



Article

Estimation of the Aboveground Biomass of Forests in Complex Mountainous Areas Using Regression Kriging

Yining Luo ^{1,2}, Lihui Yan ^{1,2,*} , Zhongfa Zhou ^{1,2,3}, Denghong Huang ^{1,2} , Lu Cai ^{1,2}, Shuanglong Du ^{1,2}, Yue Yang ^{1,2}, Youyan Huang ^{1,2} and Qianxia Li ^{2,3}

- ¹ School of Karst Science, Guizhou Normal University / State Engineering Technology Institute for Karst Desertification Control, Guiyang 550001, China; luoyining@gznu.edu.cn (Y.L.); fa6897@gznu.edu.cn (Z.Z.); hdh@gznu.edu.cn (D.H.); cailu52025@163.com (L.C.); shuanglong11@gznu.edu.cn (S.D.); yangyue123@gznu.edu.cn (Y.Y.); m15338615123@163.com (Y.H.)
- ² The State Key Laboratory Incubation Base for Karst Mountain Ecological Environment of Guizhou Province, Guiyang 550001, China; lqx2384730355@163.com
- ³ School of Geography and Environmental Sciences, Guizhou Normal University, Guiyang 550025, China
- * Correspondence: yanlihui81@163.com; Tel.: +86-13765843681

Abstract: The forest area in China's plateaus and mountainous regions accounts for as much as 43% of the country's total forest area. Accurately estimating the aboveground biomass (AGB) in these plateau and mountain forests is significant for global carbon sink assessment and climate change. However, the complexity of the natural environment poses significant challenges to the accurate estimation of forests' aboveground biomass (AGB), and the accuracy of both AGB estimation and spatial mapping needs further improvement. This study utilized support vector regression, backpropagation neural networks, and random forests to predict trends in AGB and establish an optimal original model for forest AGB estimation. Further calibration was performed using regression kriging on the optimal model. The results indicated that (1) random forests achieved the highest coefficient of determination (R^2 for cypress = 0.63, R^2 for fir = 0.66, R^2 for cryptomeria = 0.64, and R^2 for mixed forest = 0.54), showing greater potential in predicting AGB in complex mountainous mixed forests; (2) the residual kriging method significantly improved the estimation accuracy, increasing the R^2 values of the original RF model by 25%, 24%, and 22%, and improving the accuracy of mixed plot estimates from 54% to 81%; and (3) the residual kriging method effectively addressed the underestimation of high values and overestimation of low values in AGB estimates, broadening the range of AGB values and allowing for a more detailed spatial distribution of forests' aboveground biomass.

Keywords: aboveground biomass; mixed forest; regression kriging; plateaus and mountainous regions



Citation: Luo, Y.; Yan, L.; Zhou, Z.; Huang, D.; Cai, L.; Du, S.; Yang, Y.; Huang, Y.; Li, Q. Estimation of the Aboveground Biomass of Forests in Complex Mountainous Areas Using Regression Kriging. *Forests* **2024**, *15*, 1734. <https://doi.org/10.3390/f15101734>

Academic Editor: Ricardo Ruiz-Peinado

Received: 30 August 2024
Revised: 22 September 2024
Accepted: 25 September 2024
Published: 30 September 2024



Copyright: © 2024 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 closely related to carbon sources and sinks in forest ecosystems and reflects their material cycles [1]. The accurate estimation of forest biomass is helpful for understanding the whole ecosystem and climate change, provides a basis for implementing forest resource monitoring, and can provide a reference for the contribution made by the global carbon cycle [2].

Traditional forest survey requires cutting down trees, which consumes a significant amount of human, material, and financial resources, and is powerfully destructive to forest vegetation [3]. In recent years, with the development of remote sensing technology, the method of combining measured data with remote sensing data and using statistical modeling to estimate forest aboveground biomass (AGB) has been widely used to study the spatial distribution of forest biomass and its changes [4].

As representatives of active remote sensing, synthetic aperture radar (SAR) and light detection ranging (LiDAR) are able to penetrate part of the forest canopy and obtain

information about the vertical structure of the forest [5–8]. Equipped with an all-weather day and night imaging system, and with high penetrability and independence from the weather and clouds, SAR is capable of realizing all-weather ground observations, which are widely used in disaster monitoring and environmental monitoring [9]. However, the effect of changes in the terrain on the signal limits the further application of SAR in forests on complex terrain [10]. Aerial LiDAR and satellite LiDAR have the potential to detect the canopy's structure and can obtain information on the vertical structure of large forests; however, the canopy and dense thickets of complex forests are still highly disruptive, making it difficult to obtain accurate information on the surface [11]. In addition, the discontinuity of satellite-borne LiDAR has also led to its limited application [12].

The use of optical remote sensing data to extract factors related to a forest's biomass for estimating that forest's AGB has long been studied [13]. High-resolution optical data, such as China GF-6 and Quickbird, have high data density and information richness and can provide fine-grained forest distributions. However, the data coverage and cost limit their large-scale application [14]. MODIS data can provide free and relatively continuous time-series data. However, they have a low resolution, are subject to cloud cover and atmospheric disturbances, are limited in their ability to provide fine information on a forest's structural parameters, and are more suitable for inversion studies in large-scale areas [15]. Landsat series satellite images have become the most widely used in forest biomass studies over the past few decades due to the technology being more mature, their inclusion of rich spectral information and spatial texture features, and the full consideration of acquiring vegetation information [16]. Landsat satellite imagery is not only free of charge in terms of the data source, but also has a continuous observation record and a high spatial resolution, and has been widely used in monitoring land cover changes, forest disturbances, and the growth of vegetation in recent years [4].

Although research on using remote sensing data in estimating forests' aboveground biomass has made some progress, biomass modeling is still the basis for biomass prediction and correction [17]. Remote sensing data combining ground truth data with remote sensing features are the main data sources and estimation methods for constructing estimation models of regional AGB and realizing the estimation of forests' aboveground biomass on a regional scale [2,8,18,19].

In the estimation studies of forests' AGB, the main estimation methods mainly include parametric and nonparametric models [8]. Parametric models usually assume that the overall population follows a certain distribution, on the basis of which the estimation is constructed [20]. This requires a strong linear relationship between the characteristic variables and the forest's AGB, and if the relationship between the variables is too complex, the parametric model may not be able to accurately capture the nonlinear relationship between the AGB and other variables [21], limiting its ability to estimate the forest's AGB under complex forest conditions. Nonparametric models are mainly realized by establishing nonlinear relationships between the measured data and characteristic variables; they do not need to make strict assumptions on the data's distribution and relationships among the variables and are more flexible than parametric models in dealing with complex relationships [22–24]. In addition, in nonparametric models, the problem of covariance between the feature variables does not constitute a limitation in modeling AGB, which enhances the utilization of variable information and improves the models' accuracy. Support vector regression (SVR), backpropagation neural network (BPNN), and random forest (RF), as commonly used nonparametric machine learning models, have been proven to be effectively applied for estimating biomass, the growth of standing stock, and the leaf area index [10,25–27]. Estimating forests' AGB using machine learning models can effectively solve the problem of nonlinearity and the high dimensionality of multidimensional predictor variables, thus improving the AGB prediction accuracy and providing a reliable method for estimating forests' AGB by using remote sensing technology.

Kriging is a geostatistical method that can generate predictive surfaces and measure the certainty and stability of predictions to some extent [28], providing optimal unbiased

estimates. Unlike deterministic statistical interpolation methods (including inverse distance weighting and global polynomials), which may fail to fully leverage the spatial variability in data when faced with an uneven spatial distribution [29]. Kriging models, when combined with other statistical models, can adequately account for nonstationarity and trends in the data. This flexibility and applicability make kriging particularly effective in handling forest biomass data [30,31]. RK integrates regression analysis and kriging interpolation techniques to account for nonspatial variation in the data through regression modeling while capturing spatial correlation in the data using kriging techniques. Fusing regression modeling and kriging methods can provide more accurate predictions when dealing with spatially correlated data [32,33]. Previous studies have successfully combined regression kriging with random forest regression to effectively mitigate uncertainties in estimations of AGB caused by spatial heterogeneity, thereby improving the accuracy of forest AGB estimates [34]. However, it remains to be validated as to whether the machine learning-based regression kriging method can correct the issues of accuracy in estimating the aboveground biomass in the typical spatially heterogeneous forest areas of the Southwest Plateau region of China.

This study proposed a residual-based regression kriging method to improve the accuracy of forest above-ground biomass (AGB) estimation and mapping in mixed forests in the mountainous areas of the Southwest Plateau of China. Using a typical mixed forest stand in the mountainous areas of the Southwest China Plateau as an experimental study area, this study used support vector regression, inverse neural networks, and random forests to predict the trend in AGB and establish an optimal primitive model for the estimation of AGB in mixed forests, and then used kriging interpolation to improve the accuracy of these estimations. In addition, Landsat 9 imagery was used for the spatial mapping of AGB in the study area, employing random forests and regression kriging. A comparative analysis was conducted to evaluate the effectiveness of the residual-based regression kriging model in spatially mapping the aboveground biomass of mixed forests.

2. Materials and Methods

2.1. Study Area

The study area is located in the southwest of China, with a longitude of $105^{\circ}27'50''$ – $105^{\circ}36'10''$ E and latitude of $25^{\circ}47'40''$ – $25^{\circ}55'40''$ N, and covering an area of 7943 ha and characterized by typical plateau mountainous terrain and climatic features. It belongs to the subtropical humid climate zone, with an annual precipitation of 1205.1–1656.8 mm and an average annual temperature of 15°C – 20°C . The topography of the study area shows a trend of being high in the northwest and low in the southeast, with the elevation ranging from 957 m to 1807 m. It is characterized by a complex topography, a broken surface structure, and large topographic ups and downs. Coniferous forests account for 66% of the forest area in the study area, with fir (*Cunninghamia lanceolata* [35]), cryptomeria (*Cryptomeria japonica var. sinensis* Miquel [36]), and cypress (*Cupressus funebris* [35]) as the main dominant tree species. Among them, fir, cryptomeria, and cypress accounted for 52%, 11%, and 3% of the coniferous forest area, respectively, which is a typical mixed forest stand structure in the mountain forests of the Southwest China Plateau (Figure 1).

2.2. Data and Processing

2.2.1. Landsat Image Preprocessing

The Landsat 9 imagery used in this study was OLI data from 5 September 2022, downloaded from the Google Earth Engine platform as a Level T1 product, with the original product having undergone radiometric and atmospheric correction. The terrain within the forest area is complex and varies significantly. Therefore, this study employed the C correction algorithm to eliminate differences in the spectral characteristics between shaded and sunlit slopes caused by variations in the terrain [37].

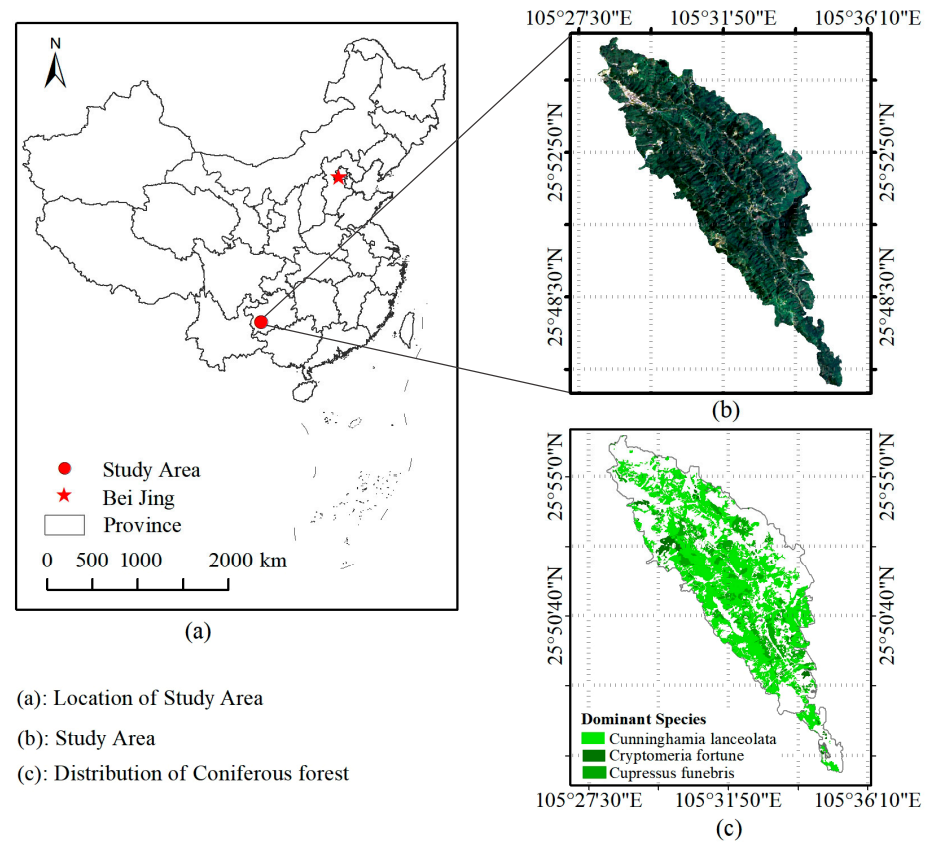


Figure 1. Location of the study area and distribution of coniferous forest.

2.2.2. Acquisition of Measurement Data

The fields were collected from July to October 2022. According to the <Technical regulations for continuous forest inventory> [38], sample plots were established with a size of 25.82 m × 25.82 m, arranged at 1 km × 1 km intervals, resulting in 66 uniformly distributed sample plots in the study area. Prior to the surveys, high-precision positioning systems (GPS) recorded the GPS coordinates of the four corners and the center of each plot, along with detailed information on the plots’ locations, measurement dates, and the dominant tree species. For each plot, trees with a diameter at breast height (DBH) greater than 5 cm were measured [39], including each tree’s height and DBH, the coordinates of each measured tree, and the total number of individuals in the plot, as illustrated in Figure 2b–e.

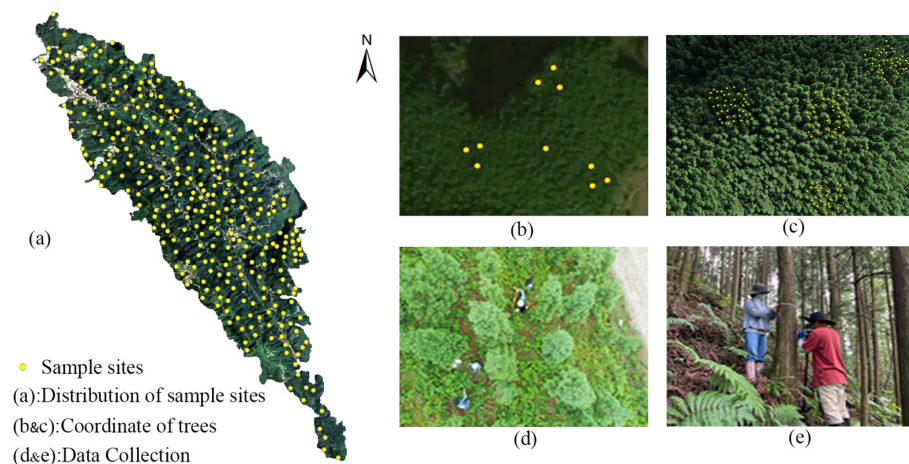


Figure 2. Distribution of sample sites and data collection.

2.2.3. Extension Plot Data Preprocessing

The study area is located in the Southwest China Plateau, characterized by complex terrain and significant changes in elevation, making many areas difficult to access for collecting extensive field data. To ensure the model's validation and calibration while enhancing its predictive performance, sample plots were expanded after verification of the accuracy with forest resource inventory data from forestry management departments. Using the "dominant tree species" field from the inventory data, forest patches were classified, and the centers of these patches were extracted to obtain the average tree height and diameter. The actual average aboveground biomass (AGB) per unit of area was calculated for each patch. Finally, the AGB values corresponding to the center points were converted to pixel values, in line with the resolution of Landsat 9 OLI imagery (pixel size of 25.82 m × 25.82 m). To ensure the reliability and accuracy of the expanded samples, outlier processing was performed on the additional sample points. The numbers of plots for cypress, cryptomeria, and fir, based on their location and proportions of area, were determined to be 11, 53, and 239, respectively, resulting in a total of 303 plots after expansion, as shown in Figure 2a.

The allometric growth equations used for calculating biomass in this study were published by the National Forestry and Grassland Administration of China (<https://www.forestry.gov.cn/> (accessed on 16 September 2022)). The parameters related to tree species (p) were derived from the basic wood densities provided by the same agency. This study first calculated the aboveground biomass (AGB) for individual trees, aggregating these results to the plot level. The plot-level AGB was then converted to biomass per hectare (Mg/ha). The corresponding formula is as follows [39]:

$$Ma = a \times D^{7/3}, \quad (1)$$

$$a = 0.3p \quad (2)$$

where Ma is the measured biomass, D is the measured diameter at breast height, a is a parameter, and p (g/cm^3) is the basic wood density associated with the tree species. It was calculated using the wood biomass of a single tree and converted to per-hectare biomass. The measured results showed that the highest value of AGB in the sample plots surveyed in the study area appeared in *Cunninghamia lanceolata*, followed by *Cupressus funebris* and *Cryptomeria japonica var. sinensis Miquel* in order; the mean values also appeared in the same order (Table 1).

Table 1. Statistical results of AGB measuring different vegetation types. Mg/ha.

Tree Species	Number	Minimum	Maximum	Averages	Standard Deviation	Coefficient of Variation (%)
<i>Cupressus funebris</i>	239	17.1803	206.7508	56.41	30.38	53.85
<i>Cryptomeria japonica var. sinensis Miquel</i>	53	16.5891	69.9303	42.5582	16.74	39.34
<i>Cunninghamia lanceolata</i>	11	28.4259	82.3241	50.3641	18.75	37.23
Mixed	303	16.5891	206.7508	53.5935	28.61	52.93

2.3. Extraction and Selection of Feature Variables

The feature variables commonly used for AGB estimation in research mainly include spectral variables and texture features. Among them, the V-type vegetation index can effectively evaluate the growth status of forest vegetation and is widely used in monitoring forest cover and the dynamics of vegetation and for estimating AGB [40,41]. In this study, nine commonly used vegetation indices were calculated, namely NDVI, GRVI, ARVI, EVI, VARI, SAVI, MSAVI, SQRT, and TNDVI (Table 2).

Models that include textural features can facilitate a more comprehensive understanding of the structure and distribution of vegetation, thus effectively improving the accuracy

of AGB estimations [31,42]. In order to screen the main textural features [43] and avoid causing redundancy in the information, this study extracted eight textural features of Band 2 [44] of Landsat 9 containing the main information as the parameters for estimating forests' biomass, which were the mean, contrast, variance, dissimilarity, homogeneity, and entropy. In the extraction of textural features, the window size affects the amount of texture information and the scale of analysis. Too small a window may result in more noise effects, while too large a window may lead to weakened heterogeneity in the texture. Finally, the textural features with a window size of 3×3 were extracted using a grayscale covariance matrix. In order to reduce the dimensionality and redundancy, this study also used principal component analysis [45] to refine the textural features, retaining at least 90% of the components of the textural information in each band.

Table 2. List of feature variables used in this study.

Variable	Feature Variable	Reference
Spectral Variable	Band reflectance (Band i , $i = 1, 2, \dots, 7$)	[46]
	Normalized difference vegetation index (NDVI)	[47]
	Red–green vegetation index (RGVI)	[48]
	Atmospherically resistant vegetation index (ARVI)	[44]
	Enhanced vegetation index (EVI)	[49]
	Visible atmospherically resistant index (VARI)	[48]
	Modified soil-adjusted vegetation index (MSAVI)	[44]
	Soil-adjusted vegetation index (SAVI)	[44]
	Transformed Normalized Difference Vegetation Index(TNDVI)	[44]
	Difference Vegetation Index (DVI)	[44]
Texture feature	Mean	[50]
	Contrast	[50]
	Variance	[50]
	Dissimilarity	[50]
	Homogeneity	[50]
	Entropy	[50]
	Correlation	[50]
	Second moment	[50]

In order to eliminate the difference in magnitude between different variables, this study carried out a uniform normalization of the extracted vegetation features, textural features, and band reflectance factors, using the following normalization formula [51]:

$$x_i = (x - x_{\min}) / (x_{\max} - x_{\min}), \quad (3)$$

where x is the original value of the feature data, x_i is the value after normalization, and x_{\max} and x_{\min} are the maximum and minimum values of the sample data, respectively.

This study analyzed the importance of all the characteristic variables, and the combination of variables with high relative importance to AGB was selected through the importance analysis [52] to participate in the construction of the AGB prediction model. In order to validate the selection of nonlinear variables, the variables were also analyzed using Pearson's correlation [53] and significance tests in SPSS26, and the linear relationship between the characteristic variables and AGB was evaluated through Pearson's correlation coefficients.

2.4. Model Selection

Nonparametric models do not require any assumptions about the distribution of the data and can describe complex nonlinear relationships, which gives them greater potential in estimating a forest's parameters. The nonparametric models SVR, BPNN, and RF have all been widely used in the estimation of forests' AGB [54–56]. SVR transforms the regression problem into an optimization problem by minimizing the regularized loss function and

maximizing the interval to solve for the model's parameters (weights corresponding to the support vectors), which, in turn, predicts the new input data, and it is suitable for different complex scenarios [57]. A BP (backpropagation) neural network is a multilayer feedforward network trained according to the error backpropagation algorithm, which does not need to determine the mathematical equations of the mapping relationship in advance and can obtain satisfactory results by its own training. RF, as a representative of nonparametric models, is widely used in the estimation of forests' AGB [58]. The random forest algorithm generates a large number of decision trees, each of which is constructed independently using a unique bootstrap sample of the training data, and the average of the predictors of all the decision trees is used as an estimate of the final target variable. In addition, RFs are randomly selected, using all the available predictors to reduce the correlations between decision trees, thus reducing noise and improving the accuracy of predictions to achieve high predictive accuracy [59]. The RF algorithm is able to deal with variables with different attributes and large differences in the range of values, can determine the importance of the variables, can avoid a large overfitting phenomenon, and has good robustness [60].

2.5. Parameters of the Prediction Model Using Regression Kriging

This study explored the trend and structural changes in the forest's AGB using a nonparametric random forest model, assessed the spatial correlations by calculating the residuals (i.e., the differences between the observations and the predictions of the regression model), and applied kriging interpolation to these residuals to make predictions. This study ultimately superimposed the predicted values of the original regression model on the predicted values of the residuals obtained by kriging interpolation to obtain the final corrected results. The formula [61] is expressed as follows:

$$\hat{z}RK(S_0) = \sum_{k=0}^p \hat{\beta}_k q_k(S_0) + \sum_{i=1}^n \lambda_i e(S_i), \quad (4)$$

where $\hat{\beta}_k$ is the estimated trend of the model's function, $q_k(S_0)$ is the predictor variable of S_0 , p is the number of auxiliary variables, $e(S_i)$ is the residuals of the regression model in, λ_i is the kriging weights determined by the autocorrelation structure of the residual space, and n is the number of known points.

2.6. Accuracy Assessment

To fully utilize the samples and improve the model's reliability, this study validated the estimation results using leave-one-out cross-validation (LOOCV) [41]. In this method, one sample was left as the test set while the remaining samples served as the training set, and this process was repeated n times to determine the final estimation result. The evaluation metrics chosen were the coefficient of determination (R^2) [62] and root mean square error (RMSE) [62]. A higher R^2 value indicates a better model fit, while a lower RMSE value signifies smaller estimation errors. The methods of calculation for the evaluation metrics are as follows:

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

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (6)$$

where y_i is the measured value of AGB, \hat{y}_i is the predicted value of the AGB model, \bar{y} is the mean value of measured AGB, and n is the number of samples.

3. Results

3.1. Relevance and Materiality Analysis

By employing a specific selection method to choose a minimal set of features with optimal model performance, a model's efficiency can be enhanced. Importance assessment, a analysis technique for nonlinear features using random forest algorithms, quantified each

variable’s contribution to the aboveground biomass (AGB) model. The Pearson’s correlation coefficient was used to evaluate the linear relationship between variables and AGB. This study combined importance analysis with Pearson’s correlation to rank all feature variables and conduct significance testing, selecting the combinations with the highest contribution and correlation coefficients for building the model. The results indicated that the spectral variables GRVI, B3, VARI, B7, B4, ARVI, and NDVI showed significant correlations with AGB at a 5% significance level, maintaining the same ranking in the importance assessment (Figure 3). Dual-validation selection provided a deep analysis of the variable impacts, significantly enhancing the model’s transparency and interpretability.

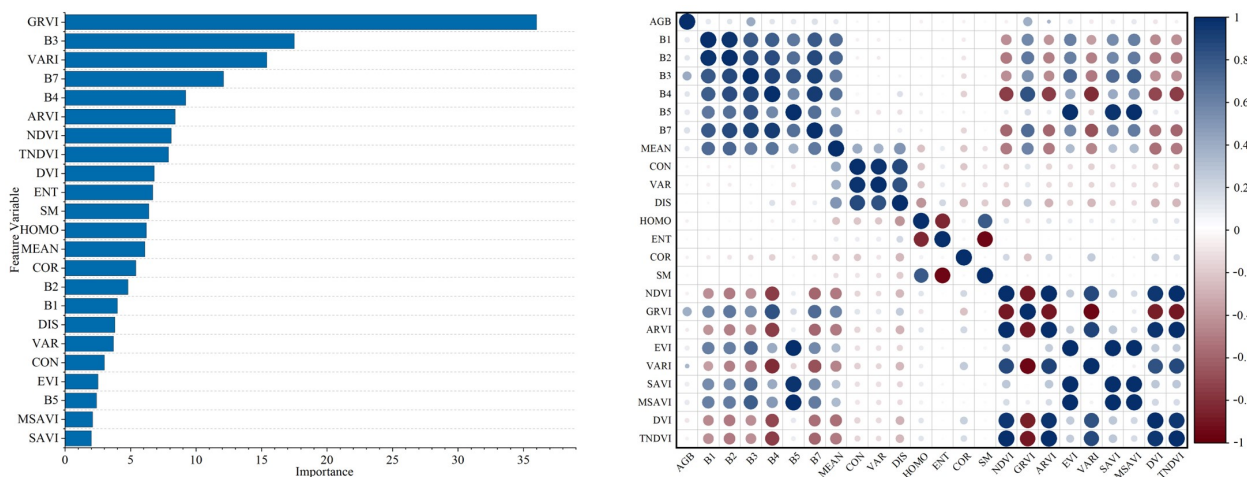


Figure 3. Correlation coefficient matrix and importance ranking between characteristic variables and AGB.

3.2. Comparison of Raw AGB Estimation Results

The performance of the model was displayed using a scatterplot showing the relationship between the observed and predicted values based on the results of leave-one-out cross-validation (Figure 4). For mixed vegetation forests, different machine learning models exhibited the lowest R^2 values (SVR: 0.34; BPNN: 0.45; RF: 0.54), with corresponding RMSE values (SVR = 17.3368; BPNN = 21.4712; RF = 21.4136). In species-based model validation, all models showed higher R^2 values compared with mixed forests, with RF achieving the highest R^2 (cypress $R^2 = 0.63$; cryptomeria $R^2 = 0.66$; fir $R^2 = 0.64$) and the lowest RMSE (cypress RMSE = 8.6289; cryptomeria RMSE = 4.9546; fir RMSE = 18.4441). The range of predicted AGB for RF was broader than those of SVR and BPNN, indicating greater predictive potential for RF in estimations of AGB for both species-specific and mixed forests (Table 3).

Table 3. Summary AGB estimation of different models.

Type	Model	Minimum	Maximum	Average	R^2	RMSE
Cypress	SVR	48.7978	61.0907	48.1854	0.37	10.0716
	BPNN	37.9780	67.5835	50.9663	0.53	12.3942
	RF	32.5075	73.7727	52.2949	0.63	8.6289
Cryptomeria	SVR	30.1387	63.8012	41.8848	0.40	5.2027
	BPNN	30.5886	64.8047	42.0657	0.55	6.1921
	RF	26.9493	65.9728	42.6753	0.78	4.9546
Fir	SVR	38.2182	156.9810	53.2810	0.46	22.9615
	BPNN	35.9082	157.5713	58.3369	0.53	21.0224
	RF	29.3553	160.3625	59.8466	0.64	18.4441
Mixed	SVR	30.1387	81.9725	51.2017	0.34	17.3368
	BPNN	30.5886	161.5713	57.2498	0.45	21.4712
	RF	26.9493	186.9406	62.8294	0.54	21.4136

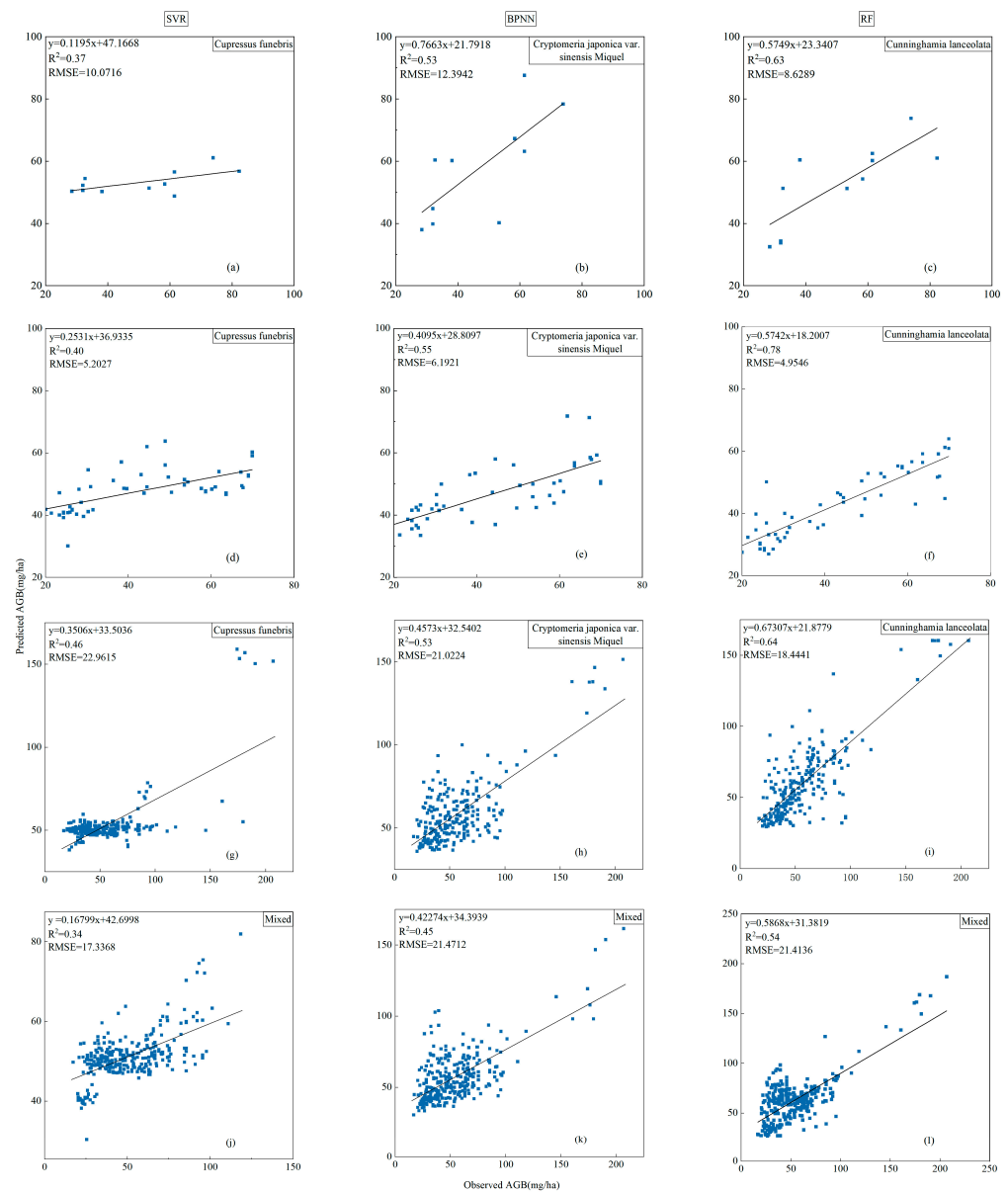


Figure 4. Scatter plot between predicted and observed AGB. (a) for cypress of SVR; (b) for cypress of BPNN; (c) for cypress of RF; (d) for cryptomeria of SVR; (e) for cryptomeria of BPNN; (f) for cryptomeria of RF; (g) for fir of SVR; (h) for fir of BPNN; (i) for fir of RF; (j) for mixed of SVR; (k) for mixed of BPNN; (l) for mixed of RF.

3.3. Regression-Kriging-Based AGB Estimation

The random forest model was used for the original prediction, the residuals of the original prediction were spatially interpolated using the simple Gaussian kernel kriging method, and the predicted AGB of the RF model was superimposed on the residual interpolation to form the improved prediction of AGB, as shown in Figure 5. Compared with the original RF model's prediction, the kriging model with superimposed residuals was of higher quality and significantly improved the accuracy of the estimated AGB of the forests. The RMSE values obtained from the regression kriging model were all lower than those of the original RF model. Among them, the RMSE of cypress decreased from 8.6289 to 8.2738, that of cryptomeria RMSE decreased from 6.9546 to 6.2763, that of fir RMSE decreased from 18.4441 to 12.4875, and the RMSE of mixed forest decreased from 21.4136 to 12.7323, with the best improvement in modeling seen in the mixed forest (Figures 4 and 5). The coefficient of determination of cypress funebris improved from 63% to 88%, that

of cryptomeria from 66% to 90%, and that of fir from 64% to 86%. The coefficient of determination of the mixed forest in the whole sample site improved from 54% to 81%. In addition, the improved model enhanced the ability to capture the spatial variability in the forests' aboveground biomass (AGB) and effectively reduced the prediction bias. After the improvement through regression kriging, the range of fir AGB values was expanded from 29.3553–160.3625 Mg/ha to 20.9844–201.7401 Mg/ha, the range of cryptomeria AGB values was expanded from 26.9493–63.9728 Mg/ha to 20.1759–70.8089 Mg/ha, the range of cypress AGB values was expanded from 32.5070–73.7728 Mg/ha to 30.2114–78.4640 Mg/ha, and the range of mixed forest's AGB values expanded from 26.9493–186.9406 Mg/ha to 15.9189–200.6346 Mg/ha. The improved AGB results were closer to the actual distribution of the AGB values in the forests of the study area.

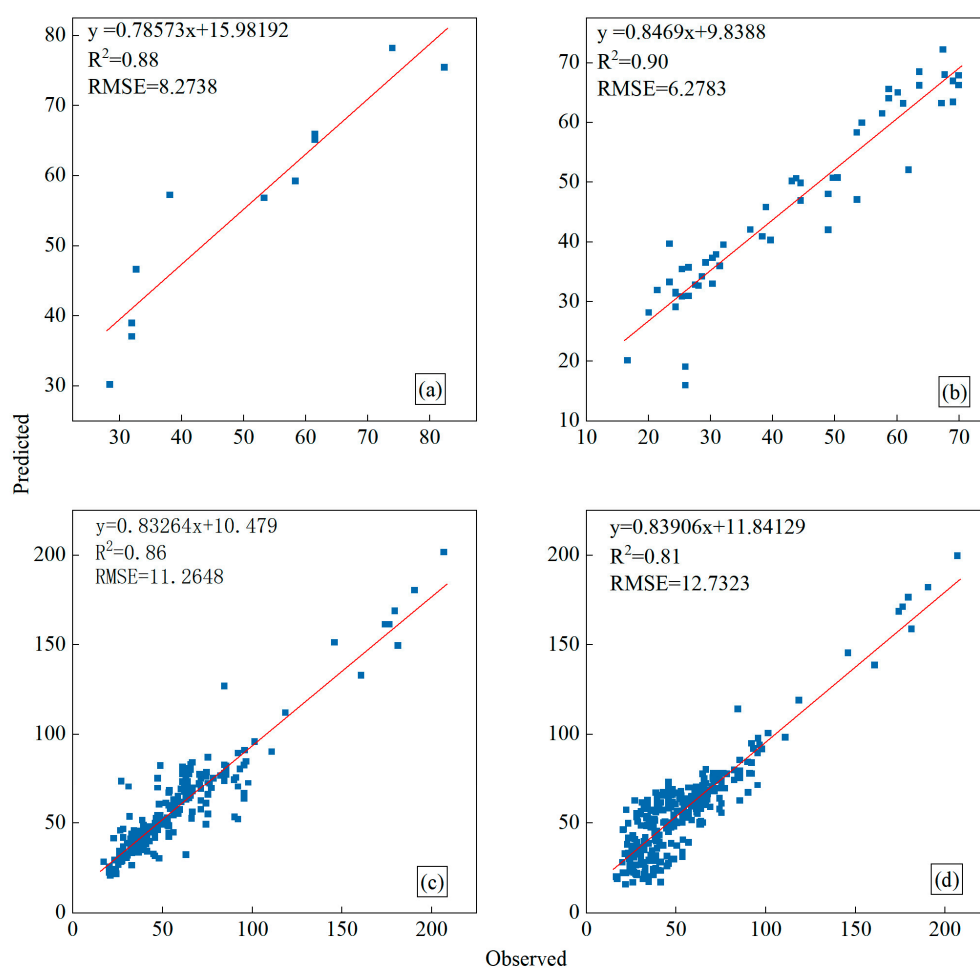


Figure 5. Scatter plots of the predicted AGB by regression kriging against the observed AGB values, (a) for cypress; (b) for cryptomeria; (c) for Fir; and (d) for mixed forests.

4. Discussion

4.1. Comparison of the Performance of Different Machine Learning Models

Different machine learning algorithms exhibit varying sensitivities to the data's distribution [63], providing diverse options for handling nonuniformly distributed data. This study focused on the forest environment of the southwestern plateau region of China, exploring the potential for the application of three advanced machine learning algorithms, namely random forest (RF), backpropagation neural network (BPNN), and support vector regression (SVR), for estimating forests' aboveground biomass (AGB). The results indicated that RF demonstrated significantly superior performance in estimating AGB in complex regional terrain compared with SVR and BPNN. Possible reasons include the sensitivity of

the SVR to kernel functions and the regularization parameter C , which may have affected the model's stability across different slope gradients, leading to decreased performance [64]. Although BPNN performs well under specific terrain conditions, it may lack sufficient generalization ability [65], resulting in unstable predictions in complex terrain. In contrast, the RF reduces the risk of overfitting by integrating multiple decision trees and requires minimal parameter adjustments, enhancing its generalization ability on new data [66] and improving its stability and predictive accuracy in complex terrains. Furthermore, RF is less sensitive to noise in the training samples and can effectively handle the complex nonlinear relationships between AGB and remote sensing, as well as accuracy reductions caused by data gaps. Therefore, RF exhibited greater robustness in estimating the AGB of forests in areas with complex terrain.

4.2. Comparison of Forest AGB Spatial Mapping

This study produced two maps of the forests' AGB based on random forest (RF) and regression kriging to compare the AGB values across the study area by tree species, namely, the original map of the forests' AGB generated from the RF model (Figure 6a) and the predicted AGB map obtained through regression kriging correction (Figure 6b). The results showed that, while the spatial distribution trends of the two AGB maps were similar, their distribution patterns differed, supporting the findings that residual-based regression kriging effectively improves AGB prediction accuracy. Furthermore, the findings in Figure 6 reveal that the range of AGB from the RF model was relatively narrow (10.0545–180.3720 Mg/ha, with a mean of 42.4348 Mg/ha), while the calibrated map of AGB exhibited a discrete distribution pattern (5.1565–219.3680 Mg/ha, with a mean of 49.6055 Mg/ha). The mean and range of predictions from the regression kriging model were superior to those of the RF model. In the same area, the calibrated AGB values were higher in regions with originally high values and lower in regions with originally low values. This supports the conclusion that the calibrated AGB values can effectively address the issues of underestimation in high-value areas and overestimation in low-value areas, thereby enhancing the accuracy of mapping the spatial distribution of forests' AGB.

4.3. Analysis of the Consistency of the Research Results

Compared with the random forest technique, simple kriging, which integrates the results of residual interpolation, exhibited higher mapping accuracy. It effectively alleviated the issues of overestimating low values and underestimating high values. All the accuracy metrics exceeded those of the RF model, supporting the findings from other researchers. For instance, Jiang Fugeng et al. [61] proposed combining random forests with ordinary kriging to enhance the accuracy of estimations of aboveground biomass in the Wangye Dian forest, increasing the coefficient of determination from 0.57 to 0.87. This indicated that the combined approach of random forests and kriging can effectively address the complexities of forest conditions, further supporting the feasibility of using this method to estimate aboveground biomass in complex forest environments.

There are also researchers whose conclusions differed from those of this study to some extent. Zhou Youfeng et al. [67] improved the estimation of forests' AGB in northern Guangdong by combining random forests with kriging, and the accuracy was improved from 0.46 to 0.57. The reasons for this discrepancy may be manifold. First, there was obvious heterogeneity among the geographic units, leading to spatial differences in the sampling points. The inconsistency of the datasets on forests' AGB used in constructing the model increased the uncertainty in the estimation process. Second, unlike studies that estimated AGB based on the type of forest, this study explores aboveground biomass at the species level and adopted a species-specific approach, which helped us to understand the spatial distribution and dynamics of the forests' AGB in a finer way. Lastly, obtaining comprehensive and accurate multivariate data is often challenging in highland mountainous areas with complex topography. Simple kriging relies on a single variable and reduces the need for large amounts of complex data, thereby reducing the burden of data

collection. In addition, in areas of high topographic relief, the biomass varies markedly between small areas, and simple kriging can provide relatively accurate estimates of these localized changes.

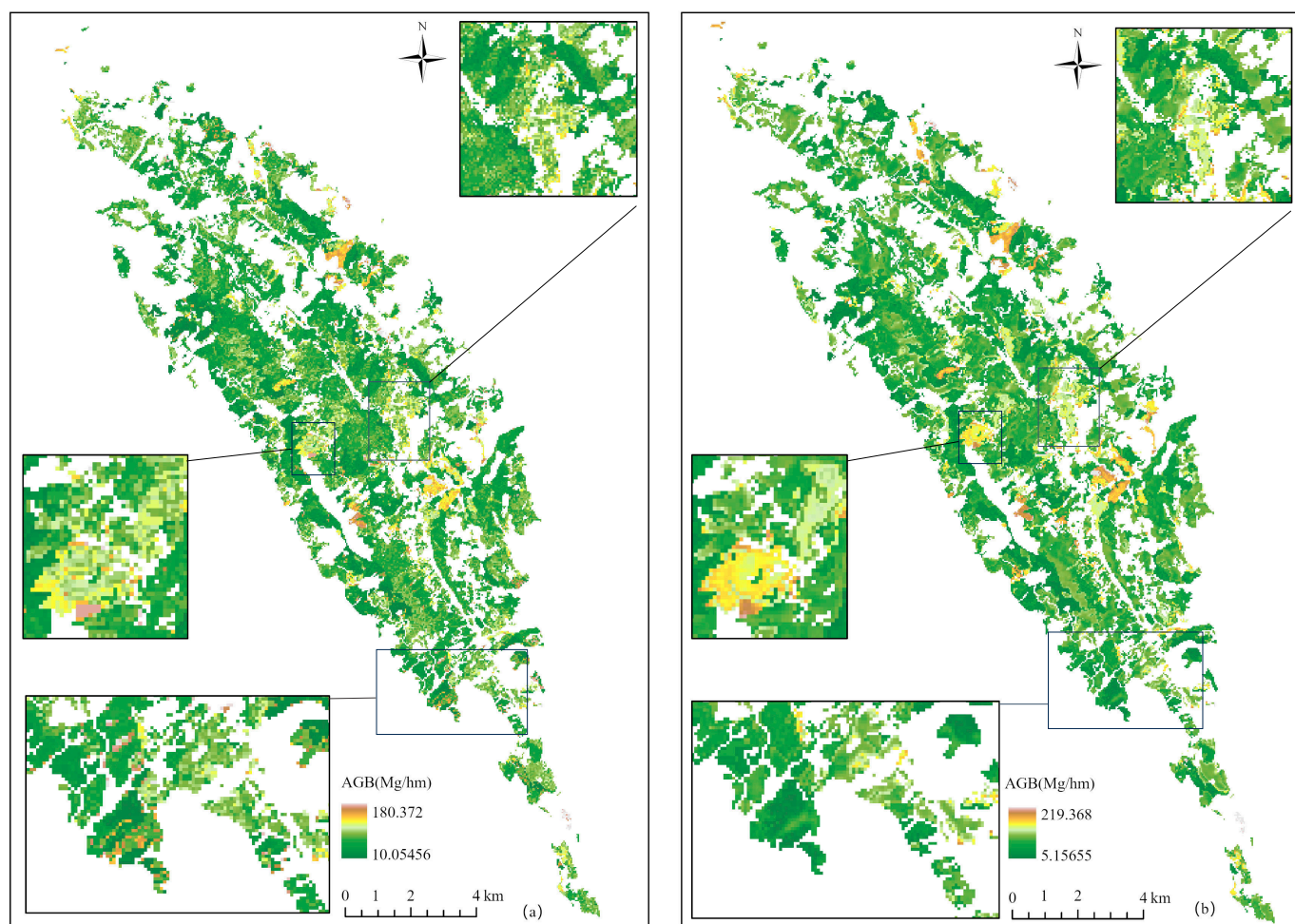


Figure 6. Spatial distribution of forest AGB. (a) For RF. (b) For regression kriging.

4.4. Shortcomings and Prospects

The forest AGB estimated using the random forest method had the phenomenon of high low values and low high values. Other scholars using optical remote sensing data to estimate forests' AGB also had this problem [52]; a possible reason is the mixed pixel effect in optical remote sensing data. In areas of forest with low AGB, the pixel values tend to be affected by the surrounding features and cannot accurately reflect the true information of the image element. On the contrary, in areas of forest with high AGB, the vegetation's reflectance tends to be saturated, resulting in the optical remote sensing data failing to capture higher biomass values, which leads to the predicted AGB values being lower than the actual ones [68]. This study area belongs to the mountainous region of the Southwest China Plateau, which has a complex and highly fragmented topography, characterized by strong spatial heterogeneity. The complex topography and vertical forest structure exacerbated the mixing of image pixels, making it difficult for the optical remote sensing data to accurately capture the true characteristics of regional forests. This may further reduce the sensitivity of optical remote sensing data to changes in the forests' AGB. In addition, studies estimating AGB for evergreen coniferous forests are also challenging. Coniferous forests are relatively homogeneous in form and structure, with a more uniform canopy structure and branch distribution. This homogeneity may hinder the effectiveness

of biomass estimation models that rely on optical data for capturing the inherent complexity of these forests.

Future research could be strengthened in several key areas. First, it has been argued that forests' AGB exhibits significant variability in stand age composition, with mature and over-mature forests typically exhibiting higher aboveground biomass (AGB) [69]. Therefore, incorporating the effect of stand age into the analysis will provide a more accurate evaluation of the current AGB status. Second, the correlations between the spectral features of optical remote sensing data and tree species diversity are unstable, spatially and temporally variable, and do not reflect the effect of the vertical structural heterogeneity of vegetation on tree species diversity. The long-wave signals of synthetic aperture radar (SAR) data have a certain ability to penetrate the forest canopy, which can mitigate the saturation effect to a certain extent. Therefore, studying the integration of optical remote sensing data and SAR data has the potential to improve the problem of optical saturation. Lastly, high plateaus and mountainous areas have a high degree of spatial heterogeneity with a large topographic relief and fragmented terrain. The current study did not incorporate topographic factors into the experiment, and the effects of the topography on the forests' AGB will be further explored in future studies to more accurately analyze forests' AGB.

5. Conclusions

This study focused on typical mixed forests in the Southwest China Plateau. It proposed a machine learning model that was suited for estimating forests' AGB in complex plateau areas. Additionally, a residual-based regression kriging method was introduced to enhance the accuracy of AGB estimation. The research results indicated the following:

- (1) The random forest model outperformed support vector regression and backpropagation neural networks in estimating AGB in mixed forests, demonstrating a higher coefficient of determination and a lower root mean square error;
- (2) The residual-based kriging regression prediction combined with Landsat 9 improved the accuracy of AGB prediction. Additionally, regression kriging effectively expanded the estimated range of AGB values, addressing the issues of underestimating high AGB values and overestimating low AGB values.

In predicting AGB in complex plateaus and mountainous mixed forests, RF demonstrated greater predictive potential among the nonparametric machine learning models. Residual-based regression kriging effectively enhanced the accuracy of AGB estimation, reducing overestimation and underestimation to some extent, and providing a more precise description of the spatial distribution of the local AGB. The methods proposed in this study are transferable to some degree and can be applied to other research areas. The findings will provide methodological references for the accurate estimation of biomass in mixed forests in complex plateaus and mountainous regions.

Author Contributions: All authors contributed to the manuscript. Conceptualization, Y.L. and L.Y.; methodology, Y.L.; validation, D.H. and L.C.; investigation, Y.L., D.H., Y.Y., S.D., and Q.L.; data curation, Y.L. and Y.Y.; writing—original draft, Y.L.; writing—review and editing, Y.L., D.H., L.Y., and Y.H.; funding acquisition, L.Y., D.H., and Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Guizhou Provincial Key Technology R&D Program ([2023] General 218 and [2023] General 211), Guizhou Provincial Basic Research Program (Natural Science) ([2021] General 194), and the Science and Technology program of Guizhou Province (Qiankehe Zhongyindi [2023]005).

Data Availability Statement: The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors gratefully acknowledge the financial support of Guizhou Normal University. We would also like to thank the editors and anonymous reviewers for their helpful and productive comments on the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

References

- Zhao, N.; Zhou, L.; Zhuang, J.; Wang, Y.; Zhou, W.; Chen, J.; Song, J.; Ding, J.; Chi, Y. Integration analysis of the carbon sources and sinks in terrestrial ecosystems, China. *Acta Ecol. Sin.* **2021**, *41*, 7648–7658.
- Fang, J.; Liu, G.; Xu, S. Biomass and Net Production of Forest Vegetation in China. *Acta Ecol. Sin.* **1996**, *16*, 497–508.
- Cao, H.; Qiu, X.; He, T. Review on Development of Forest Biomass Remote Sensing Satellites. *Acta Opt. Sin.* **2022**, *42*, 402–409.
- He, H.; Guo, Z.; Xiao, W. Application of remote sensing in forest aboveground biomass estimation. *Chin. J. Ecol.* **2007**, *26*, 1317–1322.
- Beaudoin, A.; Hall, R.J.; Castilla, G.; Filiatrault, M.; Villemaire, P.; Skakun, R.; Guindon, L. Improved k-NN Mapping of Forest Attributes in Northern Canada Using Spaceborne L-Band SAR, Multispectral and LiDAR Data. *Remote Sens.* **2022**, *14*, 1181. [[CrossRef](#)]
- Jucker, T.; Caspersen, J.; Chave, J.; Antin, C.; Barbier, N.; Bongers, F.; Dalponte, M.; van Ewijk, K.Y.; Forrester, D.I.; Haeni, M.; et al. Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *Glob. Chang. Biol.* **2017**, *23*, 177–190. [[CrossRef](#)]
- Zhang, R.; Zhou, X.; Ouyang, Z.; Avitabile, V.; Qi, J.; Chen, J.; Giannico, V. Estimating aboveground biomass in subtropical forests of China by integrating multisource. *Remote Sens. Environ.* **2019**, *232*, 111341. [[CrossRef](#)]
- Tang, X.G.; Liu, D.W.; Wang, Z.M.; Jia, M.M.; Dong, Z.Y. Estimation of forest aboveground biomass based on remote sensing data: A review. *Chin. J. Ecol.* **2012**, *31*, 1311–1318.
- Timothy, D.; Mutanga, O.; Shoko, C.; Adelabu, S.A.; Bangira, T. Remote sensing of aboveground forest biomass: A review. *Trop. Ecol.* **2016**, *57*, 125–132.
- Feng, Q.; Chen, E.; Li, Z.; Li, L.; Zhao, L. Forest Above-Ground Biomass Estimation Method for Rugged Terrain Based on Airborne P-Band PolSAR Data. *Sci. Silvae Sin.* **2016**, *52*, 10–22.
- Thiel, C.; Schmullius, C. The potential of ALOS PALSAR backscatter and InSAR coherence for forest growing stock volume estimation in Central Siberia. *Remote Sens. Environ.* **2016**, *173*, 258–273. [[CrossRef](#)]
- Shao, G.; Stark, S.C.; de Almeida, D.R.; Smith, M.N. Towards high throughput assessment of canopy dynamics: The estimation of leaf area structure in Amazonian forests with multitemporal multi-sensor airborne lidar. *Remote Sens. Environ.* **2019**, *221*, 1–13. [[CrossRef](#)]
- Ehlers, D.; Wang, C.; Coulston, J.; Zhang, Y.; Pavelsky, T.; Frankenberg, E.; Woodcock, C.; Song, C. Mapping forest aboveground biomass using multisource remotely sensed data. *Remote Sens.* **2022**, *14*, 1115. [[CrossRef](#)]
- Ahmad, A.; Gilani, H.; Ahmad, S.R. Forest aboveground biomass estimation and mapping through high-resolution optical satellite imagery—A literature review. *Forests* **2021**, *12*, 914. [[CrossRef](#)]
- Yang, T.; Wang, C.; Li, G.; Luo, S.; Xi, X.; Gao, S.; Zeng, H. Forest canopy height mapping in China based on satellite-based lidar GLAS and optical MODIS data. *Sci. Sin. Terrae* **2014**, *44*, 2487–2498.
- Wu, C.; Shen, H.; Wang, K.; Shen, A.; Deng, J.; Gan, M. Landsat imagery-based above ground biomass estimation and change investigation related to human activities. *Sustainability* **2016**, *8*, 159. [[CrossRef](#)]
- Roy, P.S.; Ravan, S.A. Biomass estimation using satellite remote sensing data—An investigation on possible approaches for natural forest. *J. Biosci.* **1996**, *21*, 535–561. [[CrossRef](#)]
- Jingcheng, L. Development and Application of Key Technologies for Stand Factor Measurement and Statistics. Ph.D. Thesis, Beijing Forestry University, Beijing, China, 2019.
- Cao, L. Estimation of Forest Stock Volume in Yanqing District Based on Sentinel-2 Images. Ph.D. Thesis, Beijing Forestry University, Beijing, China, 2019.
- Hu, R.; Chen, X.; Chen, J.; Zhang, S.; Kuang, Y.; Yu, H.; Ji, H.; Zhao, X.; Yi, S.; Meng, B.; et al. MODIS NDVI saturation assessment of alpine meadow grassland biomass estimation using remote sensing: a case study in the eastern edge of the Qinghai-Tibet Plateau. *Acta Ecol. Sin.* **2024**, *44*, 6357–6372.
- Longtao, C. Full-link forest echo simulation of domestic spaceborne LiDAR and study on inversion of forest structure parameters. *Acta Geod. Cartogr. Sin.* **2023**, *52*, 2223.
- Dai, N.; Xu, W.; Xiao, Q.; Wang, Q.; Ma, F.; Liang, H.; Duan, X. Estimation of wilted grass biomass by satellite remote sensing data in winter on Qinghai-Tibet Plateau. *Acta Ecol. Sin.* **2023**, *43*, 6033–6044.
- Di Cosmo, L.; Gasparini, P.; Tabacchi, G. A national-scale, stand-level model to predict total above-ground tree biomass from growing stock volume. *For. Ecol. Manag.* **2016**, *361*, 269–276. [[CrossRef](#)]
- Li, Y.; Li, C.; Li, M.; Liu, Z. Influence of variable selection and forest type on forest aboveground biomass estimation using machine learning algorithms. *Forests* **2019**, *10*, 1073. [[CrossRef](#)]
- Ji, Y.J.; Yang, C.R.; Zhang, W.F.; Zeng, P.; Zhang, F.X.; Qu, Y.N. Forest above ground biomass estimation using airborne P band polarimetric SAR data. *J. Zhejiang A F Univ.* **2022**, *39*, 971–980.
- Freeman, E.A.; Moisen, G.G.; Coulston, J.W.; Wilson, B.T. Random forests and stochastic gradient boosting for predicting tree canopy cover: Comparing tuning processes and model performance. *Can. J. For. Res.* **2016**, *46*, 323–339. [[CrossRef](#)]

27. Yao, J.; Yang, L.; Chen, T.; Song, C. Consecutive monitoring of the poyang lake wetland by integrating sentinel-2 with sentinel-1 and landsat 8 data. *Remote Sens. Technol. Appl.* **2021**, *36*, 760–776.
28. Diggle, P.J.; Tawn, J.A.; Moyeed, R.A. *Model-Based Geostatistics*; Springer: New York, NY, USA, 2007; Volume 21, p. Xiii+228.
29. Moasheri, S.A.; Tabatabai, S.M.; Sarani, N.; Alai, Y. Estimation Spatial distribution of Sodium adsorption ratio (SAR) in Groundwater's Using ANN and Geostatistics Methods, the case of Birjand Plain, Iran. In Proceedings of the International Conference on Latest Computational Technologies(ICLCT'2012), Bangkok, Thailand, 17 March 2012.
30. Tian, X.; Zhang, X. Estimation of forest aboveground biomass by remote sensing. *J. Beijing For. Univ.* **2021**, *43*, 137–148.
31. Li, Y.; Li, M.; Liu, Z.; Li, C. Combining Kriging Interpolation to Improve the Accuracy of Forest Aboveground Biomass Estimation using Remote Sensing Data. *IEEE Access* **2020**, *8*, 128124–128139. [[CrossRef](#)]
32. Yaw, O.S.V.; Radim, V.; Karel, N.; Vít, Š.; Věra, F.; Kateřina, N.H.; Luboš, B.; Lenka, P. Spatial Distribution of Forest Soil Base Elements (Ca, Mg and K): A Regression Kriging Prediction for Czechia. *Forests* **2024**, *15*, 1123. [[CrossRef](#)]
33. Xin, S.D.; Yan, Y.X.; Jiang, L.C. Stand biomass model for Pinus koraiensis plantation based on different additive methods in Heilongjiang Province, China. *Ying Yong Sheng Tai Xue Bao* **2020**, *31*, 3322–3330.
34. Wang, M.; Fan, C.; Gao, B.; Ren, Z.; Li, F. A spatial random forest interpolation method with semi-variogram. *Chin. J. Eco-Agric.* **2022**, *30*, 451–457.
35. iPlant. Available online: <https://www.iplant.cn/info/%E6%9D%89%E6%9C%A8> (accessed on 12 September 2024).
36. iPlant. Available online: <https://www.iplant.cn/info/%E6%9F%B3%E6%9D%89?t=p> (accessed on 12 September 2024).
37. Teillet, P.M.; Guindon, B.; Goodenough, D.G. On the Slope-Aspect Correction of Multispectral Scanner Data. *Can. J. Remote Sens.* **1981**, *8*, 84–106. [[CrossRef](#)]
38. Technical Regulations for Continuous Forest Inventory. Available online: <https://std.samr.gov.cn/gb/search/gbDetailed?id=A02801294956EBB4E05397BE0A0AB6FE> (accessed on 12 September 2024).
39. Piermattei, L.; Karel, W.; Wang, D.; Wieser, M.; Mokroš, M.; Surový, P.; Koreň, M.; Tomašík, J.; Pfeifer, N.; Hollaus, M. Terrestrial Structure from Motion Photogrammetry for Deriving Forest Inventory Data. *Remote Sens.* **2019**, *11*, 950. [[CrossRef](#)]
40. Jiang, F.; Sun, H.; Li, C.; Ma, K.; Chen, S.; Long, J.; Ren, L. Retrieving the forest aboveground biomass by combining the red edge bands of Sentinel-2 and GF-6. *Acta Ecol. Sin.* **2021**, *41*, 8222–8236.
41. 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](#)]
42. Hao, Q.; Huang, C. A review of forest aboveground biomass estimation based on remote sensing data. *Chin. J. Plant Ecol.* **2023**, *47*, 1356–1374. [[CrossRef](#)]
43. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *1973*, 610–621. [[CrossRef](#)]
44. Jingni, Y. The research on High Resolution Texture Information of Remote Sensing in Forest Volume Estimate. Master's Thesis, Xi'an University of Technology, Xian, China, 2017.
45. Jolliffe, I.T. *Principal Component Analysis*; Springer: Berlin/Heidelberg, Germany, 2002.
46. 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](#)]
47. 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* **2016**, *9*, 63–105. [[CrossRef](#)]
48. Jensen, J.R.; Lulla, K. Introductory digital image processing: A remote sensing perspective. Pearson Prentice Hall, Upper Saddle River. *Geocarto Int.* **2005**, *2*, 65. [[CrossRef](#)]
49. Dong, T.; Liu, J.; Shang, J.; Qian, B.; Ma, B.; Kovacs, J.M.; Walters, D.; Jiao, X.; Geng, X.; Shi, Y. Assessment of red-edge vegetation indices for crop leaf area index estimation. *Remote Sens. Environ.* **2019**, *222*, 133–143. [[CrossRef](#)]
50. Manjunath, B.S.; Ma, W.Y. Texture features for browsing and retrieving of large image data. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 117–128.
51. Zhang Peng Ma Qingxun Lu Jie Ji Jinliang Li, Z. Application of machine learning algorithms in estimation of above-ground biomass of forest. *Bull. Surv. Mapp.* **2021**, *12*, 28–32.
52. Yuhang, G. Study on Aboveground Carbon Storage Model of Eucalyptus in Gaofeng Forest Farm Based on Landsat 8 and Machine Learn. Master's Thesis, Beijing Forestry University, Beijing, China, 2022.
53. Sa, R.; Nie, Y.; Chumachenko, S.; Fan, W. Biomass Estimation and Saturation Value Determination Based on Multi-Source Remote Sensing Data. *Remote Sens.* **2024**, *16*, 2250. [[CrossRef](#)]
54. Yan, X.; Li, J.; Smith, A.R.; Yang, D.; Ma, T.; Su, Y.; Shao, J. Evaluation of machine learning methods and multi-source remote sensing data combinations to construct forest above-ground biomass models. *Int. J. Digit. Earth* **2023**, *16*, 4471–4491. [[CrossRef](#)]
55. Tanase, M.A.; Panciera, R.; Lowell, K.; Tian, S.; García-Martín, A.; Walker, J.P. Sensitivity of L-Band Radar Backscatter to Forest Biomass in Semiarid Environments: A Comparative Analysis of Parametric and Nonparametric Models. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 4671–4685. [[CrossRef](#)]
56. Bahadur, K.Y.; Qijing, L.; Pradip, S.; Damodar, G.; Hari, A. Estimation of Above-Ground Forest Biomass in Nepal by the Use of Airborne LiDAR, and Forest Inventory Data. *Land* **2024**, *13*, 213. [[CrossRef](#)]
57. Tian, X.; Su, Z.; Chen, E.; Li, Z.; van der Tol, C.; Guo, J.; He, Q. Reprint of: Estimation of forest above-ground biomass using multi-parameter remote sensing data over a cold and arid area. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *17*, 102–110. [[CrossRef](#)]

58. Hoover, C.M.; Ducey, M.J.; Colter, R.A.; Yamasaki, M. Evaluation of alternative approaches for landscape-scale biomass estimation in a mixed-species northern forest. *For. Ecol. Manag.* **2018**, *409*, 552–563. [[CrossRef](#)]
59. Li, H.P.; Xu, H.; Zhang, C.; Ou, G.L.; Sun, X.L. Remote Sensing estimation of Pinus yunnanensis Natural Forest Biomass based on Random Forest Model. *J. West China For. Sci.* **2022**, *51*, 60–66.
60. Yingchang, L. Optimized Method for Forest Aboveground Biomass Estimation based on Remote Sensing Data and Its Spatiotemporal Analysis. Ph.D. Thesis, Nanjing Forestry University, Nanjing, China, 2021.
61. Jiang, F.; Sun, H.; Chen, E.; Wang, T.; Cao, Y.; Liu, Q. Above-Ground Biomass Estimation for Coniferous Forests in Northern China Using Regression Kriging and Landsat 9 Images. *Remote Sens.* **2022**, *14*, 5734. [[CrossRef](#)]
62. Willmott, C.J.; Ackleson, S.G.; Davis, R.E.; Feddema, J.J.; Klink, K.M.; Legates, D.R.; O'donnell, J.; Rowe, C.M. Statistics for the evaluation and comparison of models. *J. Geophys. Res. Ocean.* **1985**, *90*, 8995–9005. [[CrossRef](#)]
63. Krawczyk, B. Learning from imbalanced data: Open challenges and future directions. *Prog. Artif. Intell.* **2016**, *5*, 221–232. [[CrossRef](#)]
64. Khan, A.; Bhat, M.A. Assessing the Effect of Terrain on Support Vector Regression Model Performance. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *104*, 102515.
65. Gonzalez, A.; Naderpour, M. Assessing the Generalization Capability of Neural Networks in Non-Stationary Environmental Data. *Environ. Model. Softw.* **2021**, *145*, 105190.
66. Tian, P.; Ren, Y.; Chen, Y. Research Progress on Identification and Extraction Methods of Soil and Water Conservation Measures. *Soil Water Conserv.* **2024**, *5*, 1–9.
67. Zhou, Y.; Xie, B.; Li, M. Mapping regional forest aboveground biomass from random forest Co-Kriging approach: A case study from north Guangdong. *J. Nanjing For. Univ. Nat. Sci. Ed.* **2024**, *48*, 169–178.
68. Li, Y.; Zhang, W.; Cui, J.; Li, C.; Ji, Y. Inversion exploration on forest aboveground biomass of optical and SAR data supported by parameter optimization method. *J. Beijing For. Univ.* **2020**, *42*, 11–19.
69. Zhu, K.; Song, Y.; Qin, C. Forest age improves understanding of the global carbon sink. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 3962–3964. [[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.