



Article

Estimating the Aboveground Biomass of Various Forest Types with High Heterogeneity at the Provincial Scale Based on Multi-Source Data

Tianbao Huang^{1,2,3}, Guanglong Ou³ , Yong Wu³, Xiaoli Zhang³, Zihao Liu³, Hui Xu³, Xiongwei Xu^{1,2}, Zhenghui Wang^{1,2} and Can Xu^{1,2,*}

- ¹ Kunming General Survey of Natural Resources Center, China Geological Survey, Kunming 650111, China; huangtianbao@swfu.edu.cn (T.H.); xuxiongwei@mail.cgs.gov.cn (X.X.); wangzhenghui@mail.cgs.gov.cn (Z.W.)
- ² Technology Innovation Center for Natural Ecosystem Carbon Sink, Ministry of Natural Resources, Kunming 650111, China
- ³ Key Laboratory of Southwest Mountain Forest Resources Conservation and Utilization, Southwest Forestry University, Ministry of Education, Kunming 650233, China; olg2007621@swfu.edu.cn (G.O.); yongwu@swfu.edu.cn (Y.W.); karojan@swfu.edu.cn (X.Z.); liuzihao@swfu.edu.cn (Z.L.); swfc213@126.com (H.X.)
- * Correspondence: xucan@mail.cgs.gov.cn

Abstract: It is important to improve the accuracy of models estimating aboveground biomass (AGB) in large areas with complex geography and high forest heterogeneity. In this study, k-nearest neighbors (k-NN), gradient boosting machine (GBM), random forest (RF), quantile random forest (QRF), regularized random forest (RRF), and Bayesian regularization neural network (BRNN) machine learning algorithms were constructed to estimate the AGB of four forest types based on environmental factors and the variables selected by the Boruta algorithm in Yunnan Province and using integrated Landsat 8 OLI and Sentinel 2A images. The results showed that (1) DEM was the most important variable for estimating the AGB of coniferous forests, evergreen broadleaved forests, deciduous broadleaved forests, and mixed forests; while the vegetation index was the most important variable for estimating deciduous broadleaved forests, the climatic factors had a higher variable importance for estimating coniferous and mixed forests, and texture features and vegetation index had a higher variable importance for estimating evergreen broadleaved forests. (2) In terms of specific model performance for the four forest types, RRF was the best model both in estimating the AGB of coniferous forests and mixed forests; the R^2 and RMSE for coniferous forests were 0.63 and 43.23 Mg ha⁻¹, respectively, and the R^2 and RMSE for mixed forests were 0.56 and 47.79 Mg ha⁻¹, respectively. BRNN performed the best in estimating the AGB of evergreen broadleaved forests; the R^2 was 0.53 and the RMSE was 68.16 Mg ha⁻¹. QRF was the best in estimating the AGB of deciduous broadleaved forests, with R^2 of 0.43 and RMSE of 45.09 Mg ha⁻¹. (3) RRF was the best model for the four forest types according to the mean values, with R^2 and RMSE of 0.503 and 52.335 Mg ha⁻¹, respectively. In conclusion, different variables and suitable models should be considered when estimating the AGB of different forest types. This study could provide a reference for the estimation of forest AGB based on remote sensing in complex terrain areas with a high degree of forest heterogeneity.

Keywords: environmental factors; forest AGB estimation; forest heterogeneity; forest types; machine learning algorithms; Yunnan Province of China



Citation: Huang, T.; Ou, G.; Wu, Y.; Zhang, X.; Liu, Z.; Xu, H.; Xu, X.; Wang, Z.; Xu, C. Estimating the Aboveground Biomass of Various Forest Types with High Heterogeneity at the Provincial Scale Based on Multi-Source Data. *Remote Sens.* **2023**, *15*, 3550. <https://doi.org/10.3390/rs15143550>

Academic Editors: Xufeng Wang, Shaobo Sun and Kai Liu

Received: 9 June 2023

Revised: 11 July 2023

Accepted: 12 July 2023

Published: 14 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Forest biomass is the basic material of forest ecosystems, which plays a vital role in addressing climate change and studying the carbon cycle [1–4]. Traditional forest biomass acquisition methods are expensive, inefficient, and ecologically damaging. However, remote sensing has become an important tool for forest biomass estimation due to its advantages such as environment friendliness, efficiency, and collection of continuous

data [5,6]. Meanwhile, uncertainties arising from remote sensing data sources, prediction models, and forest heterogeneity remain a major challenge for the accurate estimation of forest aboveground biomass (AGB) at large scales [7,8]. Especially in areas of high forest heterogeneity with complex topography, both active and passive remote sensing have a high degree of uncertainty.

Yunnan Province is located in a longitudinal ridge and valley area, with complex geological conditions; its special geographical location has led to differences in climate and soil across the province, which, in turn, have resulted in high forest heterogeneity [9–12]. Therefore, achieving accurate estimates of forest biomass in this type of area is certainly a challenge [8]. Using optical remote sensing, such as Landsat 8 OLI and Sentinel 2A, to estimate biomass has the advantages of wide coverage, easy access, high spatial and temporal resolution, and mature technology for large-scale monitoring and management of forest ecological resources. Although optical remote sensing results in data saturation as it captures the same object with different spectrum phenomena caused by mountain shadows [13–17], it is still the best choice for remote sensing estimation of forest biomass at a large scale. To improve the accuracy of remote sensing estimates of biomass, environmental factors are often used in synergistic remote sensing estimates of forest biomass. For example, Silveira et al. (2019) showed that incorporating environmental factors into remote sensing to estimate AGB reduced uncertainties in highly heterogeneous stands, such as data saturation problems [18]. Liu et al. (2021) showed that climatic heterogeneity best explained biodiversity distribution patterns in natural forests, and temperature and precipitation not only positively correlated with biodiversity but were also the main drivers of natural vegetation biodiversity patterns in Yunnan Province [19]. In addition, Yu et al. (2022) [20] showed that elevation and climate data could improve AGB estimation using remote sensing, especially for large-scale study areas with large biomass gradients.

A large number of studies have been conducted on AGB estimation using a single remote sensing variable or a combination of multiple remote sensing variables in different regions [10,21–24]. However, few studies have compared which variables are the main and the secondary influencing factors for different forest types in large-scale topographically complex areas with high forest heterogeneity to further improve estimation accuracy. Meanwhile, collaborative estimation of forest biomass by using multiple sources of remote sensing data has become a popular research topic [10,23,24]. However, integrating multiple remote sensing data sources faces problems such as data noise and interference and information redundancy. Thus, selecting the best feature variables is a key step for model construction [25,26]. For variable selection methods, Boruta is a heuristic algorithm based on a random forest learner, which is a good choice especially when the number of variables is too large or exceeds the number of sample plots [24]. Many studies have also shown that variable selection using Boruta's algorithm could solve problems such as data redundancy [27–29]. For example, Zhang et al. (2023) [30] compared Boruta with other variable screening models and showed that the estimation result was better after using the Boruta algorithm for variable screening when compared with the other algorithms.

Models play a crucial role in estimating forest biomass using remote sensing and significantly contribute to the uncertainty associated with remote sensing estimations; the selection and performance of models directly impact the accuracy and reliability of biomass estimates [31,32]. Thus, it is important to select a suitable algorithm for AGB estimation. The main advantage of machine learning algorithms is their ability to capture complex non-linear relationships between remote sensing data variables and forest AGB, which could significantly improve accuracy compared to traditional algorithms [24]. Therefore, many machine learning algorithms have been widely used for estimating AGB of various forest types. Ronoud et al. (2021) [33] compared the estimation effectiveness of different algorithms, such as k-nearest neighbors (k-NN), for AGB remote sensing estimation of broadleaved forests and found that the k-NN algorithm outperformed other algorithms. Zhang et al. (2020) [34] evaluated the estimation effectiveness of eight algorithms, including gradient boosting machine (GBM), in global-scale coniferous, broadleaved, and

mixed forests and found that the integrated algorithm had better estimation effectiveness. Jiang et al. (2021) [35] compared the estimation effectiveness of random forests (RF) with other algorithms in coniferous forests and found that RF has good estimation effectiveness. Durante et al. (2019) [36] used quantile random forest (QRF) to carry out remote sensing estimation of forests in the Region of Murcia, Spain, which remarkably improved the accuracy of estimation. Meanwhile, regularized random forests (RRF) and Bayesian regularization neural network (BRNN) algorithms have received much attention in other fields, but they have hardly been used for remote sensing forest biomass estimation. For instance, Band et al. (2020) [37] chose RRF to evaluate a model of mountain flooding vulnerability of Kalvan basin, Markazi Province, Iran, and found that RRF was superior to RF algorithm in learning. Fikret et al. (2018) [38] used extreme learning machine (ELM), BRNN, and SVM (support vector machine) to model and predict clay compression index, and found that the BRNN estimation results were better than the other algorithms' results. Although some of these algorithms have achieved good results in biomass estimation in areas of low forest heterogeneity at the regional scale or in plains [39], there is limited research on which model is more accurate when comparing different forest types with high forest heterogeneity over a large and complex terrain, and which model is more accurate for the same forest type.

There have been few studies on estimating forest AGB by comparing the importance of variables and the performance of models for different forest types in complex and heterogeneous terrains at the provincial scale. To address this gap in research, this study integrated Landsat OLI and Sentinel 2A remote sensing data, combined with ground survey data, integrated environmental factors, elevation, vegetation indices, and texture factors. Then, this study used the Boruta variable screening method to determine the main influencing factors for different forest types in highly heterogeneous areas, and compared the accuracy performance of six (RRF, QRF, BRNN, RF, GBM, and k-NN) machine learning algorithms for different forest types, among which RRF and BRNN have rarely been used for AGB remote sensing estimation.

The aims of this study were as follows:

- (1) To explore the most effective variables for AGB estimation in different forest types in large-scale areas with complex geography and high forest heterogeneity.
- (2) To analyze model accuracy for estimating AGB in different forest types and in the same forest type by comparing six machine learning models.

2. Materials and Methods

2.1. Materials

2.1.1. Study Area

Yunnan Province is located on the Yunnan–Guizhou Plateau in southwest China, and the coordinates are $97^{\circ}31'–106^{\circ}11'E$ and $21^{\circ}8'–29^{\circ}15'N$. Bordering the southeastern edge of the Tibetan Plateau, the terrain is predominantly comprised of mountains and highlands, with a total area of approximately 394,000 square kilometers [9,40]. Its altitude shows a downward trend from northwest to southeast at an altitude of 74–6457 m. Yunnan has a highland tropical monsoon climate with average summer and winter temperatures of 19–22 °C and 6–8 °C, respectively. Precipitation is very uneven across seasons and regions. The dry season is from November to April, with only 1100 mm of annual rainfall. Yunnan has rich and diverse forest resources, including tropical rainforests, seasonal rainforests, subtropical evergreen broadleaved forests, and temperate coniferous forests [41]. The study area is shown in Figure 1.

2.1.2. Remote Sensing Data Acquisition and Variable Extraction

Sentinel 2A and Landsat 8 OLI data were downloaded from Google Earth Engine (<https://code.earthengine.google.com/> (accessed on 20 January 2023)) to match the survey data. The image data are surface reflectance products that were selected with less than 3% cloud shadow and 5% cloud to calculate the median values from January to

December in 2021 for the Yunnan Province. The Landsat 8 OLI data were from “LANDSAT/LC08/C01/T1_SR” and the Sentinel 2A data were from “COPERNICUS/S2_SR” in Google Earth Engine. The image synthesis was conducted on 20 January 2023, and the images were resampled to 30 m × 30 m. Subsequently, a 30 m resolution DEM was used for terrain correction of the Sentinel 2A and Landsat 8 OLI data. The vegetation indices as well as the single band and texture features were calculated using ENVI 5.3 [39,42]. The Landsat 8 OLI data included 7 spectral bands, 17 vegetation indices, and 168 texture variables (3 × 3, 5 × 5, and 7 × 7 from the gray-level co-occurrence matrix (GLCM)), and the Sentinel 2A data included 12 spectral bands, 18 vegetation indices, and 288 texture variables (3 × 3, 5 × 5, and 7 × 7 from the gray-level co-occurrence matrix feature (GLCM)). All spectral variables are shown in Table 1.

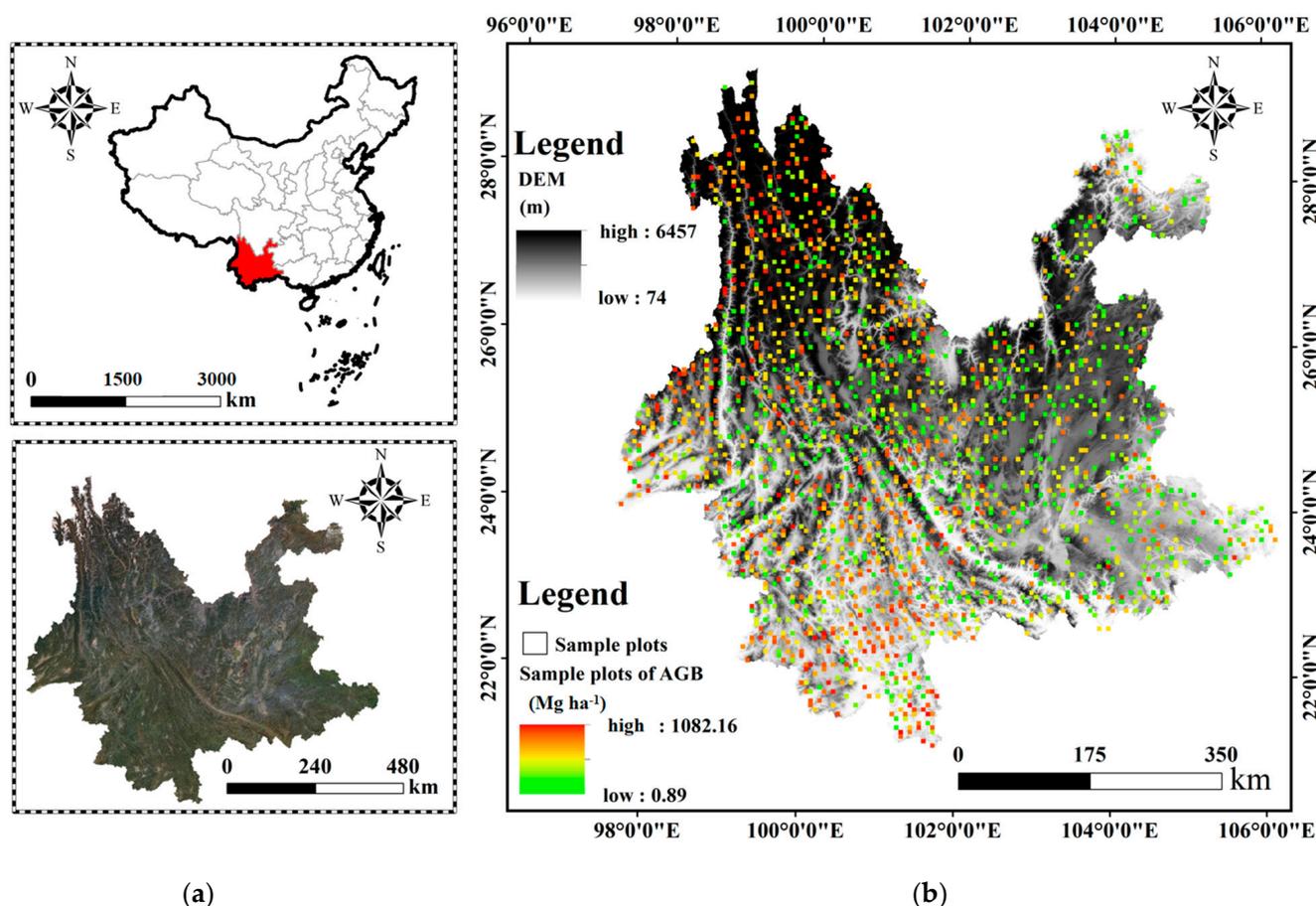


Figure 1. Location of study area: (a) the location of Yunnan Province in China, and (b) DEM data and distribution sample plots in Yunnan Province (from green to red indicating low to high).

Table 1. Spectral variables.

Image Source	Spectral Variables
Sentinel 2A	single band, RVI (ratio vegetation index), DVI (difference vegetation index), WDV (weighted difference vegetation index), IPVI (infrared vegetation index), PVI (perpendicular vegetation index), NDVI (normalized difference vegetation index), NDVI45 (NDVI with band4 and band5), GNDVI (NDVI of green band), IRECI (inverted red edge chlorophyll index), SAVI (soil-adjusted vegetation index), TSAVI (transformed soil-adjusted vegetation index), MSAVI (modified soil-adjusted vegetation index), S2REP (Sentinel-2 red edge position index), REIP (red-edge infection point index), ARVI (atmospherically resistant vegetation index), PSSRa (pigment-specific simple ratio chlorophyll index), MTCI (Meris terrestrial chlorophyll index), MCARI (modified chlorophyll absorption ratio index)

Table 1. Cont.

Image Source	Spectral Variables
Landsat 8 OLI	single band, NDVI (normalized difference vegetation index), ND43 (NDVI with band3 and band4), ND67 (NDVI with band6 and band7), ND563 (NDVI with band3 and band5 with band6), DVI (difference vegetation index), SAVI (soil-adjusted vegetation index), RVI (ratio vegetation index), B (brightness vegetation index), G (greenness vegetation index), W (temperature vegetation index), ARVI (atmospherically resistant vegetation index), MV17 (mid-infrared temperature vegetation index), MSAVI (modified soil-adjusted vegetation index), VIS234 (multiband linear combination of band2 with band3 and band4), ALBEDO (multiband linear combination), SR (simple ratio index), SAV12 (improved vegetation index), MSR (optimized simple ratio vegetation index), KT1, PC1-A, PC1-B, PC1-P
Sentinel 2A/ Landsat 8 OLI	Mean (ME), Var (VA), Homogeneity (HO), Contrast (CN), Dissimilarity (DI), Entropy (EN), Second Moment (SM), Correlation (CO)

2.1.3. Environmental Data Collection

The 19 bioclimatic factors from 1950 to 2000 were derived from World Climate (<http://www.worldclim.org/> (accessed on 20 March 2022)) at a spatial resolution of 30'' (1 km × 1 km). The 15 soil factors were produced from the 1:1 million soil data points provided by the Cold and Arid Regions Science Data Centre of the Chinese Academy of Sciences (<http://westdc.westgis.ac.cn> (accessed on 15 March 2022)) from the Nanjing Soil Institute of the Second National Land Survey, with a raster size of approximately 1 km². The DEM data were obtained from the Geospatial Data Cloud (<http://www.gscloud.cn/> (accessed on 17 January 2022)) at a spatial resolution of 30 m × 30 m. A total of 37 environmental factors were used in this study, including 19 climatic factors, 15 soil factors, and 3 topographic factors, as shown in Table 2.

Table 2. Overview of the 37 environmental factors used in this study.

Variables	Description	Variables	Description
Bio_1	Annual mean temperature	T_BS	Base saturation in the topsoil
Bio_2	Mean diurnal range	T_CEC_CLAY	Cation-exchange capacity of the clay fraction in the topsoil
Bio_3	Isothermality	T_CEC_SOIL	Cation-exchange capacity in the topsoil
Bio_4	Temperature seasonality	T_ESP	Exchangeable sodium percentage in the topsoil
Bio_5	Max. temperature of the warmest month	T_SAND	Percentage of sand in the topsoil
Bio_6	Min. temperature of the coldest month	T_SILT	Percentage of silt in the topsoil
Bio_7	Range of annual temperature	T_USDA_TEX	Topsoil texture class variable and code
Bio_8	Mean temperature of the wettest quarter	T_CLAY	Percentage of clay in the topsoil
Bio_9	Mean temperature of the driest quarter	T_OC	Percentage of organic carbon in the topsoil
Bio_10	Mean temperature of the warmest quarter	T_REF_BULK	Topsoil reference bulk density
Bio_11	Mean temperature of the coldest quarter	T_ECE	Electrical conductivity of the topsoil
Bio_12	Annual average precipitation	T_GRAVEL	Volume percentage of gravel in the topsoil
Bio_13	Precipitation of the wettest month	T_CACO ₃	Percentage of carbonate carbon in the topsoil
Bio_14	Precipitation of the driest month	T_pH_H ₂ O	Topsoil pH
Bio_15	Precipitation seasonality	T_TEB	Total exchangeable bases in the topsoil
Bio_16	Precipitation of the wettest quarter	DEM	DEM elevation
Bio_17	Precipitation of the driest quarter	SLOPE	Slope
Bio_18	Precipitation of the warmest quarter	ASPECT	Aspect
Bio_19	Precipitation of the coldest quarter		

2.1.4. Data Collection from Sample Plots and Forest AGB Calculation

The ground data were collected from systematic sampling of 1776 sample plots from the CFI (continuous forest inventory) in Yunnan Province. The plot size was 25.8 m × 25.8 m, and the sample plots were evenly distributed across Yunnan Province (Figure 1). The basic information was recorded, such as the dominant species, diameter at breast height (DBH) of individual trees, tree height, age class, average height, stand conditions, and coordinates. We calculated individual tree volume according to a table of timber volume of tree species

(groups) in Yunnan Province and calculated the timber volume of each plot according to Xu et al. (2019) [43]. The equation is as follows:

$$AGB = V \times SVD \times EBF \quad (1)$$

where *AGB* is aboveground biomass by plot (Mg ha^{-1}); *V* is volume by plot (m^3/ha); *SVD* is the basic wood density of the corresponding dominant species (Mg ha^{-1}); and *BEF* is the biomass conversion factor of the corresponding dominant species (dimensionless).

In this study, the sample plots were divided into four types, including coniferous, evergreen broadleaved, deciduous broadleaved, and mixed forests according to the dominant tree species. Table 3 shows the basic information of the samples of the four forest types. In this study, 70 percent of the samples were used for modeling and 30 percent were used as the test samples. Among them, the evergreen broadleaved forest has the largest number of samples and the widest forest AGB range, indicating that it has the highest forest heterogeneity in terms of quantity.

Table 3. Statistics of the sample plot data used in this research.

Forest Types	Total Samples	AGB Range (Mg ha^{-1})	Mean (Mg ha^{-1})	Training Samples	Testing Samples
Coniferous forest	473	1.13–593.49	110.69	330	141
Evergreen broadleaved forest	984	0.88–1082.16	107.78	668	296
Deciduous broadleaved forest	151	3.86–536.81	62.83	105	46
Mixed forest	168	5.27–359.39	105.67	117	51

2.2. Methods

The flowchart of this study is shown in Figure 2. This study used Landsat 8 OLI, Sentinel 2A, and environment factors as data sources, as well as CFI data from 1776 sample plots surveyed in Yunnan Province. Six machine learning algorithms (RRF, QRF, BRNN, RF, GBM, and k-NN) were used for AGB estimation of coniferous, evergreen broadleaved, deciduous broadleaved, and mixed forests based on the variables selected by the Boruta algorithm.

2.2.1. Variable Selection

Boruta is a heuristic algorithm based on a random forest learner whose core idea is to construct shadow features by training the original real features and aggregating the original features and shadow features into feature matrixes for training. The set of features associated with the dependent variables is selected from the original true features using the feature importance scores of the shaded features as a reference. In addition, to make it easier to qualitatively assess the importance of variables, the Boruta algorithm generates feature importance values along with three types of features (confirmed, tentative, and rejected) for qualitative evaluation, and the variables are selected based on feature confirmation [27–29]. In this study, variable selection was implemented in the R software with the Boruta package.

2.2.2. Machine Learning Algorithm

(1) Quantile Random Forest (QRF)

Quantile regression forest is a generalization of quantile regression, where, for each node in each tree, RF keeps only the average of the observations belonging to that node and ignores all other information. In contrast, quantile random forest (QRF) keeps all observations at a node and takes into account the spread of response variables, allowing the construction of prediction intervals that contain new observations with a high probability. While general regression models predict the mean, QRF models predict the distribution of data. These models could be used to predict the distribution of biomass across quartiles, and they are usually much more demanding in terms of computational power than linear

regression models [44]. In this study, QRF was implemented in the R language using a caret package, and the subsite was 0.5.

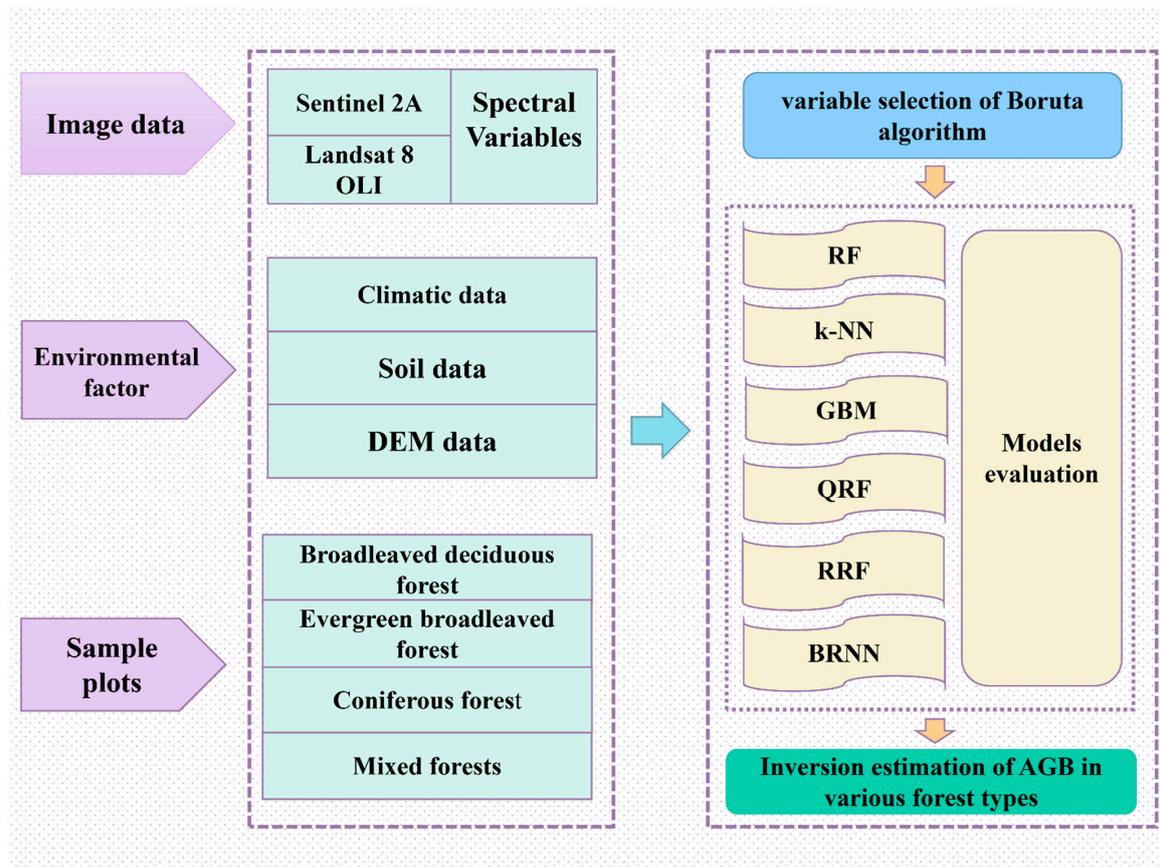


Figure 2. Flowchart of the study: Integrating environmental factors for remote sensing estimation of different forest types in Yunnan Province using multiple machine learning algorithms. Note: QRF (quantile random forest algorithm), BRNN (Bayesian regularization neural network algorithm), and RRF (regularized random forests).

(2) Bayesian Regularization Neural Network (BRNN)

BRNN is a reverse neural network for Bayesian regularization training, and one of the difficulties in designing a neural network model is determining the number of hidden neurons. Too many neurons would lead to overfitting, and conversely, networks with an insufficient number of hidden nodes would have learning difficulties; both neural network models that are too simple and those that are too complex have a poorer predictive performance. To overcome this problem, the Bayesian regularization theorem is applied to limit the scaling of thresholds and weights to improve the regularization ability of the neural network. The main advantages of the BRNN method are its ability to determine the optimal network structure, its ability to avoid overfitting and under learning, and its good robustness [45,46].

(3) Gradient Boosting Machine (GBM)

GBM combines the features of the gradient boosting algorithm system to obtain better prediction results through multiple iterations of computation, resulting in a continuous reduction in the overall loss and an increase in model performance. In addition, GBM inherits the advantages of single decision trees, including being insensitive to meaningless data and having a better learning ability for complex non-linear relationships, while also avoiding overfitting by controlling the number of iterations [34].

(4) Random Forests (RF)

A random forest model (RF) is an advanced integrated algorithm that determines the final result by constructing many decision trees and combining the average of all of them, showing excellent robustness and an easy-to-understand feature selection process. RF has been widely used in areas such as remote sensing estimation of forest biomass, and has an excellent learning effect [34].

(5) Regularized Random Forests (RRF)

Random forests form regularized random forests (RRF) utilizing a regularization strategy for the generated trees, thus selecting a subset of compressed features; the main difference from the original random forest is the application of regularized information gain [47].

(6) k-Nearest Neighbors (k-NN)

The k-NN algorithm is a common algorithm for remote sensing estimation of forest biomass. The basic principle is that the k-NN algorithm calculates the spectral distance between the spectral information value of a sample site's location and the estimated image element, and it then calculates the weighted average of the forest biomass values of the k nearest sample sites using the Euclidean distance or the Marxian distance. The more similar the image element information value of a sample site is to the estimated image element information value, the greater the weight is [48].

Six machine learning algorithm models were constructed in the R software using the CARET package, and a grid search was performed with 10-fold cross-validation to optimize the parameters.

2.2.3. Model Evaluation

The coefficient of determination (R^2) and root mean square error (RMSE) metrics were used for model evaluation. The equation is as follows:

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

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

where n is the number of sample observations; y_i is the actual value; and \hat{y}_i is the estimated value and \bar{y}_i is the mean of the observed sample.

3. Results

3.1. Importance of Variables for AGB Modeling

In total, 542 variables were obtained, including 19 single bands, 30 vegetation indices, 37 environment factors, and 456 texture features. The Boruta algorithm was used to choose the most important variables. The selected results are shown in Figure 3. A total of 15 variables were selected for the model construction of coniferous forests, and the results showed that DEM had the highest importance value, followed by the climate factors which included eight variables. The texture features and vegetation indices were in the third and fourth order, respectively. For evergreen broadleaved forests, a total of 26 variables were selected; similarly, the DEM factor had the highest importance value, and the other 25 variables were texture features and vegetation indices, which were extracted from the remote sensing images. For deciduous broadleaved forests, three variables were selected, all of which were vegetation indices from the Landsat 8 OLI data. For mixed forests, four variables were selected, and each of them came from the climate factors, topographic factors, vegetation indices, and texture features; the importance order of the four variables was DEM > b4_L8_ME7 > bio_6 > S2_WDVI. However, no soil factor was selected for

model construction for all forest types (the explanation is provided in the discussion section of this article).

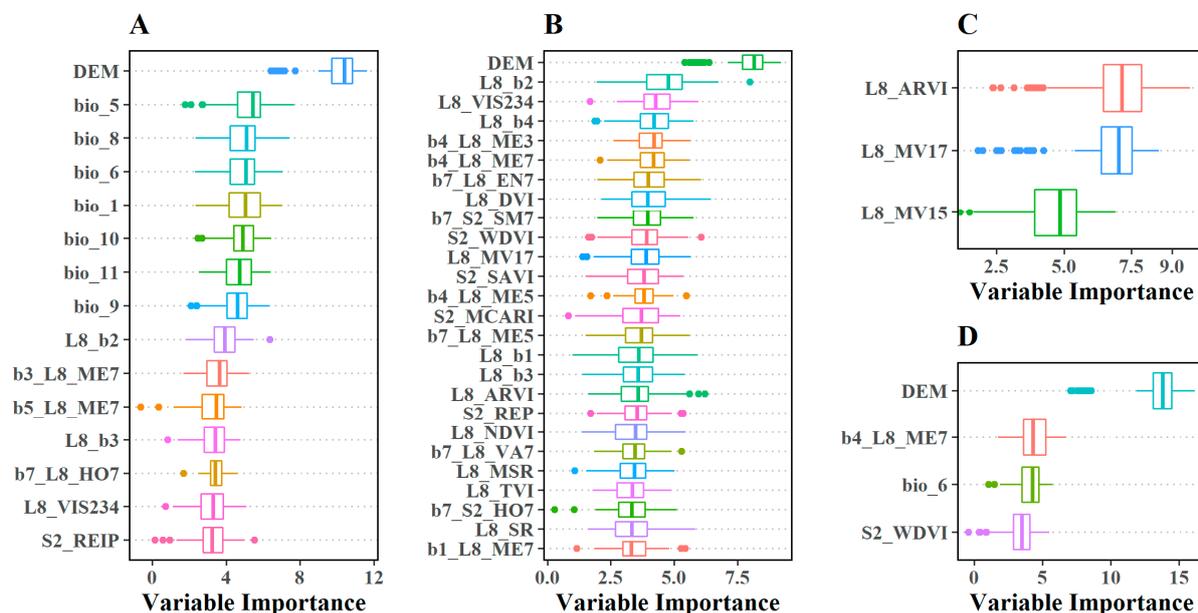


Figure 3. Variable selection results using the Boruta algorithm for different forest types (note: (A) stands for coniferous forests, (B) stands for evergreen broadleaved forests, (C) stands for deciduous broadleaved forests, (D) stands for mixed forests, L8 stands for Landsat 8 OLI, and S2 stands for Sentinel 2A).

3.2. AGB of Different Forest Types Estimated Using Remote Sensing

Six models were applied to evaluate the AGB of the four forest types based on the variables selected by the Boruta algorithm. Figure 4 shows the results of the models according to the sample independence test. R^2 and RMSE were the evaluation indicators. The results showed that (1) the performance of the four forest types differed by different models: RRF performed the best in estimating the AGB of both coniferous forests and mixed forests, while it ranked third for evergreen broadleaved forests and fourth for deciduous broadleaved forests. The R^2 and RMSE values for coniferous forests were 0.63 and 43.23 Mg ha⁻¹, and the R^2 and RMSE values for mixed forest were 0.56 and 47.79 Mg ha⁻¹. BRNN performed the best in estimating evergreen broadleaved forests, with the R^2 value being 0.53 and the RMSE being 68.16 Mg ha⁻¹; beyond that, BRNN performed the second worst for the other forest types. QRF was best in estimating deciduous broadleaved forests, with R^2 of 0.43 and RMSE of 45.09 Mg ha⁻¹; the fitting performance of QRF for mixed forests was second, while its performance was the same for coniferous forests and evergreen broadleaved forests. (2) The performance of the six models for the same forest type was as follows: For coniferous forests, the model fitting performance was RRF > RF > QRF > BRNN > k-NN > GBM; the range of R^2 was from 0.49 to 0.63, and the range of RMSE was from 43.23 to 52.53 Mg ha⁻¹. For broadleaved evergreen forests, the order was BRNN > RRF > QRF > RF > GBM > k-NN; the greatest R^2 and the smallest RMSE were 0.53 and 68.16, respectively, and the poorest R^2 was 0.42 with the highest RMSE of 77.12 Mg ha⁻¹. For deciduous broadleaved forests, the order was QRF > GBM > RRF > RF > BRNN > k-NN; the accuracy was worst at the overall level, with the range of R^2 from 0.19 to 0.43 and the range of RMSE from 45.09 to 57.85 Mg ha⁻¹. For mixed forests, the order was RRF > QRF > GBM > BRNN > RF > k-NN; the range of R^2 and RMSE was 0.42–0.56 and 47.79–55.93 Mg ha⁻¹, respectively. Except for coniferous forests, the k-NN model's fitting effect was the worst for the other three forest types. Furthermore, the mean values of the evaluation metrics calculated using the machine learning algorithms for the four forest types showed that (Table 4) the RRF, BRNN, and

QRF algorithms outperformed the RF, k-NN, and GBM algorithms, with the RRF being the best model.

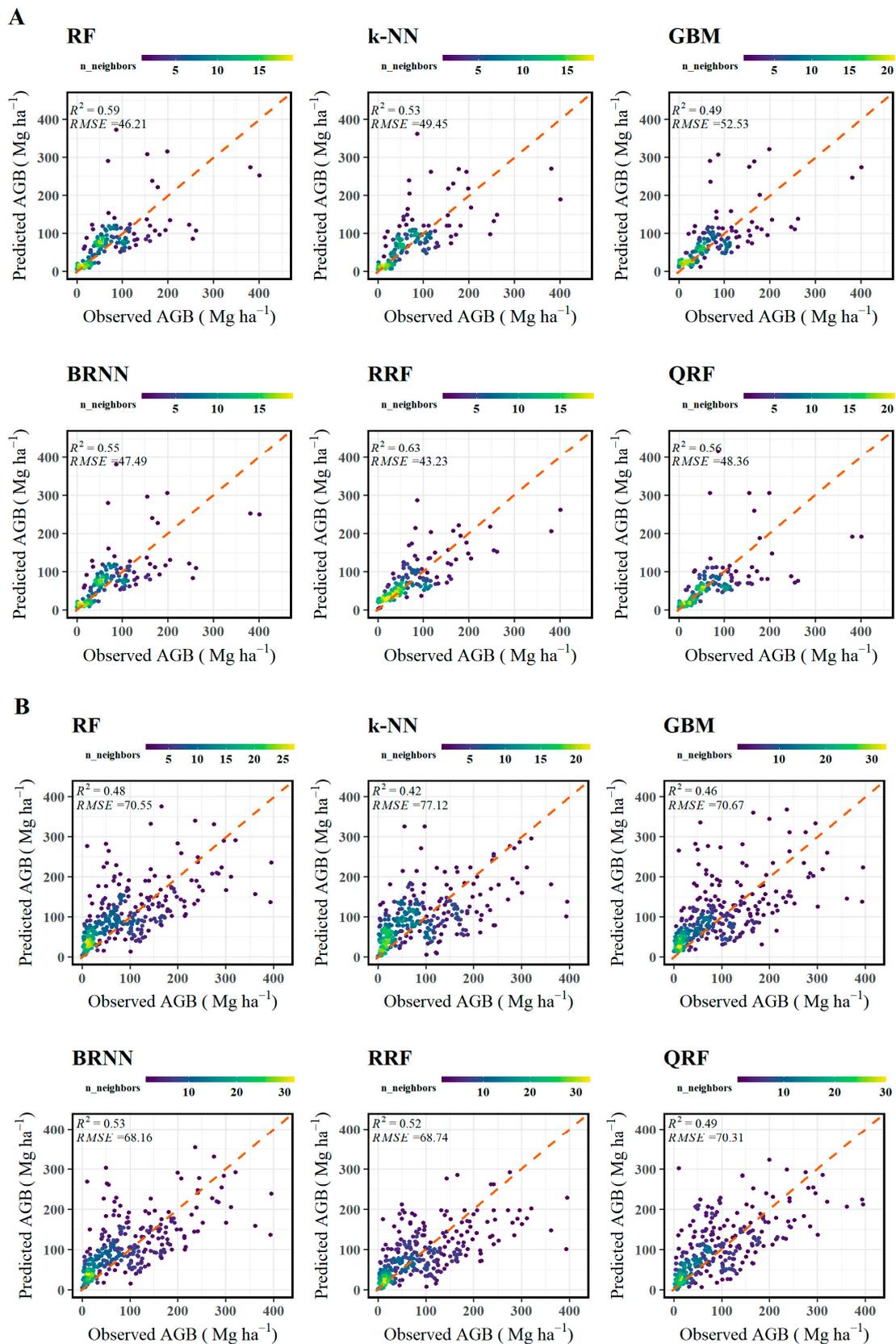


Figure 4. Cont.

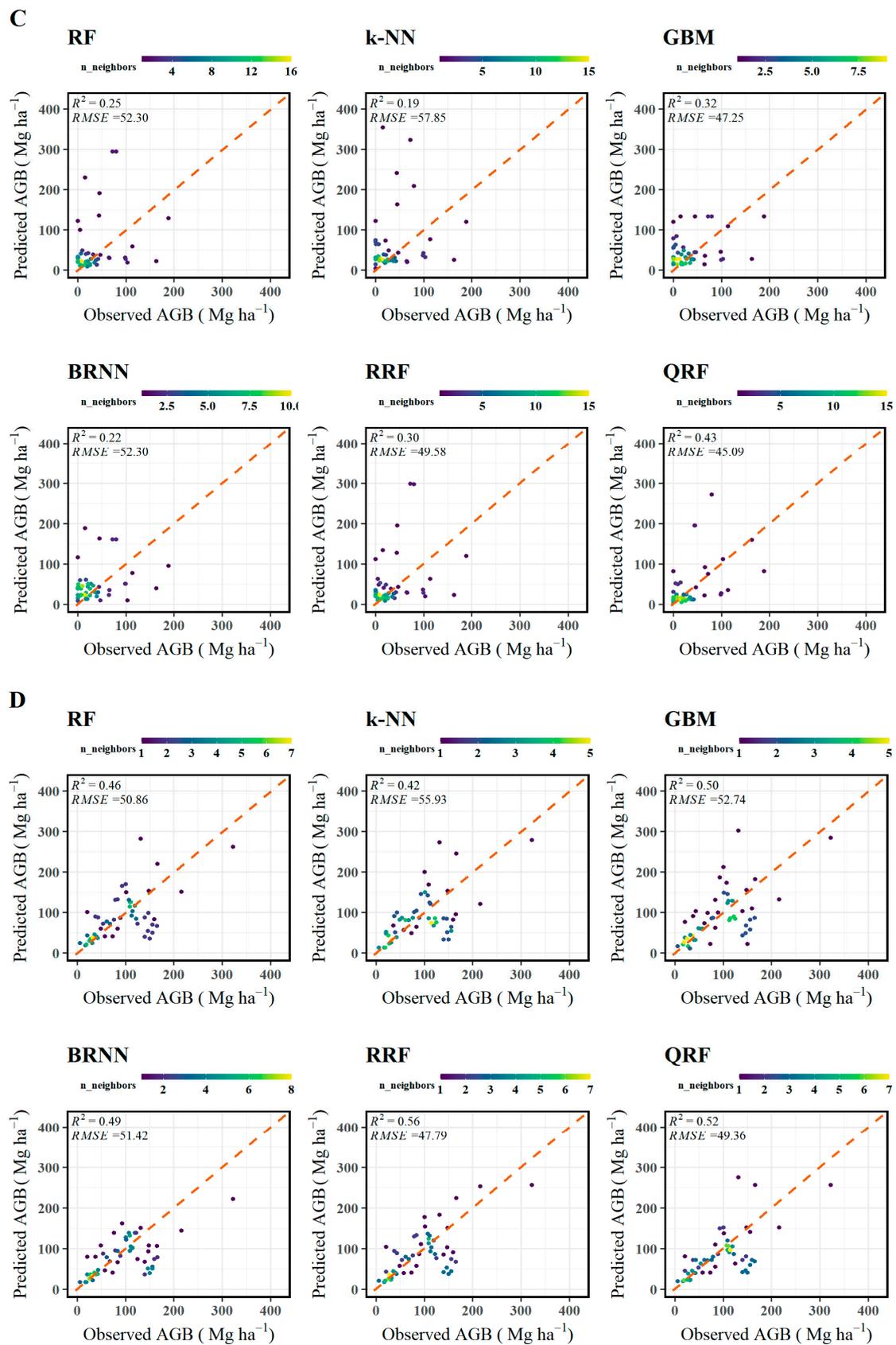


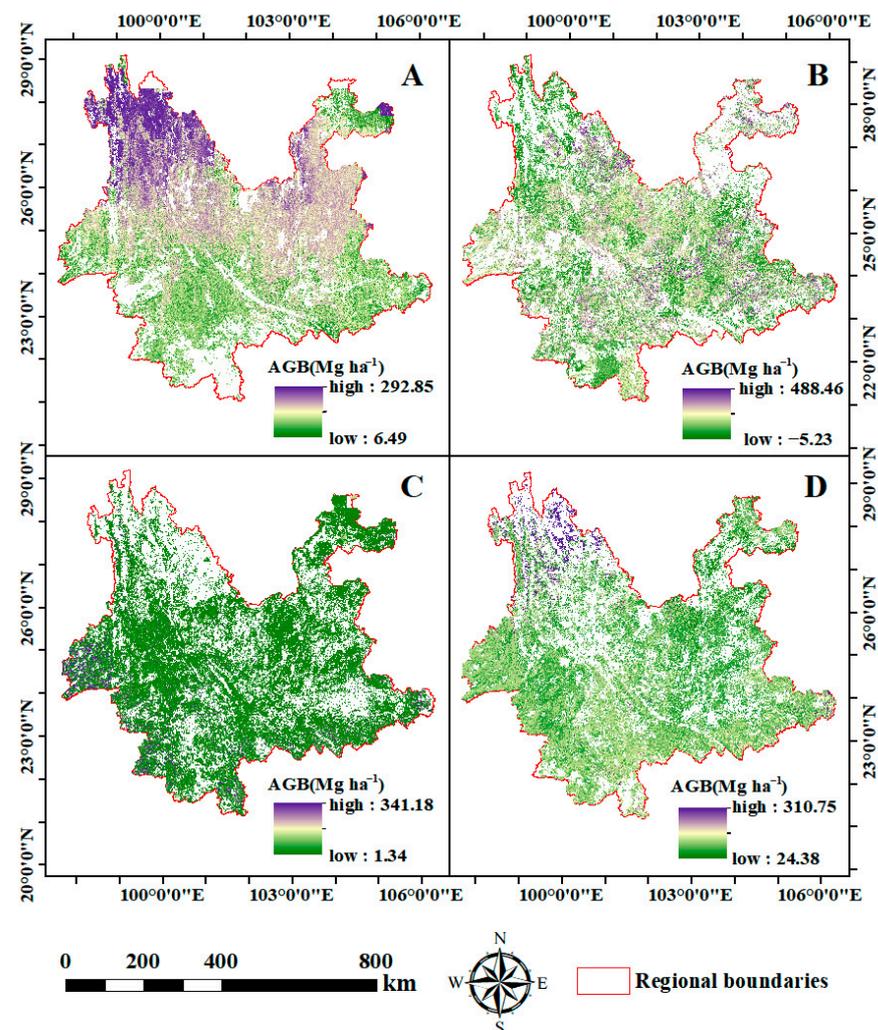
Figure 4. Evaluation results of the models' sample independence test (note: (A) stands for coniferous forests, (B) stands for evergreen broadleaved forests, (C) stands for deciduous broadleaved forests, and (D) stands for mixed forests).

Table 4. Mean values of evaluation metrics calculated using different machine learning algorithms for the four forest types.

Models	R ²	RMSE (Mg ha ⁻¹)
RF	0.445	54.978
k-NN	0.390	60.088
GBM	0.443	55.798
BRNN	0.448	54.843
RRF	0.503	52.335
QRF	0.500	53.280

3.3. Forest Biomass Inversion Estimation

Figure 5 shows the forest biomass mapping results of the four forest types based on the optimal model, with the forest sub-compartment boundary as the unit. The inversion results show that coniferous forests have the highest heterogeneity and deciduous broadleaved forests have the worst heterogeneity, which, to some extent, indicates that the integrated environmental factors based on the optical remote sensing data of Landsat 8 OLI and Sentinel 2A have a better estimation ability in estimating the AGB of coniferous forests compared to evergreen broadleaved forests, deciduous broadleaved forests, and mixed forests in Yunnan Province.

**Figure 5.** Estimated AGB inversions for the four types of forests in Yunnan Province (note: (A) stands for coniferous forests, (B) stands for evergreen broadleaved forests, (C) stands for deciduous broadleaved forests, and (D) stands for mixed forests).

4. Discussion and Conclusions

4.1. Discussion

4.1.1. Variable Selection for AGB Models

In this study, DEM had the highest importance in the remote sensing estimation model for coniferous, deciduous broadleaved, evergreen broadleaved, and mixed forests, suggesting that forest AGB has a strong correlation with DEM. The complex topography of Yunnan Province creates a large difference in altitude, which has a huge impact on the growth of its forests (the range of DEM was 74–6457 m). The reasons are that (1) forest biomass varies with microtopography and soil nutrient content. In general, AGB is lower at higher altitudes as the temperature is lower, the air is thinner, and UV light is stronger, all of which limit plant growth. In contrast, AGB is much higher at lower altitudes [49,50]. (2) AGB is significantly lower at higher altitudes as the soil moisture and nutrient conditions are poorer. The higher the altitude, the stronger the solar radiation, which leads to greater evaporation, as well as weak water-holding capability, because the lower plant richness results in less litter [4,51–53]. (3) The deficiency in soil nutrient in high elevation may be caused by high-intensity radiant heat, strong wind, and low humidity. Meanwhile, vegetation types vary according to the different DEM, mainly due to different DEM having different quantities of heat distribution, a wider range of temperature, and distinguished soil conditions [54–56]. Zhang et al. (2014) [57] also pointed out that adding DEM to vegetation remote sensing classification may be a good way to improve accuracy, and altitude determines the distribution of vegetation types in the mountainous areas of Yunnan. However, contrary to our hypothesis, the soil factors were not selected for modeling in this study. Similarly, Bennett et al.'s study also showed that soil factors did not improve the estimated model when modeling the AGB of Australian forests using climate and soil factors [58]. Soil characteristics have the potential to directly determine the type of vegetation that can be supported (e.g., grassland versus forest) and, thus, influence the structural and functional characteristics of that vegetation type. As our analysis was limited to forests, the effect of soil may be limited to its influence on forest type characteristics, rather than having a greater influence on the biomass distribution of the same forest type [58]. In addition, the soil data used in this study are simulated, coarse-resolution data. If the measured soil data were included in the model construction, the significance of the soil factors in estimating forest AGB could be improved.

Temperature variation results in forest change by affecting species diversity, CO₂, and energy exchange in the stand, thus altering vegetation types and forest boundaries [59–62]. In this study, a total of seven environmental variables were selected for the model construction of coniferous forests, indicating that the temperature factors were highly correlated with the biomass of coniferous forests. Several studies have also documented this phenomenon. For instance, Li et al. showed that among the vegetation types in Yunnan Province, cold-temperate coniferous forests are vulnerable to climatic influences because they have the highest elevation among the forest vegetation types [63]. Ma et al. (2014) [64], Zhou et al. (2018) [65], and Ni et al. (2010) [66] also showed that coniferous forests are more sensitive to temperature changes than broadleaved forests in Yunnan and elsewhere. Moreover, Dakhil et al. (2019) [67] showed that temperature is the main influence in coniferous forests in southwest China. The ecological performance and species composition of evergreen broadleaved forests in Yunnan Province are complex, with associated tree species exacerbating the complexity of the community structure, which is affected by the southwest monsoon and plateau landscape [68]. For evergreen broadleaved forests, 28 variables were selected to construct the model, of which 11 were texture features. For complex stand structures, shadows caused by terrain and spectral changes reduce the estimation accuracy. Considering that the relationship between spatial and pixels could represent a change in image gray level, it could be used to improve the recognition ability of spatial information and the AGB estimation effect [69–71]. Deciduous broadleaved forests are mainly distributed in parts of the low hills and middle mountains of Yunnan, and the area is small and sporadic with a simpler structure than that of evergreen broadleaved forests [68].

Three vegetation indices in deciduous broadleaved forests were selected to participate in the model construction, which echoed the fact that vegetation indices with infrared bands have better estimation in areas with a simple stand structure [72]. Texture characteristics, vegetation indices, DEM, and environmental factors were selected to participate in the construction of the model of mixed forests, and the characteristics of each variable were combined in the construction of the model, which could overcome the shortcomings of a single variable and improve the estimation effect to a certain extent [24].

4.1.2. Remote Sensing Estimation of Different Forest Types

According to the estimation results, the estimation effect of coniferous forests was better than that of mixed forests. The estimation for mixed forests was better than for evergreen broadleaved forest since, in Yunnan, the structure of the coniferous forest is simple, the structure of evergreen broadleaved forests is more complex, and the structure of mixed forests is between those of evergreen broadleaved forests and coniferous forests [68]. These results were consistent with Lu et al. [24], showing that the effect of AGB remote sensing estimation was better in areas with simple forest structures. However, even though the structure of deciduous broadleaved forests is relatively simple [68], the estimation effect was the worse in this study, which might be because data from the sample plots of NFI were collected over at least two seasons, and the estimation for deciduous broadleaved forests was more obvious than that of deciduous broadleaved forests. It is more difficult to estimate forest biomass via remote sensing. For example, Singh et al. (2022) [73] studied deciduous forests using AGB remote sensing estimation in India and obtained high accuracy for the rainy season; in contrast, the accuracy of the adjusted R^2 range was from -0.05 to 0.43 in the dry season.

Although environmental factors improved the estimation effect to a certain extent, the overall accuracy was not high. For example, for coniferous forests, the R^2 and RMSE values were only 0.63 and 43.23 Mg ha^{-1} , respectively, which showed the large gap in the AGB estimation accuracy. That might lead to uncertainty during the process of AGB estimation, for instance, in inventory data acquisition, remote sensing imagery, estimation of forest canopy structure and vegetation type, and data saturation issues, especially in areas with high forest heterogeneity due to the complex biophysical environment [48,74,75]. The survey data collection period lasted too long, thus making it impossible to obtain images that accurately matched the field survey data, which might be an important reason for the low estimation accuracy.

4.1.3. Limitations and Future Research

In this study, the variable screening was based on different forest types. Different forest types have different biomass accumulation processes due to different environmental and ecological processes. Therefore, variable selection for specific forest types can better reflect the characteristics of different forest types and help understand the correlation between forest types and biomass or other target variables. In addition, the estimation performance of the six different algorithms for different forest types was compared. This study provided a comprehensive exploration of the variables and algorithms of different forest types. Though this study provides an important reference value and a significant guidance for future research, there are some limitations. The classification of all forests into four types in this study was coarse on the taxonomic scale and might be one of the reasons for the low estimation accuracy, which could be improved by refining the classification of forest types by species, forest canopy structure, and geography in future study. Radar and high-resolution optical remote sensing techniques could improve AGB estimation because these techniques could provide richer vegetation spectral characteristics and vertical distribution information [76–79]. Such techniques could be used in future research to explore their suitability in regions with high heterogeneity. However, choosing the right algorithm for the AGB remote sensing estimation of specific forest types is a key step to improve accuracy. There are many excellent machine learning algorithms, such as

deep learning (long short-term memory (LSTM), convolutional neural network (CNN), group method of data handling (GMDH), adaptive neuro-fuzzy inference system (ANFIS), generalized regression neural network (GRNN), etc.), extreme gradient boosting (XGBoost), and stacking ensemble learning. The fitting performance of each type of model needs to be further researched for various forest types. In this study, we considered the DEM as a variable and performed topographic correction of the images. In future studies, we can further explore the comprehensive influence of terrain factors in complex terrain areas, such as the temperature depression effect caused by terrain, on remote sensing estimation of forest biomass, as well as hierarchical estimation of forest biomass based on terrain, elevation, slope, and slope direction.

4.1.4. Practical Applications

Considering DEM in the remote sensing estimation of forest AGB in large complex terrains with a high forest heterogeneity can improve forest biomass estimation. Texture characteristics can play a significant role in evergreen broadleaved forests with a more complex stand structure, while the correlation between vegetation indices and forest biomass are stronger in simple deciduous broadleaved forests. Coniferous forests are more sensitive to temperature, so the temperature factor should be taken into account when estimating the AGB of coniferous forests. Different models yield varying estimation effects in different forest types. Comparing several algorithms across different forest types and selecting the best algorithm for estimating forest AGB is crucial in improving the accuracy of AGB estimation using remote sensing.

4.2. Conclusions

In this study, Yunnan Province, which has a high forest heterogeneity and a complex topography, was selected as the study area. Landsat 8 OLI and Sentinel 2A images were integrated as the data source, and the Boruta algorithm was used to screen important variables. Six machine learning algorithms, including QRF, BRNN, RRF, GBM, RF, and k-NN, were applied to estimate the AGB of different forest types. The results are listed below:

(1) Among the environmental factors, the climate factors were more sensitive than the soil factors. For the topographic factors, DEM was the most important variable for estimating the AGB of coniferous, evergreen broadleaved, and mixed forests, and slope and aspect showed no significant correlation for all forest types. The vegetation indices had the highest variable importance for estimating deciduous broadleaved forests, whereas texture features along with vegetation indices provided better estimation for evergreen broadleaved forests.

(2) The performance of the six models for the same forest type was different. The model fitting performance was $RRF > RF > QRF > BRNN > k\text{-NN} > GBM$ for coniferous forests. The range of R^2 was from 0.49 to 0.63. For evergreen broadleaved forests, the order was $BRNN > RRF > QRF > RF > GBM > k\text{-NN}$, and the greatest R^2 and the smallest RMSE were 0.53 and 68.16 Mg ha^{-1} , respectively. For deciduous broadleaved forests, the order was $QRF > GBM > RRF > RF > BRNN > k\text{-NN}$, and the accuracy was the worst at an overall level, with the range of R^2 being between 0.19 and 0.43. For mixed forests, the order was $RRF > QRF > GBM > BRNN > RF > k\text{-NN}$. The range of R^2 was 0.42–0.56. Generally, the rank of fitting performance was $RRF > QRF > BRNN > RF > GBM > k\text{-NN}$, and RRF provided the best model.

In conclusion, integrating multiple sources of data and selecting suitable algorithms and variables for AGB remote sensing estimation in areas with a high forest heterogeneity and a complex geography are the key steps to improving the estimation accuracy. This research aimed to explore the suitable variables and models by integrating multiple sources of data using six models based on the Boruta algorithm to estimate the AGB of four forest types with high heterogeneity in Yunnan province. It provides an important reference value and a significant guide for future research.

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

Funding: This work was supported by a grant for the comprehensive survey of carbon sinks in typical areas of China provided by the Kunming Natural Resources Survey Center of China Geological Survey (DD20220877), and by the Expert Workstation of Yunnan Province of China (grant number 2018IC100).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Beer, C.; Reichstein, M.; Tomelleri, E.; Ciais, P.; Jung, M.; Carvalhais, N.; Rodenbeck, C.; Arain, M.A.; Baldocchi, D.; Bonan, G.B.; et al. Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate. *Science* **2010**, *329*, 834–838. [[CrossRef](#)] [[PubMed](#)]
2. Qin, S.; Nie, S.; Guan, Y.; Zhang, D.; Wang, C.; Zhang, X. Forest emissions reduction assessment using airborne LiDAR for biomass estimation. *Resour. Conserv. Recycl.* **2022**, *181*, 106224. [[CrossRef](#)]
3. Ploton, P.; Barbier, N.; Coutron, P.; Antin, C.M.; Ayyappan, N.; Balachandran, N.; Barathan, N.; Bastin, J.F.; Chuyong, G.; Dauby, G.; et al. Toward a general tropical forest biomass prediction model from very high resolution optical satellite images. *Remote Sens. Environ.* **2017**, *200*, 140–153. [[CrossRef](#)]
4. Maia, V.A.; de Souza, C.R.; de Aguiar-Campos, N.; Fagundes, N.C.A.; Santos, A.B.M.; de Paula, G.G.P.; Santos, P.F.; Silva, W.B.; de Oliveira Menino, G.C.; dos Santos, R.M. Interactions between climate and soil shape tree community assembly and above-ground woody biomass of tropical dry forests. *For. Ecol. Manag.* **2020**, *474*, 118348. [[CrossRef](#)]
5. Peng, D.; Zhang, H.; Liu, L.; Huang, W.; Huete, A.R.; Zhang, X.; Wang, F.; Yu, L.; Xie, Q.; Wang, C.; et al. Estimating the Aboveground Biomass for Planted Forests Based on Stand Age and Environmental Variables. *Remote Sens.* **2019**, *11*, 2270. [[CrossRef](#)]
6. McRoberts, R.; Tomppo, E. Remote sensing support for national forest inventories. *Remote Sens. Environ.* **2007**, *110*, 412–419. [[CrossRef](#)]
7. Friedl, M.A.; McGwire, K.C.; McIver, D.K. An Overview of Uncertainty in Optical Remotely Sensed Data for Ecological Applications. In *Spatial Uncertainty in Ecology: Implications for Remote Sensing and GIS Applications*; Springer: New York, NY, USA, 2013; pp. 258–283.
8. Lu, D.; Chen, Q.; Wang, G.; Moran, E.; Batistella, M.; Zhang, M.; Vaglio Laurin, G.; Saah, D. Aboveground Forest Biomass Estimation with Landsat and LiDAR Data and Uncertainty Analysis of the Estimates. *Int. J. For. Res.* **2012**, *2012*, 436537. [[CrossRef](#)]
9. Tang, C.Q.; Han, P.-B.; Li, S.; Shen, L.-Q.; Huang, D.-S.; Li, Y.-F.; Peng, M.-C.; Wang, C.-Y.; Li, X.-S.; Li, W.; et al. Species richness, forest types and regeneration of Schima in the subtropical forest ecosystem of Yunnan, southwestern China. *For. Ecosyst.* **2020**, *7*, 35. [[CrossRef](#)]
10. Sun, H.; Wang, J.; Xiong, J.; Bian, J.; Jin, H.; Cheng, W.; Li, A.; García Mozo, H. Vegetation Change and Its Response to Climate Change in Yunnan Province, China. *Adv. Meteorol.* **2021**, *2021*, 8857589. [[CrossRef](#)]
11. Tamme, R.; Hiiesalu, I.; Laanisto, L.; Szava-Kovats, R.; Pärtel, M. Environmental heterogeneity, species diversity and co-existence at different spatial scales. *J. Veg. Sci.* **2010**, *21*, 796–801. [[CrossRef](#)]
12. Xu, W.; Ci, X.; Song, C.; He, T.; Zhang, W.; Li, Q.; Li, J. Soil phosphorus heterogeneity promotes tree species diversity and phylogenetic clustering in a tropical seasonal rainforest. *Ecol. Evol.* **2016**, *6*, 8719–8726. [[CrossRef](#)]
13. Su, H.; Liu, H.; Heyman, W.D. Automated Derivation of Bathymetric Information from Multi-Spectral Satellite Imagery Using a Non-Linear Inversion Model. *Mar. Geod.* **2008**, *31*, 281–298. [[CrossRef](#)]
14. Welle, T.; Aschenbrenner, L.; Kuonath, K.; Kirmaier, S.; Franke, J. Mapping Dominant Tree Species of German Forests. *Remote Sens.* **2022**, *14*, 3330. [[CrossRef](#)]
15. Caughlin, T.T.; Barber, C.; Asner, G.P.; Glenn, N.F.; Bohlman, S.A.; Wilson, C.H. Monitoring tropical forest succession at landscape scales despite uncertainty in Landsat time series. *Ecol. Appl.* **2021**, *31*, e02208. [[CrossRef](#)]
16. Hudak, A.T.; Fekety, P.A.; Kane, V.R.; Kennedy, R.E.; Filippelli, S.K.; Falkowski, M.J.; Tinkham, W.T.; Smith, A.M.S.; Crookston, N.L.; Domke, G.M.; et al. A carbon monitoring system for mapping regional, annual aboveground biomass across the northwestern USA. *Environ. Res. Lett.* **2020**, *15*, 095003. [[CrossRef](#)]

17. Cooper, S.; Okujeni, A.; Pflugmacher, D.; van der Linden, S.; Hostert, P. Combining simulated hyperspectral EnMAP and Landsat time series for forest aboveground biomass mapping. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *98*, 102307. [[CrossRef](#)]
18. Silveira, E.M.d.O.; Cunha, L.I.F.; Galvão, L.S.; Withey, K.D.; Acerbi Júnior, F.W.; Scolforo, J.R.S. Modelling aboveground biomass in forest remnants of the Brazilian Atlantic Forest using remote sensing, environmental and terrain-related data. *Geocarto Int.* **2019**, *36*, 281–298. [[CrossRef](#)]
19. Liu, F.; Hu, J.; Yang, F.; Li, X. Heterogeneity-diversity Relationships in Natural Areas of Yunnan, China. *Chin. Geogr. Sci.* **2021**, *31*, 506–521. [[CrossRef](#)]
20. Yu, S.; Ye, Q.; Zhao, Q.; Li, Z.; Zhang, M.; Zhu, H.; Zhao, Z. Effects of Driving Factors on Forest Aboveground Biomass (AGB) in China's Loess Plateau by Using Spatial Regression Models. *Remote Sens.* **2022**, *14*, 2842. [[CrossRef](#)]
21. Frampton, W.J.; Dash, J.; Watmough, G.; Milton, E.J. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 83–92. [[CrossRef](#)]
22. Ou, G.; Li, C.; Lv, Y.; Wei, A.; Xiong, H.; Xu, H.; Wang, G. Improving Aboveground Biomass Estimation of *Pinus densata* Forests in Yunnan Using Landsat 8 Imagery by Incorporating Age Dummy Variable and Method Comparison. *Remote Sens.* **2019**, *11*, 738. [[CrossRef](#)]
23. Wijaya, A.; Sasmito, S.D.; Purbopuspito, J.; Murdiyarso, D. Calibration of global above ground biomass estimate using multi-source remote sensing data. In Proceedings of the Living Planet Symposium, Edinburgh, UK, 9–13 September 2013.
24. Lu, D.; Chen, Q.; Wang, G.; Liu, L.; Li, G.; Moran, E. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *Int. J. Digit. Earth* **2014**, *9*, 63–105. [[CrossRef](#)]
25. Rasel, S.M.M.; Chang, H.-C.; Ralph, T.J.; Saintilan, N.; Diti, I.J. Application of feature selection methods and machine learning algorithms for saltmarsh biomass estimation using Worldview-2 imagery. *Geocarto Int.* **2019**, *36*, 1075–1099. [[CrossRef](#)]
26. Pham, T.D.; Yokoya, N.; Xia, J.; Ha, N.T.; Le, N.N.; Nguyen, T.T.T.; Dao, T.H.; Vu, T.T.P.; Pham, T.D.; Takeuchi, W. Comparison of Machine Learning Methods for Estimating Mangrove Above-Ground Biomass Using Multiple Source Remote Sensing Data in the Red River Delta Biosphere Reserve, Vietnam. *Remote Sens.* **2020**, *12*, 1334. [[CrossRef](#)]
27. Shakhovska, N.; Melnykova, N.; Chopiyak, V.; Gregus, M.M. An ensemble methods for medical insurance costs prediction task. *Comput. Mater. Contin.* **2022**, *70*, 3969–3984. [[CrossRef](#)]
28. Li, Z.; Zan, Q.; Yang, Q.; Zhu, D.; Chen, Y.; Yu, S. Remote Estimation of Mangrove Aboveground Carbon Stock at the Species Level Using a Low-Cost Unmanned Aerial Vehicle System. *Remote Sens.* **2019**, *11*, 1018. [[CrossRef](#)]
29. Uniyal, S.; Chaurasia, K.; Purohit, S.; Rao, S.; Mahammood, V. Geo-ML Enabled Above Ground Biomass and Carbon Estimation for Urban Forests. In Proceedings of the Advanced Computing: 11th International Conference, IACC 2021, Msida, Malta, 18–19 December 2021; Revised Selected Papers. pp. 599–617.
30. Zhang, Y.; Liu, J.; Li, W.; Liang, S. A Proposed Ensemble Feature Selection Method for Estimating Forest Aboveground Biomass from Multiple Satellite Data. *Remote Sens.* **2023**, *15*, 1096. [[CrossRef](#)]
31. Shettles, M.; Temesgen, H.; Gray, A.N.; Hilker, T. Comparison of uncertainty in per unit area estimates of aboveground biomass for two selected model sets. *For. Ecol. Manag.* **2015**, *354*, 18–25. [[CrossRef](#)]
32. Lu, D. The potential and challenge of remote sensing-based biomass estimation. *Int. J. Remote Sens.* **2007**, *27*, 1297–1328. [[CrossRef](#)]
33. Ronoud, G.; Fatehi, P.; Darvishsefat, A.A.; Tomppo, E.; Praks, J.; Schaepman, M.E. Multi-Sensor Aboveground Biomass Estimation in the Broadleaved Hyrcanian Forest of Iran. *Can. J. Remote Sens.* **2021**, *47*, 818–834. [[CrossRef](#)]
34. Zhang, Y.; Ma, J.; Liang, S.; Li, X.; Li, M. An Evaluation of Eight Machine Learning Regression Algorithms for Forest Aboveground Biomass Estimation from Multiple Satellite Data Products. *Remote Sens.* **2020**, *12*, 4015. [[CrossRef](#)]
35. Jiang, F.; Kutia, M.; Ma, K.; Chen, S.; Long, J.; Sun, H. Estimating the aboveground biomass of coniferous forest in Northeast China using spectral variables, land surface temperature and soil moisture. *Sci. Total Environ.* **2021**, *785*, 147335. [[CrossRef](#)]
36. Durante, P.; Martín-Alcón, S.; Gil-Tena, A.; Algeet, N.; Tomé, J.; Recuero, L.; Palacios-Orueta, A.; Oyonarte, C. Improving Aboveground Forest Biomass Maps: From High-Resolution to National Scale. *Remote Sens.* **2019**, *11*, 795. [[CrossRef](#)]
37. Band, S.S.; Janizadeh, S.; Chandra Pal, S.; Saha, A.; Chakraborty, R.; Melesse, A.M.; Mosavi, A. Flash Flood Susceptibility Modeling Using New Approaches of Hybrid and Ensemble Tree-Based Machine Learning Algorithms. *Remote Sens.* **2020**, *12*, 3568. [[CrossRef](#)]
38. Fikret Kurnaz, T.; Kaya, Y. The comparison of the performance of ELM, BRNN, and SVM methods for the prediction of compression index of clays. *Arab. J. Geosci.* **2018**, *11*, 770. [[CrossRef](#)]
39. Li, Y.; Li, M.; Li, C.; Liu, Z. Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms. *Sci. Rep.* **2020**, *10*, 9952. [[CrossRef](#)]
40. Chen, H.; Qin, Z.; Zhai, D.-L.; Ou, G.; Li, X.; Zhao, G.; Fan, J.; Zhao, C.; Xu, H. Mapping Forest Aboveground Biomass with MODIS and Fengyun-3C VIRR Imageries in Yunnan Province, Southwest China Using Linear Regression, K-Nearest Neighbor and Random Forest. *Remote Sens.* **2022**, *14*, 5456. [[CrossRef](#)]
41. Zhang, G.; Wang, M.; Liu, K. Forest Fire Susceptibility Modeling Using a Convolutional Neural Network for Yunnan Province of China. *Int. J. Disaster Risk Sci.* **2019**, *10*, 386–403. [[CrossRef](#)]
42. Zhengqi, G.; Xiaoli, Z.; Yueting, W. Ability evaluation of coniferous forest aboveground biomass inversion using Sentinel-2A multiple characteristic variables. *J. Beijing For. Univ.* **2020**, *42*, 27–38.
43. Xu, H.; Zhang, Z.; Ou, G.; Shi, H. *A study on Estimation and Distribution for Forest Biomass and Carbon Storage in Yun-Nan Province*; Yunnan Science and Technology Press: Kunming, China, 2019.

44. Francke, T.; López-Tarazón, J.A.; Schröder, B. Estimation of suspended sediment concentration and yield using linear models, random forests and quantile regression forests. *Hydrol. Process.* **2008**, *22*, 4892–4904. [[CrossRef](#)]
45. Glória, L.S.; Cruz, C.D.; Vieira, R.A.M.; de Resende, M.D.V.; Lopes, P.S.; de Siqueira, O.H.G.B.D.; Fonseca e Silva, F. Accessing marker effects and heritability estimates from genome prediction by Bayesian regularized neural networks. *Livest. Sci.* **2016**, *191*, 91–96. [[CrossRef](#)]
46. Tien Bui, D.; Pradhan, B.; Lofman, O.; Revhaug, I.; Dick, O.B. Landslide susceptibility assessment in the Hoa Binh province of Vietnam: A comparison of the Levenberg–Marquardt and Bayesian regularized neural networks. *Geomorphology* **2012**, *171–172*, 12–29. [[CrossRef](#)]
47. Tan, K.; Wang, H.; Chen, L.; Du, Q.; Du, P.; Pan, C. Estimation of the spatial distribution of heavy metal in agricultural soils using airborne hyperspectral imaging and random forest. *J. Hazard. Mater.* **2020**, *382*, 120987. [[CrossRef](#)] [[PubMed](#)]
48. Gao, Y.; Lu, D.; Li, G.; Wang, G.; Chen, Q.; Liu, L.; Li, D. Comparative Analysis of Modeling Algorithms for Forest Aboveground Biomass Estimation in a Subtropical Region. *Remote Sens.* **2018**, *10*, 627. [[CrossRef](#)]
49. Zhu, X.; Hou, J.; Li, M.; Xu, L.; Li, X.; Li, Y.; Cheng, C.; Zhao, W.; He, N. High-resolution spatial distribution of vegetation biomass and its environmental response on Qinghai-Tibet Plateau: Intensive grid-field survey. *Ecol. Indic.* **2023**, *149*, 110167. [[CrossRef](#)]
50. Merganic, J.; Pichler, V.; Gomoryova, E.; Fleischer, P.; Homolak, M.; Merganicova, K. Modelling Impact of Site and Terrain Morphological Characteristics on Biomass of Tree Species in Putorana Region. *Plants* **2021**, *10*, 2722. [[CrossRef](#)]
51. Kobler, J.; Zehetgruber, B.; Dirnböck, T.; Jandl, R.; Mirtl, M.; Schindlbacher, A. Effects of aspect and altitude on carbon cycling processes in a temperate mountain forest catchment. *Landsc. Ecol.* **2019**, *34*, 325–340. [[CrossRef](#)]
52. Kirdeyanov, A.V.; Hagedorn, F.; Knorre, A.A.; Fedotova, E.V.; Vaganov, E.A.; Naurzbaev, M.M.; Moiseev, P.A.; Rigling, A. 20th century tree-line advance and vegetation changes along an altitudinal transect in the Putorana Mountains, northern Siberia. *Boreas* **2012**, *41*, 56–67. [[CrossRef](#)]
53. Kumar, M.; Kumar, R.; Konsam, B.; Sheikh, M.A.; Pandey, R. Above-and below-ground biomass production in *Pinus roxburghii* forests along altitudes in Garhwal Himalaya, India. *Curr. Sci.* **2019**, *116*, 1506–1514. [[CrossRef](#)]
54. Liu, S.; Dong, Y.; Sun, Y.; Li, J.; An, Y.; Shi, F. Modelling the spatial pattern of biodiversity utilizing the high-resolution tree cover data at large scale: Case study in Yunnan province, Southwest China. *Ecol. Eng.* **2019**, *134*, 1–8. [[CrossRef](#)]
55. Shen, A.; Wu, C.; Jiang, B.; Deng, J.; Yuan, W.; Wang, K.; He, S.; Zhu, E.; Lin, Y.; Wu, C. Spatiotemporal Variations of Aboveground Biomass under Different Terrain Conditions. *Forests* **2018**, *9*, 778. [[CrossRef](#)]
56. Chun, J.H.; Ali, A.; Lee, C.B. Topography and forest diversity facets regulate overstory and understory aboveground biomass in a temperate forest of South Korea. *Sci. Total Environ.* **2020**, *744*, 140783. [[CrossRef](#)]
57. Zhang, Z.; van Coillie, F.; Ou, X.; de Wulf, R. Integration of Satellite Imagery, Topography and Human Disturbance Factors Based on Canonical Correspondence Analysis Ordination for Mountain Vegetation Mapping: A Case Study in Yunnan, China. *Remote Sens.* **2014**, *6*, 1026–1056. [[CrossRef](#)]
58. Bennett, A.C.; Penman, T.D.; Arndt, S.K.; Roxburgh, S.H.; Bennett, L.T. Climate more important than soils for predicting forest biomass at the continental scale. *Ecography* **2020**, *43*, 1692–1705. [[CrossRef](#)]
59. Lenihan, J.M.; Drapek, R.; Bachelet, D.; Neilson, R.P. Climate change effects on vegetation distribution, carbon, and fire in California. *Ecol. Appl.* **2003**, *13*, 1667–1681. [[CrossRef](#)]
60. Wang, H.; Ni, J.; Prentice, I.C. Sensitivity of potential natural vegetation in China to projected changes in temperature, precipitation and atmospheric CO₂. *Reg. Environ. Chang.* **2011**, *11*, 715–727. [[CrossRef](#)]
61. Lutz, D.A.; Shugart, H.H.; White, M.A. Sensitivity of Russian forest timber harvest and carbon storage to temperature increase. *Forestry* **2013**, *86*, 283–293. [[CrossRef](#)]
62. Grant, R.; Arain, A.; Arora, V.; Barr, A.; Black, A.; Chen, J.; Wang, S.; Yuan, F.; Zhang, Y. Modelling temperature effects on CO₂ and energy exchange in temperate and boreal coniferous forests. In *AGU Spring Meeting Abstracts*; AGU: Washington, DC, USA, 2004; p. B54A-03.
63. Li, W.-J.; Peng, M.-C.; Higa, M.; Tanaka, N.; Matsui, T.; Tang, C.Q.; Ou, X.-K.; Zhou, R.-W.; Wang, C.-Y.; Yan, H.-Z. Effects of climate change on potential habitats of the cold temperate coniferous forest in Yunnan province, southwestern China. *J. Mt. Sci.* **2016**, *13*, 1411–1422. [[CrossRef](#)]
64. Ma, J.; Hu, Y.; Bu, R.; Chang, Y.; Deng, H.; Qin, Q. Predicting impacts of climate change on the aboveground carbon sequestration rate of a temperate forest in northeastern China. *PLoS ONE* **2014**, *9*, e96157. [[CrossRef](#)]
65. Zhou, R.; Li, W.; Zhang, Y.; Peng, M.; Wang, C.; Sha, L.; Liu, Y.; Song, Q.; Fei, X.; Jin, Y.; et al. Responses of the Carbon Storage and Sequestration Potential of Forest Vegetation to Temperature Increases in Yunnan Province, SW China. *Forests* **2018**, *9*, 227. [[CrossRef](#)]
66. Ni, J. Impacts of climate change on Chinese ecosystems: Key vulnerable regions and potential thresholds. *Reg. Environ. Chang.* **2010**, *11*, 49–64. [[CrossRef](#)]
67. Dakhil, M.A.; Xiong, Q.; Farahat, E.A.; Zhang, L.; Pan, K.; Pandey, B.; Olatunji, O.A.; Tariq, A.; Wu, X.; Zhang, A.; et al. Past and future climatic indicators for distribution patterns and conservation planning of temperate coniferous forests in southwestern China. *Ecol. Indic.* **2019**, *107*, 105559. [[CrossRef](#)]
68. Yunnan Institute of Botany. *Flora of Yunnan (M)*; Science Press: Beijing, China, 1977.

69. Zhang, L.; Cheng, Q.; Li, C. Improved model for estimating the biomass of *Populus euphratica* forest using the integration of spectral and textural features from the Chinese high-resolution remote sensing satellite GaoFen-1. *J. Appl. Remote Sens.* **2015**, *9*, 096010. [[CrossRef](#)]
70. López-Serrano, P.M.; López-Sánchez, C.A.; Álvarez-González, J.G.; García-Gutiérrez, J. A Comparison of Machine Learning Techniques Applied to Landsat-5 TM Spectral Data for Biomass Estimation. *Can. J. Remote Sens.* **2016**, *42*, 690–705. [[CrossRef](#)]
71. Taddese, H.; Asrat, Z.; Burud, I.; Gobakken, T.; Ørka, H.; Dick, Ø.; Næsset, E. Use of Remotely Sensed Data to Enhance Estimation of Aboveground Biomass for the Dry Afromontane Forest in South-Central Ethiopia. *Remote Sens.* **2020**, *12*, 3335. [[CrossRef](#)]
72. Lu, D. Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon. *Int. J. Remote Sens.* **2007**, *26*, 2509–2525. [[CrossRef](#)]
73. Singh, C.; Karan, S.K.; Sardar, P.; Samadder, S.R. Remote sensing-based biomass estimation of dry deciduous tropical forest using machine learning and ensemble analysis. *J. Environ. Manag.* **2022**, *308*, 114639. [[CrossRef](#)]
74. Han, H.; Wan, R.; Li, B. Estimating Forest Aboveground Biomass Using Gaofen-1 Images, Sentinel-1 Images, and Machine Learning Algorithms: A Case Study of the Dabie Mountain Region, China. *Remote Sens.* **2021**, *14*, 176. [[CrossRef](#)]
75. Zhao, P.; Lu, D.; Wang, G.; Wu, C.; Huang, Y.; Yu, S. Examining Spectral Reflectance Saturation in Landsat Imagery and Corresponding Solutions to Improve Forest Aboveground Biomass Estimation. *Remote Sens.* **2016**, *8*, 469. [[CrossRef](#)]
76. Chen, W.; Zheng, Q.; Xiang, H.; Chen, X.; Sakai, T. Forest Canopy Height Estimation Using Polarimetric Interferometric Synthetic Aperture Radar (PolInSAR) Technology Based on Full-Polarized ALOS/PALSAR Data. *Remote Sens.* **2021**, *13*, 174. [[CrossRef](#)]
77. Lee, Y.-S.; Lee, S.; Baek, W.-K.; Jung, H.-S.; Park, S.-H.; Lee, M.-J. Mapping Forest Vertical Structure in Jeju Island from Optical and Radar Satellite Images Using Artificial Neural Network. *Remote Sens.* **2020**, *12*, 797. [[CrossRef](#)]
78. Luo, S.; Wang, C.; Xi, X.; Pan, F.; Peng, D.; Zou, J.; Nie, S.; Qin, H. Fusion of airborne LiDAR data and hyperspectral imagery for aboveground and belowground forest biomass estimation. *Ecol. Indic.* **2017**, *73*, 378–387. [[CrossRef](#)]
79. Xiaoyi, W.; Huabing, H.; Peng, G.; Caixia, L.; Congcong, L.; Wenyu, L. Forest Canopy Height Extraction in Rugged Areas with ICESat/GLAS Data. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 4650–4657. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.