



Article

Reduction in Uncertainty in Forest Aboveground Biomass Estimation Using Sentinel-2 Images: A Case Study of *Pinus densata* Forests in Shangri-La City, China

Lu Li ¹, Boqi Zhou ¹, Yanfeng Liu ¹, Yong Wu ¹, Jing Tang ¹, Weiheng Xu ^{1,2} , Leiguang Wang ^{1,2} and Guanglong Ou ^{1,*}

¹ Key Laboratory of State Forestry Administration on Biodiversity Conservation in Southwest China, Southwest Forestry University, Kunming 650224, China

² Institute of Big Data and Artificial Intelligence, Southwest Forestry University, Kunming 650233, China

* Correspondence: olg2007621@swfu.edu.cn

Abstract: The uncertainty from the under-estimation and over-estimation of forest aboveground biomass (AGB) is an urgent problem in optical remote sensing estimation. In order to more accurately estimate the AGB of *Pinus densata* forests in Shangri-La City, we mainly discuss three non-parametric models—the artificial neural network (ANN), random forests (RFs), and the quantile regression neural network (QRNN) based on 146 sample plots and Sentinel-2 images in Shangri-La City, China. Moreover, we selected the corresponding optical quartile models with the lowest mean error at each AGB segment to combine as the best QRNN (QRNNb). The results showed that: (1) for the whole biomass segment, the QRNNb has the best fitting performance compared with the ANN and RFs, the ANN has the lowest R^2 (0.602) and the highest RMSE (48.180 Mg/ha), and the difference between the QRNNb and RFs is not apparent. (2) For the different biomass segments, the QRNNb has a better performance. Especially when AGB is lower than 40 Mg/ha, the QRNNb has the highest R^2 of 0.961 and the lowest RMSE of 1.733 (Mg/ha). Meanwhile, when AGB is larger than 160 Mg/ha, the QRNNb has the highest R^2 of 0.867 and the lowest RMSE of 18.203 Mg/ha. This indicates that the QRNNb is more robust and can improve the over-estimation and under-estimation in AGB estimation. This means that the QRNNb combined with the optimal quantile model of each biomass segment provides a method with more potential for reducing the uncertainties in AGB estimation using optical remote sensing images.

Keywords: Sentinel-2 images; artificial neural network; random forests; quantile regression neural network; *Pinus densata* forests



Citation: Li, L.; Zhou, B.; Liu, Y.; Wu, Y.; Tang, J.; Xu, W.; Wang, L.; Ou, G. Reduction in Uncertainty in Forest Aboveground Biomass Estimation Using Sentinel-2 Images: A Case Study of *Pinus densata* Forests in Shangri-La City, China. *Remote Sens.* **2023**, *15*, 559. <https://doi.org/10.3390/rs15030559>

Academic Editors: Huaqiang Du, Wenyi Fan, Weiliang Fan, Fangjie Mao and Mingshi Li

Received: 27 November 2022

Revised: 10 January 2023

Accepted: 10 January 2023

Published: 17 January 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 a crucial factor in carbon storage in terrestrial ecosystems and plays an essential role in protecting the ecological environment and biodiversity [1]. The biomass harvesting method is time-consuming and labor-intensive; thus, it is not available for large-scale data acquisition [2]. Along with the development of remote sensing technology, more and more researchers are using remote sensing data combined with ground survey data to estimate large-scale forest biomass [3,4].

Three types of remote sensing data are available for biomass estimation: optical images, active sensor radar data, and light detection and ranging (LiDAR) data [5,6]. The main LiDAR technology used in forest biomass estimation is backpack LiDAR and airborne LiDAR. Backpack LiDAR is hard to use for large-area assessment because the terrain and forestland accessibility easily influence it. Although airborne LiDAR is not limited by the terrain and can capture three-dimensional structure information; thus, it has a better performance for forest biomass estimation by improving the saturation problem in biomass estimation using optical remote sensing data [7,8]. However, it still needs to be more

suitable for large-area forest biomass estimation due to the limitation of the battery capacity and the increased imaging cost, etc. Moreover, LiDAR has no infrared signals, a limiting factor for vegetation analysis [9]. Radar has an intense penetration in vegetation, but the data processing is quite complicated, and the forest AGB has a different sensitivity to its wavelength [10,11]. The most accurate radar systems operate with short wavelengths (i.e., X- and C-bands). However, the radar signal does not reach the ground because it is mainly backscattered by the canopy of the upper layer [12–14]. Using long radar wavelengths (mainly L- and P-bands), the radar signal can penetrate the different layers from the top of the canopy to the ground. However, P-band radar imagery is expected to be available with the ESA BIOMASS mission to be launched in 2024 [15,16]. High- and medium-resolution optical remote sensing are used for AGB estimation commonly. Generally, high-resolution optical images are too expensive, and the images are quite hard to obtain even though they have more accurate results of AGB estimation than medium-resolution optical images [17]. Therefore, the medium-resolution satellite images (e.g., Landsat and Sentinel-2) are a better choice for forest biomass evaluation by different spatial scales due to their free accessibility and high suitability to landscape scale analysis [18]. However, reducing the uncertainties is still a significant difficulty for AGB estimation using optical remote sensing data, especially when the study area has a high canopy [19,20]. The European Space Agency launched a high-resolution and multi-spectral imaging satellite, Sentinel-2A, in 2015 and Sentinel-2B in 2017. In addition, the spatial resolutions are 10 m, 20 m, and 60 m, respectively. Sentinel-2 can revisit an area in 5 days by two satellites and it has a wide swath at 290 km with 13 multi-spectral bands, including four additional spectral bands strategically positioned in the red-edge region, which is a more sensitive band to vegetation [21,22]. It is expected to improve the uncertainties of AGB estimation [22–24].

To reduce the saturation impact on forest biomass estimation, vegetation indices (VI) have been employed in lots of research [25–27]. The VI has been shown to be related to photosynthesis to some extent and directly proportional to biomass or yield [28]. The normalized difference vegetation index (NDVI), atmospherically resistant vegetation index (ARVI), difference vegetation index (DVI), simple ratio index (RVI), etc., were extracted from images, which were used in AGB assessment [1,28,29]. With the development of the research, the researchers found that the vegetation index changed little after the biomass reached a specific value [17,18]; in particular, tropical and subtropical woodlands with high coverage and structural complexity are more likely to lead to insensitivity [30]. Moreover, the researchers found that the texture features are more sensitive to the horizontal structure of the canopy and the shadow, which may be suitable for improving the prediction precision of forest AGB biomass estimation. Some studies have been found to apply textures in forest biomass assessment [31,32], and the image texture has excellent potential to enhance the accuracy of AGB estimation [33–35]. Therefore, variables screening is vital for reducing the impact on the multi-collinearity and increasing the accuracy in the AGB remote sensing estimation [36–38].

The accuracy of forest AGB estimation is not only affected by the survey data but also impacts the methodology of the assessment model [39,40]. Two kinds of algorithms were applied for forest AGB estimation, including parametric and non-parametric algorithms [41]. The parametric method can quantitatively describe the relationship between AGB and the variables, in which h contains linear, logarithm, exponential, and other functions [42]. In contrast, the artificial neural network (ANN), K-nearest neighbor (KNN), support vector machine (SVM), random forests (RF), etc., were counted into a non-parametric model [43–46]. The relationship between AGB and variables cannot easily be analyzed by fixed quantity due to the complex relationship between AGB and forest construction. A lot of research has been conducted to compare the accuracy of parametric and non-parametric algorithms, and the result have shown that non-parametric algorithms exhibited excellent performance [46]. The artificial neural network (ANN) is a supervised learning algorithm in machine learning which has adaptability and improves the precision of updated data. It has been used widely to demonstrate the complex relationships between independent and

dependent variables [47]. The ANN frequently uses AGB estimation due to the parallelism, fault tolerance, generalization capability, and multiple input multiple output architecture. Then, it can reveal a solid ability to fit data [44]. Moreover, ANN was applied to predict the AGB in natural forest ecosystems, with it showing that it offered a higher accuracy result than traditional protocols [48–50].

The quantile regression neural network (QRNN) is a non-parametric, nonlinear model that is combined with a neural network (NN) and quantile regression (QR) approach, which was introduced by Taylor [51]. It centralizes the advantages of both the ANN and the QR. The QR was created by Koenker [52], and it can more accurately describe the influence of independent variables on the range of dependent variables and the shape of the conditional distribution. It is not impacted by abnormal data such as sharp peaks, discrete values, and heavy-tailed allocations [53,54]. When independent variables have different effects on the distribution of dependent variables in different parts, such as skewness on the left or right, it can describe the characteristics of the distribution more comprehensively [53,54]. The QRNN is a suitable methodology for predicting mixed discrete–continuous variables. It has already been applied in ecological environments [55–57]. Rarely have studies been found using the QRNN in forest AGB estimation.

In general, the forest resources of Shangri-La City are characterized by extensive forestry land and are identified as one of the species genetic pools [58,59]. Meanwhile, Yunnan is known as the kingdom of plants and animals; thus, the forest resources status of Yunnan in China or around the world is self-evident [60,61]. Given this, it is significant to emphasize the precision of forest AGB assessment in Shangri-La City to protect forest resources and improve the ecological environment. In this study, we will estimate forest AGB by combining the measured sample data, Sentinel-2 images, vegetation index, and texture value extracted from the images. We screened the correlation variables with AGB using RF, then RF, the ANN, and the QRNN were selected to compare the fitting performance. The significant contributions of this work are:

- (1) To compare different biomass estimation models—the ANN, RF, and the QRNN for estimating the biomass of *Pinus densata* forests using Sentinel-2 images in Shangri-La City.
- (2) To explore the optimal quantile model on each biomass segment to improve the AGB estimation accuracy, and then provide a method to reduce the uncertainties from over-estimation and under-estimation of forest AGB estimation.

2. Materials and Methods

2.1. Study Site

The study area is located in Shangri-La City, northwestern Yunnan, China. The coordinates of Shangri-La City are: latitude 26°52′~28°52′N and longitude 99°20′~100°29′E (Figure 1). The elevation range is from 3350 to 3696 m above sea level, the annual mean temperature is 4.7–16.5 °C, and the extreme maximum and minimum temperature are 25.1 and −20.1 °C, respectively. The dry and wet seasons are distinct, and the four seasons are not apparent in Shangri-La City. For the rainfall time concentrates from June to September, the mean annual precipitation is 607 mm and the average annual evaporation is 1643.6 mm [62]. The particular geographical environment and complex ecological conditions have created a unique natural landscape and rich natural resources. The original forest area with the sub-alpine coniferous forest is the main forest area that is well preserved in China.

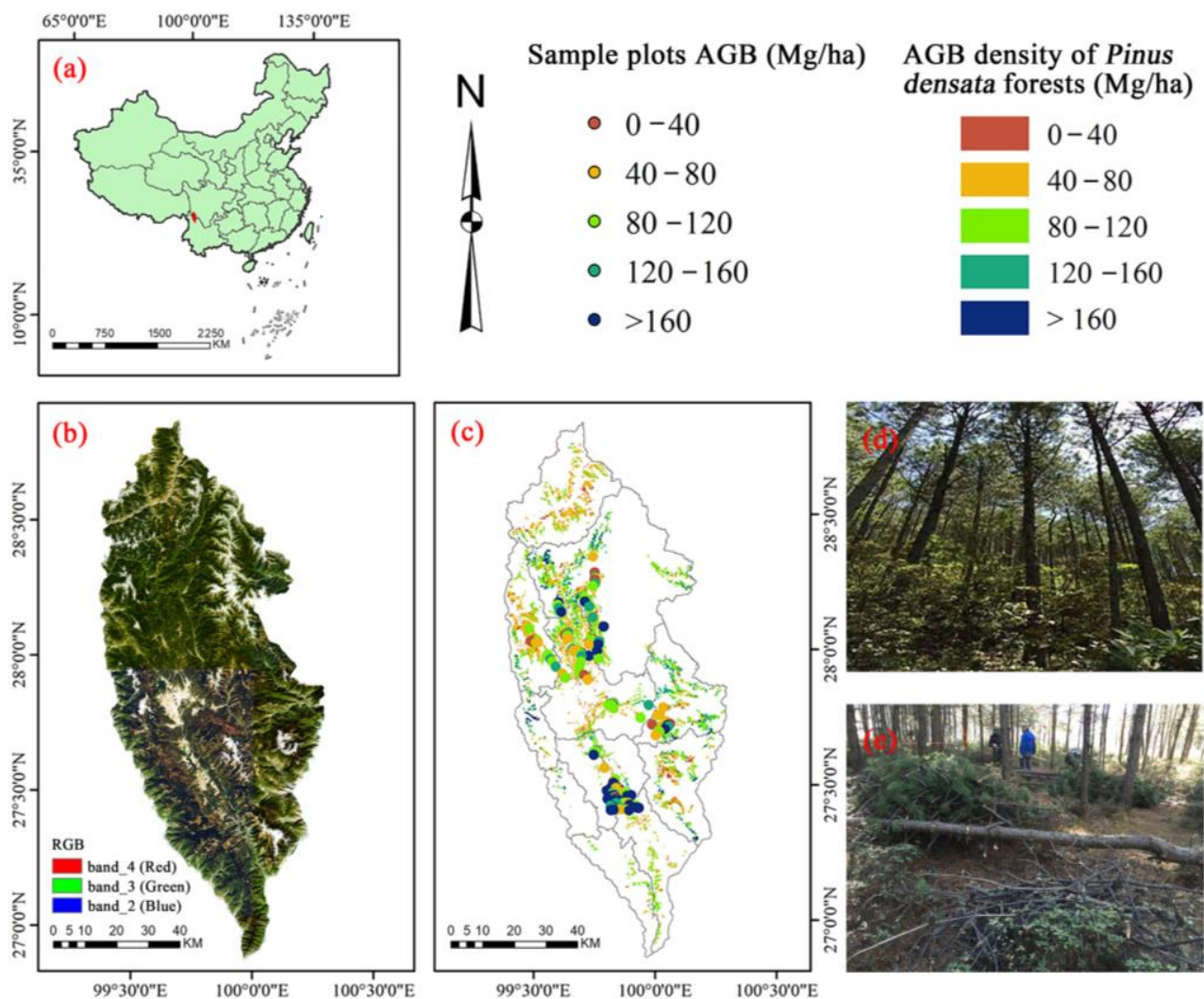


Figure 1. (a) Location of Shangri-La City in China; (b) The Sentinel-2 images of the study area; (c) the spatial distribution of *Pinus densata* forests according to the forest management inventory (FMI) data in 2016 and the sample plots investigated in 2016; (d) the typical stand structure of *Pinus densata* forests in the study area; and (e) the field investigation of AGB.

Pinus densata, one of the barren tolerance pioneer tree species of sub-alpine coniferous, is light-loving and cold-resistant in Shangri-La City. *Pinus densata* forests are single-storied stands with even age in common, and most of the study areas were conducted in pure *Pinus densata* forests [60,61,63] (Figure 1).

2.2. Flow Chart

In Figure 2, the methodological framework of this study was described in the following steps: (1) collecting the sample plots and tree biomass data and the Sentinel images data; (2) calculating the plot AGB; (3) pre-processing of the Sentinel images; (4) correlation between spectral variables and AGB; (5) developing the model: the artificial neural network (ANN), random forests (RF), and the quantile regression neural network (QRNN); and (6) assessing the models.

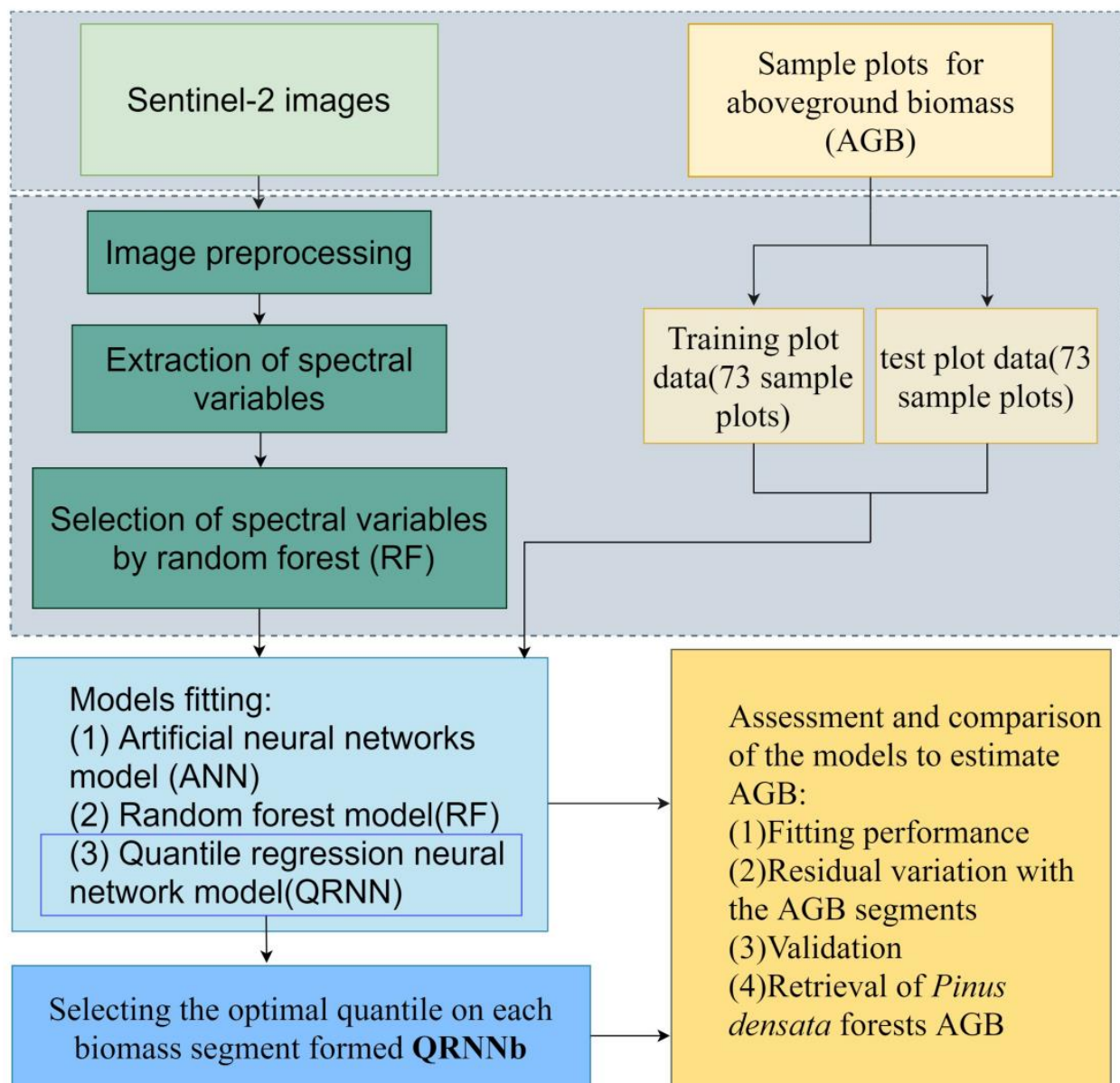


Figure 2. The methodological framework of estimating the forest above-ground biomass (AGB). RF is the random forests, ANN is the artificial neural network, QRNN is the quantiles regression neural network, and QRNNb is the quantile regression neural network with the best fitting performance in each biomass segment.

2.3. Field Data Collection and Aboveground Biomass Calculation

Field data collection work was conducted in August 2016, and in situ data from over 146 sample plots were collected. The plot size was 30 m × 30 m. A GPS was used to measure and record the coordinates and elevation. All of the trees of each plot with a diameter at breast height 1.3 m above ground (DBH) >5 cm were measured. The trees on the south and west boundary of the sample plot were recorded. Three to five trees with a similar average stand DBH were chosen, and the height of the selected trees was measured to calculate the average height of the stand in each plot. The other information in the plot needed to be recorded, such as forest site conditions, origin, age, soil, and the trees' health situation, etc.

The process of investigation, sampling, determination, and individual tree biomass construction has been detailed in the literature [63]. The equation for the tree AGB was as below:

$$AGB_i = 0.073 \cdot DBH^{1.739} \cdot H^{0.880} \quad (1)$$

where DBH is the tree diameter at the breast height >5 cm, H is the tree height, and AGB_i is the AGB of the individual tree in the plot (kg).

Equation (2), as below, was the sample plot AGB (Mg/ha). To ensure enough comparable sample plot datasets at each biomass segment for the fitting test and validation test, the sample numbers of the two datasets were the same; then, 146 sample plots were randomly divided into a fitting dataset of 73 plots and a test dataset of 73 plots, and the statistical information is listed in Table 1. In addition, there were no significant statistical differences in the mean and standard deviation values between the fitted and the test datasets.

$$AGB_s = \frac{\sum_{i=1}^n AGB_i}{900} \cdot 10,000/1000 \quad (2)$$

where AGB_s is the AGB of a plot, AGB_i is the AGB of individual trees, and n is the number of trees within each plot.

Table 1. The statistical parameters of sample plot datasets. H is the average tree height, Dg is the average diameter at breast height (1.3 m), and AGB the is above-ground biomass.

Variables		Fitting Data ($n = 73$)	Test Data ($n = 73$)	All Data ($n = 146$)
Minimum	H (m)	2.2	2.9	2.2
	Dg (cm)	2.9	4.9	2.9
	AGB (Mg/ha)	2.1	11.1	2.1
Maximum	H (m)	24.3	19.5	24.3
	Dg (cm)	41.3	24.7	41.3
	AGB (Mg/ha)	335.9	344.4	344.4
Mean	H (m)	10.0	10.3	10.1
	Dg (cm)	14.6	15.0	14.8
	AGB (Mg/ha)	120.7	122.2	121.5
Standard deviation	H (m)	3.8	3.7	3.7
	Dg (cm)	6.3	4.5	5.5
	AGB (Mg/ha)	67.5	79.9	73.7

2.4. Remote Sensing Data and Variables

2.4.1. Pre-Processing of Sentinel-2 Images

Five Sentinel-2 images obtained from the European Space Agency (ESA) were used in this study (Table 2). Since there were no level-2A products before May 2017, level-1C products with UTM/WGS 84 ortho-images were downloaded, and they were orthorectified top-of-atmosphere reflectance products. Bottom-of-atmosphere reflectance product L2A needed to be obtained by atmospheric correction. Thereby, the Sen2Cor (version 02.05) plugin under the toolbox in SNAP was installed to create L2A products, and the open-access software of SNAP was downloaded from <http://step.esa.int/main/download/snap-download/> and accessed on 10 October 2022. Then, we resampled all of the bands with a 10 m resolution under cubic convolution interpolation by using the resample tool in SNAP. Finally, we resampled all of the bands with a 30 m resolution to meet the plot size of the field AGB survey, and the images were cropped and spliced in ENVI.

Table 2. The parameters of five Sentinel-2 images.

Image ID	Acquisition Date	Central Longitude (Degree)	Central Latitude (Degree)	Solar Elevation	Solar Azimuth	Mean Cloud Amount (%)
S2A_MSIL1C_20161124T040102_N0204_R004_T47RNK_20161124T040118	24 November 2016	99.5513	26.6257	1.0249	162.1176	12.6
S2A_MSIL1C_20161124T040102_N0204_R004_T47RNL_20161124T040118	24 November 2016	99.5557	27.5287	1.0249	162.2853	25.6
S2A_MSIL1C_20161124T040102_N0204_R004_T47RNM_20161124T040118	24 November 2016	99.5604	28.4315	1.0249	162.4446	41.7
S2A_MSIL1C_20161124T040102_N0204_R004_T47RPL_20161124T040118	24 November 2016	100.5684	27.5209	1.0249	163.4582	15.1
S2A_MSIL1C_20161124T040102_N0204_R004_T47RPL_20161124T040118	24 November 2016	100.5815	28.4235	1.0249	163.6144	38.5

2.4.2. Extraction Feature Variables from Remote Sensing

The vegetation index and conversion factor have been widely used to estimate forest AGB [27,28]. The texture feature is an essential feature of remote sensing images, and it reflects the properties of the object itself and helps to distinguish two different objects [28]. First and foremost, texture features have been proven to have essential contributions to increasing the accuracy of AGB estimation because they can describe complex forest structures with high accuracy [17,28,31]. Therefore, this study extracted 134 remote sensing variables, including 11 spectral bands, 21 vegetation indices, 6 image conversion algorithms, and 96 texture measurements (Table 3).

Table 3. Spectral variables derived from Sentinel-2 images.

Data Sources	SV	Definitions of SV	Number of SV
Sentinel-2	Original band	b2—blue, b3—green, b4—red, b5—vegetation red edge, b6—vegetation red edge, b7—vegetation red edge, b8—NIR, b9—water vapor, b10—SWIR-cirrus, b11—SWIR, b12—SWIR	11
	Vegetation indices	Normalized difference vegetation index (NDVI), atmospherically resistant vegetation index (ARVI), difference vegetation index (DVI), ratio vegetation index (RVI), vegetation index of soil adjustment ratio (SARV), oil adjusted vegetation index (SAVI), modified soil vegetation index (MSAVI), short infrared temperature vegetation index (MVI5), mid-infrared temperature vegetation index (MVI7), transformation vegetation index (TVI), nonlinear vegetation index (NLI), perpendicular vegetation index (PVI), infrared vegetation index (II), optimization simple ratio index (MSR), simple vegetation index (SR), brightness vegetation index (B), temperature vegetation index (W), greenness vegetation index (G), normalized difference vegetation index using R and G bands (ND43), normalized difference vegetation index using band 6 and band 7 (ND67), normalized difference vegetation index using band 5, band 6, and band 3 (ND563)	21
	Image transformations	The first three components from the tasseled cap transform (K T transform) and the first three principal components of principal component analysis (PCA)	6
	Texture measures	Grey-level co-occurrence matrix-based texture measures including the mean, angular second moment, contrast, correlation, dissimilarity, entropy, homogeneity, and variance using moving window sizes of 3×3 , 5×5 , and 7×7 pixels	96

2.4.3. Variables Screening

A total of 134 variables were extracted, but not all of them were sensitive to AGB. Random forests (RF) technology was chosen to analyze the correlation between derived variables and the field-based AGB data to gain a set of parsimonious and valid variables for the AGB model. Then, the final vital performed variables were selected for building the regression model. RFs is an ensemble machine-learning algorithm that was first proposed by Breiman [36]. The keys to an RF construction include the random selection of the decision tree and features (subset). Two thirds of variables are randomly selected from the original dataset by the bootstrap sampling method to avoid over-fitting so that the training dataset of the decision tree and the data amount of all training datasets is consistent with the amount of original data [64]. The features (the other 1/3 of the original data) are selected as the nodes of each decision tree and they were also chosen randomly. The features are split based on the Gini criterion. The remaining features are the out-of-bag (OOB) data used as validation samples. OOB data can be used to calculate the unbiased estimate of prediction error by comparing the dataset with the out-of-bag data. Meanwhile, they can also be applied to determine the importance of the variables. The optimal solution is obtained by voting according to the principle that the minority is subordinate to the majority. Moreover, the quality of the RFs model is related to the mean square errors (MSE) between the decision tree and the features, and the smaller the MSE is, the better [36]. In this study, 80% of the original data was used as the training dataset, and the left data were the test dataset. Random forest recursive feature elimination (RF-RFE) was used to remove variables that did not contribute significantly to model accuracy. This experiment was conducted in the sklearn.assembly module of Python 3.7, used the RandomizedSearchCV and GridSearchCV functions to optimize the model parameters and variables screening.

2.5. Modeling Methods

2.5.1. Random Forests Modeling (RF)

Random forests (RF) are an accurate methodology for classification and a validation way to predict the AGB [65]. The two parameters that must be set are the number of trees for growing (ntree) and the split variables for selecting randomly. Balancing the two parameters is the most critical work for avoiding the lowest generalization error [36]. Different numbers for the ntree and minimum sample split (mtry) and the other factors, such as the max-depth (the sample depth that contains the minimum sample) and min-sample-leaf (the minimum number of samples at the leaf node), were chosen to compare the R^2 in Python. The highest R^2 was finally obtained. The parameters were set as follows: the maximum number of iterations was 200, the max-depth was 10, the min-samples-leaf was 1, and the min-samples-split was 2. In this study, 80% of the original data was used as the training dataset, and the other 20% was the test dataset. Ten-fold cross-validation was applied to prevent over-fitting and to prevent it affecting the accuracy and stability of the model.

2.5.2. Artificial Neural Networks Model (ANN)

The artificial neural network (ANN) is a mathematical model for information processing using similar structures to the synaptic connections in the brain [66]. It consists of a large number of nodes (or neurons) that are connected. Each node represents a specific output function called the activation function. Each connection between two nodes represents the weighted value of the signal passing through the link. The training of the neural network model is the process of modifying the connection weight between the neuron and the neuron deviation according to the training data. ANN comprises three essential elements: the processing unit, network topology, and training rules. The processing unit is the basic unit of artificial neural network operation. A processing unit has multiple input and output paths. The network topology determines the information transmission between each processing unit and each layer, generally composed of an input layer, a hidden layer, and an output layer [44]. The number of hidden layer nodes has been paid a lot attention to

because the network cannot have the necessary learning ability and information prediction ability if the number of hidden layer nodes is too small. On the contrary, it will increase the complexity of the network structure and make the network fall into a partial minimum or lead to overfitting [67]. Training rules are trained and adjusted repeatedly to achieve the required accuracy. It mainly uses the transformation function to weigh and sum the processed data and trains the network system to carry on the pattern recognition.

The back propagation neural network (BPNN), one of the most widely used neural network algorithms, was applied in this study. The BP neural network model was constructed by using the R language package. Firstly, the initially hidden node was set to 4. It was found that the average error decreased at the beginning and then increased with the number of hidden nodes increasing. When the number of hidden nodes was 7, the average error was minimum. Ten-fold cross-validation was used to test the accuracy; 80% of the modeling samples was used as a training set, and 20% was used as a test set. Each test data will yield a corresponding rate of accuracy (or error rate). The average of the accuracy (or the error rate) of the ten modeling results was used as an estimate of the accuracy of the algorithm [68,69].

2.5.3. Quantile Regression Neural Network (QRNN)

The nonlinear relationship between a dependent and independent variable is very complex and challenging to describe. Taylor used neural network structure to establish neural network quantile regression (QRNN) [51]. The model combines the nonparametric, nonlinear quantile regression method and achieves a nonlinear mapping of conditional quantiles from dependent variables to independent variables. As the artificial neural networks, the number of hidden layer nodes has an essential effect on the complexity of the model. The number of hidden layer nodes should be manageable because it would cause the fitting time to be too long, which may add non-regular content and this leads to over-fitting [55].

QRNN was constructed with the QRNN function package in R software. Three hidden layers and seven hidden nodes were used that were the same as the ANN. At the same time, 10-fold cross-validation was also carried out to prevent over-fitting or reduce errors from affecting the accuracy and stability of the model. The scale of the training and test dataset was the same as in the ANN.

Moreover, the corresponding optimal quartile models with the lowest mean error at each AGB segment were combined as the best QRNN (QRNNb). In addition, the AGB segments were 0–40 Mg/ha, 40–80 Mg/ha, 80–120 Mg/ha, 120–160 Mg/ha, and greater than 160 Mg/ha. Therefore, the QRNNb represents a complete biomass estimation model formed by selecting the highest precision of the five quantitative models corresponding to the QRNN on each of the five biomass segments.

2.6. Assessment and Validation of the Models

It is critical to obtain the AGB model and the assessment values during the process of AGB model building. Coefficient determination (R^2) and mean square root error (RMSE) were used to estimate the AGB prediction model and assessment value. R^2 and RMSE were applied to compare the accuracy of prediction values from different estimate models based on fitting plots data in Table 1.

Linear regression between AGB predicted values of different biomass segments and the observed data was used to assess models' performance using 73 test plots. In addition, except for R^2 and RMSE, the mean absolute error (MAE) and mean error (ME) were added to test the validation of each model by using the test dataset. The AGB segments were 0–40 Mg/ha, 40–80 Mg/ha, 80–120 Mg/ha, 120–160 Mg/ha, and greater than 160 Mg/ha [41].

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

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (4)$$

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

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

where $\hat{y}_i$ and y_i are the predicted AGB and the corresponding AGB in the sample plot, $\bar{y}$ is the mean AGB of the sample plots, and n is the number of sample plots.

3. Results

3.1. Results of Spectral Variables Screening

The prediction accuracy would decrease if all of the biomass prediction variables extracted from the images were applied to build the AGB estimation model, and information redundancy would also occur. The function of AGB assessment would be reduced as some variables may have a weak association with biomass. Thus, screening suitable and strong correlation variables was the critical step. In this study, RFs was used to screen the characteristic variables according to the sort of variable importance. In addition, the first ten important variables were VA3_2, VA5_12, CO7_8, DI5_8, VA5_2, HO5_3, VA3_12, VA7_12, ME5_12, and SE5_3. To prevent the selected variables from displaying multi-collinearity and thus reducing the accuracy of biomass estimation, we performed collinearity tests on selected variables using the Kappa functions in R software. The results indicated less collinearity between the variables as the Kappa coefficient value was 11.72709, which was less than 100. The correlation between forest AGB and the characteristic variables of *Pinus densata* forests is shown in Figure 3.

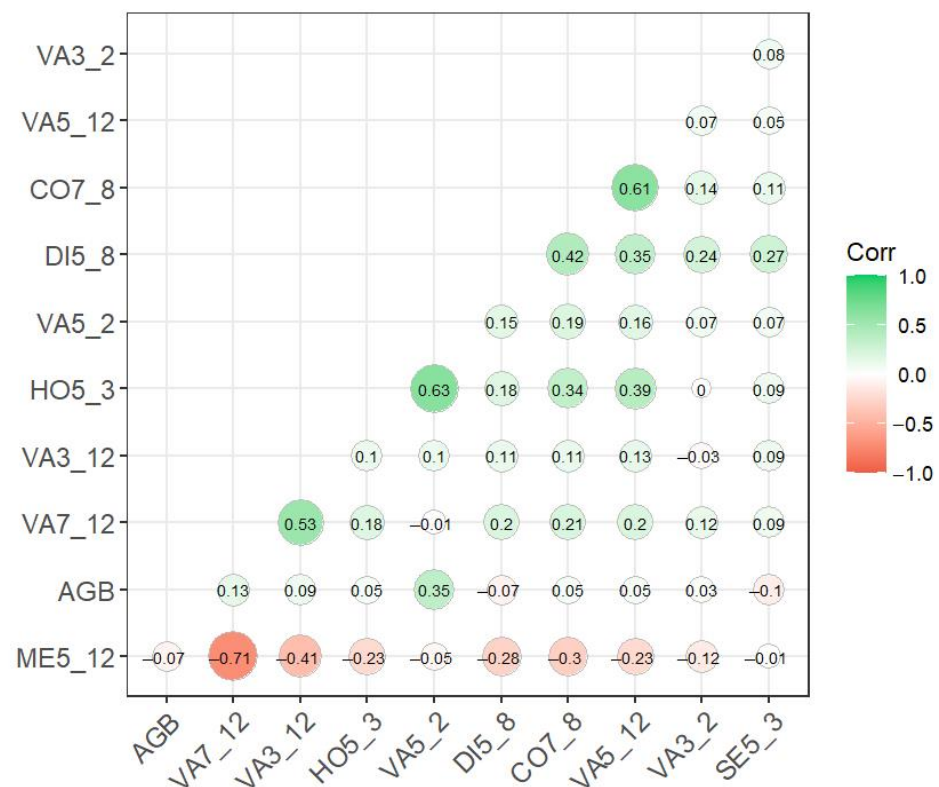


Figure 3. The correlation between AGB and the characteristic variables of *Pinus densata* forests. Corr is the correlation coefficient between the characteristic variables and AGB. VA3_2 is the variance on band 2 with the window size 3×3 , VA5_12 is the variance on band 12 with the window size 5×5 ,

CO7_8 is the correlation on band 8 with the window size 7×7 , DI5_8 is the dissimilarity on band 8 with the window size 5×5 , VA5_2 is the variance on band 2 with the window size 5×5 , HO5_3 is the third homogeneity on band 3 with the window size 5×5 , VA3_12 is the variance on band 12 with the window size 3×3 , VA7_12 is the variance on band 12 with the window size 7×7 , ME5_12 is the mean on band 12 with the window size 5×5 , and SE5_3 is the second moment on band 5 with the window size 3×3 .

3.2. Model Comparison of the Model

3.2.1. Model Fitting

Scatter plots of AGB and predicted biomass based on the ANN and QRNN models based on ten variables are shown in Figure 4. It was shown that the ANN's fitting performance was not significantly different from that of QRNN at 0.1, 0.25, and 0.5 percentiles. However, when the optimal quantile model for each biomass segment was integrated into a complete QRNNb (Figure 4c), the fitting accuracy of the QRNNb was significantly improved. The R^2 and RMSE of the ANN were 0.722 and 31.0689, respectively. In addition, the R^2 and RMSE of the QRNNb were 0.962 and 13.9326, respectively. The results also demonstrated that the fitting performance of RFs ($R^2 = 0.934$, RMSE = 11.3305) was quite similar to that of the QRNNb. RFs and the QRNNb had a better fitting performance than the ANN and the QRNN, which means both RFs and the QRNNb had a higher accuracy than the ANN and the QRNN.

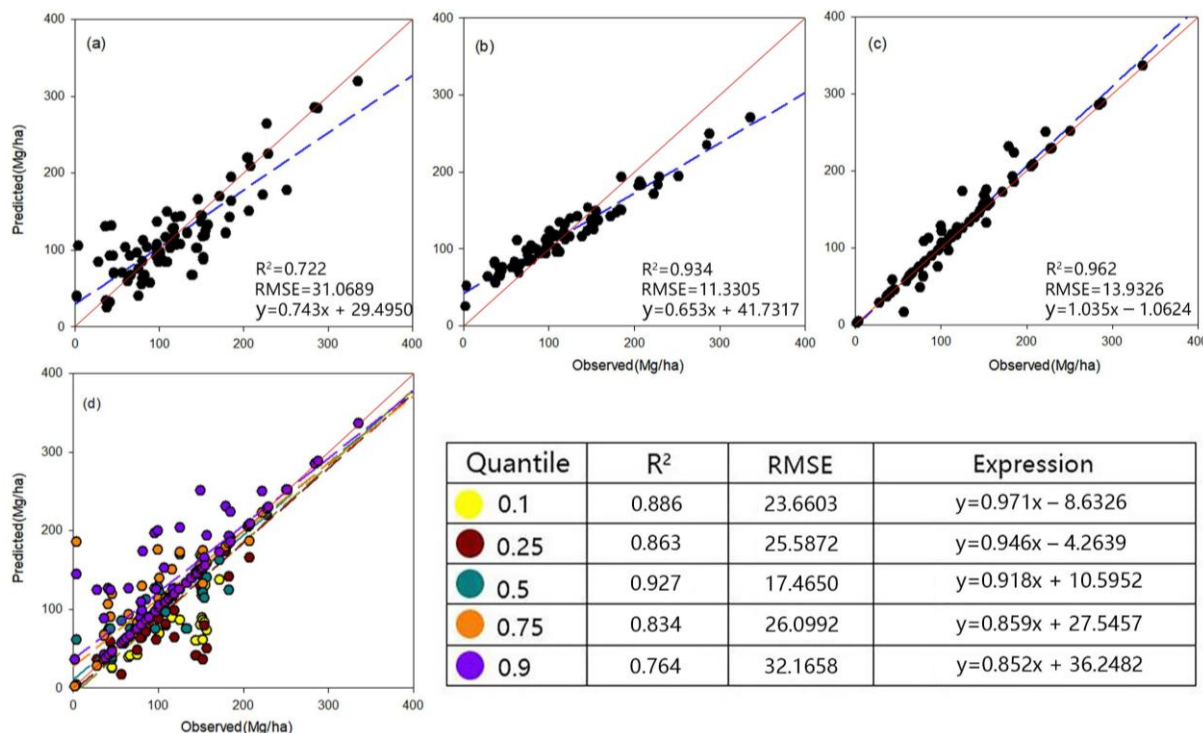


Figure 4. Scatter plots of the ground-observed and estimated biomass values for the (a) artificial neural network model (ANN); (b) the random forests model (RF); (c) the quantile regression neural network model (QRNN), and the quantiles groups are 0.1, 0.25, 0.5, 0.75, and 0.9; and (d) the quantile regression neural network with the best fitting performance in each biomass segment (QRNNb).

Compared to the scatter plots of the ANN, RFs, and the QRNN at each quantile, the scatter plot of the QRNNb was narrower and looked more similar to the line of $y = x$. The ranked absolute intercept value of each model was: RFs (41.7317) > QRNN0.9 (36.2482) > ANN (29.4950) > QRNN0.75 (27.5457) > QRNN0.5 (10.5952) > QRNN0.1 (8.6326) > QRNN0.25 (4.2639) > QRNNb (1.0624). The larger the intercept, the greater the angle with $y = x$, indicating the greater the degree of deviation. Figure 4a shows that the ANN had an

excellent fitting performance at the middle biomass level. Still, it would overestimate the lower biomass value and underestimate the higher one as it had a greater intercept value. Similarly, RFs (Figure 4b) and the QRNN (Figure 4d) at each quantile showed the same phenomenon. Figure 4c indicates that the QRNNb had an excellent accuracy because it had a smaller intercept.

3.2.2. Method Validation

The biomass prediction accuracy of each model for each biomass segment was further validated by comparing the R^2 and RMSE (Table 4). These results indicated that QRNNb has a higher R^2 (0.943) and a higher RMSE (18.203) in the three models, especially as the AGB segment was 0–40 Mg/ha and >160 Mg/ha.

Table 4. Summary of R^2 , RMSE, ME, and MAE at the different AGB segments based on the test dataset. ANN is the artificial neural network, RFs is the random forests, and QRNNb is the best quantile regression neural network in each biomass segment.

Indices		Models		
		ANN	RFs	QRNNb
R^2	0–40	0.105	0.402	0.961
	40–80	0.043	0.094	0.757
	80–120	0.167	0.598	0.430
	120–160	0.277	0.385	0.671
	>160	0.480	0.857	0.867
	Total	0.602	0.936	0.943
RMSE (Mg/ha)	0–40	8.341	6.818	1.733
	40–80	11.948	11.624	6.019
	80–120	10.421	7.242	9.851
	120–160	11.915	10.987	8.034
	>160	43.555	23.215	22.052
	Total	48.180	19.396	18.203
ME (Mg/ha)	0–40	−44.364	−30.845	1.035
	40–80	−33.623	−19.38	7.029
	80–120	−0.338	2.093	2.683
	120–160	13.741	8.230	−6.861
	>160	44.386	34.321	−11.617
	Total	−1.507	1.927	−1.419
MAE (Mg/ha)	0–40	48.400	30.846	1.035
	40–80	36.041	19.438	7.090
	80–120	11.213	5.720	5.926
	120–160	18.874	18.482	9.202
	>160	47.465	34.321	11.618
	Total	32.066	21.271	8.357

For the ME values, there were significant differences among the three models in different biomass segments, and QRNNb had no significant difference from zero at each biomass segment. The ANN and RFs showed negative mean errors in the 0–40 Mg/ha biomass segment. They were significantly different from zero at the significance level of 0.01, which means significant overestimation in the AGB segment. The ANN and RFs had a positive mean error as the segment was at 80–120 Mg/ha and >160 Mg/ha, demonstrating a significantly different value from zero, which would give a lower estimate at a higher biomass value, especially as the AGB was greater than 160 Mg/ha.

The MAE values showed that QRNNb was not significantly different from zero, and the MAE value was 8.359. QRNNb had a small MAE at the lower and higher biomass segments, which means the prediction values at these two segments were close to the observed value. The MAE showed that the prediction value from RFs and the ANN models at 80–120 Mg/ha had a minor error compared with the other biomass segments, while

the MAE for the RFs and ANN models at 0–40 Mg/ha and >160 Mg/ha showed that the prediction value had a substantial deviation from zero. The bias of the ANN and RFs models for all of the biomass segments except for the 80–120 Mg/ha segment was relatively high. The highest MAE of ANN and RFs was 48.400 Mg/ha and 30.846 Mg/ha at the biomass segment of 0–40 Mg/ha, and 47.465 Mg/ha and 34.321 Mg/ha at the AGB >160 Mg/ha, respectively. In addition, RFs and the ANN showed significant deviations from zero.

In sum, QRNNb was more accurate than the ANN and RFs in biomass estimation, especially in the low-biomass segment and the high-biomass segments. QRNNb can improve the problem of low-value overestimation and high-value underestimation and it has a very stable prediction effect.

The AGB maps of the *Pinus densata* forests are shown in Figure 5, which was inverted by using three models. The high heterogeneity of the AGB distribution can be seen using the model of QRNNb, which means the model of QRNNb has an excellent prediction of AGB biomass value at each of the AGB segments. On the contrary, the ANN had a higher count at the segment with 120–160 Mg/ha and 40–80 Mg/ha, which means that the ANN cannot capture the AGB at the lower biomass segment, which would lead to an over-estimation of the low AGB biomass. Meanwhile, some of the higher AGB biomass values (>160 Mg/ha) may be counted into 120–160 Mg/ha, leading to an under-estimation of the high AGB biomass. The prediction AGB biomass values of RFs were more concentrated at 40–80 Mg/ha and 80–120 Mg/ha than the ANN. This proved that the high precision of RFs was at the cost of discarding high accuracy.

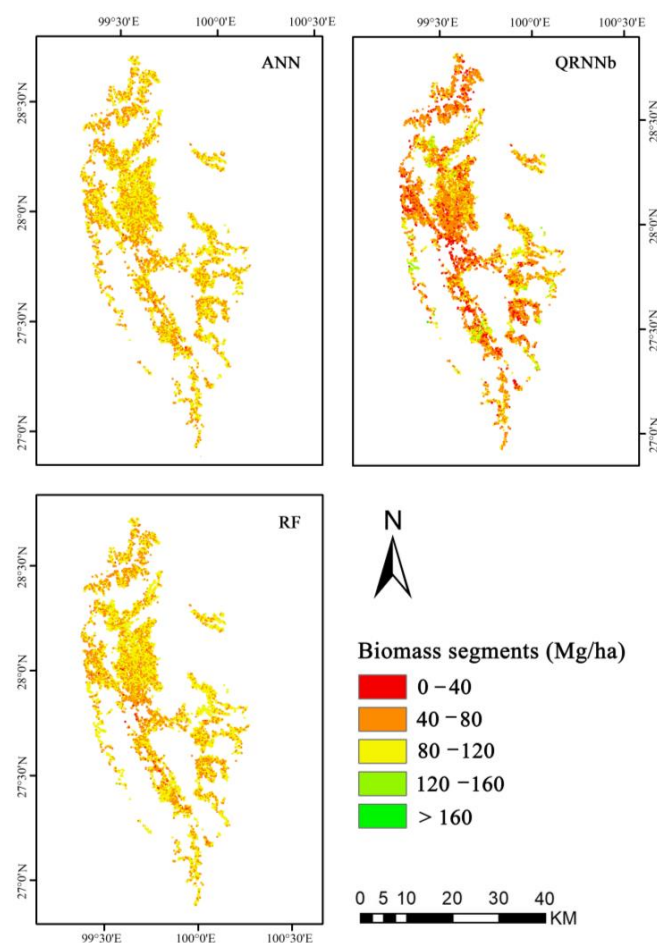


Figure 5. The spatial distributions of the predicted aboveground biomass (AGB) values of the *Pinus densata* forests using four models. ANN is the artificial neural network, RF is the random forests, and QRNNb is the best quantile regression neural network in each biomass segment.

4. Discussion

4.1. Accuracy Comparison

Shettles et al. [69] found that model uncertainty is the main element affecting the accuracy of AGB estimation, and model uncertainty accounts for 55% of total uncertainty. Thus, improving model accuracy is still the main challenge for AGB estimation using optical remote sensing data. This research attempted to promote the accuracy of biomass assessment by comparing three non-parametric model regression models. The results have shown that the ranked fitting performance for the three models was the QRNNb > RFs > the ANN. From the values of R^2 and RMSE in the fitting model using the observed and predicted values, the accuracy for the RFs was slightly better than the QRNNb. Still, the intercept for the QRNNb was 1.0624 Mg/ha, which means the prediction value was much closer to the observed value. In contrast, RFs had apparent phenomena of under-estimation at higher biomass segments and over-estimation at lower biomass segments with a high intercept. The value reached 41.7317 Mg/ha, affecting the entire forest AGB assessment value. Thus, the QRNNb has the best performance among the three models. Moreover, RFs has a higher R^2 value and a lower RMSE in this study. Many studies have shown that RFs exhibited excellent performance [70,71]. Then, RFs was the most optimal model with the highest accuracy under the premise of considering only the overall situation.

Furthermore, for the different biomass segments, the results showed RFs at the lower and higher biomass segment was significantly worse than the QRNNb, the R^2 values for RFs at AGB < 40 Mg/ha and >160 Mg/ha were lower than for the QRNNb, and the RMSE values at both biomass segments for RFs were extremely larger than the QRNNb. This reveals that the QRNNb could promote biomass estimation accuracy, especially at the lower and higher biomass segments. The QRNNb could describe the complete conditional distribution of biomass with more stability and it is not easily affected by the extreme value. Then, the QRNNb would be an excellent method to reduce the uncertainties from over-estimation and under-estimation in the AGB estimation using optical remote sensing data.

In addition, the Sentinel-2 images were resampled with 30 m × 30 m corresponding with the plot size of the field survey in this study. The mismatch between the former image spatial resolution and field size would affect the AGB estimation accuracy. We performed AGB estimation using the resampled Sentinel-2 image product with a spatial resolution of 10 m. Similar fitting and validation results for the three models were obtained, and the QRNNb was more accurate than the ANN and RFs in biomass estimation, especially in the low-biomass segments and the high-biomass segments (see Appendices A and B). This further illustrates the availability of the proposed method for reducing the uncertainties of AGB estimation using optical remote sensing.

4.2. Data Resource and Variables

The information extracted from optical remote sensing is the radiation information of the canopy surface, which is easily affected by the complexity of forest crown layers. Therefore, the precision problem is the biggest challenge of optical remote sensing in current remote sensing biomass estimation [19,20]. Using high-resolution and hyperspectral remote sensing images will enhance biomass estimation accuracy, but the high price limits such data being widely utilized [17,72]. Researchers prefer to choose free, open-source data, such as Landsat or Sentinel-2. Even though those two are both optical remote sensing, Sentinel-2 has a double-satellite orbit and has four more bands than Landsat. It is the unique one with three bands of data in the red edge range, which can efficiently obtain more rich geographical information [21]. Studies have shown that Sentinel-2 is more suitable than Landsat for improving estimation accuracy [73]. Although the vegetation index will bring a saturation problem, the vegetation index extracted from near-infrared and red edge can strengthen the estimation accuracy [74]. This study found that band 2 (blue), band 3 (green), band 5 (vegetation red edge), band 8 (NIR), and band 12 (SWIR) of Sentinel-2 had a strong correlation with biomass. Because the vegetation index is affected by the saturation value in biomass estimation, the texture feature has been introduced as a variable. Then, the biomass

value is more sensitive to the texture feature [75,76]. This study also extracted the textural features of different window sizes (3×3 , 5×5 , and 7×7) to model. After screening and analysis, it was found that the texture information of entropy and the correlation with various window sizes and bands (VA5_2 and VA7_12) strongly correlated with biomass.

Moreover, Shangri-La City has a cold-temperate monsoon climate with altitudes ranging from 3350 to 3696 m above sea level. The cloud and snow significantly affect the spectral bands of optical remote sensing [77]. Lacking high-quality images with a lower cover of cloud and snow corresponding to the field investigation date, we only obtained five Sentinel-2 images from the ESA. The image acquisition date is 24 November 2016. The time difference between the survey data (August) and the remote sensing data (November) is about three months. To avoid or reduce the impact of the time mismatch between image acquisition and the field survey, we obtained the bottom-of-atmosphere reflectance product by atmospheric correction to normalize as a common reference [78]. Furthermore, *Pinus densata* is an evergreen coniferous tree distributed in the alpine and sub-alpine areas in China, and it grows slowly within one to two years [79]. Therefore, the tree growth and forest structure are almost unchanged; then, the change of image reflectance caused by forest growth in the three months will have a negligible impact on the AGB estimation in this study.

4.3. Limitation and Future Research

Although QRNNb obtained a high-accuracy estimation in the different biomass segments, this study still has some limitations. Firstly, Sentinel-2 can yield an accurate biomass estimation [23]. Still, some studies have shown that mixed remote sensing data are more precise than single-source data, especially in tropical and subtropical regions where the stand structures and tree species are complex [80–82]. Secondly, the accuracy of AGB estimation is highly dependent on prediction methods [83]. Therefore, other models for biomass estimation in subsequent studies should be considered to improve the precision of biomass estimation, for instance, combining quantile regression and random forests to form quantile random forests (QRF), the convolutional neural networks (CNN), the gradient boost regression tree (GBRT) [84–86], etc. Thirdly, the best combination of different vegetation indices is expected to predict the AGB of vegetation at different stages [87].

Moreover, we only selected *Pinus densata* forests as the research area. They are mainly distributed over the subalpine and alpine areas in southern Qinghai, western Sichuan, northwestern Yunnan, and southeastern Tibet in China. In addition, *Pinus densata* forests are single-storied stands with even age in common [60,61]. Therefore, the proposed method can be applied to improve the forest AGB estimation for even-aged or single-storied forests. The applicability in the multi-storied stands or the uneven-aged forests with complex stand structures would be further explored.

5. Conclusions

To reduce uncertainties from under-estimation and over-estimation, optical remote sensing was applied to assess forest AGB. In this study, Sentinel-2 was used to explore the potential and capability of three non-parametric models of the ANN, RF, and the QRNN for *Pinus densata* in Shangri-La City. In addition, the biomass was segmented, and the quantile regression neural network with the best fitting performance in each biomass segment was selected to combine an integrity model named QRNNb. The results showed: (1) from the whole biomass data, the performance of QRNNb and RFs was a priority over the ANN. The corresponding R^2 and RMSE were QRNNb: 0.943, 18.203 Mg/ha; RF: 0.936, 19.396 Mg/ha; ANN: 0.602, 48.180 Mg/ha. (2) The prediction accuracy of QRNNb at different biomass segments was higher than the ANN and RF. It had the highest R^2 and the smallest RMSE when $AGB < 40$ Mg/ha and $AGB > 160$ Mg/ha. The R^2 at values those two biomass segments were 0.961 and 0.867, and the RMSE values for those two were 1.733 Mg/ha and 22.052 Mg/ha. This demonstrated that QRNNb could efficiently improve the under-estimation at higher biomass values and the over-estimation at lower biomass

values compared with the ANN and RF. QRNNb was sensitive to extreme values and could express low biomass values and high biomass values wholly and effectively. This means that QRNNb combined with the optimal quantile model of each biomass segment provides a more suitable method for estimating AGB for even-aged or single-storied forests.

Author Contributions: L.L. and B.Z. participated in the collection of the field data, conducted the data analysis, and wrote the draft of the paper; Y.L. and J.T. helped with the data analysis and the writing of the paper; Y.L., Y.W., W.X. and L.W. participated in collecting and analyzing the field data; G.O. supervised and coordinated the research project, designed the experiment, and revised the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant numbers 31770677 and 31760206) and the Ten-Thousand Talents Program of Yunnan Province, China (YNWR-QNBJ-2018-184).

Data Availability Statement: The data used in this study can be acquired from the corresponding author upon reasonable request.

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

Appendix A

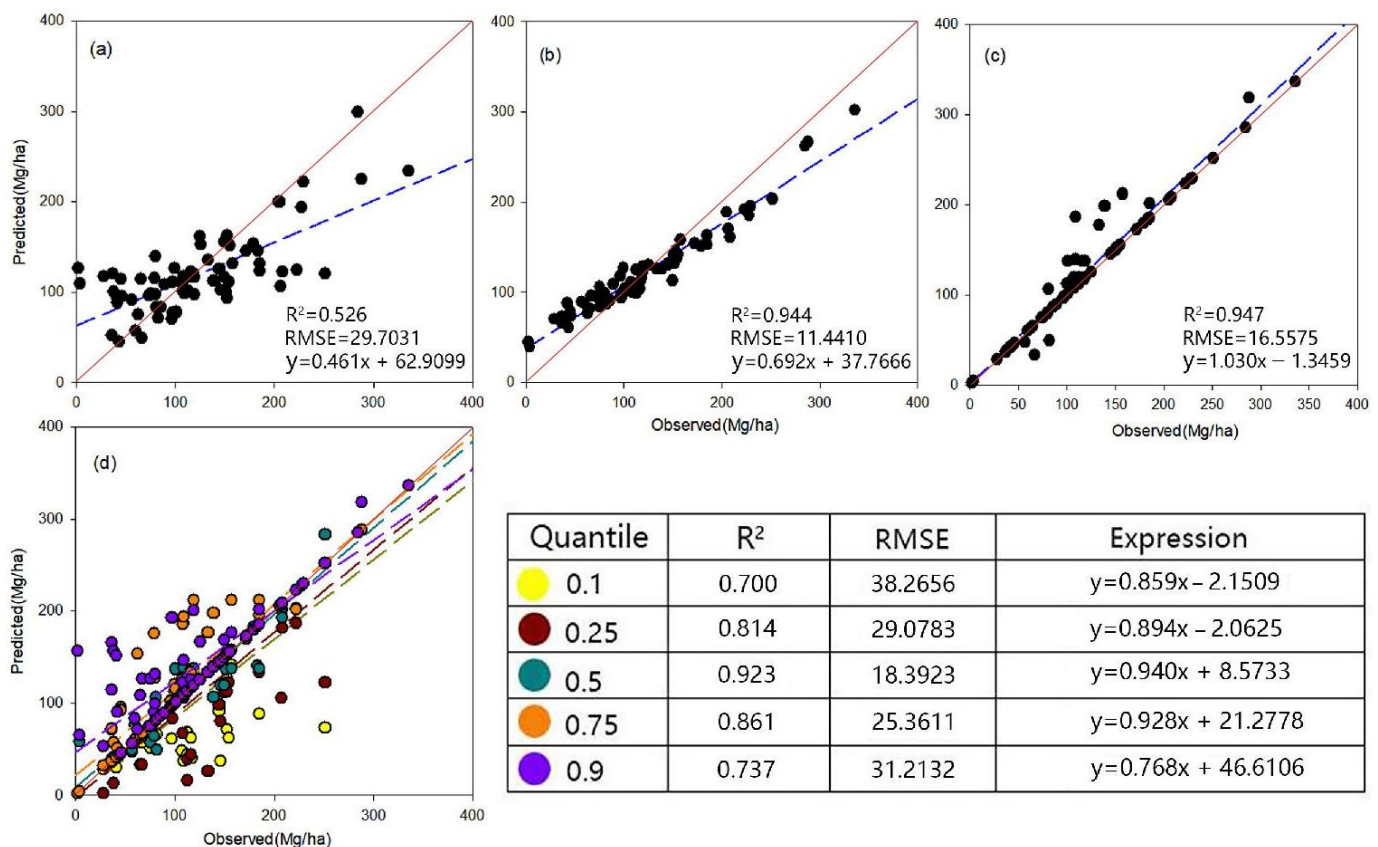


Figure A1. Scatter plots of the ground-observed and estimated biomass values using the resampled Sentinel-2 image product with a spatial resolution of 10 m. (a) The artificial neural network model (ANN); (b) the random forests model (RF); (c) the quantile regression neural network model (QRNN), and the quantiles groups are 0.1, 0.25, 0.5, 0.75, and 0.9, respectively; (d) and the quantile regression neural network with the best fitting performance in each biomass segment (QRNNb).

Appendix B

Table A1. Summary of R^2 , RMSE, ME, and MAE at the different AGB segments based on the test dataset using the resampled Sentinel-2 image product with a spatial resolution of 10 m. ANN is the artificial neural network, RFs is the random forests, and QRNNb is the best quantile regression neural network in each biomass segment.

Indices		Models		
		ANN	RFs	QRNNb
R^2	0–40	0.384	0.076	0.958
	40–80	0.022	0.200	0.889
	80–120	0.050	0.241	0.430
	120–160	0.031	0.302	0.234
	>160	0.257	0.861	0.968
	Total	0.549	0.932	0.956
RMSE (Mg/ha)	0–40	6.919	8.474	1.799
	40–80	12.075	10.926	4.068
	80–120	11.050	9.952	8.621
	120–160	13.586	11.708	12.258
	>160	52.062	22.510	2.774
	Total	51.310	19.960	16.063
ME (Mg/ha)	0–40	−58.615	−31.676	−0.597
	40–80	−37.525	−18.002	1.489
	80–120	−7.131	−1.365	−7.131
	120–160	10.183	7.239	−4.231
	>160	60.937	42.327	0.200
	Total	−0.454	2.275	−1.211
MAE (Mg/ha)	0–40	61.077	31.676	0.601
	40–80	39.856	18.243	0.941
	80–120	20.327	6.905	1.455
	120–160	25.946	10.825	5.084
	>160	65.955	42.327	0.308
	Total	42.060	22.555	2.396

References

- Houghton, R.A.; Hackler, J.L.; Lawrence, K.T. The U.S. Carbon budget: Contributions from land-use change. *Science* **1999**, *285*, 574–578. [\[CrossRef\]](#) [\[PubMed\]](#)
- Feng, H.; Chen, Q.; Hu, Y.; Du, Z.; Lin, G.; Wang, C.; Huang, Y. Estimation of forest aboveground biomass by using a mixed-effects model. *Int. J. Remote Sens.* **2021**, *42*, 8675–8690. [\[CrossRef\]](#)
- Sun, S.; Wang, Y.; Song, Z.; Chen, C.; Zhang, Y.; Chen, X.; Chen, W.; Yuan, W.; Wu, X.; Ran, X.; et al. Modelling aboveground biomass carbon stock of the Bohai rim coastal wetlands by integrating remote sensing, terrain, and climate data. *Remote Sens.* **2021**, *13*, 4321. [\[CrossRef\]](#)
- Wulder, M.A.; Roy, D.P.; Radeloff, V.C.; Loveland, T.R.; Anderson, M.C.; Johnson, D.M.; Healey, S.; Zhu, Z.; Scambos, T.A.; Pahlevan, N.; et al. Fifty years of Landsat science and impacts. *Remote Sens. Environ.* **2022**, *280*, 113195. [\[CrossRef\]](#)
- Foody, G.M.; Boyd, D.S.; Cutler, M.E.J. Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. *Remote Sens. Environ.* **2003**, *85*, 463–474. [\[CrossRef\]](#)
- Puliti, S.; Solberg, S.; Næsset, E.; Gobakken, T.; Zahabu, E.; Mauya, E.; Malimbwi, R. Modelling aboveground biomass in *Tanzanian miombo* woodlands using TanDEM-X world DEM and field data. *Remote Sens.* **2017**, *9*, 984. [\[CrossRef\]](#)
- Xue, B.W. Lidar and Machine Learning Estimation of Hardwood Forest Biomass in Mountainous and Bottomland Environments. Master's Thesis, University of Arkansas, Fayetteville, NC, USA, 2015; p. 1274.
- Tian, Y.; Huang, H.; Zhou, G.; Zhang, Q.; Tao, J.; Zhang, Y.; Lin, J. Aboveground mangrove biomass estimation in Beibu Gulf using machine learning and UAV remote sensing. *Sci. Total Environ.* **2021**, *781*, 146816. [\[CrossRef\]](#)
- Vaglio, L.G.; Chen, Q.; Lindsell, J.A.; Coomes, D.A.; Frate, F.D.; Guerriero, L.; Pirotti, F.; Valentini, R. Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS J. Photogramm. Remote Sens.* **2014**, *89*, 49–58. [\[CrossRef\]](#)
- Listopad, C.M.C.S.; Drake, J.B.; Masters, R.E.; Weishampel, J.F. Portable and airborne small footprint LiDAR: Forest canopy structure estimation of fire managed plots. *Remote Sens.* **2011**, *3*, 1284–1307. [\[CrossRef\]](#)

11. Spriggs, R.; Coomes, D.; Jones, T.; Caspersen, J.; Vanderwel, M. An alternative approach to using LiDAR remote sensing data to predict stem diameter distributions across a temperate forest landscape. *Remote Sens.* **2017**, *9*, 944. [\[CrossRef\]](#)
12. Cutler, M.E.J.; Boyd, D.S.; Foody, G.M.; Vetrivel, A. Estimating tropical forest biomass with a combination of SAR image texture and Landsat TM data: An assessment of predictions between regions. *ISPRS J. Photogramm. Remote Sens.* **2012**, *70*, 66–77. [\[CrossRef\]](#)
13. El Hage, M.; Villard, L.; Huang, Y.; Ferro-Famil, L.; Koleček, T.; Le Toan, T.; Polidori, L. Multicriteria accuracy assessment of digital elevation models (DEMs) produced by airborne P-band polarimetric SAR tomography in tropical rainforests. *Remote Sens.* **2022**, *14*, 4173. [\[CrossRef\]](#)
14. Quegan, S.; Le Toan, T.; Chave, J.; Dall, J.; Exbrayat, J.-F.; Minh, D.H.T.; Lomas, M.; D' Alessandro, M.M.; Paillou, P.; Papathanassiou, K.; et al. The European space agency BIOMASS Mission: Measuring forest above-ground biomass from space. *Remote Sens. Environ.* **2019**, *227*, 44–60. [\[CrossRef\]](#)
15. El Idrissi Essebtey, S.; Villard, L.; Borderies, P.; Koleček, T.; Burban, B.; Le Toan, T. Long-term trends of P-Band temporal decorrelation over a tropical dense forest-experimental results for the BIOMASS mission. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–15. [\[CrossRef\]](#)
16. Le Toan, T.; Quegan, S.; Davidson, M.W.J.; Balzter, H.; Paillou, P.; Papathanassiou, K.; Plummer, S.; Rocca, F.; Saatchi, S.; Shugart, H.; et al. The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle. *Remote Sens. Environ.* **2011**, *115*, 2850–2860. [\[CrossRef\]](#)
17. Papathanassiou, K.P.; Cloude, S.R.; Pardini, M.; Quiñones, M.J.; Hoekman, D.; Ferro-Famil, L.; Goodenough, D.; Chen, H.; Tebaldini, S.; Neumann, M.; et al. Forest Applications. In *Polarimetric Synthetic Aperture Radar: Principles and Application*; Remote Sensing and Digital Image, Processing; Hajnsek, I., Desnos, Y.-L., Eds.; Springer International Publishing: Cham, Germany, 2021; pp. 59–117. ISBN 978-3-030-56504-6.
18. López-Serrano, P.M.; Cárdenas, D.; José, L.; Corral-Rivas, J.J.; Jiménez, E.; López-Sánchez, C.A.; Vega-Nieva, D.J. Modeling of aboveground biomass with Landsat 8 OLI and machine learning in temperate forests. *Forests* **2019**, *11*, 11. [\[CrossRef\]](#)
19. Zeng, N.; Ren, X.L.; He, H.L.; Zhang, L.; Zhao, D.; Ge, R.; Li, P.; Niu, Z.G. Estimating grassland aboveground biomass on the Tibetan Plateau using a random forest algorithm. *Ecol. Indic.* **2019**, *102*, 479–487. [\[CrossRef\]](#)
20. Sagang, L.B.T.; Ploton, P.; Sonké, B.; Poilvé, H.; Couteron, P.; Barbier, N. Airborne Lidar sampling pivotal for accurate regional AGB predictions from multispectral images in forest-Savanna landscapes. *Remote Sens.* **2020**, *12*, 1637. [\[CrossRef\]](#)
21. SUHET. *Sentinel-2 User Handbook*; European Space Agency: Paris, France, 2013.
22. Chen, Y.; Guerschman, J.; Shendryk, Y.; Henry, D.; Harrison, M.T. Estimating pasture biomass using Sentinel-2 imagery and machine learning. *Remote Sens.* **2021**, *13*, 603. [\[CrossRef\]](#)
23. Castillo, J.A.A.; Apan, A.A.; Maraseni, T.N.; Salmo, S.G. Estimation and mapping of above-ground biomass of mangrove forests and their replacement land uses in the Philippines using Sentinel imagery. *ISPRS J. Photogramm. Remote Sens.* **2017**, *134*, 70–85. [\[CrossRef\]](#)
24. Abdullah, H.; Skidmore, A.K.; Darvishzadeh, R.; Heurich, M.; Pettorelli, N.; Disney, M. Sentinel-2 accurately maps green-attack stage of European spruce bark beetle (*Ips typographus* L.) compared with Landsat-8. *Remote Sens. Ecol. Conserv.* **2018**, *5*, 87–106. [\[CrossRef\]](#)
25. Huete, A.R.; Liu, H.Q.L.; Batchily, K.B.; Leeuwen, W.V. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sens. Environ.* **1997**, *59*, 440–451. [\[CrossRef\]](#)
26. Kalaitzidis, C.; Zianis, D.; Heinzl, V. A Review of Vegetation Indices for the Estimation of Biomass. In *Proceedings of the 29th Symposium of the European Association of Remote Sensing Laboratories, Chania, Greece*; IOS Press Ebook: Amsterdam, The Netherlands, 2009; pp. 201–208. [\[CrossRef\]](#)
27. Safari, A.; Sohrabi, H. Integration of synthetic aperture radar and multispectral data for aboveground biomass retrieval in Zagros oak forests, Iran: An attempt on Sentinel imagery. *Int. J. Remote Sens.* **2020**, *41*, 8069–8095. [\[CrossRef\]](#)
28. Luan, P.V.; Everardo, C.M.; do Amaral, C.H.; Christopher, M.U.; NealeZution, G.I.; Filgueiras, R.; Fernando, C.E. Potential of using spectral vegetation indices for corn green biomass estimation based on their relationship with the photosynthetic vegetation sub-pixel fraction. *Agric. Water Manag.* **2020**, *236*, 106155. [\[CrossRef\]](#)
29. David, R.M.; Rosser, N.J.; Donoghue, D.N.M. Improving above ground biomass estimates of Southern Africa dryland forests by combining Sentinel-1 SAR and Sentinel-2 multispectral imagery. *Remote Sens. Environ.* **2022**, *282*, 113232. [\[CrossRef\]](#)
30. Lu, D.S. The potential and challenge of remote sensing-based biomass estimation. *Int. J. Remote Sens.* **2007**, *27*, 1297–1328. [\[CrossRef\]](#)
31. Kuplich, T.M.; Curran, P.J.; Atkinson, P.M. Relating SAR image texture to the biomass of regenerating tropical forests. *Int. J. Remote Sens.* **2011**, *26*, 4829–4854. [\[CrossRef\]](#)
32. Wang, X.Q.; Pang, Y.; Zhang, Z.J.; Yuan, Y. Forest Aboveground Biomass Estimation Using SPOT-5 Texture Indices and Spectral Derivatives. In *Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec, QC, Canada, 13–18 July 2014*; pp. 2830–2833. [\[CrossRef\]](#)
33. Ghosh, S.M.; Behera, M.D. Aboveground biomass estimation using multi-sensor data synergy and machine learning algorithms in a dense tropical forest. *Appl. Geogr.* **2018**, *96*, 29–40. [\[CrossRef\]](#)
34. Jolliffe, I.T.; Cadima, J. Principal component analysis: A review and recent developments. *Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **2016**, *374*, 20150202. [\[CrossRef\]](#)

35. Su, H.Y.; Shen, W.J.; Wang, J.R.; Ali, A.; Li, M. Machine learning and geostatistical approaches for estimating aboveground biomass in Chinese subtropical forests. *For. Ecosyst.* **2020**, *7*, 64. [\[CrossRef\]](#)
36. Breiman, L. Random forests. *Mach. Learn.* **2001**, *34*, 5–32. [\[CrossRef\]](#)
37. Belgiu, M.; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [\[CrossRef\]](#)
38. Arsalan, G.; Soheil, Z.; Reza, M.A.; Meisam, A.; Mohammadzadeh, A.; Sadegh, J. Ghrobian-Mangrove ecosystem mapping using Sentinel-1 and Sentinel-2 Satellite images and random forest algorithm in Google Earth Engine. *Remote Sens.* **2021**, *13*, 2565. [\[CrossRef\]](#)
39. 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. [\[CrossRef\]](#)
40. Zeng, P.; Zhang, W.; Li, Y.; Shi, J.; Wang, Z. Forest total and component above-ground biomass (AGB) estimation through C- and L-band polarimetric SAR data. *Forests* **2022**, *13*, 442. [\[CrossRef\]](#)
41. Lourenço, P.; Godinho, S.; Sousa, A.; Gonçalves, A.C. Estimating tree aboveground biomass using multispectral satellite-based data in Mediterranean agroforestry system using random forest algorithm. *Remote Sens. Appl. Soc. Environ.* **2021**, *23*, 100560. [\[CrossRef\]](#)
42. Ou, G.L.; Lv, Y.Y.; Xu, H.; Wang, G.X. Improving forest aboveground biomass estimation of *Pinus densata* forest in Yunnan of southwest China by spatial regression using Landsat 8 images. *Remote Sens.* **2019**, *11*, 2750. [\[CrossRef\]](#)
43. Axelsson, C.; Skidmore, A.K.; Schlerf, M.; Fauzi, A.; Verhoef, W. Hyperspectral analysis of mangrove foliar chemistry using PLSR and support vector regression. *Int. J. Remote Sens.* **2012**, *34*, 1724–1743. [\[CrossRef\]](#)
44. Wang, S.; Wang, D.; Sun, J. Artificial neural network-based ionospheric delay correction method for satellite-based augmentation systems. *Remote Sens.* **2022**, *14*, 676. [\[CrossRef\]](#)
45. 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. [\[CrossRef\]](#)
46. Yadav, S.; Padalia, H.; Sinha, S.K.; Srinet, R.; Chauhan, P. Above-ground biomass estimation of Indian tropical forests using X band Pol-InSAR and Random Forest. *Remote Sens. Appl. Soc. Environ.* **2021**, *21*, 100462. [\[CrossRef\]](#)
47. Joshua, O.L.; Chinenye, A.L.; Adewale, G.A. Multi-layer perceptron artificial neural network (MLP-ANN) prediction of biomass higher heating value (HHV) using combined biomass proximate and ultimate analysis data. *Model. Earth Syst. Environ.* **2022**, *8*, 3177–3191. [\[CrossRef\]](#)
48. Vahedi, A.A. Artificial neural network application in comparison with modeling allometric equations for predicting above-ground biomass in the Hyrcanian mixed-beech forests of Iran. *Biomass Bioenergy* **2016**, *88*, 66–76. [\[CrossRef\]](#)
49. Xie, Y.; Sha, Z.; Yua, M.; Bai, Y.; Zhang, L. A comparison of two models with Landsat data for estimating above-ground grassland biomass in Inner Mongolia, China. *Ecol. Model.* **2009**, *220*, 1810–1818. [\[CrossRef\]](#)
50. Yang, S.; Feng, Q.; Liang, T.; Liu, B.; Zhang, W.; Xie, H. Modeling grassland above-ground biomass based on artificial neural network and remote sensing in the Three-River Headwaters Region. *Remote Sens. Environ.* **2018**, *204*, 448–455. [\[CrossRef\]](#)
51. Taylor, J.W. A quantile regression neural network approach to estimating the conditional density of multiperiod returns. *J. Forecast.* **2000**, *19*, 299–311. [\[CrossRef\]](#)
52. Koenker, R.; Bassett, G. Regression Quantiles. *Econometrica* **1978**, *46*, 33–50. [\[CrossRef\]](#)
53. Cade, B.S.; Noon, B.R. A gentle introduction to quantile regression for ecologists. *Front. Ecol. Environ.* **2003**, *1*, 412–420. [\[CrossRef\]](#)
54. Julien, L. A quantile regression study of climate change in Chicago, 1960–2010. *SIAM Undergrad. Res. Online* **2012**, *5*, 148–165. [\[CrossRef\]](#)
55. Cannon, A.J. Quantile regression neural networks: Implementation in R and application to precipitation downscaling. *Comput. Geosci.* **2011**, *37*, 1277–1284. [\[CrossRef\]](#)
56. Cannon, A.J. Non-crossing nonlinear regression quantiles by monotone composite quantile regression neural network, with application to rainfall extremes. *Stoch. Environ. Res. Risk Assess.* **2018**, *32*, 3207–3225. [\[CrossRef\]](#)
57. He, Y.; Li, H. Probability density forecasting of wind power using quantile regression neural network and kernel density estimation. *Energy Convers. Manag.* **2018**, *164*, 374–384. [\[CrossRef\]](#)
58. Dong, M.; Wu, D.; Fu, X.; Deng, H.; Wu, G. Regional-scale analysis on the strengths, weaknesses, opportunities, and threats in sustainable development of Shangri-La County. *Int. J. Sustain. Dev. World Ecol.* **2014**, *22*, 171–177. [\[CrossRef\]](#)
59. Guo, Y.; Li, Y.; Huang, Y.; Jarvis, D.; Sato, K.; Kato, K.; Tsuyuzaki, H.; Chen, L.; Long, C. Genetic diversity analysis of hulless barley from Shangri-la region revealed by SSR and AFLP markers. *Genet. Resour. Crop Evol.* **2011**, *59*, 1543–1552. [\[CrossRef\]](#)
60. Wang, B.; Mao, J.F.; Jie, G.; Wei, Z.; Wang, X.R. Colonization of the Tibetan plateau by the homoploid hybrid pine *Pinus densata*. *Mol. Ecol.* **2011**, *20*, 3796–3811. [\[CrossRef\]](#) [\[PubMed\]](#)
61. Compilation Committee of Yunnan Forest. *Yunnan Forest*; Yunnan Science and Technology Press: Kunming, China; China Forestry Publishing House: Beijing, China, 1986.
62. Zhang, J.; Lu, C.; Xu, H.; Wang, G. Estimating aboveground biomass of *Pinus densata*-dominated forests using Landsat time series and permanent sample plot data. *J. For. Res.* **2019**, *30*, 1689–1706. [\[CrossRef\]](#)

63. Ou, G.L.; Li, C.; Lv, Y.Y.; Wei, A.C.; Xiong, H.X.; Xu, H.; Wang, G.X. 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. [\[CrossRef\]](#)
64. Ghosh, A.; Fassnacht, F.E.; Joshi, P.K.; Koch, B. A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *26*, 49–63. [\[CrossRef\]](#)
65. Wang, Z.; Zhang, X.Q.; Chhin, S.; Zhang, J.; Duan, A. Disentangling the effects of stand and climatic variables on forest productivity of Chinese fir plantations in subtropical China using a random forest algorithm. *Agric. For. Meteorol.* **2021**, *304–305*, 108412. [\[CrossRef\]](#)
66. Roy, N.; Pal, A.; Dey, A.; Das, S. Applications of artificial intelligence in machine learning: Review and Prospect. *Int. J. Comput. Appl.* **2015**, *115*, 31–41. [\[CrossRef\]](#)
67. Stathakis, D. How many hidden layers and nodes? *Int. J. Remote Sens.* **2009**, *30*, 2133–2147. [\[CrossRef\]](#)
68. Tiryaki, S.; Aydın, A. An artificial neural network model for predicting compression strength of heat treated woods and comparison with a multiple linear regression model. *Constr. Build. Mater.* **2014**, *62*, 102–108. [\[CrossRef\]](#)
69. Babcock, C.; Finley, A.O.; Bradford, J.B.; Kolka, R.; Birdsey, R.; Ryan, M.G. LiDAR-based prediction of forest biomass using hierarchical models with spatially varying coefficients. *Remote Sens. Environ.* **2015**, *169*, 113–127. [\[CrossRef\]](#)
70. Wang, Y.; Wu, G.; Deng, L.; Tang, Z.; Wang, K.; Sun, W.; Shangguan, Z. Prediction of aboveground grassland biomass on the Loess Plateau, China, using a random forest algorithm. *Sci. Rep.* **2017**, *7*, 6940. [\[CrossRef\]](#) [\[PubMed\]](#)
71. Sun, X.; Li, G.; Wang, M.; Fan, Z. Analyzing the uncertainty of estimating forest aboveground biomass using optical imagery and spaceborne LiDAR. *Remote Sens.* **2019**, *11*, 722. [\[CrossRef\]](#)
72. Zhu, X.L.; Liu, D.S. Improving forest aboveground biomass estimation using seasonal Landsat NDVI time-series. *J. Photogramm. Remote Sens.* **2015**, *102*, 222–231. [\[CrossRef\]](#)
73. Sibanda, M.; Mutanga, O.; Rouget, M. Examining the potential of Sentinel-2 MSI spectral resolution in quantifying above-ground biomass across different fertilizer treatments. *ISPRS J. Photogramm. Remote Sens.* **2015**, *110*, 55–65. [\[CrossRef\]](#)
74. Mutanga, O.; Adam, E.; Cho, M.A. High-density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 399–406. [\[CrossRef\]](#)
75. Pandit, S.; Tsuyuki, S.; Dube, T. Exploring the inclusion of Sentinel-2 MSI texture metrics in above-ground biomass estimation in the community forest of Nepal. *Geocarto Int.* **2019**, *35*, 1832–1849. [\[CrossRef\]](#)
76. Kelsey, K.; Neff, J. Estimates of aboveground biomass from texture analysis of Landsat imagery. *Remote Sens.* **2014**, *6*, 6407–6422. [\[CrossRef\]](#)
77. Xu, H.; Yue, C. *Study on Forest Landscape Change and Forest Biomass Estimation in Shangri-La Based on Remote Sensing Technology*; Yunnan Science and Technology Press: Kunming, China, 2014.
78. Powell, S.L.; Cohen, W.B.; Healey, S.P.; Kennedy, R.E.; Moisen, G.G.; Pierce, K.B.; Ohmann, J.L. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sens. Environ.* **2010**, *114*, 1053–1068. [\[CrossRef\]](#)
79. Xie, F.; Shu, Q.; Zi, L.; Wu, R.; Wu, Q.; Wang, H.; Liu, Y.; Ji, Y. Remote sensing estimation of *Pinus densata* aboveground biomass based on k-NN nonparametric model. *Acta Agric. Univ. Jiangxiensis* **2018**, *40*, 743–750.
80. Li, L.; Zhou, X.S.; Chen, L.; Chen, L.; Zhang, Y.; Liu, Y. Estimating urban vegetation biomass from Sentinel-2A image data. *Forests* **2020**, *11*, 125. [\[CrossRef\]](#)
81. Lu, D.S.; Batistella, M.; Moran, E. Satellite estimation of aboveground biomass and impacts of forest stand structure. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 967–974. [\[CrossRef\]](#)
82. Chang, J.S.; Shoshany, M. Mediterranean Shrublands Biomass Estimation Using Sentinel-1 and Sentinel-2. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016. [\[CrossRef\]](#)
83. Masjedi, A.; Zhao, J.Q.; Thompson, A.M.; Yang, K.W.; Flatt, J.E.; Crawford, M.M.; Ebert, D.S.; Tuinstra, M.R.; Hammer, G.; Chapman, S. Sorghum Biomass Prediction Using UAV-Based Remote Sensing Data and Crop Model Simulation. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018.
84. Freeman, E.A.; Moisen, G.G. An Application of Quantile Random Forests for Predictive Mapping of Forest Attributes. In Proceedings of the New Directions in Inventory Techniques & Applications Forest Inventory & Analysis (FIA) Symposium, Portland, OR, USA, 8–10 December 2015; p. 362.
85. Pham, T.D.; Le, N.N.; Ha, N.T.; Nguyen, L.V.; Xia, J.; Yokoya, N.; To, T.T.; Trinh, H.X.; Kieu, L.Q.; Takeuchi, W. Estimating Mangrove Above-Ground Biomass Using Extreme Gradient Boosting Decision Trees Algorithm with Fused Sentinel-2 and ALOS-2 PALSAR-2 Data in Can Gio Biosphere Reserve, Vietnam. *Remote Sens.* **2020**, *12*, 777. [\[CrossRef\]](#)
86. Kattenborn, T.; Leitloff, J.; Schiefer, F.; Hinz, S. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 24–49. [\[CrossRef\]](#)
87. Xu, M.; Liu, R.; Chen, J.M.; Liu, Y.; Shang, R.; Ju, W.; Wu, C.; Huang, W. Retrieving leaf chlorophyll content using a matrix-based vegetation index combination approach. *Remote Sens. Environ.* **2019**, *224*, 60–73. [\[CrossRef\]](#)

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.