



## Research Article

# Nondestructive individual tree aboveground biomass estimation using a hierarchical Bayesian approach in combination with individual tree competition indices

Zengrui Zhang <sup>a</sup>, Yuting Zhao <sup>a</sup>, Zhen Zhen <sup>a</sup>, Yinghui Zhao <sup>a,\*</sup>, Jun Li <sup>b</sup>, Yuan Zhou <sup>c</sup><sup>a</sup> School of Forestry, Key Laboratory of Sustainable Forest Ecosystem Management-Ministry of Education, Northeast Forestry University, Harbin, 150040, PR China<sup>b</sup> Heilongjiang Geological Science Institute, Harbin, 150040, PR China<sup>c</sup> Ecosystem Big Data Research and Development Center, School of Forestry, Northeast Forestry University, Harbin, 150040, PR China

## ARTICLE INFO

## ABSTRACT

## Keywords:

Individual tree AGB  
 Hierarchical Bayesian spatial modeling  
 INLA-SPDE  
 Nondestructive AGB estimation

Ecological variables like aboveground biomass (AGB) are often spatially autocorrelated, and AGB prediction may be underestimated if spatial correlations are ignored in remote sensing-based models. Thus, incorporating spatial correlations into AGB prediction models is crucial for accurate AGB estimation, especially in natural secondary forests with complex structures and intense competition. Terrestrial laser scanning (TLS) enables fine-scale and nondestructive measurements of individual trees while reconstructing the complete spatial structure and competitive relationships of the forest. Consequently, the utilization of TLS data for developing an individual tree AGB model that considers competition permits nondestructive AGB estimation at both the tree and plot levels. Focusing on 13 natural secondary forest sample plots located in northeast China, this study combined UAV and TLS LiDAR data to explore the applicability of the hierarchical Bayesian spatial approach (INLA-SPDEs) to nondestructive individual tree AGB estimation in natural secondary mixed forests. The analyses also considered the effect of the individual tree competition indices. This study used the INLA-SPDEs method to construct four models (a base model, a Bayesian spatial model, a hierarchical Bayesian model, and a hierarchical Bayesian spatial model) to estimate individual tree AGB. The results showed that relative to the base model ( $R^2 = 0.836$ ), the model fitting accuracy of the models incorporating random effects were improved, while the hierarchical Bayesian spatial model that included two random effects had the best estimation results ( $R^2$  was increased by 13.52 %, and the RMSE was decreased by 53.34 %). The results of the study indicate that the INLA-SPDE method that considers spatial autocorrelation is both efficient and robust for biomass estimation. Integrating Bayesian, spatial correlation, and individual tree competition factors allowed us to implement effective AGB estimation for complex forest ecosystems with significant hierarchical structures. The results thus provide strong support for spatial modeling and the analysis of ecological processes.

## 1. Introduction

Forests are one of the largest terrestrial carbon reservoirs on Earth [1]. As the smallest unit of a forest, the accurate estimation of individual tree aboveground biomass (AGB) is important for estimating forest carbon storage, understanding the forest carbon cycle, and predicting future carbon dynamics [1,2]. Natural secondary forests are one of the most important forest types in northeast China, as they may have recovered naturally after clearing or other forms of disturbance [3]. The abundance of tree species and the high density of trees in natural secondary forests lead to significant competition between trees, and this in

turn affects the growth of natural secondary forests [4]. Species-specific allometric equations for various forest types have been extensively developed and implemented, owing to the strong relationships between individual tree AGB and structural characteristics such as diameter at breast height (DBH), tree height (TH), and crown width (CW) [5,6]. Natural secondary forests are produced following the degradation of primary forests by forest fires, historical intensive logging, and other factors [7]. Secondary forest accounts for 46.2 % of China's total forest area and is the basis of forest resources, especially in northeast China [8]. Natural secondary forests have greater tree species richness and closer spacing between trees than primary forests [9], and the greater

\* Corresponding author.

E-mail address: [yinghuizhao@nefu.edu.cn](mailto:yinghuizhao@nefu.edu.cn) (Y. Zhao).

stand density leads to further increases in the variation in tree size within the forest. Tree size anisotropy regulates aboveground biomass and species diversity through plant interactions that reflect the impact of competition on the growth of natural secondary forests [10]. Therefore, it is worth exploring how the competitive relationships among trees used in forest biomass models substantially affect the estimation of AGB.

Conventional forest inventory techniques that rely on ground-based methods are often destructive and present challenges in accurately measuring tree height and other crown-related structural parameters due to restricted viewing angles [11,12]. In recent years, light detection and ranging (LiDAR) technology, particularly near-surface LiDAR platforms such as unmanned aerial vehicle laser scanning (ULS), terrestrial laser scanning (TLS), backpack laser scanning (BLS), and mobile laser scanning (MLS) have shown significant advantages in obtaining individual tree information (e.g., DBH, TH) [6]. This can be attributed to LiDAR's ability to obtain three-dimensional (3D) structural data [13], and thus the methods have been used extensively in the study of individual tree AGB [14]. In particular, multiple LiDAR methods can be combined to fully utilize the advantages of different data for nondestructive calculation of the competitive relationships between individual trees in natural secondary forests.

The inter-tree relationships in natural secondary forests are highly complex, where the distribution of resources is contingent upon the type, size, and spatial distribution of competing trees [15]. A competition index is a quantitative indicator of the intensity of competition among trees [16]. Competition indices can be roughly categorized into two types: distance-independent competition indices and distance-dependent competition indices [17]. A distance-independent individual tree competition index generally reflects the degree of dominance among individuals in forests; the index does not correlate with stand growth, and thus cannot objectively reflect the actual level of stand competition (e.g., through relative DBH or relative basal area) [18]. A distance-dependent individual tree competition index reflects the actual level of competition between individuals and neighboring trees, and thus can more objectively reflect the size of the growth space of trees (examples include Hegyi's competition index, and the competition index of Martin and Ek and its variants) [19,20]. Currently, a large number of competition indices calculated using individual tree structural indicators (e.g., tree height, DBH, and crown width) have been applied to models (e.g., individual tree growth models and tree crown width models) [21,22]. However, few studies have applied these indicators to individual-tree AGB models, especially combining the two types of competition index and for typical tree species in natural secondary forests. In addition, in research related to the competition between trees, it is important to accurately identify competing trees and quantify the interactions between trees, depending on the research objective and environmental conditions [23]. Previous research has shown that trees will compete for resources (e.g., light, water, and nutrients), and trees that are near the target tree can have a significant effect on tree growth, usually manifested as an inhibitory effect [24]. This inhibitory effect may weaken with increasing distance [25]. Therefore, determining an appropriate distance to quantify tree competition in natural secondary forests is crucial for developing accurate models of individual tree biomass.

Typically, the traditional statistical method assumes that model parameters are fixed but unknown constants, without considering multiple biotic or abiotic factors. However, individual tree data with a hierarchical structure (e.g., varying geographical sites, forest categories, environmental conditions, or growth phases) exhibit high spatial heterogeneity and may produce local effects [26,27]. Thus, the use of a classical statistical approach may affect the accuracy of biomass estimation [28]. Mixed-effects models are the standard approach to analyzing random variation [12]. However, traditional mixed-effects models do not provide reliable estimates of parameter uncertainty and only give point estimates of the model parameters [29]. In contrast, the hierarchical Bayesian approach using probability distributions and

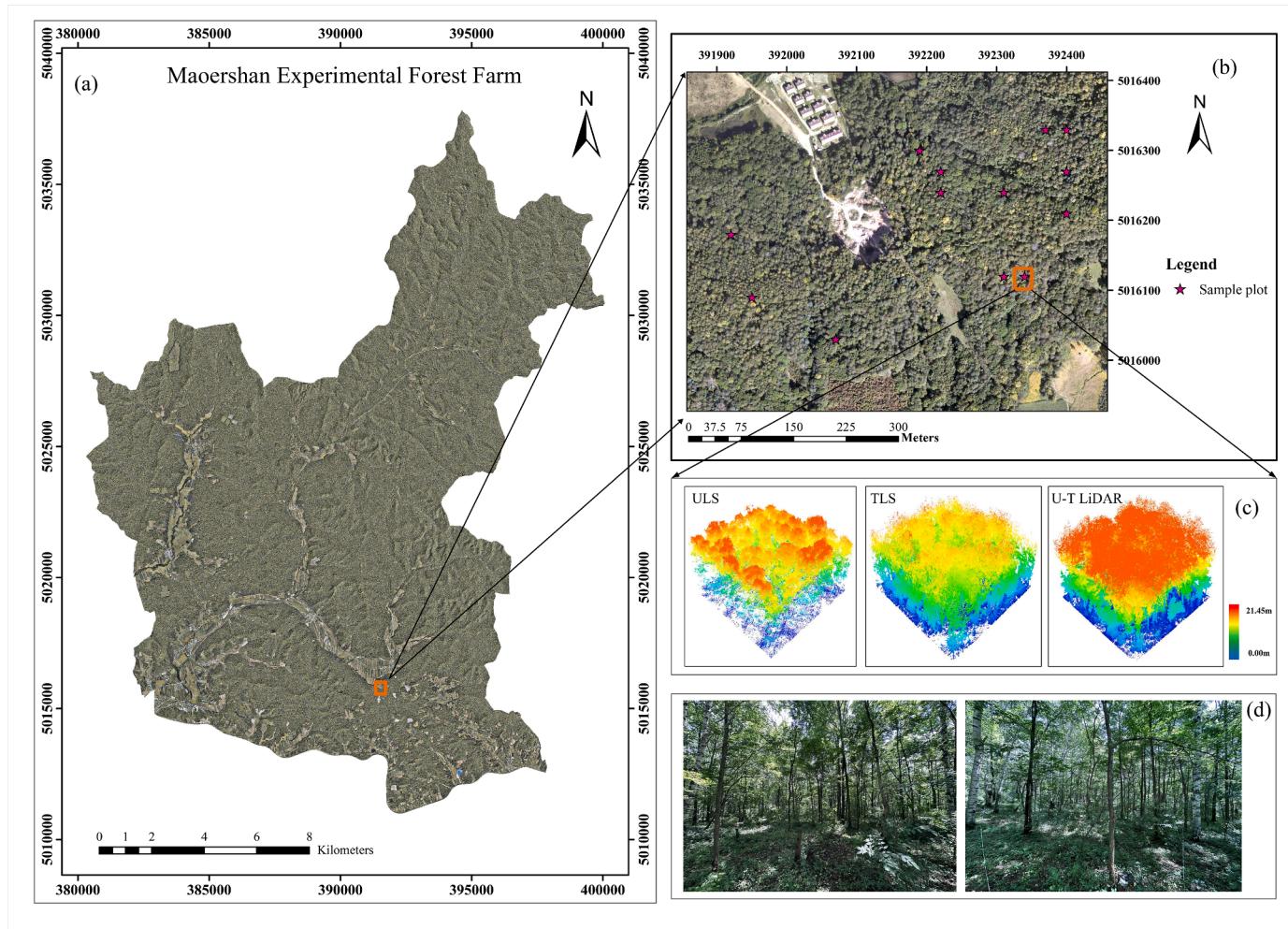
sample information can reduce the uncertainty in model parameter estimates while considering the complex dependencies between parameters [30]. However, when using the Bayesian framework (e.g., the Markov Chain Monte Carlo [MCMC] method) for fitting applications, there may be many model parameters due to the amount and complexity of the data [31]. This makes the model fitting prone to difficulty in the convergence of the posterior distribution of the model parameters, and this increases the computational burden [32]. The integrated nested Laplace approximation-stochastic partial differential equation (INLA-SPDE) method is a Bayesian approach developed over the past decade to rapid spatial modeling and analysis of geostatistical data [33]. By approximating the continuous Gaussian random field (GF) in geostatistics with a discrete Gaussian Markov random field (GMRF) using stochastic partial differential equations (SPDEs), it is possible to perform more rapid Bayesian inference using the INLA method [34]. Corresponding studies have been conducted in various fields using the INLA-SPDE method with good achievements [35]. But there are few studies using INLA-SPDE method in forestry, especially in the estimation of individual tree AGB [36]. The application of the INLA-SPDE method in forestry presents significant potential. Furthermore, INLA-SPDE method can not only process geostatistical data but also model spatial stochastic point processes using Bayesian spatial models [37]. Many studies have shown that model fitting and prediction accuracy can be improved by separating the error uncertainty into spatial and nonspatial components and introducing spatial variation and observation location information [38,39]. In terms of the spatial random effect, few studies have estimated multi-species individual tree AGB based on the hierarchical Bayesian spatial approach, particularly for typical natural secondary forests in northeast China.

Therefore, this study aimed to explore the viability of using the hierarchical Bayesian spatial approach (i.e., INLA-SPDE) to nondestructive estimation of individual tree AGB for natural secondary forests in northeastern China, while simultaneously considering individual tree competition. Specifically, this study aimed to (1) nondestructively estimate individual tree parameters (i.e., DBH and tree height) by an individual tree segmentation algorithm based on the fusion of UAV and TLS LiDAR data (hereinafter U-T LiDAR data); (2) determine the competing trees based on the fixed radius method and extract distance-independent and distance-dependent individual tree competition indices after analyzing the variation in competition intensity within different radius ranges; (3) identify the optimal explanatory variables and establish a basic individual tree AGB model, a hierarchical Bayesian model, and a hierarchical Bayesian spatial model using the INLA-SPDE approach, and compare the performance between models. This study demonstrates the effectiveness of Bayesian approach to individual tree AGB estimation by integrating the advantages of individual tree competition indices and the hierarchical Bayesian approach, providing an efficient, flexible and interpretable spatial modeling framework through the use of the INLA-SPDE methodology, thus providing novel insights for individual tree AGB modeling within multilevel forest structures.

## 2. Materials and methods

### 2.1. Study area

The study area was situated in China's Heilongjiang Province, specifically within the Maoershan Experimental Forestry Farm in Shangzhi City, from 127°29'E to 127°44'E and 45°14'N to 45°29'N (Fig. 1). Located on the western slope of Zhangguangcailing in the Changbai Mountains, the study area has an average elevation of approximately 300 m. The area experiences a temperate continental monsoon climate, characterized by an annual average temperature of about 2.5 °C and yearly precipitation averaging 700 mm [40]. The vegetation of the Maoershan region belongs to the Changbai plant flora, a typical natural secondary forest in northeastern China. The main broad-leaved tree species include white birch (*Betula platyphylla* Suk), Manchurian ash



**Fig. 1.** Location of the study area. (a) Maoershan Experimental Forest Farm in Heilongjiang Province, P.R. China, with a background of Digital Orthophoto Map (DOM); (b) Locations of 13 sample plots; (c) LiDAR point clouds of one sample plot (d) Stand conditions (i.e., mixed broadleaf forest); in the same sample plot from different angles.

(*Fraxinus mandshurica* Ruqr.), Manchurian walnut (*Juglans mandshurica* Maxim.), Mongolian oak (*Quercus mongolica* Fisch. ex Ledeb.), elm (*Ulmus pumila* L.), and coniferous tree species such as Changbai larch (*Larix olgensis* Henry), Korean pine (*Pinus koraiensis* Sieb.), and Mongolian scots pine (*Pinus sylvestris* var. *mongolica* Litv.) [12].

## 2.2. Data and preprocessing

### 2.2.1. LiDAR data and preprocessing

In this study, two types of LiDAR data were used, unmanned aerial vehicle laser scanning (ULS) data and terrestrial laser scanning (TLS) data. The ULS data were collected in August 2022 using a DJI M600 Pro drone flying at an altitude of 200 m. The flight path had a 30 % overlap to ensure comprehensive coverage. The LiDAR sensor used was the Riegl miniVUX-1UAV operating at a wavelength of 905 nm and emitting laser pulses at a frequency of 100 kHz. The average point cloud density was 270 pt/m<sup>2</sup>. The TLS data were acquired in August of 2022 and 2023 using a RIEGL VZ-400i terrestrial laser scanner (RIEGL Laser Measurement Systems, Horn, Austria), under the coverage of ULS. The average point cloud density for each sample plot was 350,000 pt/m<sup>2</sup> (max = 464000 pt/m<sup>2</sup> and min = 276000 pt/m<sup>2</sup>). The projection coordinate system for the two LiDAR datasets was WGS 1984 UTM Zone 52N. Table S1 in the Supplementary Materials lists the main parameters of the UAV and terrestrial LiDAR data.

The preprocessing of ULS and TLS data included: (1) Noise removal, which involves eliminating noise points from the raw point cloud data to improve the accuracy and efficiency of subsequent processing. Including the removal of high-level roughness caused by low-flying objects during data acquisition and low-level roughness caused by instruments during measurement; (2) Use an improved progressive triangulated irregular network (TIN) densification method to classify ground points and non-ground points [41]. This algorithm first generates a sparse triangulated network through seed points, and then iteratively processes to densify layer by layer until all ground points are classified. (3) Normalize the point cloud to eliminate the impact of terrain undulations on the elevation values of the point cloud data. The above preprocessing steps were implemented using LiDAR360 V5.2 of GreenValley International.

### 2.2.2. Field inventory data

We investigated a total of 13 30 m × 30 m sample plots dominated by the main tree species of the study area in August 2022 and 2023. In total, 971 trees with a diameter at breast height (DBH) equal to or greater than 5 cm were recorded. The data included tree height (m), DBH (cm), tree species, and location. The tree height and DBH were measured using a Vertex IV ultrasound instrument system and a perimeter ruler, and the location of each tree was recorded using a real-time kinematic (RTK) global positioning system receiver with a positional error within 5 cm.

The descriptive statistics for the main variables in the 13 sample plots are listed in [Table 1](#).

In this study, the values of individual tree AGB were calculated based on the summation of the stems, branches, and foliage biomass of each tree using species-specific additive biomass equations for northeastern China [\[40\]](#). The additive biomass equations were structured as follows:

$$\begin{aligned} \ln W_s &= a_s^* + b_s^* \cdot \ln D + c_s^* \cdot \ln H + \varepsilon_s \\ \ln W_b &= a_b^* + b_b^* \cdot \ln D + c_b^* \cdot \ln H + \varepsilon_b \\ \ln W_f &= a_f^* + b_f^* \cdot \ln D + c_f^* \cdot \ln H + \varepsilon_f \\ \ln W_a &= \ln (W_s + W_b + W_f) \end{aligned} \quad (1)$$

where  $W_s$ ,  $W_b$ ,  $W_f$  and  $W_a$  represent stems, branches, foliage, and aboveground biomass (kg), respectively;  $D$  represents the diameter at breast height (cm);  $H$  represents tree height (m);  $a_i^*$ ,  $b_i^*$ ,  $c_i^*$  ( $i = [s, b, f]$ , representing stems, branches, and foliage, respectively) are the regression coefficients listed in [Table S2](#) in the Supplementary Materials for different tree species, and  $\varepsilon_i$  is the model error term.

### 2.3. Methods

The preprocessed ULS and TLS data were initially registered and combined to generate U-T LiDAR data, and then the CSP algorithm was employed to identify individual trees and extract LiDAR-based parameters (i.e., DBH and TH) (object 1) for each tree. The fixed-radius method was employed to identify competing trees based on several factors, including the location, height, and DBH of each detected tree in each sample plot, then four distance-independent competition indices and four distance-dependent competition indices were calculated. Stepwise regression was used to select the best features (object 2). Subsequently, a basic individual tree AGB model was established based on LiDAR-derived individual tree parameters and individual tree competition indices, and the random effect of tree species and a spatial random effect were incorporated into the basic model to develop a hierarchical Bayesian model, a Bayesian spatial model, and a hierarchical Bayesian spatial model. Finally, we compared the models and obtained the best prediction results (object 3). The flow chart for the study is shown in [Fig. 2](#).

#### 2.3.1. Tree crown delineation and individual tree parameters

In this study, the two datasets (i.e., ULS and TLS) were first registered under the same projected coordinate system to realize coarse registration. Then, the ULS data were set as the reference points cloud, and the iterative closest points (ICP) algorithm [\[41,42\]](#) was employed to conduct a fine-scale registration of the two sets of LiDAR data to minimize the difference between the two point clouds and obtain U-T LiDAR data for each sample plot.

Following this, the normalized fused LiDAR (U-T LiDAR) data was used to identify individual trees through the application of the

**Table 1**

Descriptive statistics for the main tree species of the 13 sample plots.

Tree species	N	Tree height (m)		DBH (cm)		Reference AGB (kg)	
		Mean	Std	Mean	Std	Mean	Std
BP	304	20.42	3.09	19.58	6.50	207.79	181.20
QM	129	16.12	5.21	18.55	10.96	299.07	505.43
AM	111	12.00	3.61	11.40	5.38	62.56	93.42
FM	68	19.90	3.65	18.36	7.78	212.70	252.57
PD	64	20.38	5.56	24.47	13.08	359.37	424.19
UP	130	13.75	4.70	13.71	8.55	97.00	219.57
Others*	165	17.78	5.51	18.61	8.17	156.51	181.35
Total	971	—	—	—	—	—	—

Note: BP - white birch; QM - Mongolian oak; AM - Maple; FM - Manchurian ash; PD - Dahurian poplar; UP - elm; \*Others included other remaining tree species; N is the number of trees of each tree species.

comparative shortest-path (CSP) [\[43\]](#) segmentation algorithm. The CSP algorithm is a bottom-up region-growing approach that can effectively differentiate tree trunks and thus obtain the DBH of individual trees. Additionally, the algorithm considers the highest point within each tree crown as the tree height. The distance between the target tree and the adjacent trees can be calculated at the same time.

In this study, the accuracy of individual tree segmentation was evaluated using 1:1 matched trees defined as the tree with the minimum difference in tree height and DBH from the reference tree among all the trees detected in a 3-m buffer zone around the reference tree [\[44\]](#). The precision evaluation of individual tree segmentation was based on three indices, the recall ( $r$ ), precision ( $p$ ), and F-score ( $F$ ) [\[45\]](#). The equations were as follows:

$$r = \frac{TP}{(TP + FN)} \quad (2)$$

$$p = \frac{TP}{(TP + FP)} \quad (3)$$

$$F = 2 \times \frac{r \times p}{(r + p)} \quad (4)$$

where  $TP$  (true positive) represents trees being detected correctly, i.e., 1:1 matched trees;  $FN$  (false negative) represents that trees were not detected, and  $FP$  (false positive) represents that trees were detected incorrectly.

To assess the accuracy of individual tree parameter estimation based on the 1:1 matched trees, this study employed the coefficient of determination ( $R^2$ ) for the regression between estimated and measured parameters (i.e., tree height and DBH), as well as the root mean square error (RMSE) and relative root mean square error (rRMSE) of the estimated parameters [\[12\]](#).

#### 2.3.2. Estimation of individual tree competition indices

Previous studies have shown that individual tree AGB is not only strongly correlated with tree height and DBH but is also closely related to the competition index [\[18\]](#). The determination of competing trees is thus a prerequisite for calculating the competition index. There are four main methods for determining competing trees: fixed radius method, dynamic radius method, adjacent tree method and control tree number method [\[23,46–48\]](#). Due to the complexity of natural secondary forests, this study employed the fixed radius method to identify competing trees based on several factors, including the location, height, and DBH of each detected tree in each sample plot. Mao proposed using 2.5 m as the radius of the sample circle for the target tree, calculating the competitive intensity of the target tree at different radius, and analyzing how the competitive intensity changes with the increase of the radius [\[23\]](#). If the increase in the competitive intensity is negligible, this means that the increase in the number of competing trees has little effect on the radius determined by the target tree. Since the trees competing with the target tree at the edge of the sample plot may actually be outside the sample plot, the eight-neighborhood translation method was used in this study to eliminate edge effects [\[49\]](#).

Based on the detected tree height and DBH above, we calculated four distance-independent competition indices and four distance-dependent competition indices [\[50\]](#). Among these, the four distance-independent competition indices included the ratio of the target tree height to the maximum tree height and the average tree height in the sample plot (i.e.,  $RH_{max}$  and  $RH_{mean}$ ) and the ratio of the target tree DBH to the maximum DBH and the average DBH in the sample plot (i.e.,  $RD_{max}$  and  $RD_{mean}$ ). The distance-dependent competition indices were calculated based on the detected tree height, the DBH, and the locations of the target tree and competing trees, where all adjacent trees within a fixed radius of the target tree were considered as competing trees. The candidate variables used in this study are listed in [Table 2](#), and the parameters extracted

**Table 2**

Two types of LiDAR-derived individual tree competition index.

Type	Variable	Equations
The distance-independent competition index	RH <sub>max</sub>	$\frac{h_i}{H_{max}}$
	RH <sub>mean</sub>	$\frac{h_i}{\bar{h}}$
	RD <sub>max</sub>	$\frac{d_i}{D_{max}}$
	RD <sub>mean</sub>	$\frac{d_i}{\bar{D}}$
The distance-dependent competition index	CI <sub>1</sub>	$\sum_{j \neq i} \frac{d_j}{d_i \times dist_{ij}}$
	CI <sub>2</sub>	$\sum_{j \neq i} \arctan \left( \frac{d_j}{dist_{ij}} \right)$
	CI <sub>3</sub>	$\sum_{j \neq i} \frac{h_j}{h_i \times dist_{ij}}$
	CI <sub>4</sub>	$\sum_{j \neq i} \left( \frac{h_j}{h_i} \right) \times \arctan \left( \frac{h_j}{dist_{ij}} \right)$

Note:  $d_i$  is the DBH of ith target tree;  $d_j$  is the DBH of jth competing tree;  $D_{max}$  and  $\bar{D}$  are the maximum DBH and the average DBH in the sample plot, respectively;  $h_i$  is the TH of ith target tree;  $h_j$  is the TH of jth competing tree;  $H_{max}$  and  $\bar{H}$  are the maximum tree height and the average tree height in the sample plot, respectively;  $dist_{ij}$  is the distance between ith target tree and jth competing tree.

based on U-T LiDAR are shown in Fig. S1 in the Supplementary Materials.

### 2.3.3. Establishment of a U-T LiDAR individual tree AGB model

Previous experience showed that the tree AGB followed a gamma ( $\gamma$ ) distribution. Therefore, this study used log-transformed data (i.e., log (AGB)) as the response variable to establish the nonspatial model (M1) as the base model. To identify the optimal explanatory variables, the stepwise regression method was employed after extracting the above variables based on the U-T LiDAR data. The variables were simultaneously tested for multicollinearity, and those with a variance inflation factor (VIF) below 10 were retained.

In this study we used the integrated nested Laplace approximation (INLA) to fit the model within a Bayesian framework [33], while a stochastic partial differential equation (SPDE) was used to assess the spatial correlations that may affect individual tree AGB. Because the Bayesian individual tree AGB model established by INLA-SPDE is a hidden Gaussian model, it belongs to the subclass of structural additive regression models. Therefore, M1 is a linear regression model form.

### 2.3.4. Establishment of a hierarchical Bayesian spatial AGB model

Due to the hierarchical structure (i.e., sample trees of different species) of the secondary forest data, a hierarchical Bayesian method was used to investigate the effect of tree species on individual tree AGB, where tree species was included as a random effect based on the nonspatial model (i.e., M1). The following three models (i.e., M2, M3, and M4) of individual tree AGB were established using INLA-SPDE:

- (1) M2: a model including a random effect of tree species.
- (2) M3: a model including a spatial random effect.
- (3) M4: a model including a random effect of tree species and a spatial random effect.
- M2: Hierarchical Bayesian model

The hierarchical Bayesian model M2 was designed using INLA-SPDE with tree species as the random effect. This model can describe the dataset as a hierarchical structure [33], thereby achieving information sharing among different tree species. The model (M2) consists of a data model, a process model, and a parameter model, and its framework in this study is represented by Equations (5) and (6). The data model follows a normal distribution (Eq. (5));  $y_{ij}$  represents the log(AGB) of the ith tree in the jth species;  $\bar{y}_{ij}$  is the mean of  $y_{ij}$ , and  $\sigma$  is the standard deviation.

$$y_{ij} \sim N(\bar{y}_{ij}, \sigma^2) \quad (5)$$

In the process model, we assumed that the parameters of the likelihood family ( $\psi$ ) connect the mean value of the individual tree AGB  $y_{ij}$  to the structured additive predictor  $\eta$  with a link function  $g(\cdot)$ , such as  $g(\psi) = \eta$  [51]. Then, the structured additive predictor  $\eta$  can be defined as:

$$\eta_i = \beta_0 + \sum_{m=1}^M [\beta_m X_m] + \mu(T_i) + \varepsilon_i \quad (6)$$

where  $\eta_i$  is the linear predictor of the hidden Gaussian model, i.e., the expected value of the response variable  $y_{ij}$  (or the posterior predicted mean,  $\bar{y}_{ij} = \eta_i$ ).  $\beta_0$  is the intercept;  $\beta_m X_m$  ( $m = 1, \dots, M$ ) is the covariance matrix,  $X_m$  ( $m = 1, \dots, M$ ) is the vector of covariates,  $\beta_m$  ( $m = 1, \dots, M$ ) is the vector of the coefficients of the covariates;  $\mu(T_i)$  is the random effect of the tree species, and  $\varepsilon_i$  is the Gaussian white noise for independent spatial locations, with a mean of 0 and a variance of  $\sigma^2$  [52].

#### ● M3: Bayesian spatial model

M3 (the Bayesian spatial model) was produced by adding a spatial random effect to the base model (M1). We considered the response variable log(AGB) as a sample of spatial processes within a fixed region, denoted by  $y(s_i)$  ( $i = 1, \dots, n$ ). The structured additive predictor  $\eta(s_i)$  of  $y(s_i)$  consists of four components:

$$\eta(s_i) = \beta_0 + \sum_{m=1}^M [\beta_m X_m] + \xi(s_i) + \varepsilon_i \quad (7)$$

where  $\xi(s_i)$  is a continuous spatially indexed Gaussian random field (GF) [51].

#### ● M4: Hierarchical Bayesian spatial model

M4 (the hierarchical Bayesian spatial model) was developed by adding a spatial random effect to M2. The structured additive predictor  $\eta(s_i)$  can be defined as follows:

$$\eta(s_i) = \beta_0 + \sum_{m=1}^M [\beta_m X_m] + \xi(s_i) + \mu(T_i) + \varepsilon_i \quad (8)$$

where the parameter descriptions in Eq. (8) are consistent with the above equations.

All Bayesian modeling processes in this study were performed in R (v4.2.2), and the “R-INLA” package (v. 22.12) was used for model fitting.

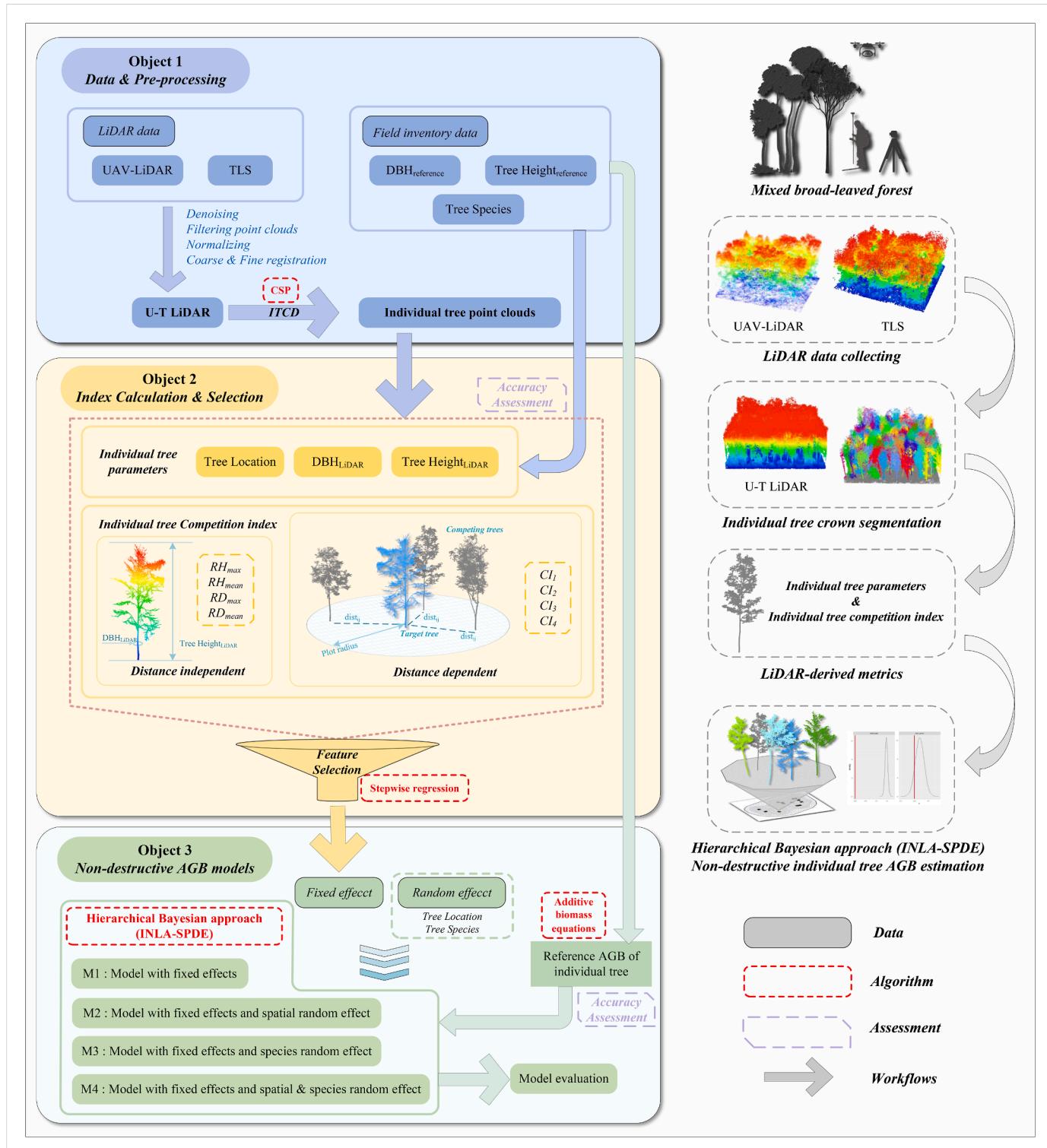


Fig. 2. Flowchart for the present study, indicating each of the research objectives (left side) and key points (right side) of the study.

### 2.3.5. Model evaluation

The Deviance Information Criterion (DIC) [53] and the widely applicable Bayesian information criterion (WAIC) were used to evaluate the goodness-of-fit of the Bayesian models (M1-M4). The formulas are as follows:

$$DIC = E(D) + p_D \quad (9)$$

$$WAIC = -2^*(lppd - pWAIC) \quad (10)$$

where  $D$  is the deviation;  $E(D)$  is the posterior mean of the deviation, which reflects the model's goodness-of-fit; the model with smaller  $E(D)$  has a better goodness-of-fit;  $p_D$  is the model complexity, which is summarized by the effective number of parameters;  $lppd$  is the log pointwise predictive density, and  $pWAIC$  is the effective number of parameters [54].

The root mean square error (RMSE) and the coefficient of determination ( $R^2$ ) were applied to compare the predictive performance of the models as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (11)$$

$$R^2 = \frac{\sum_{i=1}^n (y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{\sqrt{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}}_i)^2}} \quad (12)$$

where  $y_i$  represents the reference parameters of individual trees;  $\hat{y}_i$  represents the estimated parameters;  $\bar{y}_i$  represents the mean value of the reference parameters;  $\bar{\hat{y}}_i$  represents the mean value of the estimated parameters, and  $n$  is the number of samples.

### 3. Results

#### 3.1. Tree crown delineation and individual tree parameters

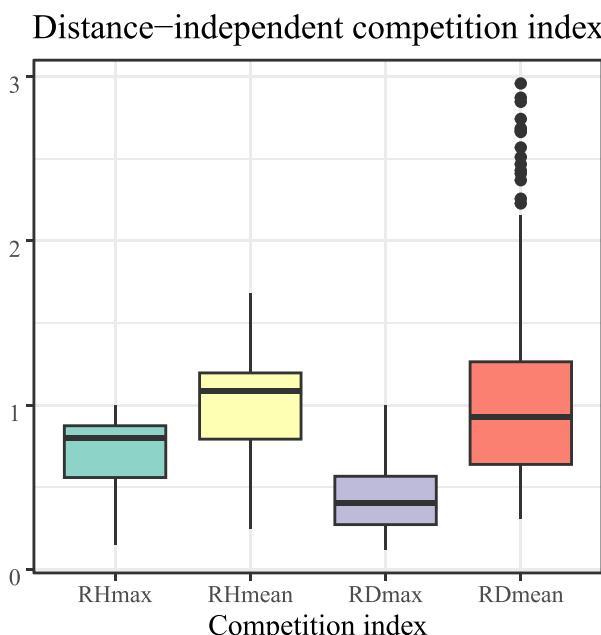
Based on U-T LiDAR data, the average  $F$  (F-score) value of 13 sample plots was 0.88 (Table S3 of Supplementary Materials). Out of 971 reference trees, 856 sample trees were successfully identified and segmented, which were subsequently utilized for modeling purposes.

Based on the 1:1 matched trees (i.e., 856 sampled trees), the accuracy of DBH and tree height estimates was evaluated using the field inventory data. The accuracy of the estimates of DBH and tree height were very similar ( $R^2$ : 0.985 and 0.926; RMSE: 1.564 and 1.819; rRMSE: 9.85 % and 10.61 %).

#### 3.2. Estimation of individual tree competition indices

##### 3.2.1. Distance-independent competition index

The data distribution of tree height and DBH in the 13 sample plots of this study is listed in Table S4 of Supplementary Materials. The average value of tree height in each sample plot ranged from 15 to 18 m, and the average value of DBH ranged from 15 to 20 cm. The distribution of the four distance-independent competition indices calculated from the U-T LiDAR data is shown in Fig. 3; the two variables related to tree height



**Fig. 3.** Distribution of the distance-independent competition index.

had skewed distributions, while the competition index related to the average DBH of the sample plots contained more outliers, a result that may be attributed to the presence of taller trees in some sample plots (e.g., sample plots No. 5 and No. 8).

##### 3.2.2. Distance-dependent competition index

Table S5 in the Supplementary Materials presents the tree spacing distances in the 13 sample plots based on the locations of the detected trees. The mean value of the distance between trees in the 13 plots was about 15 m. Therefore, this study considered all the adjacent trees within a fixed radius (i.e., 2.5–15 m) of the target tree as competing trees, and extracted the competition indices of individual trees within the corresponding distance.

Fig. 4 shows the average competition intensity for all target trees within each radius (specific parameters are in Table S6 of Supplementary Materials). The results showed that when the plot radius was 2.5–10 m, the average competitive intensity had a significant trend, with a large increase. When the radius was greater than 10 m, the increase in competitive intensity was relatively stable and gradually decreased, with an inflection point at 10 m. Therefore, this study used 10 m as a fixed radius to determine the competing trees in the sample plots and calculated the competition index for the individual tree AGB model.

#### 3.3. Establishment of an individual tree AGB model

Stepwise regression and multicollinearity testing obtained five optimal variables (i.e., DBH, RHmax, RDmax, CI2 and CI4) used for subsequent individual-tree AGB modeling and model fitting in this study.

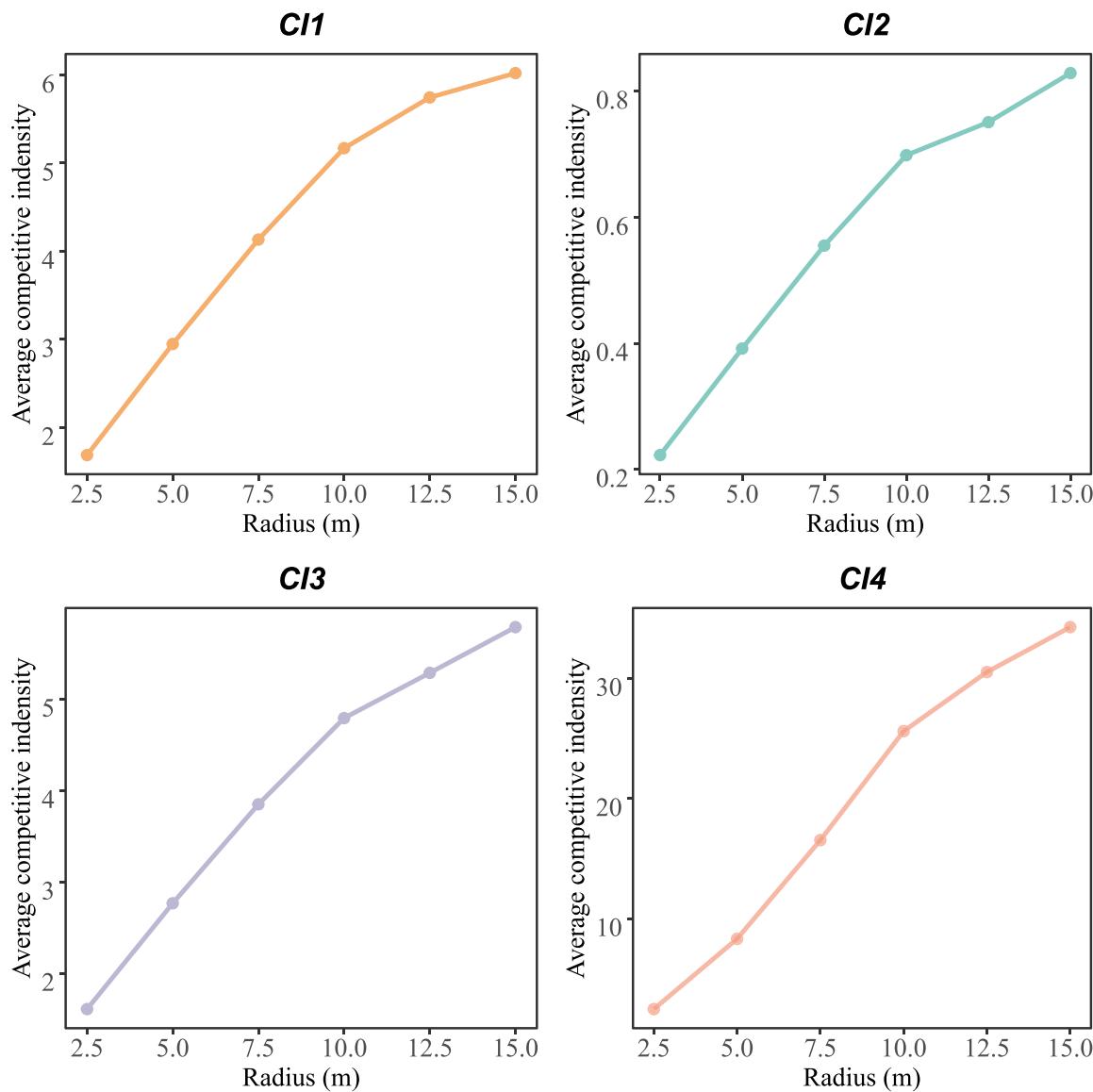
There were seven tree species (white birch, Mongolian oak, maple, Manchurian ash, Dahurian poplar, elm, and others) in the 13 sample plots. Based on M1 (the basic U-T LiDAR individual-tree AGB model), tree species were added to the model as a random effect to construct a mixed-effects model. There were six parameters ( $\beta_0$ – $\beta_5$ ) in the base model, with 63 different mixed-effect model combinations (see Table S7 in Supplementary Materials). The best-performing model was selected according to the DIC and WAIC values of all parameters and combinations to obtain M2 (the hierarchical Bayesian model) (Table 3).

Using the base model and the hierarchical Bayesian model, the spatial random effect consisting of log(AGB) hidden Gaussian fields was added to construct M3 (the Bayesian spatial model) and M4 (hierarchical Bayesian spatial model), as shown in Table 3.

#### 3.4. Model comparison

To validate the effects of tree species and spatial correlation on the estimation of individual tree AGB in natural secondary forests, we used the INLA-SPDE method to construct four models with different random effects. Table 4 presents the parameter estimates and fitting performances of M1 (the base model), M2 (the hierarchical Bayesian model), M3 (the Bayesian spatial model), and M4 (the hierarchical Bayesian spatial model). The parameters of all models were statistically significant, among which the hierarchical Bayesian spatial model (M4) exhibited the highest goodness of fit and the lowest DIC value and WAIC value. Compared to M1, the model considering the random effect of tree species (M2) improved the value of  $R^2$  by 11.84 % and reduced the RMSE by 42.06 %; the model considering the spatial random effect (M3) increased the value of  $R^2$  by 4.67 % and reduced the RMSE by 21.98 %, and the model considering both random effects (M4) significantly improved the model fitting performance, with an increase in  $R^2$  by 13.52 % and a reduction in the RMSE by 53.34 %. This suggests that the inclusion of a tree species random effect can produce a better model fit than the spatial random effect; however, the spatial random effect is also an important auxiliary variable for model fitting.

Fig. 5 shows the 97.5 % confidence intervals for the parameters estimated by the four models. The 97.5 % confidence intervals for the



**Fig. 4.** Average competitive intensity in the different radii.

parameter estimates became wider after adding random effects to the base model. The parameter estimates of the base model M1 were similar to those of M2 (which only considered tree species random effect), but differed significantly from those of M3 (which only considered the

spatial random effect), and the parameters with greater variation were correlated with DBH ( $\beta_2$  and  $\beta_3$ ), indicating that DBH was more sensitive to spatial variations. Compared to M2 and M3, M4 included both a tree species random effect and a spatial random effect. Moreover, the M4 estimates had smaller standard errors, suggesting that the model estimates incorporating both random effects were more robust.

Fig. 6 illustrates the relationship between the reference and the predicted individual tree AGB values derived from the four models. There was a significant improvement in the prediction accuracy of individual tree AGB for the hierarchical Bayesian and hierarchical Bayesian spatial models that included random effects compared to the base model ( $R^2 = 0.886$  vs. 0.935 and 0.897 vs. 0.935). The hierarchical Bayesian spatial model produced the best results ( $R^2 = 0.935$ , RMSE = 98.29 kg). When the individual tree AGB values were large, the model including the tree species random effect yielded better prediction results.

**Table 3**  
The model forms for M1, M2, M3 and M4.

Model	Model Forms
M1	$y_{ij} = \beta_0 + \beta_1 DBH_{ij} + \beta_2 RDmax_{ij} + \beta_3 RHmax_{ij} + \beta_4 CI2_{ij} + \beta_5 CI4_{ij} + \epsilon_{ij}$
M2	$y_{ij} = (\beta_0 + u_{0i}) + (\beta_1 + u_{1i})DBH_{ij} + \beta_2 RDmax_{ij} + (\beta_3 + u_{3i})RHmax_{ij} + \beta_4 CI2_{ij} + \beta_5 CI4_{ij} + \epsilon_{ij}$
M3	$y_{ij} = \beta_0 + \beta_1 DBH_{ij} + \beta_2 RDmax_{ij} + \beta_3 RHmax_{ij} + \beta_4 CI2_{ij} + \beta_5 CI4_{ij} + \xi(s_i) + \epsilon_{ij}$
M4	$y_{ij} = (\beta_0 + u_{0i}) + (\beta_1 + u_{1i})DBH_{ij} + \beta_2 RDmax_{ij} + (\beta_3 + u_{3i})RHmax_{ij} + \beta_4 CI2_{ij} + \beta_5 CI4_{ij} + \xi(s_i) + \epsilon_{ij}$

Note:  $y_{ij}$ ,  $DBH_{ij}$ ,  $RHmax_{ij}$ ,  $RDmax_{ij}$ ,  $CI2_{ij}$  and  $CI4_{ij}$  are the log-converted AGB, DBH and competition index of the  $j$ th tree of the  $i$ th tree species;  $\beta_0 - \beta_5$  are the model parameters to be determined;  $u_{0i} - u_{3i}$  are the random effects caused by the  $i$ th tree species on  $\beta_0 - \beta_3$ , respectively;  $\xi(s_i)$  is the spatial random effect;  $\epsilon_{ij}$  is the error term.

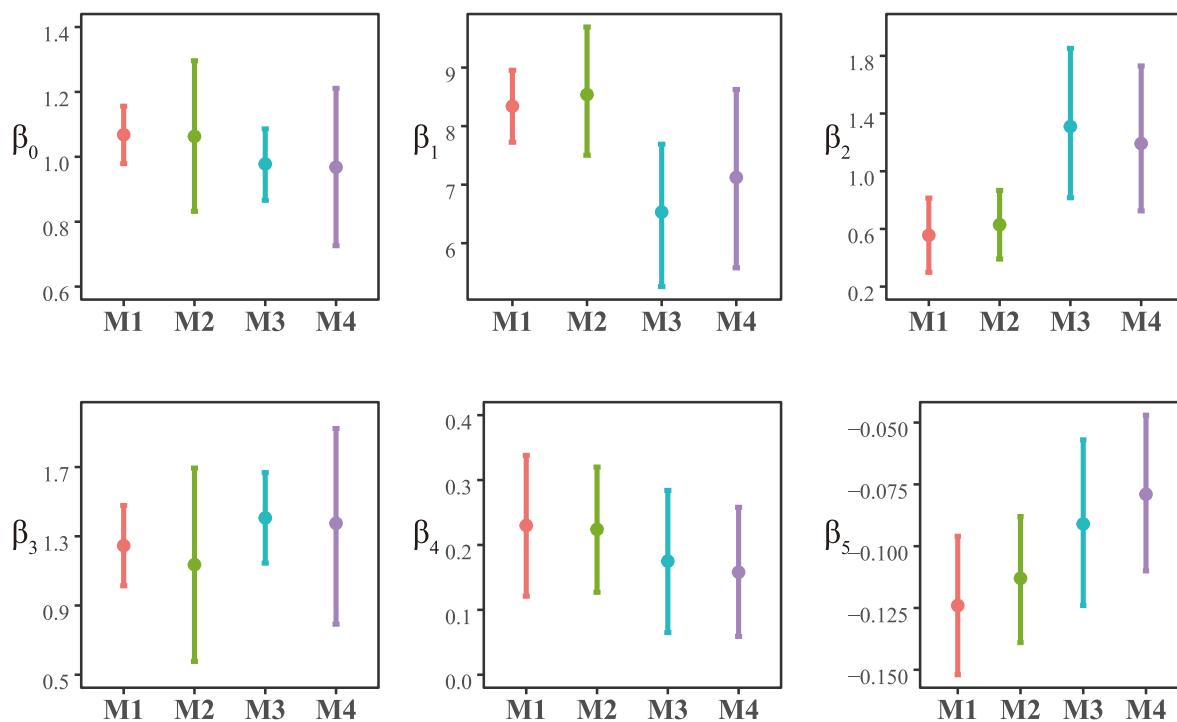


Fig. 5. Comparison of 97.5 % confidence intervals for parameter estimates ( $\beta_0$ - $\beta_5$ ) derived from four model specifications.

#### 4. Discussion

##### 4.1. Individual tree AGB estimation with LiDAR data

In recent years, near-surface LiDAR data (e.g., ULS, TLS) have become increasingly employed for nondestructive forest inventory and

AGB estimation [13]. In terms of individual tree AGB, U-T LiDAR data provides a more comprehensive view than traditional field measurements, and can accurately capture tree trunk parameters (e.g., DBH), tree height, and canopy characteristics, resulting in an ideal accuracy for estimates of individual tree AGB [14]. In this study, we examined the accuracy of the estimates for basic features (i.e., DBH and tree height) of

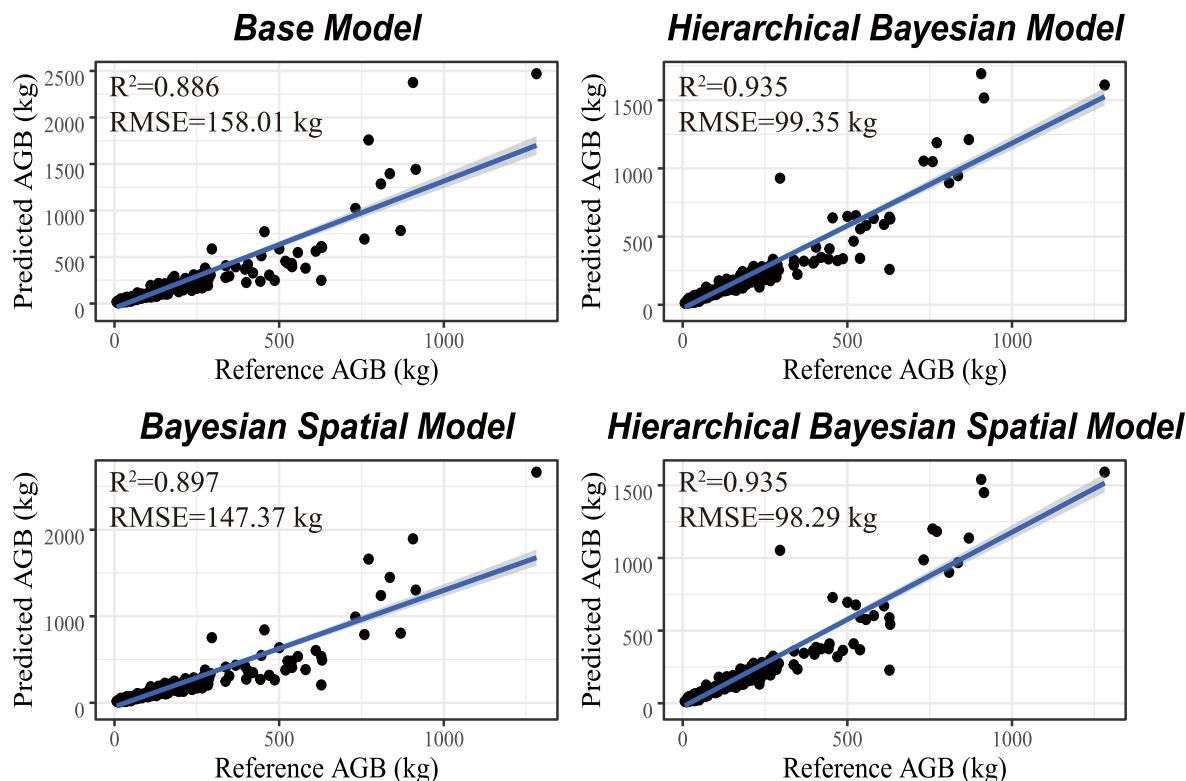


Fig. 6. Predicted and reference individual tree AGB values calculated by different models.

the sampled trees after 1:1 matching. The results showed that the accuracy of estimation of DBH and tree height was satisfactory ( $R^2$ : 0.985 and 0.926, RMSE: 1.564 and 1.819; rRMSE: 9.85 % and 10.61 %), indicating that the individual tree parameters obtained from the U-T LiDAR measurements could be used in place of the actual measurement results [55]. To verify whether there were significant differences in the prediction results, a hierarchical Bayesian spatial model was used to calculate the individual tree AGB values based on the detected and measured data, and a paired *t*-test was indicated no significant difference between the individual tree AGB values ( $P = 0.2373 > 0.05$ ). Consequently, the LiDAR detection results could be used to model the individual tree AGB.

In order to eliminate any doubt that the number of 13 sample plots was minimal. This study attempted the same experiment using 30 sample plots (2152 matched trees) and 83 sample plots (7058 matched trees), even though this study focused on AGB at the individual tree scale. None of the models exhibited overfitting, but regrettably, the accuracy of the experimental results diminished compared to those obtained with the original 13 sample plots (856 matched trees), as detailed in Table S8 of the Supplementary Materials. This may be due to the fact that the Bayesian approach yields better predictions when the sample size is small [56]. As the number of sample plots increases, the number and types of individual trees will also change, and extreme values will appear (such as large trees with DBH exceeding 60 cm). The imbalance in tree species distribution and the increase in data heterogeneity will lead to a decrease in experimental accuracy. Prior information plays an important role in Bayesian analysis and may also affect model accuracy. The prior distribution settings in this study were based on the method proposed by Wang et al. [12], but the relationship between prior information and sample size, as well as its impact on model fitting, requires further investigation in subsequent studies. Therefore, for this study, when the sample size is small, using the hierarchical Bayesian spatial model to estimate the AGB of individual trees can yield good results. However, modeling with large sample sizes requires further research.

#### 4.2. Individual tree AGB estimation with competition index

Most studies have explored the effect of individual tree competition indices on the establishment of growth models and crown width models, while fewer studies have used individual tree competition indices to establish individual tree AGB models [21,22]. By quantifying the competitive intensity between trees in forest stands, the individual tree competition index reflects the relationship between the growth of individual trees and their living situation. The index represents the aggregate competitive pressure exerted on the growth and development of the target tree in a stand by all surrounding competing trees [19]. In our previous studies, the fitting accuracy of the model using only the individual tree parameters without the addition of individual tree competition indices ( $R^2 = 0.794$ ) was lower than that of the model including competition index ( $R^2 = 0.836$ ), suggesting that the addition of individual tree competition index facilitates the estimation of AGB for individual trees. Consequently, analyzing the effect of adding competition indices to individual tree AGB models can provide a foundation for forest management [21].

This study employed eight distinct competition indices (i.e., four distance-independent competition indices and four distance-dependent competition indices) to quantify the degree of competition among adjacent trees. The distance-independent individual tree competition index reflects the relative size of the forest stand by calculating the relationship between individual trees and the forest stand. The distance-dependent individual tree competition index indicates the effect of inter-tree distance on individual tree competition [50]. The methods of determining competing tree are generally fixed radius method, dynamic radius method, adjacent tree method and control tree number method [23,46–48]. However, using the dynamic radius method and the

neighbor tree method to study tree competition usually uses a small amount of sample data and contains only a single tree species [47,48], whereas the object of this study is a natural secondary forest with many tree species and a large amount of sample data, so the above two methods are not applicable to this study. Our preliminary experiment tried to use the fixed radius method and the control tree number method to determine the competing trees [23,46]. The accuracy of the individual tree AGB model constructed from the competition indices that obtained from two methods (i.e. fixed radius method and control tree number method) was comparable, but the fitting accuracy of the control tree number method was lower and the RMSE was higher than those of the fixed radius method, as detailed in Table S9 of the Supplementary Materials. Therefore, the fixed radius method was finally adopted in this study to determine the competing trees, calculate the distance-dependent competition index and participate in the subsequent modeling.

To verify whether 10m is the optimal competitive radius, this study constructed multiple base models using individual tree competitive indices obtained from different competitive radii (ranging from 2.5 to 15 m), to test the robustness of the statistical results. In the base model format of this study, we controlled other variables and only changed the distance-dependent competition indices  $C2$  and  $C4$  to observe their impact on model fitting. Experiments revealed that the model fitting performance and  $R^2$  gradually improved from 2.5m to 7.5m, while they gradually decreased from 10m to 15m (detailed results are provided in Table S10 in the Supplementary Materials). The model fitting performance,  $R^2$  and RMSE are similar for 7.5m and 10m, but considering the tree height and crown diameter characteristics of the data used in this study, we prefer to use 10m as the competition radius for subsequent experimental studies.

The results of this study indicate that incorporating competition index into the individual tree AGB model can improve the model fitting, and the fixed radius method performs well in determining competing trees. However, for mixed forests with complex structures, the fixed radius method is relatively simple and does not consider the role of tree species in competition. Consequently, it is necessary to consider more rigorous methods for selecting competing trees and constructing competitive factors, improve the descriptive capacity of forest competitive factors, and explore the influence of such competitive factors on individual tree AGB estimation.

#### 4.3. Individual tree AGB estimation based on INLA-SPDE

In this study, we employed a hierarchical Bayesian approach (i.e., INLA-SPDE) to calculate the AGB of individual trees and considered potential spatial autocorrelation in the environment by incorporating a spatial random effect (i.e., the Bayesian spatial model and the hierarchical Bayesian spatial model). In addition, the INLA-SPDE approach was also used to compare the results with the models that did not account for spatial autocorrelation. The global Moran's index for the data sampled in this study was 0.1, indicating a weak spatial aggregation effect that was consistent with the results of this study, i.e., that spatial random effects had little correlation with individual tree AGB (Table 4 and Fig. S2). In this study, a constrained subdivided Delaunay triangular mesh was constructed using the finite element method, and the continuous spatial processes within the region represented by the mesh (containing the observations) were extended beyond the study area. The *a posteriori* predictive distributions of the spatial effects at a given location were computed while avoiding boundary effects [34]. The posterior predictive means of the spatial random effect of the Bayesian spatial model and the hierarchical Bayesian spatial model are shown in Fig. S2 of the Supplementary Materials. The effect was positive in the central part of the study area, and negative in the eastern and western edges, whereas the mean of the posterior distribution of spatial effects for the whole study area showed a trend of increasing and decreasing from east to west.

**Table 4**

Parameter estimates with 97.5 % credible intervals (CI) and fitting statistics for M1 (the base), M2 (Hierarchical Bayesian), M3 (Bayesian spatial), and M4 (Hierarchical Bayesian Spatial).

	Parameters	M1	M2	M3	M4
Fixed effects (with 97.5 % CI)	$\beta_0$	1.068 (0.979, 1.156)	1.063 (0.832, 1.296)	0.978 (0.866, 1.086)	0.968 (0.726, 1.211)
	$\beta_1$	8.339 (7.724, 8.954)	8.539 (7.502, 9.693)	6.531 (5.258, 7.691)	7.123 (5.579, 8.625)
	$\beta_2$	0.556 (0.299, 0.814)	0.629 (0.391, 0.867)	1.310 (0.816, 1.852)	1.192 (0.725, 1.730)
	$\beta_3$	1.246 (1.014, 1.478)	1.136 (0.577, 1.694)	1.406 (1.145, 1.668)	1.375 (0.792, 1.923)
	$\beta_4$	0.230 (0.121, 0.338)	0.224 (0.127, 0.320)	0.175 (0.065, 0.284)	0.158 (0.059, 0.258)
	$\beta_5$	-0.124 (-0.152, -0.096)	-0.113 (-0.139, -0.088)	-0.091 (-0.124, -0.057)	-0.079 (-0.110, -0.047)
Random effects (with 97.5 % CI)	$StdDev(u_{0i})$		0.559 (0.321, 0.943)		0.567 (0.318, 0.954)
	$StdDev(u_{1i})$		1.217 (0.518, 2.911)		1.159 (0.530, 2.299)
	$StdDev(u_{3i})$		0.628 (0.348, 1.054)		0.616 (0.332, 1.032)
	$StdDev(s_i)$			0.306 (0.288, 0.326)	0.275 (0.258, 0.292)
Fitting statistics	$R^2$	0.836	0.935	0.875	0.949
	$RMSE (kg)$	221.61	128.41	172.90	103.41
	DIC	398.80	249.12	323.85	204.72
	WAIC	403.25	276.62	331.92	231.92

After adding spatial and tree species random effects to the models, the model fitting accuracy reached 0.949, indicating that INLA-SPDE can capture the spatial correlations in AGB, thereby improving the prediction accuracy. This suggests that additional explanatory variables need to be included in the estimation. In addition, we compared the computational time of the traditional MCMC method and the INLA-SPDE method for the same data. The traditional MCMC method (ran 138587 s) was about 163 times slower than the INLA-SPDE method (ran 848 s). In contrast, INLA simplifies the computational process of complex statistical models through a series of mathematical transformations and approximations, allowing the model to produce reliable predictions in a reasonable amount of time [33]. Furthermore, as the complexity of the model increases, the advantages of alteration and modification become more apparent [57].

INLA-SPDE is not a direct tool for measuring AGB, but rather serves as a robust methodology for processing and analyzing AGB data with significant spatial complexity [33]. This method captures the spatial dynamics and distributional characteristics of AGB by constructing stochastic partial differential equation models of continuous spatial processes that reflect the interdependence and spatial autocorrelation of AGB across geographic locations [26,27]. Specifically, the AGB status of one location may be affected by neighboring locations.

This study demonstrated that superior prediction results can be attained through the hierarchical Bayesian approach (INLA-SPDE) to estimating individual tree AGB from LiDAR-derived individual tree parameters and individual tree competition indices. The INLA-SPDE method is an innovative and effective tool for ecological research and ecosystem management [34]. Integrating spatial stochastic effects with empirical observation data enables high-precision prediction and analysis of AGB spatial distributions, thereby providing robust support for ecological conservation and sustainable development [26].

#### 4.4. Limitations and future work

Although this study assessed the spatial trend of individual tree AGB in the study area, other factors related to spatial factors were not considered. Given the limited scope of the study area, the relatively consistent environmental factors related to topography, soil, and climate, these variables were not included in the present analysis. Nonetheless, the spatial distribution and dynamics of forest AGB are significantly influenced by a range of environmental factors, with topographic, soil, and climatic characteristics being particularly critical [58]. These variables shape ecosystem productivity patterns by affecting processes such as water cycling, nutrient partitioning, light distribution, and plant physiology [59]. We will consider more relevant environmental factors in our future work to achieve more accurate modeling and prediction of forest biomass by systematically quantifying the

synergistic effects of topography-soil-climate, which will provide a scientific basis for global carbon cycle research and ecosystem management [60].

The biomass prediction accuracy was high based on the fusion of TLS and ULS data, but it was difficult to estimate biomass at a larger scale due to the limitation of the use of ground-based data and the small study area. In this study, the dimension of the 13 sample plots was determined based on the nominal spatial resolution of Landsat 8 OLI imagery (30 m × 30 m) to correspond with the pixels of Landsat 8 OLI images. On this basis, it is advisable to consider expanding the study's scope by employing UAV LiDAR data as an intermediary between optical remote sensing data (i.e., Landsat data) and inventory data. For example, expanding from individual tree scale to sample plot scale or even forest scale, and from single climate zone to multi-climate zone. The application of hierarchical Bayesian spatial modeling in large-scale forest biomass studies should be further investigated by increasing the diversity of data sources.

#### 5. Conclusion

Close-range LiDAR (i.e., TLS) provides the possibility of nondestructive and detailed biomass estimation for individual trees, further quantifying the spatial correlation and heterogeneity of the forest structure. The method produced more accurate and precise AGB predictions when individual tree attributes were coupled with and competitive factors using a hierarchical Bayesian method (INLA-SPDE) with a spatial random effect.

This study completed individual tree segmentation, extracted individual tree parameters, calculated competition indices based on U-T LiDAR data, and used the INLA-SPDE to establish an AGB model that included both species random effects and spatial random effects. The analysis has provided an in-depth discussion of the feasibility of nondestructive estimation of AGB for individual trees in natural secondary forests of northeastern China. Compared to the base model, considering the tree species random effect improved the model fitting accuracy ( $R^2$  increased by 11.84 %, and RMSE decreased by 42.06 %); adding the spatial random effect improved the model fitting, but the accuracy was lower than the tree species random effect ( $R^2$  increased by 4.67 % and RMSE decreased by 21.98 %); and the inclusion of the two random effects (the tree species random effect and the spatial random effect) in the hierarchical Bayesian spatial model improved the model fitting performance, with  $R^2$  increased by 13.52 % and RMSE reduced by 53.34 %. The results indicate that both tree species and spatial random effects are important auxiliary variables in the individual tree AGB model, but the tree species random effect has a greater impact on the model.

The INLA-SPDE in this study can provide a basis for multi-tree species and high-precision nondestructive individual tree AGB modeling, and as an efficient Bayesian computational framework, the INLA-SPDE algorithm can significantly improve the accuracy of forest dynamics prediction, carbon stock assessment, and ecological restoration planning through fast computation of complex spatial models, quantification of uncertainty, and integration of multi-source data, which can provide a powerful support for spatial modeling and ecological process analysis.

## Author contributions

Zengrui Zhang: conceptualization, software, methodology, validation and writing—original draft. Yuting Zhao: conceptualization and visualization. Yinghui Zhao: supervision, funding acquisition, writing—review and editing. Zhen Zhen: funding acquisition, writing—review and editing. Jun Li: project administration. Yuan Zhou: project administration. All authors have read and agreed to the published version of the manuscript.

## Funding

This work was supported by the Key Project of National Key Research and Development Plan [2023YFF1304003]; National Natural Science Foundation of China [32071677]; and National Forestry and Grassland Data Center-Heilongjiang platform [2005DKA32200-OH].

## Data availability

Data available on request from the corresponding author.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.plaphe.2025.100120>.

## References

- [1] L.F. Xie, F.R. Li, L.J. Zhang, F.R.A. Widagdo, L.H. Dong, A bayesian approach to estimating seemingly unrelated regression for tree biomass model systems, *Forests* 11 (12) (2020) 1–30, <https://doi.org/10.3390/f11121302>.
- [2] J.C. Jenkins, D.C. Chojnacky, L.S. Heath, R.A. Birdsey, National-scale biomass estimators for United States tree species, *For. Sci.* 49 (1) (2003) 12–35, <https://doi.org/10.1093/forests/49.1.12>.
- [3] K. Stimm, M. Heym, E. Uhl, S. Tretter, H. Pretzsch, Height growth-related competitiveness of oak (*Quercus petraea* (Matt.) Liebl. and *Quercus robur* L.) under climate change in Central Europe. Is silvicultural assistance still required in mixed-species stands? *For. Ecol. Manag.* 482 (2021) <https://doi.org/10.1016/j.foreco.2020.118780>.
- [4] H.X. Liang, J.G. Huang, Q.Q. Ma, J.Y. Li, Z. Wang, X.L. Guo, H.X. Zhu, S.W. Jiang, P. Zhou, B.Y. Yu, et al., Contributions of competition and climate on radial growth of *Pinus massoniana* in subtropics of China, *Agric. For. Meteorol.* 274 (2019) 7–17, <https://doi.org/10.1016/j.agrformet.2019.04.014>.
- [5] B.R. Parresol, Assessing tree and stand biomass: a review with examples and critical comparisons, *For. Sci.* 45 (4) (1999) 573–593, <https://doi.org/10.1093/forests/45.4.573>.
- [6] B. Brede, L. Terryn, N. Barbier, H.M. Bartholomeus, R. Bartolo, K. Calders, G. Derroire, S.M. Krishna Moorthy, A. Lau, S.R. Levick, et al., Non-destructive estimation of individual tree biomass: allometric models, terrestrial and UAV laser scanning, *Rem. Sens. Environ.* 280 (2022), <https://doi.org/10.1016/j.rse.2022.113180>.
- [7] X.L. Liu, S. Liu, L. Cai, X.H. Li, Z.J.R. Xu, H.L. Pan, Q.H. Feng, L. Zhang, Q.L. Liu, Research progress and trends of natural secondary forests in China, *J. Sichuan For. Sci. Technol.* 43 (1) (2022) 1–11, <https://doi.org/10.21272/202112140001>.
- [8] H.R. Zhang, C.P. Yang, Z.G. Liu, *Tending and Regeneration Research of Natural Secondary Forest in Northeastern China*. Beijing (China), Science Press, 2022.
- [9] R.T. Corlett, What is secondary forest? *J. Trop. Ecol.* 10 (3) (1994) 445–447, <https://doi.org/10.1017/S0266467400008129>.
- [10] J. Weiner, O.T. Solbrig, The meaning and measurement of size hierarchies in plant populations, *Oecologia* 61 (3) (1984) 334–336, <https://doi.org/10.1007/BF00379630>.
- [11] T. Nakajima, Y. Hirata, T. Hiroshima, N. Furuya, S. Tatsuhara, S. Tsuyuki, N. Shiraishi, A growth prediction System for local stand volume derived from LiDAR data, *GIScience Remote Sens.* 48 (3) (2011) 394–415, <https://doi.org/10.2747/1548-1603.48.3.394>.
- [12] M. Wang, J. Im, Y.H. Zhao, Z. Zhen, Multi-Platform LiDAR for non-destructive individual aboveground biomass estimation for Changbai larch (*Larix olgensis* Henry) using a hierarchical Bayesian approach, *Remote Sens.* 14 (17) (2022) 4361, <https://doi.org/10.3390/rs14174361>.
- [13] F.G. Jiang, H. Sun, K.S. Ma, L.Y. Fu, J. Tang, Improving aboveground biomass estimation of natural forests on the Tibetan Plateau using spaceborne LiDAR and machine learning algorithms, *Ecol. Indic.* 143 (2022), <https://doi.org/10.1016/j.ecolin.2022.109365>.
- [14] G.P. Fan, Z.Y. Xu, J.H. Wang, L.L. Nan, H.J. Xiao, Z.M. Xin, F.X. Chen, Plot-level reconstruction of 3D tree models for aboveground biomass estimation, *Ecol. Indic.* 142 (2022), <https://doi.org/10.1016/j.ecolin.2022.109211>.
- [15] R.O. Pedersen, O.M. Bollandsås, T. Gobakken, E. Naesset, Deriving individual tree competition indices from airborne laser scanning, *For. Ecol. Manag.* 280 (2012) 150–165, <https://doi.org/10.1016/j.foreco.2012.05.043>.
- [16] Y. Wang, H. Yang, Y.L. Li, S. Qiu, Tree competition index based on the structural equation model, *J. Beijing For. Univ.* 37 (4) (2015) 28–37, <https://doi.org/10.13332/j.1000-1522.20140075>.
- [17] C. Kuehne, A.R. Weiskittel, J. Waskiewicz, Comparing performance of contrasting distance-independent and distance-dependent competition metrics in predicting individual tree diameter increment and survival within structurally-heterogeneous, mixed-species forests of Northeastern United States, *For. Ecol. Manag.* 433 (2019) 205–216, <https://doi.org/10.1016/j.foreco.2018.11.002>.
- [18] G.S. Biging, M. Dobbertin, Evaluation of competition indices in individual tree growth models, *For. Sci.* 41 (2) (1995) 360–377, <https://doi.org/10.1093/forescience/41.2.360>.
- [19] F. Hegyi, A simulation model for managing jack-pine stands, in: *Growth Models for Tree and Stand Simulation*, Royal College of Forestry., 1974, pp. 74–90.
- [20] G.L. Martin, A.R. Ek, A comparison of competition measures and growth models for predicting plantation red pine diameter and height growth, *For. Sci.* 30 (3) (1984) 731–743, <https://doi.org/10.1093/forescience/30.3.731>.
- [21] Y.P. Qin, X. He, X.D. Lei, L.Y. Feng, Z. Zhou, J. Lu, Tree size inequality and competition effects on nonlinear mixed effects crown width model for natural spruce-fir-broadleaf mixed forest in northeast China, *For. Ecol. Manag.* 518 (2022) 120291, <https://doi.org/10.1016/j.foreco.2022.120291>.
- [22] T.P. Pitkänen, S. Bianchi, A. Kangas, Quantifying the effects of competition on the dimensions of Scots pine and Norway spruce crowns, *Int. J. Appl. Earth Obs. Geoinf.* 112 (2022) 102941, <https://doi.org/10.1016/j.jag.2022.102941>.
- [23] L. Mao, D.Q. Yang, D.M. Wang, X.H. Yang, Analyses of intraspecific and interspecific competition of *Pinus sylvestris* var. *mongolica* natural forest in Honghuaerji Nature Reserve of Inner Mongolia, *J. Plant Resour. Environ.* 17 (2) (2008) 9–14.
- [24] Y. Tang, Y.W. Tong, Y.G. Han, W.M. Zhou, L. Zhou, L.M. Dai, D.P. Yu, Effect of neighborhood competition on key tree species growth in broadleaved-Korean pine mixed forest in Changbai Mountain, China, *Chin. J. Appl. Ecol.* 30 (5) (2019) 1479–1486, <https://doi.org/10.13287/j.1001-9332.201905.021>.
- [25] W.P. Hua, T. Qiu, X.D. Jiang, X.M. Gai, L.Z. Huang, Study on individual tree competition analysis and development model of the *Pinus taiwanensis*, *J. For. Environ.* 37 (2) (2017) 201–206, <https://doi.org/10.13324/j.cnki.jfcf.2017.02.012>.
- [26] Planche NR. Ver, A.O. Finley, J.A. Kershaw, A.R. Weiskittel, M.C. Kress, Hierarchical Bayesian models for small area estimation of forest variables using LiDAR, *Rem. Sens. Environ.* 204 (2018) 287–295, <https://doi.org/10.1016/j.rse.2017.10.024>.
- [27] T.S. Chinembiri, O. Mutanga, T. Dube, Hierarchical Bayesian geostatistics for C stock prediction in disturbed plantation forest in Zimbabwe, *Ecol. Inform.* 73 (2023), <https://doi.org/10.1016/j.ecoinf.2022.101934>.
- [28] S. Saatchi, M. Marlier, R.L. Chazdon, D.B. Clark, A.E. Russell, Impact of spatial variability of tropical forest structure on radar estimation of aboveground biomass, *Rem. Sens. Environ.* 115 (11) (2011) 2836–2849, <https://doi.org/10.1016/j.rse.2010.07.015>.
- [29] R. Li, B. Stewart, A. Weiskittel, A Bayesian approach for modelling non-linear longitudinal/hierarchical data with random effects in forestry, *Forestry* 85 (1) (2012) 17–25, <https://doi.org/10.1093/forestry/cpr050>.
- [30] X.Q. Zhang, A.G. Duan, J.G. Zhang, C.W. Xiang, Estimating tree height-diameter models with the bayesian method, *Sci. World J.* 2014 (2014), <https://doi.org/10.1155/2014/683691>.
- [31] A.O. Finley, S. Banerjee, B.D. Cook, J.B. Bradford, Hierarchical Bayesian spatial models for predicting multiple forest variables using waveform LiDAR, hyperspectral imagery, and large inventory datasets, *Int. J. Appl. Earth Obs. Geoinf.* 22 (1) (2013) 147–160, <https://doi.org/10.1016/j.jag.2012.04.007>.
- [32] J.S. Speagle, *A Conceptual Introduction to Markov Chain Monte Carlo Methods*. arXiv: Other Statistics, 2019.
- [33] P. Moraga, S.M. Cramb, K.L. Mengersen, M. Pagano, A geostatistical model for combined analysis of point-level and area-level data using INLA and SPDE, *Spatial Stat.* 21 (2017) 27–41, <https://doi.org/10.1016/j.spasta.2017.04.006>.
- [34] F. Lindgren, H. Rue, J. Lindström, An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach, *J. Roy. Stat. Soc. B* 73 (4) (2011) 423–498, <https://doi.org/10.1111/j.1467-9868.2011.00777.x>.

- [35] M. Saez, M.A. Barceló, Spatial prediction of air pollution levels using a hierarchical Bayesian spatiotemporal model in Catalonia, Spain, *Environ. Model. Software* 151 (2022), <https://doi.org/10.1016/j.envsoft.2022.105369>.
- [36] H.Y. Wang, C. Daas, E.O. Coul, K.J. Jonas, MSM with HIV: improving prevalence and risk estimates by a Bayesian small area estimation modelling approach for public health service areas in the Netherlands, *Spatial and Spatio-Temporal Epidemiol.* 45 (2023), <https://doi.org/10.1016/j.sste.2023.100577>.
- [37] L.D. Williamson, B.E. Scott, M. Laxton, J.B. Illian, V.L.G. Todd, P.I. Miller, K. L. Brookes, Comparing distribution of harbour porpoise using generalized additive models and hierarchical Bayesian models with integrated nested laplace approximation, *Ecol. Model.* 470 (2022), <https://doi.org/10.1016/j.ecolmodel.2022.110011>.
- [38] Z.W. Li, L. Ding, B.B. Shen, J.Q. Chen, D.W. Xu, X. Wang, W. Fang, A.S. Pulatov, M. Kussainova, et al., Quantifying key vegetation parameters from Sentinel-3 and MODIS over the eastern Eurasian steppe with a Bayesian geostatistical model, *Sci. Total Environ.* 909 (2024), <https://doi.org/10.1016/j.scitotenv.2023.168594>.
- [39] L. Poggio, A. Gimona, L. Spezia, M.J. Brewer, Bayesian spatial modelling of soil properties and their uncertainty: the example of soil organic matter in Scotland using R-INLA, *Geoderma* 277 (2016) 69–82, <https://doi.org/10.1016/j.geoderma.2016.04.026>.
- [40] L.H. Dong, L.J. Zhang, F.R. Li, Developing additive systems of biomass equations for nine hardwood species in Northeast China, *Trees Struct. Funct.* 29 (4) (2015) 1149–1163, <https://doi.org/10.1007/s00468-015-1196-1>.
- [41] X.Q. Zhao, Q.H. Guo, Y.J. Su, B.L. Xue, Improved progressive TIN densification filtering algorithm for airborne LiDAR data in forested areas, *ISPRS J. Photogrammetry Remote Sens.* 117 (2016) 79–91, <https://doi.org/10.1016/j.isprsjprs.2016.03.016>.
- [42] P.J. Besl, N.D. McKay, A method for registration of 3-D shapes, *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (2) (1992) 239–256, <https://doi.org/10.1109/34.121791>.
- [43] S.L. Tao, F.F. Wu, Q.H. Guo, Y.C. Wang, W.K. Li, B.L. Xue, X.Y. Hu, P. Li, D. Tian, C. Li, et al., Segmenting tree crowns from terrestrial and mobile LiDAR data by exploring ecological theories, *ISPRS J. Photogrammetry Remote Sens.* 110 (2015) 66–76, <https://doi.org/10.1016/j.isprsjprs.2015.10.007>.
- [44] J. Murray, D. Gullick, G.A. Blackburn, J.D. Whyatt, C. Edwards, ARBOR: a new framework for assessing the accuracy of individual tree crown delineation from remotely-sensed data, *Rem. Sens. Environ.* 231 (2019), <https://doi.org/10.1016/j.rse.2019.111256>.
- [45] W.K. Li, Q.H. Guo, M.K. Jakubowski, M. Kelly, A new method for segmenting individual trees from the lidar point cloud, *Photogramm. Eng. Rem. Sens.* 78 (1) (2012) 75–84, <https://doi.org/10.14358/PERS.78.1.75>.
- [46] Z. Sun, Y.F. Wang, L. Pan, Y.J. Sun, Hegyi competition index decomposition to improve estimation accuracy of Larix olgensis crown radius, *Ecol. Indic.* 143 (2022), <https://doi.org/10.1016/j.ecolind.2022.109322>.
- [47] S.S. Long, S.Q. Zeng, H.H. Xiao, F.L. Liu, M. Hu, Analysis on the competitive status of cyclobalanopsis glaucaSecondary Forest based on the improved hegvi model, *Forest Res. Manag.* (1) (2018) 50–56, <https://doi.org/10.13466/j.cnki.lyzygl.2018.01.008>.
- [48] X.Z. Yang, W.H. Zhang, Q.Y. He, Effects of intraspecific competition on growth, architecture and biomass allocation of Quercus liaotungensis, *J. Plant Interact.* 14 (1) (2019) 284–294, <https://doi.org/10.1080/17429145.2019.1629656>.
- [49] M.P. Tang, Y.G. Chen, Y.J. Shi, G.M. Zhou, M.S. Zhao, Intraspecific and Interspecific competition analysis of community dominant plant populations based on Voronoi diagram, *Acta Ecol. Sin.* 27 (11) (2007), <https://doi.org/10.0000/j.1000-0933.2007271147074716>, 4707-471.
- [50] S.K. Bhandari, E.J. Veneklaas, L. McCaw, R. Mazanec, M. Renton, Investigating the effect of neighbour competition on individual tree growth in thinned and unthinned eucalypt forests, *For. Ecol. Manag.* 499 (2021), <https://doi.org/10.1016/j.foreco.2021.119637>.
- [51] H. Bakka, H. Rue, G.A. Fuglstad, A. Riebler, D. Bolin, J. Illian, E. Krainski, D. Simpson, F. Lindgren, Spatial modeling with R-INLA: a review, *Wiley Interdiscip. Rev.: Comput. Stat.* 10 (2018) e1443, <https://doi.org/10.1002/wics.1443>.
- [52] N.A.C. Cressie, Statistics for spatial data, revised edition, *Biometrics* 50 (1) (1994) 319, <https://doi.org/10.2307/2533238>.
- [53] D.J. Spiegelhalter, N.G. Best, B.P. Carlin, A. van der Linde, Bayesian measures of model complexity and fit, *J. Roy. Stat. Soc. B* 64 (4) (2002) 583–639. <http://www.jstor.org/stable/3088806>.
- [54] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, D.B. Rubin, *Bayesian Data Analysis*, third ed., Chapman and Hall/CRC, 2013 <https://doi.org/10.1201/b16018>.
- [55] X.L. Liang, J. Hyppä, Automatic stem mapping by merging several terrestrial laser scans at the feature and decision levels, *Sensors (Switzerland)* 13 (2) (2013) 1614–1634, <https://doi.org/10.3390/s130201614>.
- [56] M. Zapata-Cuertas, C.A. Sierra, L. Alleman, Probability distribution of allometric coefficients and Bayesian estimation of aboveground tree biomass, *For. Ecol. Manag.* 277 (2012) 173–179, <https://doi.org/10.1016/j.foreco.2012.04.030>.
- [57] H. Rue, S. Martino, N. Chopin, Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, *J. Roy. Stat. Soc. B Stat. Methodol.* 71 (2) (2009) 319–392, <https://doi.org/10.1111/j.1467-9868.2008.00700.x>.
- [58] M. Soubeiran, P. Marchand, L. Duchesne, Y. Bergeron, F. Gennaretti, Interactions between climate, soil and competition drive tree growth in Quebec forests, *For. Ecol. Manag.* 555 (2024), <https://doi.org/10.1016/j.foreco.2024.121731>.
- [59] J. Zhang, S.M. Huang, F.L. He, Half-century evidence from western Canada shows forest dynamics are primarily driven by competition followed by climate, *Proc. Natl. Acad. Sci. U. S. A* 112 (13) (2015) 4009–4014, <https://doi.org/10.1073/pnas.1420844112>.
- [60] Y.W. Yao, H.R. Ren, Estimation of grassland aboveground biomass in northern China based on topography-climate-remote sensing data, *Ecol. Indic.* 165 (2024) 112230, <https://doi.org/10.1016/j.ecolind.2024.112230>.