biomass and sapwood area allometric equations for six boreal tree
611 species of northern Manitoba. *Canadian Journal of Forest Research*, 32(8), 1441-
612 1450.
- 613 9. Dong, L., Zhang, L., & Li, F. (2016). Developing two additive biomass equations
614 for three coniferous plantation species in Northeast China. *Forests*, 7(7), 136.
- 615 10. Brown, S. (1997). Estimating biomass and biomass change of tropical forests: a
616 primer (Vol. 134). Food & Agriculture Org..
- 617 11. Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., ... &
618 Yamakura, T. (2005). Tree allometry and improved estimation of carbon stocks and
619 balance in tropical forests. *Oecologia*, 145(1), 87-99.
- 620 12. Clark, D. B., & Kellner, J. R. (2012). Tropical forest biomass estimation and the
621 fallacy of misplaced concreteness. *Journal of Vegetation Science*, 23(6), 1191-1196.
- 622 13. García, M., Riaño, D., Chuvieco, E., & Danson, F. M. (2010). Estimating biomass
623 carbon stocks for a Mediterranean forest in central Spain using LiDAR height and
624 intensity data. *Remote sensing of environment*, 114(4), 816-830.
- 625 14. Zheng D, Rademacher J, Chen J, et al. Estimating aboveground biomass using
626 Landsat 7 ETM+ data across a managed landscape in northern Wisconsin, USA.
627 *Remote sensing of environment*, 2004, 9(3): 402-411.
- 628 15. Cui G S, Lee W K, Zhu W H, et al. Vegetation classification and biomass
629 estimation using IKONOS imagery in Mt. ChangBai mountain area. *Journal of*
630 *Korean Society of Forest Science*, 2012, 101(3): 356-364.
- 631 16. Lefsky, M. A., Cohen, W. B., Harding, D. J., Parker, G. G., Acker, S. A., & Gower,
632 S. T. (2002). Lidar remote sensing of above-ground biomass in three biomes.
633 *Global ecology and biogeography*, 11(5), 393-399.
- 634 17. Luo, S., Wang, C., Xi, X., Pan, F., Peng, D., Zou, J., ... & Qin, H. (2017). Fusion
635 of airborne LiDAR data and hyperspectral imagery for aboveground and
636 belowground forest biomass estimation. *Ecological Indicators*, 73, 378-387.

18. Pajares, G. (2015). Overview and current status of remote sensing applications based on unmanned aerial vehicles (UAVs). *Photogrammetric Engineering & Remote Sensing*, 81(4), 281-330.
19. Hame, T., Salli, A., Andersson, K., & Lohi, A. (1997). A new methodology for the estimation of biomass of coniferdominated boreal forest using NOAA AVHRR data. *International Journal of Remote Sensing*, 18(15), 3211-3243.
20. White, M. A., Thornton, P. E., & Running, S. W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global biogeochemical cycles*, 11(2), 217-234.
21. Mutanga, O., & Skidmore, A. K. (2004). Narrow band vegetation indices overcome the saturation problem in biomass estimation. *International journal of remote sensing*, 25(19), 3999-4014.
22. Eriksson, L. E., Santoro, M., Wiesmann, A., & Schmullius, C. C. (2003). Multitemporal JERS repeat-pass coherence for growing-stock volume estimation of Siberian forest. *IEEE Transactions on Geoscience and Remote Sensing*, 41(7), 1561-1570.
23. Fuchs, H., Magdon, P., Kleinn, C., & Flessa, H. (2009). Estimating aboveground carbon in a catchment of the Siberian forest tundra: Combining satellite imagery and field inventory. *Remote Sensing of Environment*, 113(3), 518-531.
24. Liu, C., Chen, D., Zou, C., Liu, S., Li, H., Liu, Z., ... & Ye, L. (2022). Modeling biomass for natural subtropical secondary forest using multisource data and different regression models in Huangfu Mountain, China. *Sustainability*, 14(20), 13006.
25. Song, C. (2013). Optical remote sensing of forest leaf area index and biomass. *Progress in Physical Geography*, 37(1), 98-113.
26. Lu, D., & Batistella, M. (2005). Exploring TM image texture and its relationships with biomass estimation in Rondônia, Brazilian Amazon. *Acta Amazonica*, 35, 249-257.
27. Foody, G. M., Boyd, D. S., & Cutler, M. E. (2003). Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. *Remote sensing of environment*, 85(4), 463-474.
28. Zheng, D., Rademacher, J., Chen, J., Crow, T., Bresee, M., Le Moine, J., & Ryu, S. R. (2004). Estimating aboveground biomass using Landsat 7 ETM+ data across a managed landscape in northern Wisconsin, USA. *Remote sensing of environment*, 93(3), 402-411.
29. Wang, Z., & Menenti, M. (2021). Challenges and opportunities in Lidar remote sensing. *Frontiers in Remote Sensing*, 2, 641723.
30. Guo, Q., Su, Y., Hu, T., Guan, H., Jin, S., Zhang, J., ... & Coops, N. C. (2020). Lidar boosts 3D ecological observations and modelings: A review and perspective. *IEEE Geoscience and Remote Sensing Magazine*, 9(1), 232-257.
31. Malambo, L., Popescu, S. C., Murray, S. C., Putman, E., Pugh, N. A., Horne, D. W., ... & Bishop, M. (2018). Multitemporal field-based plant height estimation

- using 3D point clouds generated from small unmanned aerial systems high-resolution imagery. *International Journal of Applied Earth Observation and Geoinformation*, 64, 31-42.
32. Shen, X., Cao, L., Ma, Y., Coops, N. C., Muise, E. R., Wang, G., & Cao, F. (2025). Estimating structure of understory bamboo for giant panda habitat by developing an advanced vertical vegetation classification approach using UAS-LiDAR data. *International Journal of Applied Earth Observation and Geoinformation*, 136, 104304.
33. E Carvalho, A., Portugal, D., & Peixoto, P. (2025). On terrain traversability analysis in unstructured environments: recent advances in forest applications. *Intelligent Service Robotics*, 1-19.
34. 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 sensing of environment*, 114(4), 816-830.
35. Wang, D., Momo Takoudjou, S., & Casella, E. (2020). LeWoS: A universal leaf-wood classification method to facilitate the 3D modeling of large tropical trees using terrestrial LiDAR. *Methods in Ecology and Evolution*, 11(3), 376-389.
36. Fassnacht, F. E., White, J. C., Wulder, M. A., & Næsset, E. (2024). Remote sensing in forestry: current challenges, considerations and directions. *Forestry: An International Journal of Forest Research*, 97(1), 11-37.
37. Brovkina, O., Novotny, J., Cienciala, E., Zemek, F., & Russ, R. (2017). Mapping forest aboveground biomass using airborne hyperspectral and LiDAR data in the mountainous conditions of Central Europe. *Ecological Engineering*, 100, 219-230.
38. Popescu, S. C., Wynne, R. H., & Scrivani, J. A. (2004). Fusion of small-footprint lidar and multispectral data to estimate plot-level volume and biomass in deciduous and pine forests in Virginia, USA. *Forest Science*, 50(4), 551-565.
39. Li, P. F., Wu, H. E., Jing, J. F., & Li, R. Z. (2016). Noise classification denoising algorithm for point cloud model. *Computer Engineering and Application*, 52(20), 188-192.
40. Zhang W, Qi J, Wan P, et al. An easy-to-use airborne LiDAR data filtering method based on cloth simulation. *Remote sensing*, 2016, 8(6): 501.
41. Changsai, Z., Zhengjun, L., Shuwen, Y., & Zhiquan, Z. (2018). Applicability analysis of cloth simulation filtering algorithm based on LiDAR data. *Laser technology*, 42(3), 410-416.
42. Yang Q, Niu C, Liu X, et al. Mapping high-resolution forest aboveground biomass of China using multisource remote sensing data. 2023, 60(1): 2203303.
43. Zeng, N., Ren, X., He, H., Zhang, L., Zhao, D., Ge, R., ... & Niu, Z. (2019). Estimating grassland aboveground biomass on the Tibetan Plateau using a random forest algorithm. *Ecological Indicators*, 102, 479-487.
44. Nation-wide Technical Standardization Committee of Geographic Information. Technical specification of multisensor consistency testing for unmanned aerial vehicle(UAV)borne low altitude remote sensing monitoring GB/T 41450-2022.

721 Beijing: Standards Press of China, 2022: 4.

722 Acknowledgments

723 We thank the National Field Scientific Observation and Research Station of the
724 Forest Ecosystem in Jixian County, Shanxi, for supporting this study.

725 Funding

726 This research was funded by the National Key R&D Program of China (grant
727 number: 2022YFE0104700).

728 Author contributions

729 Conceptualization, Jiaqi Cheng; methodology, Jiaqi Cheng; software, Jiaqi Cheng;
730 validation, Jiaqi Cheng; formal analysis, Jiaqi Cheng; investigation, Jiaqi Cheng, Yawei
731 Hu, Yawei Hu, Jiongchang Zhao and Yang Li; data curation, Jiaqi Cheng; writing—
732 original draft, Jiaqi Cheng; writing—review and editing, Xuexia Zhang, Yawei Hu and
733 Yang Li; visualization, Jiaqi Cheng; supervision, Jianjun Zhang; project administration,
734 Xuexia Zhang and Jianjun Zhang; funding acquisition, Jianjun Zhang.

735 Ethics Declarations

736 Ethics approval and consent to participate

737 Not applicable.

738 Consent to publish

739 Not applicable.

740 Competing interests

741 The authors declare no competing interests.

742 Availability of data and materials

743 All relevant data supporting the findings of this study are available on request.

745 Table 1 Basic information of the samples

Plot	Longitude and latitude	Aspect/°	Slope/°	Altitude/m	stand density /(number·hm ⁻²)	AGB /(Mg·hm ⁻²)
1	36°16'18.09"N 110°46'11.18"E	70	31	1027	1950	106.89
2	36°16'18.89"N 110°46'08.99"E	50	15	1041	1400	76.17
3	36°16'11.10"N 110°46'29.01"E	0	13	922	2200	94.82
4	36°16'16.96"N 110°46'08.76"E	130	25	1047	2250	75.15
5	36°16'18.07"N 110°46'08.47"E	40	15	1052	2050	72.74
6	36°16'17.58"N 110°46'07.01"E	22	40	1050	2667	74.79
7	36°16'29.87"N 110°45'53.94"E	230	25	1084	2000	70.17
8	36°16'29.76"N 110°45'55.43"E	50	20	1080	1350	59.47
9	36°16'32.75"N 110°45'52.91"E	35	0	1073	2444	95.12
10	36°16'32.24"N 110°45'50.76"E	23	22.5	1084	1100	60.77
11	36°16'24.81"N 110°45'46.37"E	90	25	1098	2125	40.23
12	36°16'20.35"N 110°45'41.41"E	150	14	1066	3022	102.38
13	36°16'20.58"N 110°45'40.73"E	200	20	1068	2667	51.77

Continued Table 1

Plot	Longitude and latitude	Aspect/°	Slope/°	Altitude/m	stand density /(number·hm ⁻²)	AGB /(Mg·hm ⁻²)
14	36°16'22.44"N 110°45'44.28"E	200	18	1076	2425	68.19
15	36°16'22.20"N 110°45'45.38"E	130	20	1070	1675	64.87
16	36°16'20.04"N 110°45'54.40"E	90	37	1054	1950	70.13
17	36°16'18.71"N 110°45'53.24"E	180	36	1066	3300	59.35
18	36°16'30.31"N 110°45'50.77"E	90	15	1097	1500	56.64
19	36°16'29.67"N 110°45'50.46"E	110	15	1097	1600	64.99
20	36°16'30.85"N 110°45'50.11"E	50	16	1099	1350	49.63
21	36°16'26.52"N 110°45'47.71"E	270	30	1115	1325	59.14
22	36°16'27.11"N 110°45'46.41"E	105	28	1134	1675	71.92
23	36°16'26.87"N 110°45'44.99"E	180	31	1126	2711	58.92
24	36°16'29.67"N 110°45'47.42"E	65	27	1130	1475	61.17
25	36°16'15.34"N 110°45'27.16"E	160	27	1103	1156	46.99
26	36°16'14.86"N 110°45'29.68"E	240	30	1107	1911	55.72
27	36°16'14.43"N 110°45'31.09"E	45	30	1112	1733	54.56
28	36°16'22.66"N 110°45'26.15"E	140	24	1114	1825	55.05

Continued Table 1

Plot	Longitude and latitude	Aspect/°	Slope/°	Altitude/m	stand density /(number·hm ⁻²)	AGB /(Mg·hm ⁻²)
29	36°16'18.89"N 110°45'26.59"E	75	26	1115	1475	37.25
30	36°16'17.99"N 110°45'24.32"E	240	30	1105	1600	59.76
31	36°16'27.9210"N 110°45'04.8540"E	240	23	1135	2050	63.96
32	36°16'37.2766"N 110°44'49.0756"E	210	20	1158	1675	51.34
33	36°16'26.7970"N 110°44'49.1841"E	230	30	1185	1275	81.82
34	36°16'28.2848"N 110°44'49.4810"E	335	25	1181	950	38.75
35	36°16'27.1420"N 110°44'51.3284"E	270	32	1178	1875	67.75
36	36°16'36.4821"N 110°44'47.8818"E	310		1059	1350	97.25
37	36°16'40.2540"N 110°44'44.7138"E	285	26	1158	1875	84.25
38	36°16'30.0838"N 110°44'31.3110"E	225	23	1148	1525	77.17
39	36°16'28.9921"N 110°44'33.9137"E	100	23	1142	1250	74.25
40	36°16'32.1703"N 110°44'24.8988"E	180	20	1166	1150	63.70
41	36°16'42.0954"N 110°44'32.5029"E	196	22	1168	2475	78.92
42	36°16'45.0980"N 110°44'36.4999"E	140	24	1186	1525	65.55
43	36°16'45.5066"N 110°44'39.0658"E	255	30	1186	1575	73.62
44	36°16'48.5846"N 110°45'12.3821"E	150	20	1170	2075	70.80

45	36°16'53.1245"N	206	19	1140	1950	60.15
	110°45'18.7089"E					

746