remote sensing

MDPI

Article

Effects of Driving Factors on Forest Aboveground Biomass (AGB) in China's Loess Plateau by Using Spatial Regression Models

Shichuan Yu [1,2], Quanping Ye [1,2], Qingxia Zhao [3], Zhen Li [1,2], Mei Zhang [1,2], Hailan Zhu [1,2] and Zhong Zhao [1,2,*]

[1] State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Northwest A&F University, Yangling District, Xianyang 712100, China; 2017060253@nwafu.edu.cn (S.Y.); yeqp@nwafu.edu.cn (Q.Y.); yunyuan@nwsuaf.edu.cn (Z.L.); 2020060255@nwafu.edu.cn (M.Z.); zhuhailan@nwafu.edu.cn (H.Z.)
[2] College of Forestry, Northwest A&F University, Yangling District, Xianyang 712100, China
[3] Shaanxi Institute of Zoology, Xi'an 710032, China; zhaoqingxia@nwafu.edu.cn
* Correspondence: zhaozh@nwafu.edu.cn; Tel.: +86-29-87082801

Abstract: Forests are the main body of carbon sequestration in terrestrial ecosystems and forest aboveground biomass (AGB) is an important manifestation of forest carbon sequestration. Reasonable and accurate quantification of the relationship between AGB and its driving factors is of great importance for increasing the biomass and function of forests. Remote sensing observations and field measurements can be used to estimate AGB in large areas. To explore the applicability of the panel data models in AGB and its driving factors, we compared the results of panel data models (spatial error model and spatial lag model) with those of geographically weighted regression (GWR) and ordinary least squares (OLS) to quantify the relationship between AGB and its driving factors. Furthermore, we estimated the tree height, diameter at breast height, canopy cover (CC) and species diversity index (Shannon–Wiener index) of Robinia pseudoacacia plantations in Changwu on the Loess Plateau using field data and remote sensing images by a random forest model and estimated soil organic carbon (SOC) contents using laboratory data by ordinary kriging (OK) interpolation. We estimated AGB using the already estimated tree height and diameter at breast height combined with the allometric growth equation. In this study, we estimated SOC contents by OK interpolation, and the accuracy R^2 values for each soil layer were greater than 0.81. We estimated diameter at breast height (DBH), CC, SW and tree height (TH) using the random forest, and the accuracy R^2 values were 0.85, 0.82, 0.76 and 0.68, respectively. We estimated AGB with random forest and the allometric growth equation and found that the average AGB was 55.80 t/ha. The OLS results showed that the residuals of the OLS regression exhibited obvious spatial correlations and rejected OLS applications. GWR, SEM and SLM were used for spatial regression analysis, and SEM was the best model for explaining the relationship between AGB and its driving factors. We also found that AGB was significantly positively correlated with CC, SW, and 0–60 cm SOC content ($p < 0.05$) and significantly negatively correlated with slope aspect ($p < 0.01$). This study provides a new idea for studying the relationship between AGB and its driving factors and provides a basis for practical forest management, increasing biomass, and giving full play to the role of carbon sequestration.

Keywords: above-ground biomass (AGB); estimation; driving factors; spatial regression analysis; *Robinia pseudoacacia*

Citation: 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. https://doi.org/10.3390/rs14122842

Academic Editor: Henning Buddenbaum

Received: 24 March 2022
Accepted: 10 June 2022
Published: 14 June 2022

1. Introduction

As the largest carbon stock in terrestrial ecosystems, forests play an important role in tackling climate change and improving the ecological environment [1,2]. The aboveground biomass (AGB) of forests is an important manifestation of forest growth through years of accumulation and represents 70–90% of the total forest biomass [1,3,4]. Forest AGB estimation is the basis of understanding the carbon storage of terrestrial ecosystems and

ecological function. At present, there are various methods for estimating forest AGB, and each method has its advantages and disadvantages. Field measurements are the most accurate method used to estimate forests for small forest stands; however, this method can be time-consuming, labor-intensive and costly, and is not suitable for estimating AGB in large areas [2,5]. However, the combination of remote sensing observations and field measurements can be used to estimate AGB in large areas [6,7]. Previous studies have indicated that machine learning techniques can be applied to estimate AGB and have exhibited excellent performance [8–12]. Zhao (2019) compared different machine learning algorithms for estimating forest AGB on the Loess Plateau and showed that random forest exhibited the highest accuracy [13]. We used random forest to estimate the forest AGB of the study area.

One of the direct indicators reflecting the structure and function of the forest ecosystem, AGB is the result of the joint effect of the forest stand and the surrounding environment. Analyzing the driving factors of forest AGB can provide a theoretical basis for sustainable forestry management. The driving factors of forest AGB have always been a popular topic among researchers. Different perspectives have been used to analyze the driving factors of forest AGB. For example, the climate usually affects forest AGB on larger temporal and spatial scales. Chen (2015) studied climate change impacts on the net AGB of four forest types in western Canada's boreal forests and found that persistent warming and reduction in water availability had deep negative impacts on forest biomass [14]. Bordin (2021) thought that the annual temperature range had a negative effect on forest AGB in subtropical forests in southern Brazil [15]. Sullivan (2020) found that maximum temperature may reduce woody productivity and become the most important factor affecting aboveground biomass in tropical forests [16]. Similarly, Li (2022) found that warming and drying caused by deforestation lead to an additional 5.1 $\pm$ 3.7% loss of aboveground biomass in the Amazon [17]. Longitude, latitude and altitude generally affect biomass on larger spatial scales, but aspect, slope position and slope affect biomass on smaller spatial scales. Xu (2018) found that moso bamboo biomass decreased with decreasing latitude in the subtropical region of China [18]. Sheikha (2020) found that the AGB of Cedrus deodara forests decreased with altitude in the Central Himalayas [19]. Shen (2018) studied the distribution of AGB under different terrain conditions and found that AGB was positively correlated with slope [20]. Pariyar (2019) found that stand AGB was greater on northern aspects than on southern aspects [21]. The stand factor is another important driving factor for forest AGB accumulation. Previous studies have shown that stand density and stand age directly adjust forest AGB [22,23], and crown density and species diversity affect forest AGB by adjusting stand structure [24,25]. Soil creates conditions for the growth of trees by providing nutrients. Ali (2019b) demonstrated that soils indirectly affected AGB via stand structural complexity in tropical forests [25].

The driving factors for forest AGB have always been a popular topic among researchers, and traditional regression analysis methods (for example, principal component analysis, redundancy analysis, correlation analysis and ordinary least squares) have been widely used to establish the relationship between forest AGB and driving factors [26–29]. However, traditional regression analysis methods do not consider the spatial relationship between AGB and its driving factors. To study the spatial relationship, some scholars used geographically weighted regression (GWR) models to explain the effect of driving factors on forest AGB, and provided a spatially explicit technique [30]. Panel data models, as spatial regression models, reduce the effects of multiple nonlinearities among variables and provide accurate estimations [31]. Yuan (2018) studied the impact of urban form on haze pollution using spatial regression analysis in China and found that urban form can impact the concentration of PM2.5 [32]. Sun (2020) analyzed the COVID-19 period prevalence in US counties by spatial regression model and thought that spatial regression models could better estimate the period prevalence for counties [33]. Panel data models have been widely used in economics, sociology and the environment [34,35], but panel data models have not been applied to assess the relationships between AGB and its driving factors. Gaofen-1

(GF-1) images have high spatial resolution and abundant spatial structure and texture information. GF-1 images have been applied to estimating biomass, leaf area index, chlorophyll and vegetation coverage, and GF-1 images exhibited a good potential for application in various fields [36–38]. We chose GF-1 as the data source. To explore the applicability of the panel data models in AGB and driving factors, we put our plans into a technology roadmap as shown in Figure 1: First, we estimated the AGB, canopy cover (CC), species diversity index (Shannon–Wiener index) and soil organic carbon (SOC) contents of *Robinia pseudoacacia* plantations in Changwu on the Loess Plateau. Second, to reflect combined with the situation in the study area, AGB was used as the dependent variable, and slope, aspect, SOC content, CC and Shannon–Wiener index (SW) were used as independent variables. Third, we planned to compare the panel data models with the GWR model and ordinary least squares (OLS) and applied the results to quantify the relationship between AGB and its driving factors to obtain the results of the panel data model application. Finally, the panel data models, GWR model and OLS were used to quantify the regression relationship between AGB and its driving factors to choose the most suitable regression model. We hope to provide a new idea for studying the relationship between biomass and driving factors.

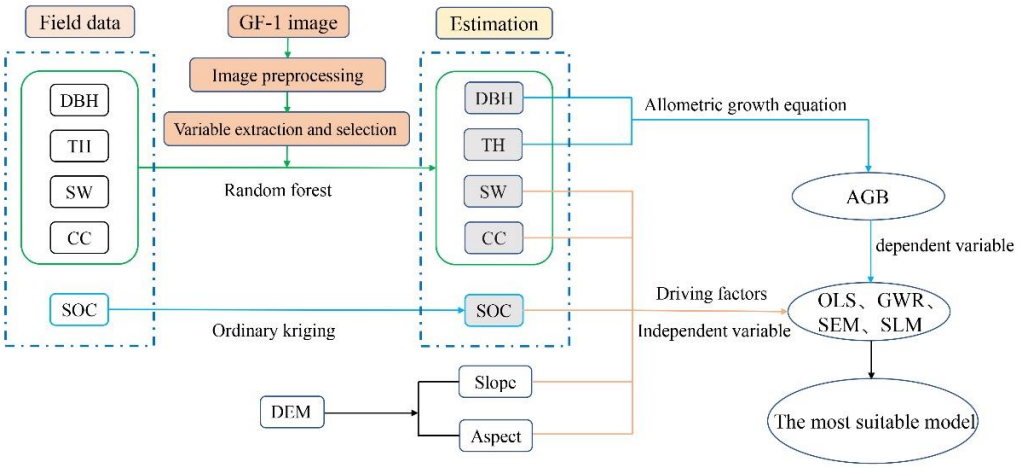

Figure 1. The technology roadmap for this paper. DBH, TH, SW, CC and SOC represent diameter at breast height, tree height, Shannon–Wiener index, canopy cover and soil organic carbon, respectively. DEM represents a digital elevation model, and AGB represents aboveground biomass. OLS, GWR, SEM and SLM represent ordinary least squares, geographically weighted regression model, spatial error model and spatial lag model.

2. Materials and Methods

2.1. Study Region

The study area covers 567.1 km^2, and is situated in Changwu County, Shaanxi Province, in the south of the Loess Plateau of China (34°59′N—35°18′N and 107°38′E—107°58′E) (Figure 2). The elevation of the study area ranges from 847 to 1274 m above sea level. The study area is located in the loess gully region of the Loess Plateau, and the dominant soil types are dark loessial soil and loessial soil [39,40].

The study area is located in a warm temperate semihumid continental monsoon climate zone. The annual precipitation is 587.8 mm (the annual precipitation concentration period is in July–September, and precipitation can reach 321.4 mm, accounting for 54.9% of the total annual precipitation), and the average annual temperature is 9.1 °C. Furthermore, the frost-free period is 171 days [39]. The study area was dominated by planted forest, and afforestation began in the 1950s. The main tree species are *R. pseudoacacia*, *Pinus tabuliformis*, *Platycladus orientalis* and *Populus* spp. Among these, *R. pseudoacacia* makes up more than 90% of the forested area [40–42]. The main shrub species are *Rubus mesogaeus*, *Ziziphus jujuba* var. *spinose* and *Lespedeza bicolor*, and the main herb species are *Imperata cylindrical*, *Elymus dahuricus* and *Humulus scandens*.

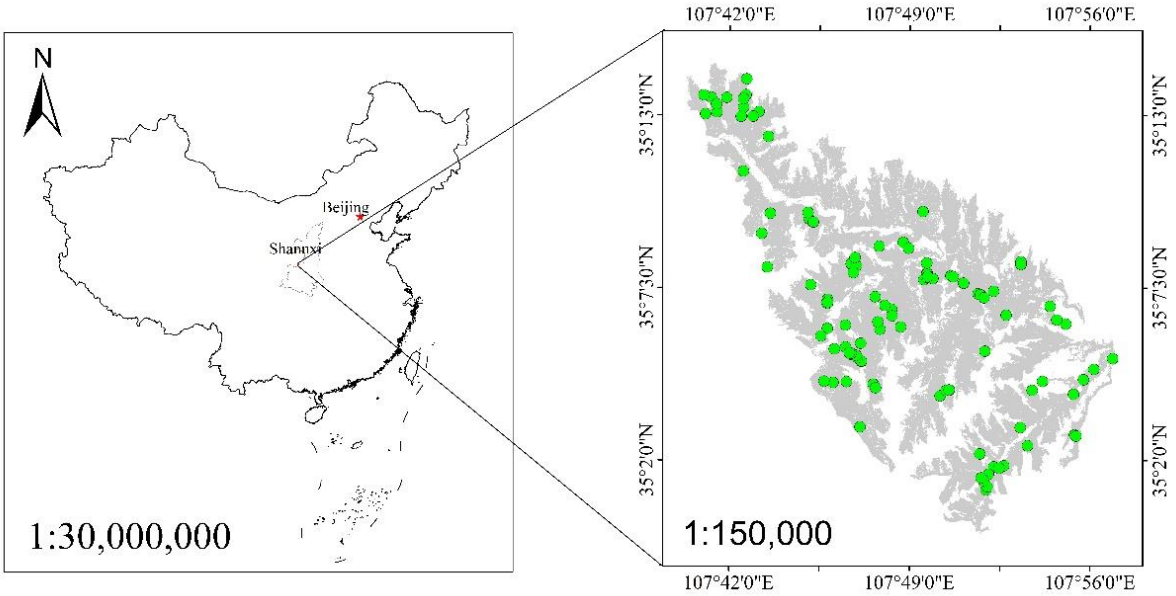

Figure 2. Geographical location and field sample plots of the study area.

2.2. Field and Laboratory Data

The field sample plots were set from July to August 2019. We randomly selected 104 sample plots in the study area. The plots were set as evenly as possible in the accessible area (Figure 2). Each plot was located in a pure *R. pseudoacacia* plantation, and the plot size was 30 × 30 m. In each plot, we measured trees with a diameter at breast height (DBH) greater than 5 cm. The DBH, tree height (TH), species and number of species were documented. SW was calculated by species and number of species. The CC was surveyed with a digital hemispherical photograph. Then, we used Gap Light Analyzer 2.0 to process digital hemispherical photographs [43]. In each plot, the 0–80 cm soil layers were selected and divided into the 0–20 cm, 20–40 cm, 40–60 cm, and 60–80 cm layers. We used a 5 cm diameter stainless steel corer to select nine soil cores along an S-shaped pattern, and then we mixed the nine soil samples into one composite soil sample [44]. We removed the stones, litter and plant roots from the soil samples and divided them into three subsamples. Then, we brought them back to the laboratory to air-dry. Finally, soil samples were treated by lapping and screening a 2 mm sieve to measure the SOC content using the dichromate oxidation method [44,45]. The analyses were performed using three subsamples, and the three subsample values were averaged to generate a mean value that was used as the calculated plot SOC content.

2.3. Estimation of SOC

We used ordinary kriging (OK) interpolation to estimate SOC based on four soil layers (0–20 cm, 20–40 cm, 40–60 cm and 60–80 cm) in the study area by ArcGIS 10.7. We used 70% of the sample plot datasets for OK interpolation and 30% of the sample plot datasets for accuracy assessment. OK interpolation calculates the weights based on the distances between tested points and predicted points by the semivariogram model, which is the most popular spatial analysis method and has been applied in SOC cartography [46]. We used the root mean square error (RMSE), mean error (ME) and coefficient of determination (R^2) of the tested and predicted values to evaluate the results (Table S1). Smaller RMSE and ME values represent a higher interpolation accuracy [47,48].

2.4. Spatial Data Collection

As the first satellite of China's high-resolution earth observation system, the GF-1 image is configured with multispectral (8 m spatial resolution) and panchromatic (2 m spatial resolution) images [49]. We acquired a GF-1 image of the study area from 18 June 2020 from

the China Center for Resources Satellite Data and Application (http://www.cresda.com/CN/, accessed on 28 June 2020). Then, each image was radiometrically, atmospherically and geometrically corrected and registered to the Xian_1980 datum and Xian_1980_3_Degree_GK_CM_108E projections using ArcGIS 10.7(Copyright © 1992–1999 Tom Sawyer Software, Published in the United States of America) and ENVI. 5.3 (source from Exelis Visual Information Solutions in America) image processing software [50,51].

We extracted slope and aspect from the digital elevation model (DEM).We divided the slope into six classes (1–6 represented 0–5°, 6–15°, 16–25°, 26–35°, 36–45°, ≥46°) (Figure 3a) and divided the aspect into four classes (1–4 represented shade (315–45°), semi-shady (45–90°, 270–315°), semi-sunny (90–135°, 225–270°), and sunny (135–225°)) (Figure 3b), according to the technical regulations for inventory for forest management planning and design (GB/T 26424—2010). The extraction and division of the slope and aspect were completed in ArcGIS 10.7.

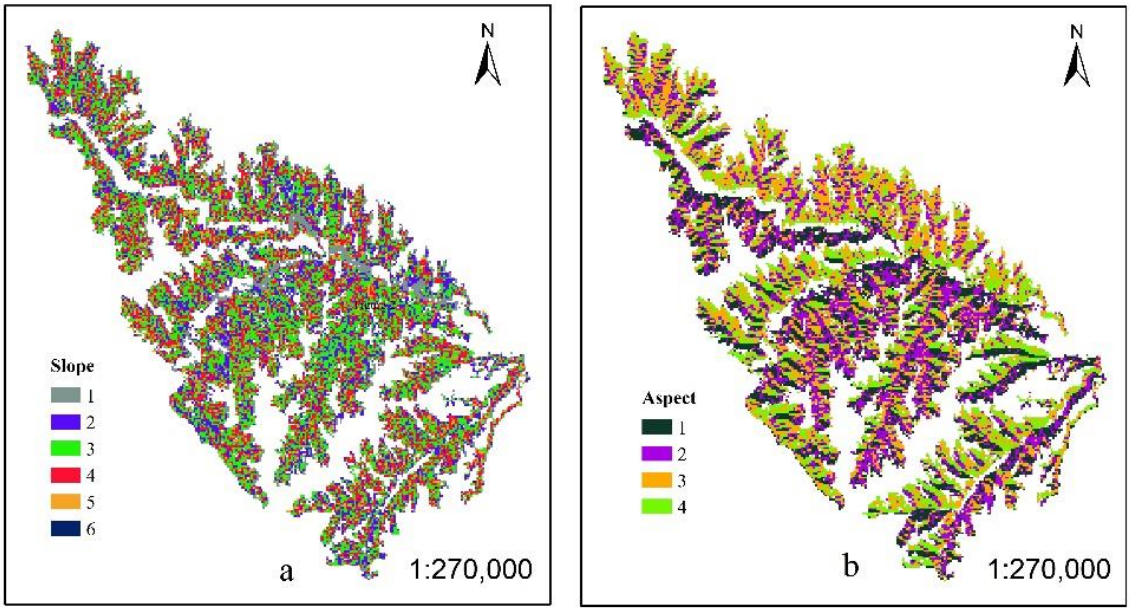

Figure 3. (**a**) Slope classes, where 1–6 represent 0–5°, 6–15°, 16–25°, 26–35°, 36–45° and ≥46°, respectively. (**b**) Aspect classes, where 1–4 represent shaded (315–45°), semi-shady (45–90°, 270–315°), semi-sunny (90–135°, 225–270°) and sunny (135–225°), respectively.

2.5. DBH, TH, SW, CC and AGB Estimation

2.5.1. Predictor Variables

We selected four multiple spectral bands (blue-b1, green-b2, red-b3 and near infrared-b4) and one panchromatic band (P) from the Gaofen 1 multispectral image. Furthermore, we calculated 18 vegetation indices from four multiple spectral bands that are widely used in forest studies (Table S2) [13,52]. Then, we used eight gray level co-occurrence matrixes (GLCMs) [53] to extract the texture variables (Table S3). Finally, we calculated the eight GLCM measures with ten different window sizes (3 × 3, 5 × 5, 7 × 7, 9 × 9, 11 × 11, 13 × 13, 15 × 15, 17 × 17, 19 × 19, and 21 × 21) from the four multiple spectral bands, the eighteen vegetation indices and one panchromatic band. All of the above data processing and analysis steps were completed in ArcMap 10.7 software and ENVI 5.3 software.

2.5.2. Variable Selection

Variable selection is often an important step in the application of machine learning methods. There are often too many variables for practical model building, but not all variables are related to the response variable. Therefore, choosing the best set of variables is essential for model building. The Boruta algorithm developed by Kursa and Rudnicki [54] is an all-relevant feature selection wrapper algorithm around the random forest algorithm.

The important variables are selected by comparing the Z scores obtained by the original features to the Z scores obtained by the randomly generated shadow attributes. The Boruta algorithm is a powerful algorithm for variable selection [55].

In this study, we implemented the Boruta algorithm using the "Boruta" package (Kursa and Rudnicki, 2010) in R software (version 4.1.0, source from a GNU project) to screen the full set of variables from the Gaofen 1 images, which included 20 spectral variables and 3528 texture variables.

2.5.3. Random Forest and Accuracy Assessment

Random forest was developed by Breiman by constructing many regression trees that can correct model overfitting [56]. Random forest has the advantage of high accuracy and few generalization errors, making the method appropriate for solving problems with nonlinear multivariable constraints or without prior knowledge and incomplete data. The random forest algorithm requires the input parameters mtry and ntree. The mtry parameter is the number of independent variables supplied to each predictor tree, and the ntree parameter is the number of regression trees grown based on a bootstrap sample of the measures [57]. Through tuning parameters, random forest attempts to maintain prediction strength while inducing diversity among the trees [13,56].

In each plot, the mean value was used for model testing and training. The mean value from each plot investigation indicator dataset was randomly divided into testing samples (30%) and training samples (70%). We established the regression models by the training data and selected variables and used the testing data to validate the model accuracy. We randomly repeated the training samples 100 times to lessen the change and obtained the mean value of the 100 accuracy values [13]. We calculated the RMSE, the R^2 and the relative RMSE (rRMSE) to evaluate the model performances using the "randomForest" package in R software (version 4.1.0) (Table S1) [58]. We used the estimated DBH and TH to calculate the AGB of the R. pseudoacacia forest in the study area using the allometric growth equation (Table S4) [59]. The AGB calculation was completed in ArcGIS 10.7.

2.6. Regression Analysis

2.6.1. Ordinary Least Squares (OLS)

OLS is widely used in linear regression parameter estimation. It takes the minimum value of the sum of the squares of the difference between the actual value and the model estimate, and the minimum value is selected as the parameter estimate. In this study, we used OLS as the reference for the spatial regression models. The formula is as follows:

$$Y_i = \alpha + \beta_1 X_1 + \cdots + \beta_n X_n + \varepsilon \quad (1)$$

where Y_i is the dependent variable, X_1 to X_n are independent variables, α is the intercept, β_1 to β_n are estimated regression coefficients, and ε is the error term.

2.6.2. Spatial Regression Model

Geographically Weighted Regression Model (GWR)

The GWR model was proposed by Fotheringham and Brunsdon [60,61] and allows the adjustment of sets of spatially limited models based on global regression models, thus yielding local regression outputs. The GWR model not only expresses the relationship between dependent and independent variables but also accounts for spatial heterogeneity. The spatial position of the data is introduced into the model as the regression coefficient, and the spatial regression coefficient of each spatial position is given by the nonparametric estimation method. According to the estimated values of the regression coefficients at each spatial location, the spatial nonstationarity of the parameters and the heterogeneity of the location are explored and analyzed.

The formula is as follows:

$$Y_i = \beta_0(u_i, v_i) + \sum_{n=1}^{m} \beta_n(u_i, v_i) X_{in} + \varepsilon_i \quad (2)$$

where Y_i is the dependent variable at position i, X_{in} ($n = 1, \ldots, m$) are independent variables at position i, and $\beta_0(\mu_i, \nu_i)$ is the intercept. μ_i and ν_i represent the coordinates of the sample i in space, β_n are estimated regression coefficients and ε is the error term.

$$\beta(u_i, v_i) = \left(A^T W(u_i, v_i)\mathrm{A}\right)^{-1} A^T W(u_i, v_i) C \tag{3}$$

where β (μ_i, ν_i) is the regression parameter vector in location (μ_i, ν_i), X is a sampling matrix of the independent variable, C is a sampling vector of the dependent variable, and $W(\mu_i, \nu_i)$ is a diagonal matrix, whose diagonal elements represent the geographical weightings of each observation around point i.

$$W_i = \begin{bmatrix} w_{i1} & 0 & \cdots & 0 \\ 0 & w_{i2} & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{in} \end{bmatrix} \tag{4}$$

where w_{in} is the weight assigned to the observation at location i.

Spatial Lag Model (SLM)

The SLM is a panel data model that mainly tests whether there are any spillover effects of the variables or whether there is spatial diffusion. The SLM assumes a close association between the dependent variable and explanatory variables and incorporates spatial dependence into the regression model with a "spatially lagged dependent variable" [62]. The formula is as follows:

$$Y_{it} = \delta \sum_{j=1}^{N} W_{ij} y_{jt} + \beta X_{it} + \mu_i + \varepsilon_{it} \tag{5}$$

where t is an indicator for time, with $t = 1, 2, 3 \ldots, T$. i is an individual in cross section, with $I = 1, 2, 3 \ldots, N$. δ represents the coefficient of spatial autoregressive, W_{ij} is the spatial weight matrix, Y_{it} is the dependent variable at i and t, X_{it} is the vector of the independent variables at i and t, β represents the coefficient of the vector, and μ_i is the spatial-specific effect for i, while ε_{it} represents an error term for i and t.

$$W_{ij} = \begin{bmatrix} W_{11} & W_{12} & \cdots & W_{1j} \\ W_{21} & W_{22} & \cdots & W_{2j} \\ \vdots & \vdots & \vdots & \vdots \\ W_{i1} & W_{i2} & \cdots & W_{ij} \end{bmatrix} \tag{6}$$

where W_{ij} is the spatial weight matrix, if i is adjacent to j, $W_{ij} = 1$; else, $W_{ij} = 0$.

Spatial Error Model (SEM)

The SEM is a panel data model that considers that the dependent variable depends on a set of observed local characteristics and assumes that the error terms have an impact on spatial dependence [62]. The formula is as follows:

$$Y_{it} = \beta X_{it} + \mu_i + \varphi_{it} \tag{7}$$

$$\varphi_{it} = \rho \sum_{j=1}^{N} W_{ij} \varphi_{jt} + \varepsilon_{it} \tag{8}$$

where φ_{it} is the spatially auto-correlated error term and ρ is the spatial autocorrelated coefficient. $\sum_{j=1}^{N} W_{ij}\varphi_{jt}$ denotes the interaction effects among the disturbance terms of the different units and λ is the coefficient of spatial autocorrelation.

In this study, we used AGB as the dependent variable and CC, SW, SOC content (0–20 cm, 20–40 cm, 40–60 cm, 60–80 cm), aspect and slope as the independent variables to complete regression analysis. The OLS model, SLM and SEM were run in GeoDa 1.18

software (Developed by Dr. Luc Anselin and his team, http://geodacenter.github.io/formats.html, accessed on 15 October 2020). The GWR model was run in GWR4 software (Developed at NCG (National Centre for Geocomputation, National University of lreland Maynooth) and Department of Geography, Ritsumeikan University, Japan Development Team, https://sgsup.asu.edu/sparc/gwr4, accessed on 15 October 2020).

2.6.3. Model Selection and Evaluation

Moran's I value is widely used to test spatial autocorrelation [63]. Moran's I value ranges from −1 to 1. If the value is greater than 0, it represents positive spatial dependence; if the value is less than 0, it represents negative spatial dependence; if the value is equal to 0, it represents spatial randomness of the target variables [64]. Robust LM and Lagrange multiplier (LM) tests were used to evaluate the SEM and SLM and determine which spatial regression model should be used. The R^2, log likelihood (LIK) and Akaike information criterion (AIC) were used to evaluate model performance. Therefore, a regression model is well fitted if it has high values of LIK and R^2 and low values of AIC [64,65].

3. Results

3.1. SOC Estimation

The best model was selected for each soil layer SOC by comparing the semivariogram models, and the estimation accuracy results are shown in Table 1. The R^2 values for all soil layers were greater than 0.81, and RMSE and ME values for all soil layers were smaller. The RMSE, ME and R^2 values met the accuracy requirements. The spatial distribution of the SOC content in different soil layers in the study area is shown in Figure 4. The mean (min–max) SOC contents from the 0–20 cm, 0–40 cm, 40–60 cm, and 60–80 cm soil layers in the study area were 10.15 (6.23–21.02) g/kg, 5.34 (3.95–10.50) g/kg, 4.35 (2.90–7.37) g/kg and 3.67 (2.78–5.39) g/kg, respectively. As the soil depth increased, the SOC contents decreased. The SOC contents in the study area were mostly aggregated according to similar values. In the study area, the SOC contents were higher in the midland and lower in the northwest and southeast.

Table 1. The estimation accuracy results for SOC in different soil layers by ordinary kriging.

Accuracy	0–20 cm	20–40 cm	40–60 cm	60–80 cm
RMSE	0.79	1.33	1.06	0.86
ME	−0.017	0.002	−0.016	0.001
R^2	0.88	0.81	0.84	0.85

3.2. DBH, TH, SW, CC and AGB Estimation

We obtained explanatory variables related to the dependent variables by the Boruta algorithm, as shown in Figure 5. The number of texture variables was more than that of spectral variables among the variables we obtained. The spectral variables b2 and b4 had the highest importance values relevant to TH and SW, respectively. The texture variables MSAVIMEA3×3 and DVICOR5×5 had the highest importance values relevant to CC and DBH, respectively.

The random forest regression models were used to estimate the DBH, TH, SW and CC and compare the results with the measured values in the field, as shown in Figure 6. The DBH exhibited the highest accuracy, with an R^2 value of 0.85 (rRMSE = 0.09). The DBH was followed by CC and SW, with R^2 values of 0.82 and 0.76 and rRMSE values of 0.11 and 0.21, respectively. The TH exhibited lower accuracy, with an R^2 value of 0.68 and an rRMSE value of 0.10.

Figure 7a–d display the spatial distributions of the DBH, TH, SW and CC values in the study area estimated by random forest, respectively. The results show that the values of DBH, TH, SW and CC were unevenly distributed in the study area. The higher values of DBH and TH were mainly distributed in the middle and south of the study area. The

values of SW and CC had a similar distribution trend, and the values increased from the edge to the inside of the study area.

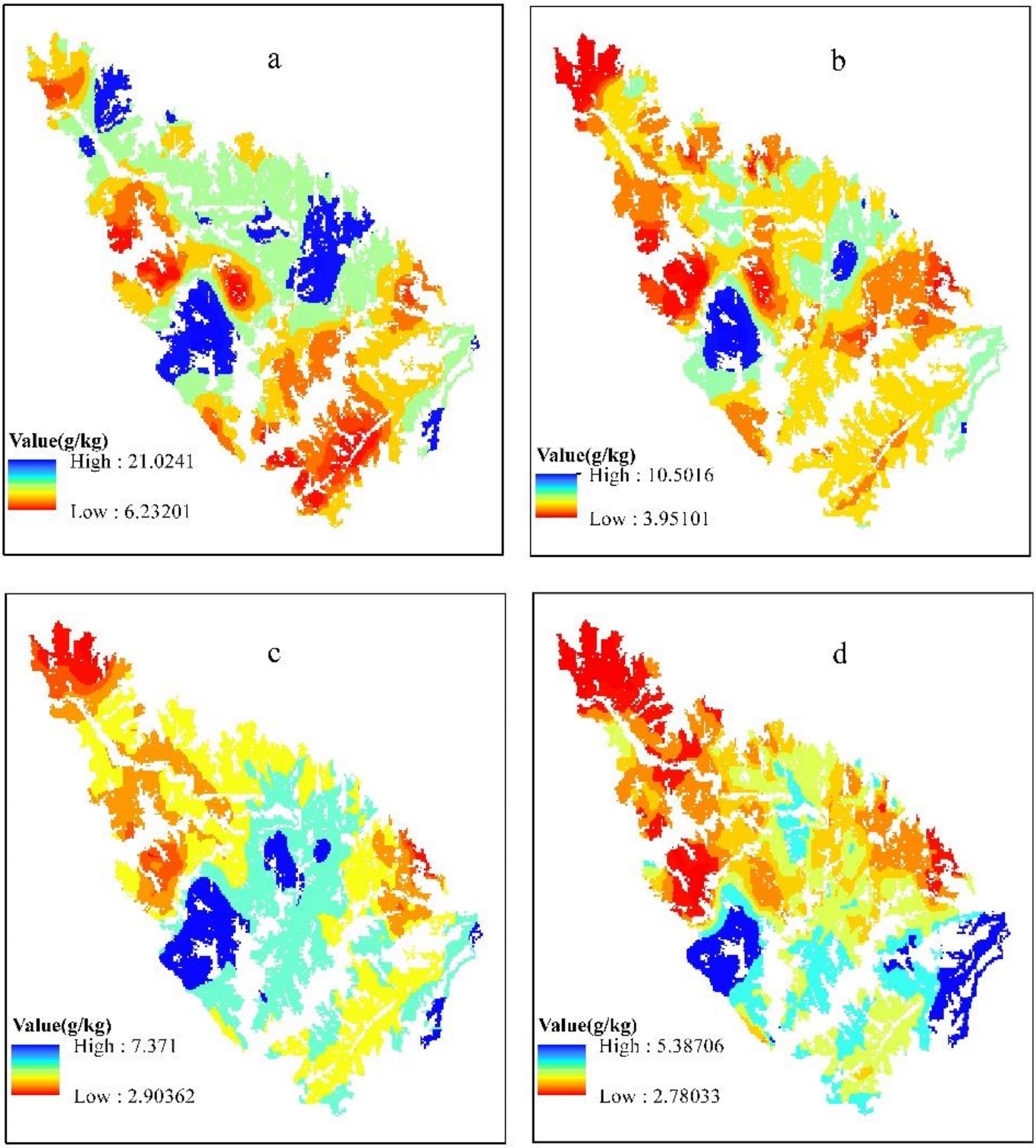

Figure 4. (**a**–**d**) represent the distribution of the SOC content in the 0–20 cm, 20–40 cm, 40–60 cm and 60–80 cm soil layers in the study.

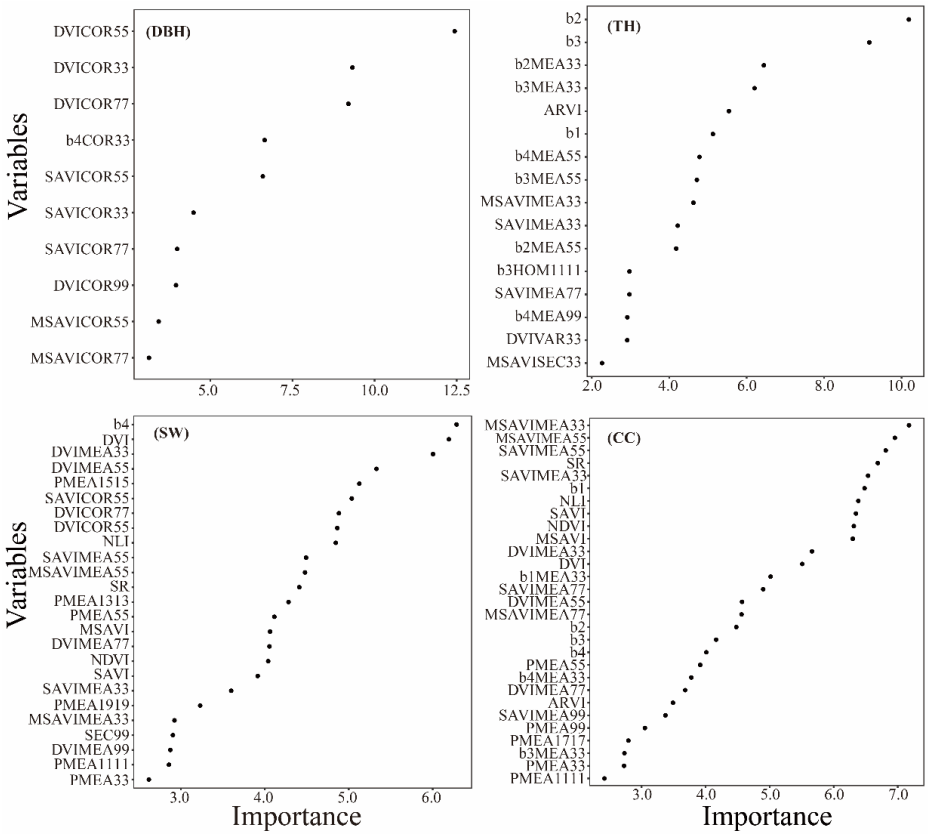

Figure 5. Variables were selected by the Boruta algorithm and their importance values.

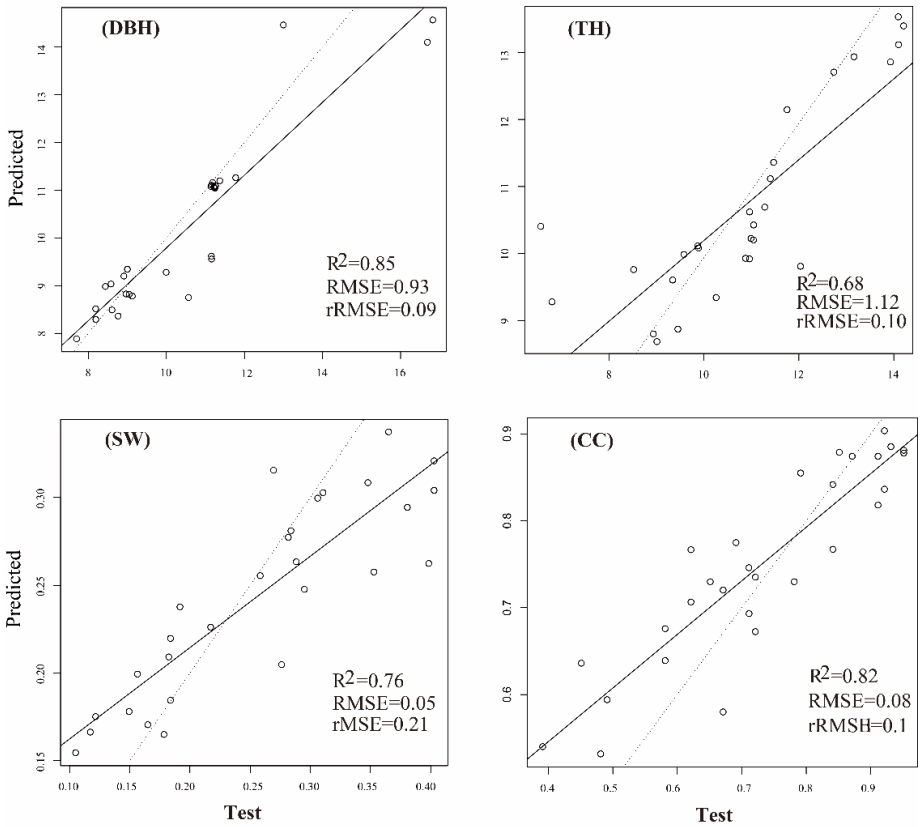

Figure 6. Validation of field-test and predicted DBH, TH, CC and SW by the random forest regression models.

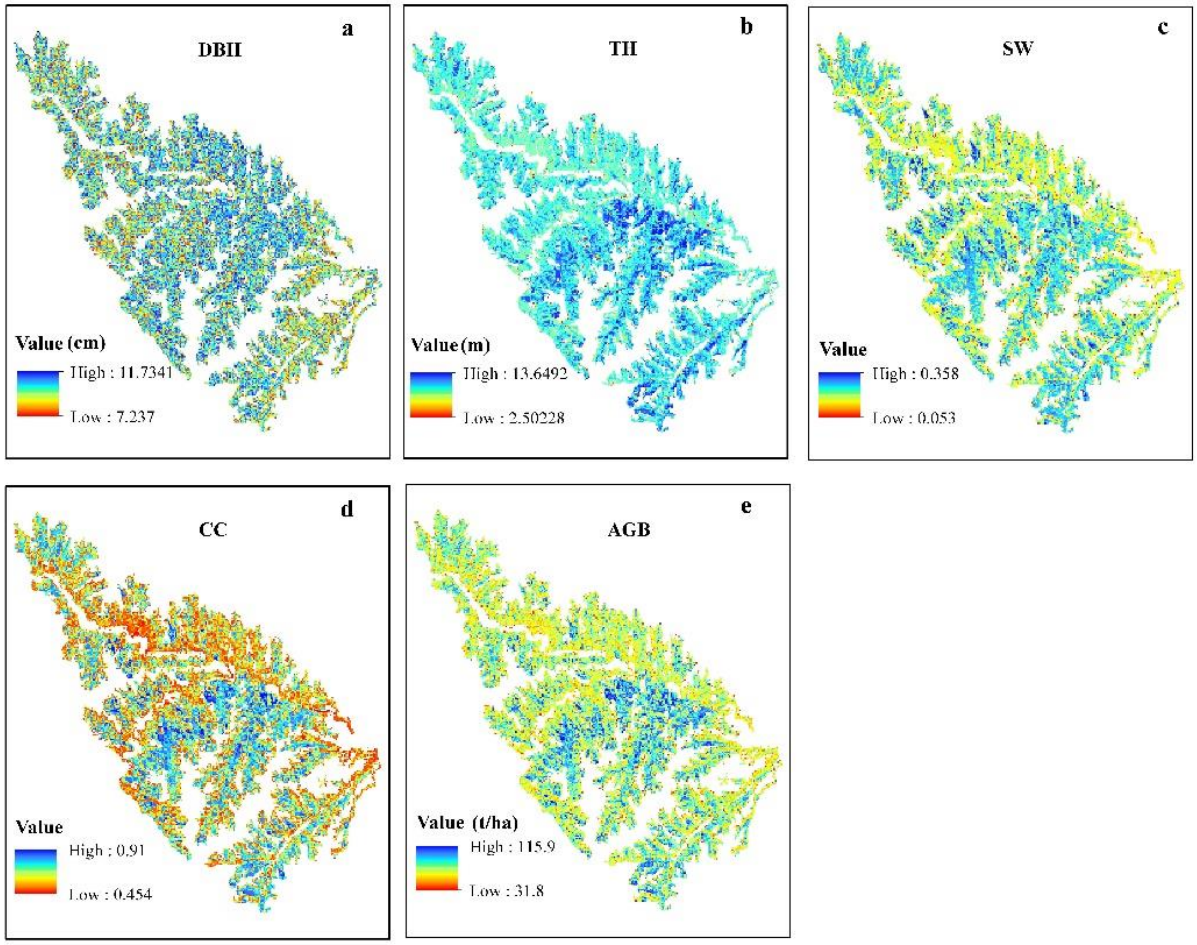

Figure 7. (**a**–**e**) Distributions of the DBH, TH, SW, CC and AGB in the study, respectively.

Figure 7e shows the spatial distribution of AGB estimated by the random forest model and the allometric growth equation in the study area. The distribution of AGB values had obvious differences. The average AGB was 55.80 t/ha, and the higher values were mainly distributed in the middle and south of the study area.

3.3. Regression Analysis

Moran's I tests were implemented to test spatial correlation. As shown in Table 2, Moran's I value was 0.5047, and $p < 0.01$, meaning that the original hypothesis of "the residuals do not have spatial correlation" was strongly rejected at the significance level of 1%, indicating that the residuals of the OLS model regression exhibited obvious spatial correlations. Moreover, the LM and robust LM test results were significant at the 1% level, indicating that the regression results of the OLS model were unreliable, and spatial regression models needed to be established. The AIC, log-likelihood and R^2 values were used to select appropriate panel data models. Our research indicated that the AIC value of SEM was 1,461,430, which was lower than that of SLM, and the log-likelihood and R^2 values of SEM were −730,704.64 and 0.87, respectively, which were higher than those of SLM (Table 3). Thus, SEM was more suitable than SLM as an explanatory model. The GWR model results showed that the AIC and R^2 values were 1,861,513 and 0.74, respectively. The results indicate that the GWR model was worse than the SEM. Thus, SEM was the best model for explaining the relationship between AGB and its driving factors in this study.

Table 2. Results of OLS regression and results evaluation.

Index	Value
Moran's I	0.5047 **
Robust LM-lag	393.8851 **
Robust LM-error	288,840.2266 **
LM-lag	249,662.4473 **
L M-error	538,108.7888 **
R^2	0.74
LIK	−930,747
AIC	1,861,510

** denotes $p < 0.01$, extremely significant.

Table 3. Evaluation of SEM, SLM and GWR regression results.

Index	SEM	SLM	GWR
R^2	0.87	0.82	0.74
LIK	−730,704.64	−827,284	−1,861,493.15
AIC	1,461,430	1,654,590	1,861,513.15

AGB was extremely significantly positively correlated with SW and CC (Table 4). AGB was extremely significantly positively correlated with the SOC at the 0–20 cm level and the SOC at the 40–60 cm level ($p < 0.01$), and AGB was significantly positively correlated with the SOC at the 20–40 cm level ($p < 0.05$); however, the relationship between AGB and the SOC at the 60–80 cm level was not significant. This result shows that AGB was extremely significantly negatively correlated with aspect ($p < 0.01$), which indicated that shady slopes and semi-shady slopes were more conducive to AGB accumulation than semi-sunny slopes and sunny slopes were. The relationship between AGB and slope was not significant.

Table 4. Results of SEM regression.

Independent Variable	Regression Coefficient
SW	23.109 **
CC	4.97327 **
SOC 0–20 cm level	0.00989803 **
SOC 20–40 cm level	0.0137214 *
SOC 40–60 cm level	0.13154 **
SOC 60–80 cm level	−0.0115614
Aspect	−0.0106707 **
Slope	−0.00145979

* denotes $p < 0.05$, significant. ** denotes $p < 0.01$, extremely significant.

4. Discussion

Traditional survey methods are time-consuming and laborious and obtain only limited point data. However, estimation by modeling allows us to quickly determine the status of an indicator in an area and provides a broad perspective for research. OK interpolation has been widely used to estimate the distribution of SOC. Gouri (2017) estimated the spatial variation in SOC in the Medinipur Block, West Bengal, India, and found that OK was a superior method for the interpolation of SOC spatial distribution by comparing different interpolation techniques [46]. This research provides support for our choice of OK interpolation for estimating the distribution of SOC. Zhao (2017) used OK to estimate the spatial distribution of SOC in a hilly gully watershed of the Loess Plateau, China [66]. This result further showed that OK interpolation could be applied to determine the distribution of SOC on the Loess Plateau. Similarly, our research found that the accuracy of OK interpolation met the accuracy requirements and was consistent with the results of previous studies. Our research found that the similar SOC values in the study area were mostly concentrated, indicating that SOC was spatially related.

Zhao (2019) found that random forest was most suitable for forest parameter estimations on the Loess Plateau [13]. We used random forest to estimate the DBH, TH, SW and CC and obtained satisfactory results. The estimation accuracies (R^2) of the DBH, CC and SW were all greater than 0.75, but TH had a lower accuracy, with an R^2 of 0.68. High estimation accuracy was closely related to variable selection using the Boruta algorithm. Shu (2021) estimated soil and plant analyzer development values for maize leaves and used the Borura algorithm to filter variables to improve evaluation accuracy [67]. High estimation accuracy was also related to high spatial resolution remote sensing images, which could provide more detailed spatial information [68]. We used Gaofen-1 satellite data with a high spatial resolution, which were successfully applied in our study. Han (2022) estimated forest AGB using GF-1 satellite data and Sentinel-1 satellite data in the Dabie Mountain Region in China, and found that GF-1 satellite data were better than Sentinel-1 satellite data [38]. This result may support our use of GF-1 as being correct. The low estimation accuracy of TH may be related to the optical remote sensing that was used because optical remote sensing has weak recognition of the vertical structure of a forest. Furthermore, the correlation between TH and canopy size in the study area was low, resulting in a slightly lower estimation accuracy. To solve this problem, some scholars have used radar remote sensing and believe that radar remote sensing can better recognize the vertical structure of forest stands [69,70]. This approach could be applied in our future research; however, the estimation accuracy still met our needs. We used the allometric growth equation to calculate the forest AGB. The allometric growth equation was used to estimate the AGB of stumpage by easily measuring stumpage factors. Compared with using direct logging to measure AGB, the allometric growth equation has the advantages of protecting trees, avoiding damage to the ecological environment in the study area, and conserving manpower, material resources and financial resources.

We found that the explanatory variable and the dependent variable were spatially correlated through the spatial correlation test and analysis and judged that the least squares method was not suitable for analyzing the relationship between AGB and its driving factors. Therefore, we chose spatial regression models (GWR, SEM, SLM) to explain the relationship between AGB and driving factors. The research results showed that the SEM was most suitable to explain the spatial relationship between AGB and its driving factors compared to the SLM and GWR. Wu (2020) used spatial regression models to research the affecting factors leading to the spatial variation in heavy metals in suburban soil and found that the SLM had the best applicability [64]. This result shows that the performance of the spatial regression model differs depending on the research objects. Our research results may provide a new idea for understanding the relationship between biomass and its driving factors. In this study, AGB was significantly positively correlated with SOC at depths of 0–20 cm, 20–40 cm and 40–60 cm ($p < 0.05$) and had no significant relationship with SOC at depths of 60–80 cm, which may be related to the distribution of *R. pseudoacacia* roots. Zhang (2011) studied the root distribution characteristics of *R. pseudoacacia* in Yan'an Yangou on the Loess Plateau and found that the fine roots of *R. pseudoacacia* were concentrated in the 0–60 cm soil layer [71]. This result explains our research results well. A large number of studies have shown that biodiversity promotes the accumulation of biomass [72,73]. This may be due to positive relationships among species that make full use of natural resources. The niche complementarity hypothesis supports this result [74]. With the continuous growth of the forest stand, the AGB continues to accumulate, and the stand CC continues to increase. Therefore, there was a significantly positive correlation between the AGB and CC before the stand was completely closed. AGB was significantly related to the slope aspect, and AGB was better on shady slopes and semi-shaded slopes than on semi-sunny slopes and sunny slopes. This result may be because water is the dominant factor for plant growth in the Loess Plateau. Shady slopes and semi-shady slopes have better water conditions than do semi-sunny slopes and sunny slopes, which is conducive to the accumulation of AGB in the *R. pseudoacacia* forest.

5. Conclusions

In this study, we estimated the DBH, TH, SW, CC and SOC. We used different regression models to fit the relationships between AGB and SW, CC, SOC, slope and aspect, and we found that AGB was obviously spatially correlated with driving factors. Therefore, the OLS model was not suitable for constructing the relationship between AGB and its driving factors, and a suitable spatial regression model needed to be selected. After further comparison, we found that SEM was most suitable for constructing the spatial relationship between AGB and its driving factors. The study also found that AGB was significantly positively correlated with CC, SW and the 0–60 cm SOC contents, and significantly negatively correlated with slope aspect. This study provides a reference for studying the relationship between biomass and driving factors and provides a basis for rational forest management, increasing biomass, and giving full play to the role of carbon sequestration.

The SEM model had the best application effect in this study, but the application of the model has limitations. It is uncertain whether it can have the same effect in other regions, and further research needs to explore this further.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs14122842/s1, Table S1: Accuracy assessment indices were selected to use for OK and random forest model accuracy verification; Table S2: Vegetation indices were selected to use for forest parameter estimation; Table S3: Formula for extracting the texture variables in this study; Table S4: Regression equation for biomass of each organ of *Robinia pseudoacacia*.

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

Funding: This research was funded by the National Key Research and Development Program of China, grant number 2016YFC0501706-1.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank Binhou Yu from Changwu Forestry Bureasu for his help in field investigation.

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

References

1. Chen, L.; Ren, C.; Zhang, B.; Wang, Z.; Xi, Y. Estimation of Forest Above-Ground Biomass by Geographically Weighted Regression and Machine Learning with Sentinel Imagery. *Forests* **2018**, *9*, 582. [CrossRef]
2. Nguyen, T.H.; Jones, S.; Soto-Berelov, M.; Haywood, A.; Hislop, S. Landsat Time-Series for Estimating Forest Aboveground Biomass and Its Dynamics across Space and Time: A Review. *Remote Sens.* **2020**, *12*, 98. [CrossRef]
3. Shen, G.; Wang, Z.; Liu, C.; Han, Y. Mapping aboveground biomass and carbon in Shanghai's urban forest using Landsat ETM+ and inventory data. *Urban For. Urban Green.* **2020**, *51*, 126655. [CrossRef]
4. Hao, M.; Messier, C.; Geng, Y.; Zhang, C.; Zhao, X.; Von Gadow, K. Functional traits influence biomass and productivity through multiple mechanisms in a temperate secondary forest. *Forstwiss. Cent.* **2020**, *139*, 959–968. [CrossRef]
5. Timothy, D.; Onisimo, M.; Cletah, S.; Adelabu, S.; Tsitsi, B. Remote sensing of aboveground forest biomass: A review. *Trop. Ecol.* **2016**, *57*, 125–132.
6. Muukkonen, P.; Heiskanen, J. Estimating biomass for boreal forests using ASTER satellite data combined with standwise forest inventory data. *Remote Sens. Environ.* **2005**, *99*, 434–447. [CrossRef]
7. Lefsky, M.; Turner, D.; Guzy, M.; Cohen, W. Combining lidar estimates of aboveground biomass and Landsat estimates of stand age for spatially extensive validation of modeled forest productivity. *Remote Sens. Environ.* **2005**, *95*, 549–558. [CrossRef]
8. Morais, T.G.; Teixeira, R.F.; Figueiredo, M.; Domingos, T. The use of machine learning methods to estimate aboveground biomass of grasslands: A review. *Ecol. Indic.* **2021**, *130*, 108081. [CrossRef]
9. Moradi, F.; Darvishsefat, A.A.; Pourrahmati, M.R.; Deljouei, A.; Borz, S.A. Estimating Aboveground Biomass in Dense Hyrcanian Forests by the Use of Sentinel-2 Data. *Forests* **2022**, *13*, 104. [CrossRef]
10. Purohit, S.; Aggarwal, S.P.; Patel, N.R. Estimation of forest aboveground biomass using combination of Landsat 8 and Sentinel-1A data with random forest regression algorithm in Himalayan Foothills. *Trop. Ecol.* **2021**, *62*, 288–300. [CrossRef]

11. Dyderski, M.K.; Pawlik, L. Drivers of forest aboveground biomass and its increments in the Tatra Mountains after 15 years. *Catena* **2021**, *205*, 105468. [CrossRef]
12. Torre-Tojal, L.; Bastarrika, A.; Boyano, A.; Lopez-Guede, J.M.; Graña, M. Above-ground biomass estimation from LiDAR data using random forest algorithms. *J. Comput. Sci.* **2022**, *58*, 101517. [CrossRef]
13. Zhao, Q.; Yu, S.; Zhao, F.; Tian, L.; Zhao, Z. Comparison of machine learning algorithms for forest parameter estimations and application for forest quality assessments. *For. Ecol. Manag.* **2019**, *434*, 224–234. [CrossRef]
14. Chen, H.Y.H.; Luo, Y. Net aboveground biomass declines of four major forest types with forest ageing and climate change in western Canada's boreal forests. *Glob. Chang. Biol.* **2015**, *21*, 3675–3684. [CrossRef]
15. Bordin, K.M.; Esquivel-Muelbert, A.; Bergamin, R.S.; Klipel, J.; Picolotto, R.C.; Frangipani, M.A.; Zanini, K.J.; Cianciaruso, M.V.; Jarenkow, J.A.; Jurinitz, C.F.; et al. Climate and large-sized trees, but not diversity, drive above-ground biomass in subtropical forests. *For. Ecol. Manag.* **2021**, *490*, 119126. [CrossRef]
16. Sullivan, M.J.P.; Lewis, S.L.; Affum-Baffoe, K.; Castilho, C.; Costa, F.; Sanchez, A.C.; Ewango, C.E.N.; Hubau, W.; Marimon, B.; Monteagudo-Mendoza, A.; et al. Long-term thermal sensitivity of Earth's tropical forests. *Science* **2020**, *368*, 869–874. [CrossRef]
17. Li, Y.; Brando, P.M.; Morton, D.C.; Lawrence, D.M.; Yang, H.; Randerson, J.T. Deforestation-induced climate change reduces carbon storage in remaining tropical forests. *Nat. Commun.* **2022**, *13*, 1–13. [CrossRef]
18. Xu, M.; Ji, H.; Zhuang, S. Carbon stock of Moso bamboo (Phyllostachys pubescens) forests along a latitude gradient in the subtropical region of China. *PLoS ONE* **2018**, *13*, e0193024. [CrossRef]
19. Sheikh, M.A.; Kumar, M.; Todaria, N.; Pandey, R. Biomass and soil carbon along altitudinal gradients in temperate Cedrus deodara forests in Central Himalaya, India: Implications for climate change mitigation. *Ecol. Indic.* **2020**, *111*, 106025. [CrossRef]
20. Shen, A.; Wu, C.; Jiang, B.; Deng, J.; Yuan, W.; Wang, K.; He, S.; Zhu, E.; Lin, Y.; Wu, C. Spatiotemporal Variations of Aboveground Biomass under Different Terrain Conditions. *Forests* **2018**, *9*, 778. [CrossRef]
21. Pariyar, S.; Volkova, L.; Sharma, R.P.; Sunam, R.; Weston, C.J. Aboveground carbon of community-managed Chirpine (Pinus roxburghii Sarg.) forests of Nepal based on stand types and geographic aspects. *PeerJ* **2019**, *7*, e6494. [CrossRef] [PubMed]
22. Becknell, J.M.; Powers, J.S. Stand age and soils as drivers of plant functional traits and aboveground biomass in secondary tropical dry forest. *Can. J. For. Res.* **2014**, *44*, 604–613. [CrossRef]
23. Liu, L.; Zeng, F.; Song, T.; Wang, K.; Du, H. Stand Structure and Abiotic Factors Modulate Karst Forest Biomass in Southwest China. *Forests* **2020**, *11*, 443. [CrossRef]
24. Xu, L.; Shi, Y.; Fang, H.; Zhou, G.; Xu, X.; Zhou, Y.; Tao, J.; Ji, B.; Xu, J.; Li, C.; et al. Vegetation carbon stocks driven by canopy density and forest age in subtropical forest ecosystems. *Sci. Total Environ.* **2018**, *631–632*, 619–626. [CrossRef]
25. Ali, A.; Lin, S.; He, J.; Kong, F.-M.; Yu, J.-H.; Jiang, H.-S. Climate and soils determine aboveground biomass indirectly via species diversity and stand structural complexity in tropical forests. *For. Ecol. Manag.* **2019**, *432*, 823–831. [CrossRef]
26. Yang, Y.; Dou, Y.; An, S.; Zhu, Z. Abiotic and biotic factors modulate plant biomass and root/shoot (R/S) ratios in grassland on the Loess Plateau, China. *Sci. Total Environ.* **2018**, *636*, 621–631. [CrossRef]
27. Sun, J.; Niu, S.; Wang, J. Divergent biomass partitioning to aboveground and belowground across forests in China. *J. Plant Ecol.* **2018**, *11*, 484–492. [CrossRef]
28. Yang, Y.; Dou, Y.; An, S. Environmental driving factors affecting plant biomass in natural grassland in the Loess Plateau, China. *Ecol. Indic.* **2017**, *82*, 250–259. [CrossRef]
29. Li, Z.; Liang, M.; Li, Z.; Mariotte, P.; Tong, X.; Zhang, J.; Dong, L.; Zheng, Y.; Ma, W.; Zhao, L.; et al. Plant functional groups mediate effects of climate and soil factors on species richness and community biomass in grasslands of Mongolian Plateau. *J. Plant Ecol.* **2021**, *14*, 679–691. [CrossRef]
30. Ahmed, M.A.A.; Abd-Elrahman, A.; Escobedo, F.J.; Cropper, W.P.; Martin, T.A.; Timilsina, N. Spatially-explicit modeling of multi-scale drivers of aboveground forest biomass and water yield in watersheds of the Southeastern United States. *J. Environ. Manag.* **2017**, *199*, 158–171. [CrossRef]
31. Hsiao, C. Rejoinder on: Panel data analysis—Advantages and challenges. *Test* **2007**, *16*, 56–57. [CrossRef]
32. Yuan, M.; Huang, Y.; Shen, H.; Li, T. Effects of urban form on haze pollution in China: Spatial regression analysis based on PM2.5 remote sensing data. *Appl. Geogr.* **2018**, *98*, 215–223. [CrossRef]
33. Sun, F.; Matthews, S.A.; Yang, T.-C.; Hu, M.-H. A spatial analysis of the COVID-19 period prevalence in U.S. counties through June 28, 2020: Where geography matters? *Ann. Epidemiol.* **2020**, *52*, 54–59.e1. [CrossRef] [PubMed]
34. Resende, G.M.; de Carvalho, A.X.Y.; Sakowski, P.A.M.; Cravo, T.A. Evaluating multiple spatial dimensions of economic growth in Brazil using spatial panel data models. *Ann. Reg. Sci.* **2016**, *56*, 1–31. [CrossRef]
35. Lončar, D.; Paunković, J.; Jovanović, V.; Krstić, V. Environmental and social responsibility of companies cross EU countries–Panel data analysis. *Sci. Total Environ.* **2019**, *657*, 287–296. [CrossRef]
36. Jia, K.; Liang, S.; Gu, X.; Baret, F.; Wei, X.; Wang, X.; Yao, Y.; Yang, L.; Li, Y. Fractional vegetation cover estimation algorithm for Chinese GF-1 wide field view data. *Remote Sens. Environ.* **2016**, *177*, 184–191. [CrossRef]
37. Zhang, Y.; Yang, J.; Liu, X.; Du, L.; Shi, S.; Sun, J.; Chen, B. Estimation of Multi-Species Leaf Area Index Based on Chinese GF-1 Satellite Data Using Look-Up Table and Gaussian Process Regression Methods. *Sensors* **2020**, *20*, 2460. [CrossRef] [PubMed]
38. Han, H.; Wan, R.; Li, B. Estimating Forest Aboveground Biomass Using Gaofen-1 Images, Sentinel-1 Images, and Machine Learning Algorithms: A Case Study of the Dabie Mountain Region, China. *Remote Sens.* **2022**, *14*, 176. [CrossRef]

39. Yu, S.; Wang, F.; Qu, M.; Yu, B.; Zhao, Z. The Effect of Land Use/Cover Change on Soil Erosion Change by Spatial Regression in Changwu County on the Loess Plateau in China. *Forests* **2021**, *12*, 1209. [CrossRef]
40. Zhang, X.-C. A comparison of explicit and implicit spatial downscaling of GCM output for soil erosion and crop production assessments. *Clim. Chang.* **2007**, *84*, 337–363. [CrossRef]
41. Zhang, X.-C.; Liu, W. Simulating potential response of hydrology, soil erosion, and crop productivity to climate change in Changwu tableland region on the Loess Plateau of China. *Agric. For. Meteorol.* **2005**, *131*, 127–142. [CrossRef]
42. Wu, H.; Qian, H. Innovative trend analysis of annual and seasonal rainfall and extreme values in Shaanxi, China, since the 1950s. *Int. J. Clim.* **2017**, *37*, 2582–2592. [CrossRef]
43. Frazer, G.W.; Canham, C.D.; Lertzman, K.P. *Gap Light Analyzer (GLA), Version 2.0: Imaging Software to Extract Canopy Structure and Gap Light Transmission Indices from True-Colour Fisheye Photographs, Users Manual and Program Documentation*; Simon Fraser University: Burnaby, BC, Canada; The Institute of Ecosystem Studies: Millbrook, NY, USA, 1999; p. 36.
44. Liu, J.; Dang, P.; Gao, Y.; Zhu, H.; Zhu, H.; Zhao, F.; Zhao, Z. Effects of tree species and soil properties on the composition and diversity of the soil bacterial community following afforestation. *For. Ecol. Manag.* **2018**, *427*, 342–349. [CrossRef]
45. Jia, B.; Niu, Z.; Wu, Y.; Kuzyakov, Y.; Li, X.G. Waterlogging increases organic carbon decomposition in grassland soils. *Soil Biol. Biochem.* **2020**, *148*, 107927. [CrossRef]
46. Göl, C.; Bulut, S.; Bolat, F. Comparison of different interpolation methods for spatial distribution of soil organic carbon and some soil properties in the Black Sea backward region of Turkey. *J. Afr. Earth Sci.* **2017**, *134*, 85–91. [CrossRef]
47. Das, S. Extreme rainfall estimation at ungauged sites: Comparison between region-of-influence approach of regional analysis and spatial interpolation technique. *Int. J. Clim.* **2019**, *39*, 407–423. [CrossRef]
48. Ding, Q.; Wang, Y.; Zhuang, D. Comparison of the common spatial interpolation methods used to analyze potentially toxic elements surrounding mining regions. *J. Environ. Manag.* **2018**, *212*, 23–31. [CrossRef]
49. Xie, B.; Zhang, H.K.; Xue, J. Deep Convolutional Neural Network for Mapping Smallholder Agriculture Using High Spatial Resolution Satellite Image. *Sensors* **2019**, *19*, 2398. [CrossRef]
50. Jaafar, W.S.W.M.; Maulud, K.N.A.; Kamarulzaman, A.M.M.; Raihan, A.; Sah, S.; Ahmad, A.; Saad, S.N.M.; Azmi, A.T.M.; Syukri, N.K.A.J.; Khan, W.R. The Influence of Deforestation on Land Surface Temperature—A Case Study of Perak and Kedah, Malaysia. *Forests* **2020**, *11*, 670. [CrossRef]
51. Lara, M.J.; Chen, Y.; Jones, B.M. Recent warming reverses forty-year decline in catastrophic lake drainage and hastens gradual lake drainage across northern Alaska. *Environ. Res. Lett.* **2021**, *16*, 124019. [CrossRef]
52. Ye, Q.; Yu, S.; Liu, J.; Zhao, Q.; Zhao, Z. Aboveground biomass estimation of black locust planted forests with aspect variable using machine learning regression algorithms. *Ecol. Indic.* **2021**, *129*, 107948. [CrossRef]
53. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *6*, 610–621. [CrossRef]
54. Kursa, M.B.; Rudnicki, W.R. Feature Selection with theBorutaPackage. *J. Stat. Softw.* **2010**, *36*, 1–13. [CrossRef]
55. Degenhardt, F.; Seifert, S.; Szymczak, S. Evaluation of variable selection methods for random forests and omics data sets. *Briefings Bioinform.* **2019**, *20*, 492–503. [CrossRef] [PubMed]
56. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
57. Srinet, R.; Nandy, S.; Patel, N. Estimating leaf area index and light extinction coefficient using Random Forest regression algorithm in a tropical moist deciduous forest, India. *Ecol. Informatics* **2019**, *52*, 94–102. [CrossRef]
58. RColorBrewer, S.; Liaw, M.A. *Package 'Randomforest'*; University of California, Berkeley: Berkeley, CA, USA, 2018.
59. Zhang, B.L.; Chen, C.G. Biomass and production of robinia pseudoacacia planatation in Hongxing tree farm of Changwu county, Shaanxi province. *Shaanxi For. Sci. Technol.* **1992**, *3*, 13–17.
60. Stewart Fotheringham, A.; Charlton, M.; Brunsdon, C. The geography of parameter space: An investigation of spatial non-stationarity. *Int. J. Geogr. Inf. Syst.* **1996**, *10*, 605–627. [CrossRef]
61. Brunsdon, C.; Fotheringham, A.S.; Charlton, M.E. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geogr. Anal.* **1996**, *28*, 281–298. [CrossRef]
62. Anselin, L. *Spatial Econometrics: Methods and Models*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 1988; Volume 4.
63. Hankach, P.; Gastineau, P.; Vandanjon, P.-O. Multi-scale spatial analysis of household car ownership using distance-based Moran's eigenvector maps: Case study in Loire-Atlantique (France). *J. Transp. Geogr.* **2022**, *98*, 103223. [CrossRef]
64. Wu, Z.; Chen, Y.; Han, Y.; Ke, T.; Liu, Y. Identifying the influencing factors controlling the spatial variation of heavy metals in suburban soil using spatial regression models. *Sci. Total Environ.* **2020**, *717*, 137212. [CrossRef] [PubMed]
65. You, H.; Zhou, D.; Wu, S.; Hu, X.; Bie, C. Social Deprivation and Rural Public Health in China: Exploring the Relationship Using Spatial Regression. *Soc. Indic. Res.* **2020**, *147*, 843–864. [CrossRef]
66. Zhao, B.; Li, Z.; Li, P.; Xu, G.; Gao, H.; Cheng, Y.; Chang, E.; Yuan, S.; Zhang, Y.; Feng, Z. Spatial distribution of soil organic carbon and its influencing factors under the condition of ecological construction in a hilly-gully watershed of the Loess Plateau, China. *Geoderma* **2017**, *296*, 10–17. [CrossRef]
67. Shu, M.; Zuo, J.; Shen, M.; Yin, P.; Wang, M.; Yang, X.; Tang, J.; Li, B.; Ma, Y. Improving the estimation accuracy of SPAD values for maize leaves by removing UAV hyperspectral image backgrounds. *Int. J. Remote Sens.* **2021**, *42*, 5862–5881. [CrossRef]

68. Chen, Y.; Ming, D.; Zhao, L.; Lv, B.; Zhou, K.; Qing, Y. Review on High Spatial Resolution Remote Sensing Image Segmentation Evaluation. *Photogramm. Eng. Remote Sens.* **2018**, *84*, 629–646. [CrossRef]
69. Chen, W.; Zheng, Q.; Xiang, H.; Chen, X.; Sakai, T. Forest Canopy Height Estimation Using Polarimetric Interferometric Synthetic Aperture Radar (PolInSAR) Technology Based on Full-Polarized ALOS/PALSAR Data. *Remote Sens.* **2021**, *13*, 174. [CrossRef]
70. Lee, Y.-S.; Lee, S.; Baek, W.-K.; Jung, H.-S.; Park, S.-H.; Lee, M.-J. Mapping Forest Vertical Structure in Jeju Island from Optical and Radar Satellite Images Using Artificial Neural Network. *Remote Sens.* **2020**, *12*, 797. [CrossRef]
71. Zhang, L.D.; Xu, X.X. Distribution Characters of Robinia pseudoacacia root in Yangou Watershed in Yanan. *J. Northwest For. Univ.* **2011**, *26*, 9–14.
72. Ali, A.; Chen, H.Y.; You, W.-H.; Yan, E.-R. Multiple abiotic and biotic drivers of aboveground biomass shift with forest stratum. *For. Ecol. Manag.* **2019**, *436*, 1–10. [CrossRef]
73. Ali, A.; Sanaei, A.; Li, M.; Nalivan, O.A.; Ahmadaali, K.; Pour, M.J.; Valipour, A.; Karami, J.; Aminpour, M.; Kaboli, H.; et al. Impacts of climatic and edaphic factors on the diversity, structure and biomass of species-poor and structurally-complex forests. *Sci. Total Environ.* **2020**, *706*, 135719. [CrossRef]
74. Ali, A.; Yan, E.R. Relationships between biodiversity and carbon stocks in forest ecosystems: A systematic literature review. *Trop. Ecol.* **2017**, *58*, 1–14.