

Article

Coniferous Forests Aboveground Biomass Inversion in Typical Regions of China with Joint Sentinel-1 and Sentinel-2 Remote Sensing Data Supported by Different Feature Optimizing Algorithms

Fuxiang Zhang ¹, Armando Marino ², Yongjie Ji ^{1,*}  and Wangfei Zhang ³ 

¹ School of Geography and Ecotourism, Southwest Forestry University, Kunming 650224, China; zfx13211@163.com

² Biological and Environmental Sciences, The University of Stirling, Stirling FK94LA, UK; armado.marino@stir.ac.uk

³ Forestry College, Southwest Forestry University, Kunming 650224, China; mewhff@163.com

* Correspondence: jiyongjie@live.cn

Abstract: Multispectral remote sensing (RS) data and synthetic aperture radar (SAR) data can provide horizontal and vertical information about forest AGB under different stand conditions. With the abundance of RS features extracted from multispectral and SAR datasets, a key point for accurate forest AGB estimation is to use suitable feature optimization inversion algorithms. In this study, feature optimization inversion algorithms including multiple linear stepwise regression (MLSR), K-nearest neighbor with fast iterative feature selection (KNN-FIFS), and random forest (RF) were explored, with a total of 93 RS features working as inversion model input for forest AGB inversion. The results showed that KNN-FIFS with the combination of Sentinel-1 and Sentinel-2 performed best at both test sites ($R^2 = 0.568$ and $RMSE = 15.05$ t/hm² for Puer and $R^2 = 0.511$ and $RMSE = 32.29$ t/hm² for Genhe). Among the three feature optimization inversion algorithms, RF performed worst for forest AGB estimation with $R^2 = 0.348$ and $RMSE = 18.06$ t/hm² for Puer and $R^2 = 0.345$ and $RMSE = 35.98$ t/hm² for Genhe using the feature combination of Sentinel-1 and Sentinel-2. The results indicated that a combination of features extracted from Sentinel-1 and Sentinel-2 can improve the inversion accuracy of forest AGB, and the KNN-FIFS algorithm has robustness and transferability in forest AGB inversions.

Keywords: forest AGB; different feature optimizing algorithms; Sentinel-1; Sentinel-2; KNN-FIFS



Citation: Zhang, F.; Marino, A.; Ji, Y.; Zhang, W. Coniferous Forests Aboveground Biomass Inversion in Typical Regions of China with Joint Sentinel-1 and Sentinel-2 Remote Sensing Data Supported by Different Feature Optimizing Algorithms. *Forests* **2024**, *15*, 56. <https://doi.org/10.3390/f15010056>

Forests **2024**, *15*, 56. <https://doi.org/10.3390/f15010056>

Academic Editor: Rafael María Navarro-Cerrillo

Received: 20 November 2023

Revised: 24 December 2023

Accepted: 25 December 2023

Published: 28 December 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

Forests are crucial for global terrestrial ecosystem carbon cycling and climate change research [1]. Forest ecosystems are the largest and most important natural ecosystems in terrestrial ecosystems and play an important role in maintaining global ecological balance [2,3]. Forest AGB is an important variable in evaluating the capacity of carbon sequestration and maintaining carbon balance. Accurate forest AGB estimation is significant for studying carbon cycling in regional terrestrial ecosystems [4,5].

Currently, traditional field measurement and remote sensing (RS) inversion are the two main ways of obtaining forest AGB. Although the field measurement method is the most accurate way for forest AGB calculation, it is not feasible for regional or global-scale forest AGB calculation. It is also time cost and labor cost [6]. RS, on the other hand, is capable of measuring and monitoring forest AGB on a regional or global scale. As a result, RS images acquired from various sensors, such as passive and active sensors, are widely used for forest AGB estimation. Multispectral RS datasets have been widely used in forest AGB estimation, and the correlation between forest AGB and RS features extracted from multispectral RS images has been explored and established. Spectral reflectance,

vegetation index, and texture characteristics show a good correlation with changes in forest AGBs. However, multispectral RS datasets have poor vegetation canopy penetration, which prevents the acquisition of accurate information on the vertical structure of forests and then results in saturation problems, especially when forestry AGB densities are high. Synthetic aperture radar (SAR), operated at microwave bands, can penetrate the vegetation canopy and interact with the main biomass components, for example, trunks and branches, showing great potential for increasing the accuracy of forest AGB estimation [7]. Meanwhile, the penetration capability increases with the growing wavelength, which means that a stronger correlation exists between forest AGB and SAR reflected signals at longer microwave bands [8].

Relevant studies have pointed out that there are some limitations in the estimation of forest AGB using a single SAR dataset or a single optical dataset, and the inversion accuracy and efficiency can be improved when both are used jointly [9–11]. Shao et al. estimated coniferous forest AGB using Radarsat-2 and Landsat-8 OLI in Genhe, Inner Mongolia, and showed that combined optical and microwave information could estimate forest AGB more accurately [12]. Li et al. found poor performance in forest AGB estimation when only SAR backscatter coefficients were used. While the combination of Radarsat-2 and Landsat-5 TM improved the estimation accuracy with $R^2 = 0.768$ and $RMSE = 19.14 \text{ t/hm}^2$ [13], Pan et al. estimated forest AGB using a combination of GF-3 PolSAR and Landsat-8 OLI with KNN-SFS in Guangxi, and the results showed that the accuracy of the inversion results was higher than that of the estimation using a single data source with $R^2 = 0.75$ and $RMSE = 21.05 \text{ t/hm}^2$ [14].

Although the combination of optical and SAR datasets can improve forest biomass, there are still significant limitations in data access, and some of the data is not available to the public, which results in certain costs. Sentinel-1 SAR data and Sentinel-2 multispectral data are not only free of charge but also timely revisits in global coverage and provide new opportunities for inversion and monitoring of forest AGBs at a regional or global scale. Pan et al. used backscatter coefficients and texture features extracted from Sentinel-1 with the multiple linear stepwise regression (MLSR) algorithm for forest AGB estimation at the regional scale [15]. The study by Guo et al. showed that Sentinel-2 data can be better used for regional forest AGB inversion [16]. David et al. combined Sentinel-1 and Sentinel-2 to estimate forest AGB and achieved a good result with $R^2 = 0.95$ and $RMSE = 0.25 \text{ t/hm}^2$ [17]. However, only a few studies have explored the potential of using a combination of Sentinel-1 and Sentinel-2 to estimate forest AGBs.

Another key issue in forest AGB estimation is selecting suitable inversion models, especially with the capability of feature optimization algorithms. Tian et al. classified the estimation methods used in forest AGB into traditional estimation models and machine learning models [7]. Traditional model estimation methods have been widely used for forest AGB estimation, among which the MLSR algorithm has become a more commonly used algorithm due to its simple principle and easy operation. MLSR has been used for forest AGB estimation in tropical, subtropical, and temperate zones. Machine learning algorithms can solve the problems of data nonlinearity and high dimensionality, significantly improving the estimation accuracy of forest AGB. However, since the “black-box” operation, machine learning algorithms find it difficult to reflect the mechanistic process between RS features and forest AGB [15]. Exploring how to select an appropriate forest AGB inversion model under the support of optimal RS feature selection algorithms in different test sites, which aims to improve the estimation accuracy of forest AGB and the generalization capability of the inversion model in different forest scenes, has become one of the most hot research topics.

The north-eastern and south-western regions of China, as the main sources of timber supply and forest products in China, have made an important contribution to China’s timber industry and development [18,19]. Larch is one of the major afforestation species in the Northeast, while Simao Pine is one of the major high-yielding resin harvesting species in the Southwest, and turpentine is an important industrial raw material. Both species

are the main species in the National Natural Forest Protection Project [20]. Therefore, accurate estimation of the biomass of larch and Simao pine is not only beneficial to forest resource management and conservation but can also accurately calculate carbon stock. Currently, fewer studies are using Sentinel-1 SAR data and Sentinel-2 multispectral data for the inversion of AGB in larch and Simao pine, and there are uncertainties in the inversion results. In this study, two types of coniferous forests, Xing'an larch in Inner Mongolia and Simao pine in Yunnan Province, were selected to explore the suitable RS feature optimization inversion algorithms for their forest AGB estimation. RS features extracted from Single Sentinel-1, Single Sentinel-2, and combinations of Sentinel-1 and Sentinel-2 were input for the forest AGB inversion. The objective of this study is to explore the advantages of coniferous forest AGB estimation in typical regions in China under different combinations of the two RS data supported by different feature optimization inversion algorithms. Meanwhile, explore the robustness and transferability of these inversion algorithms to enhance monitoring and decision-making capabilities for forest resource management.

2. Materials and Methods

2.1. A General Description of the Test Sites

2.1.1. Yunnan Puer Test Site

This test site is located at part of Wanzhang Mountain covered by the aerial flight flies area (Figure 1). It is located in Puer City, Yunnan Province ($100^{\circ}29'–101^{\circ}7' E$, $22^{\circ}33'–23^{\circ}1' N$). In the test site, the forest-covered area is about $17,440 \text{ hm}^2$, with an elevation ranging from 570 m to 2010 m. The climate here is a low-latitude, mountainous, southern subtropical plateau monsoon climate. The average annual temperature is $19.4^{\circ} C$, and the annual precipitation is 1340.9 mm. The main dominant tree species is Simao pine (*Pinus kesiya var. langbianensis*). Among Simao pines, red cone (*Castanopsis hystrix* Miq), short-slabbbed quebracho (*C. echidnocarpa*), southwestern birch (*Betula alnoides* Buch), and so on are sporadically distributed.

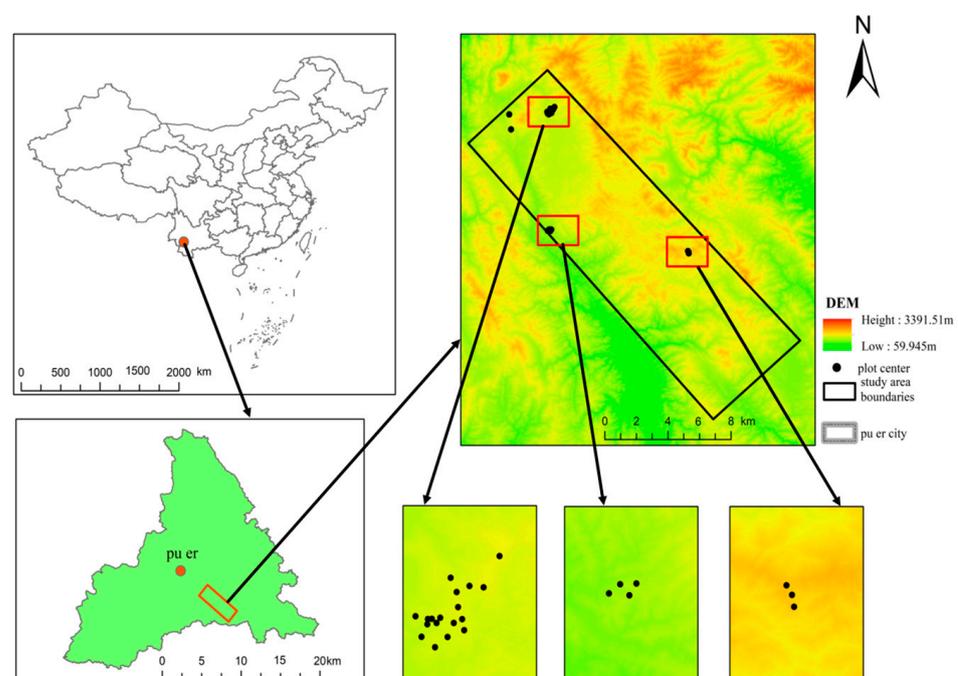


Figure 1. The location of the Puer test site. The red rectangle in the left-bottom sub-figure is the boundary of the aerial flight-flying area. The right sub-figures show the distribution of field-collected samples located in the aerial flight-flied area.

2.1.2. Genhe Test Site in Inner Mongolia

The test site is located at the National Field Scientific Observatory for Forest Ecosystems of the Greater Khingan Mountains in Genhe City, Inner Mongolia ($121^{\circ}30'–121^{\circ}31' E$, $50^{\circ}49'–50^{\circ}51' N$; Figure 2). The area of the test site is around $10,200 \text{ hm}^2$, and the elevation ranges from 810 m to 1116 m. It is currently the highest latitude for a forest ecosystem field scientific observatory in China. The test site has a temperate continental monsoon climate with an annual average temperature of $-5.3 \text{ }^{\circ}\text{C}$ and an annual temperature difference of $47.4 \text{ }^{\circ}\text{C}$. The test site is a typical area of high-latitude perennial permafrost and cold-temperate forest ecosystems. About 75% of the area in the test site is covered by forest, and the dominant tree species is Xing'an larch (*Larix gmelinii*), in which camphor pine (*Pinus sylvestris var. mongolica*), white birch (*Betula platyphylla*), and *Populus davidiana* (*Populus davidiana*) are sporadically distributed.

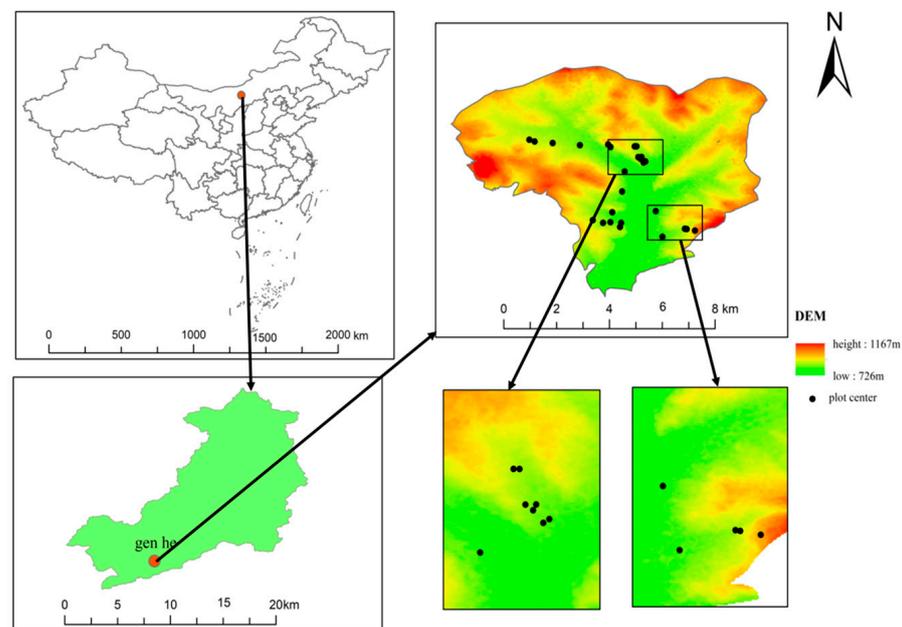


Figure 2. The location of the Genhe test site. The left-bottom sub-figure is the coverage of Genhe City. The right sub-figures show the distribution of field-collected samples located at the General Field Scientific Observatory Station.

2.2. Research Data Acquisition and Pre-Processing

2.2.1. Field Campaign Conducted in Puer

The ground survey was conducted in November and December 2020. A total of 27 pure Simao pine forest plots were collected. All of the sample plots are $20 \text{ m} \times 20 \text{ m}$ in size and were shaped as square plots. Differential GPS was used to locate the four corner points and center coordinates of each sample plot during the field sample collection. The location accuracy was controlled within 20 cm. Diameter at breast height (DBH), tree height, crown spread, and crown coverage rate were recorded during the field campaign, and the trees with DBH greater than 5 cm in each plot were recorded. The AGB of each tree in the sample plots was calculated with Equation (1) [21], and the AGB value of each tree in each plot was totaled to obtain the AGB of each plot, then the area of each sample plot (0.04 ha) was divided to obtain the AGB of each plot.

$$W_1 = 0.0582D^{2.1203}H^{0.4668} \quad (1)$$

where: D is the measured DBH, H is the measured tree height, and W_1 is the AGB of Simao pine.

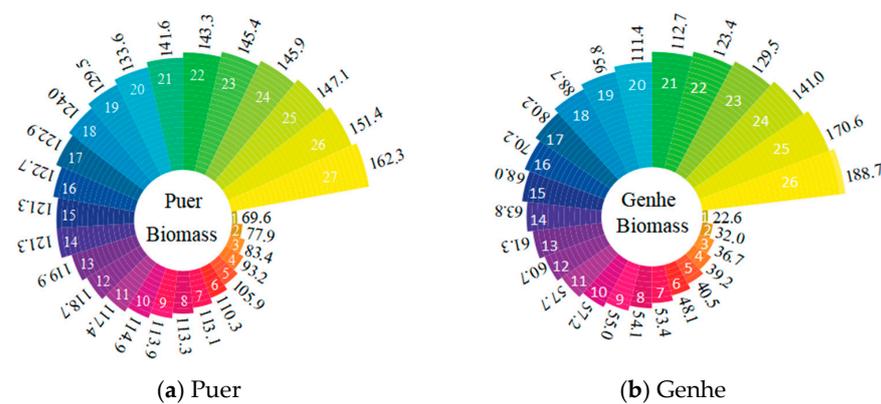
2.2.2. Field Campaign Conducted in Genhe

The field sample plots in Genhe were investigated from 31 July to 28 August 2016. A total of 26 pure forest plots of Xing'an larch were collected. The sample plots were 45 m × 45 m and 30 m × 30 m. Differential GPS was used in the same way as the Puer test site to locate plots. DBH, tree height, height under the first active branch, and crown spread were recorded. The AGB of dominated tree species was calculated by (2), as proposed by Chen and Zhu [22].

$$W_2 = 0.0277D_{BH}^{2.7930} \quad (2)$$

where W_2 is the above-ground biomass of Xing'an larch and D_{BH} is the measured diameter at breast height.

Figure 3a,b show the detailed information of the forest AGBs at the two test sites.



2.3. RS Features Extraction

2.3.1. Sentinel-1 RS Features Extraction

- (1) Sentinel-1 backscattering coefficients. After pre-processing the Sentinel-1 data in SNAP, the images were converted into sigma backscatter coefficient images in dB; the formula used for conversion is shown as (3); the point targets of the images were quantitatively measured and analyzed; and the resulted backscatter coefficient values were used for AGB modeling.

$$dB = 10 * \log(P/P_0) \quad (3)$$

where P and P_0 denote the target and reference quantities, respectively.

- (2) Sentinel-1 texture features. Texture features are an important source of information for high-resolution SAR data, and image texture has certain advantages in the identification of stand structure, such as stand age, stand density, and leaf area index [24]. Texture, the fine structure of an image, refers to the frequency of tonal changes in the image. Visually finer textures indicate less spatial variation in image luminance values over regions [23,25]. Visually coarser textures have more drastic variations in pixel values across regions [25]. Although texture parameters are used to classify land-use types and vegetation, the relationship between image texture and forest AGB has not been fully explored [15].

Firstly, the backward scattering coefficient image is converted to a normalized backward scattering coefficient image as it can provide better dynamic range on low backward scattering targets. Secondly, the Frost filtering process is performed to suppress speckle noise [26]. Common texture features can be classified into first-order texture and second-order texture; the first-order texture is the spectral statistics value in the neighborhood, and the second-order texture characterizes the spatial distribution of image elements by describing the interdependence between the image elements. The commonly used method is the gray level co-occurrence matrix (GLCM) [27]. In this study, a total of 64 second-order texture variables extracted based on the GLCM of the VV and VH polarization images of the Sentinel-1 data were extracted using four window sizes of 3×3 , 5×5 , 7×7 , and 9×9 , respectively [25]. Table 1 shows the detailed information on the extracted features.

Table 1. Sentinel-1 remote sensing features.

Types	Features	Descriptions
Texture variables (window: 3×3 , 5×5 , 7×7 , 9×9)	Mean_vv, Mean_vh	Reflects the degree of uniformity in the distribution of pixel values, and texture regularity is positively correlated with the mean value.
	Variance_vv, Variance_vh	Capable of reflecting the degree of change in the gray value of remotely sensed images.
	Entropy_vv, Entropy_vh	Indicates the amount of information contained in the remote sensing image, which can indicate the complexity of the remote sensing image; the larger the value, the higher the level of complexity, and the greater the amount of information.
	Contrast_vv, Contrast_vh	Able to reflect the clarity of remotely sensed images.
	Homogeneity_vv, Homogeneity_vh	The magnitude of its value reflects, to some extent, the magnitude of the local homogeneity of remotely sensed images.
	Dissimilarity_vv, Dissimilarity_vh	Characterize the degree of similarity of texture information in remotely sensed images; the higher the degree of dissimilarity, the stronger the uniqueness of the texture information.
	Correlation_vv, Correlation_vh	Show the similarity of matrix elements in rows and columns on the grayscale covariance matrix.
	Angular second moment_vv, Angular second moment_vh	Indicates the degree of gray scale distribution uniformity and texture coarseness of the image; the larger the angular second-order moments, the clearer the image texture.

Table 1. Cont.

Types	Features	Descriptions
backward scattering coefficient variables	VV	Backward scattering coefficient for VV polarization mode.
	VH	Backward scattering coefficient for VH polarization mode.
	$(VH - VV)/(VH + VV)$	Standard dual-polarization rate
	VV/VH	Ratio of backward scattering coefficients.

2.3.2. Sentinel-2 RS Feature Extraction

- (1) Sentinel-2 band information. The multi-spectral data acquired by the Sentinel-2 satellite contained 13 spectral bands, of which 4 bands have a resolution of 10 m, 6 bands have a resolution of 20 m, and 3 bands have a resolution of 60 m. Ten bands with 10 m and 20 m resolution were selected for this study, and six bands with 20 m resolution were resampled to 10 m resolution using the nearest neighborhood method.
- (2) Sentinel-2 vegetation indices. A total of 15 vegetation indices were selected as independent variables in this study, and the ENVI software was used to extract each vegetation index based on Sentinel-2 imagery [28,29]. The vegetation indices and their calculation formulas are shown in Table 2.

Table 2. Sentinel-2 Vegetation indices features.

Types	Remote Sensing Features	Calculation Formula	Description
Vegetation indices	DVI	$B8 - B4$	Difference Vegetation Index (DVI), which is very sensitive to changes in soil and can be used to monitor the environmental conditions of forest stands.
	WDVI	$B8 - 0.5 \times B4$	Weighted difference vegetation index.
	ARVI	$B8 - (2 \times B4 - B2)/B8 + (2 \times B4 - B2)$	Atmospheric modification of vegetation index.
	NDVI	$(B8 - B4)/(B8 + B4)$	The Normalized Difference Vegetation Index (NDVI), which effectively distinguishes between vegetation and other objects, determines the relationship between vegetation and other objects and monitors the vitality of vegetation.
	GNDVI	$(B7 - B3)/(B7 + B3)$	Green band normalized vegetation indices can be used to identify different rates of chlorophyll concentration, and chlorophyll concentration is strongly correlated with nitrogen content.
	IPVI	$B8/(B8 + B4)$	Infrared percent vegetation index
	MCARI	$[(B5 - B4) - 0.2 \times (B5 - B3)] \times (B5 - B4)$	Modified chlorophyll uptake ratio index.
	MTCI	$(B6 - B5)/(B5 - B4)$	Medium-resolution terrestrial chlorophyll index.
	NDI45	$(B5 - B4)/(B5 + B4)$	Optimized Normalized Vegetation Index (NDVI) with higher linearity and higher saturation compared to the NDVI chlorophyll index.
	PSSRa	$B7/B4$	Red-edge bending position index, which can be used for biomass and nitrogen uptake measurement or management in heterogeneous fields, and reflectance around the red edge are sensitive to changes in chlorophyll content, nitrogen content, vegetation cover, and biomass of crops.
	REIP	$700 + 40 \times [(B4 + B7)/2 - B5]/(B6 - B5)$	Ratio Vegetation Index (RVI), which can reduce the influence of soil background on vegetation indices and improve confidence.
	RVI	$B8/B4$	Ratio Vegetation Index (RVI), which can reduce the influence of soil background on vegetation indices and improve confidence.
	S2REP	$705 + 35 \times [(B4 + B7)/2 - B5]/(B6 - B5)$	Sentinel-2 red edge position index, based on linear interpolation. Sentinel-2 has a key advantage in linear interpolation methods, where the red edge is the inflection point of strong red absorption on near-infrared reflection.

Table 2. Cont.

Types	Remote Sensing Features	Calculation Formula	Description
Vegetation indices	TNDVI	$\frac{[(B8 - B4)/(B8 + B4) + 0.5]}{1/2}$	Transformed Normalized Vegetation Index, an algorithm interpreting the relationship between green biomass found in pixels.
	TSAVI	$\frac{0.5 \times (B8 - 0.5 \times B4 - 0.5)}{(0.5 \times B8 + B4 - 0.15)}$	Converted Soil Regulation Vegetation Index.

2.4. Research Methodology

2.4.1. MLSR

Linear regression analysis is one of the most commonly used methods to construct relationships between forest AGB and RS features. To avoid data redundancy and determine the appropriate RS features, in this study we adopted MLSR as one of the feature optimization inversion models to estimate forest AGB. MLSR is a commonly used feature variable optimization method based mainly on the magnitude of the effect of the independent variables on the regression equation. It alternately introduces and excludes the variables under the set level of the F-test, selects the optimal features, and establishes the corresponding regression equation [30]. The features were selected by choosing RS features with a high correlation at a probability of significance P less than 0.05. The MLSR regression model for the forest AGB estimation used in this study is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_j \quad (4)$$

where y is the forest AGB to be estimated; β_0 is a constant term; $\beta_1, \beta_2, \dots, \beta_i$ is the coefficient; x_1, x_2, \dots, x_j is the remote sensing eigen-factor.

2.4.2. KNN-FIFS

The K-nearest neighbor (KNN) algorithm is a non-parametric method that can be used to solve classification and regression problems and does not depend on a specific function distribution. It is more flexible, explicit, and has a strong generalization ability in forest AGB estimation, but when the RS feature dimensions are high, it may result in low efficiency. To improve the shortcomings of KNN, like inefficient feature combination and poor model estimation, in this study we adopt the KNN-FIFS method proposed by Han et al. to estimate forest AGB using optical RS features using a fast iterative algorithm to find out the best combination of RS features that are the most relevant for the forest AGB estimation [31]. The basic principle of KNN-FIFS is shown in Figure 4:

For the KNN-FIFS method, when the distance metric is certain, the value of k will have a certain impact on the estimation results, and the value of k is affected by the number of sample plots, the geographic location of the sample plots, and the changes in forest AGB. In this study, the range of k is set as 1~11, and the variation range of the RS extraction window is also 1~11.

2.4.3. RF

RF is a machine learning algorithm proposed by Breiman Leo and Adele Cutler in 2001 for classification, regression, and survival analysis [32]. The basic principle is that the sample set with the same sample capacity as the original training set is first extracted using the bootstrap sampling method. Then, the same number of decision trees for each of the K samples are separated, and each of the decision trees predicts each sample to obtain K sets of predicted values. Then K sets of the predicted values are averaged and selected as the predicted value of each record. This algorithm can be understood as calculating the importance of features and ranking them in order of importance [33]. In this paper, the algorithm is implemented in R software.

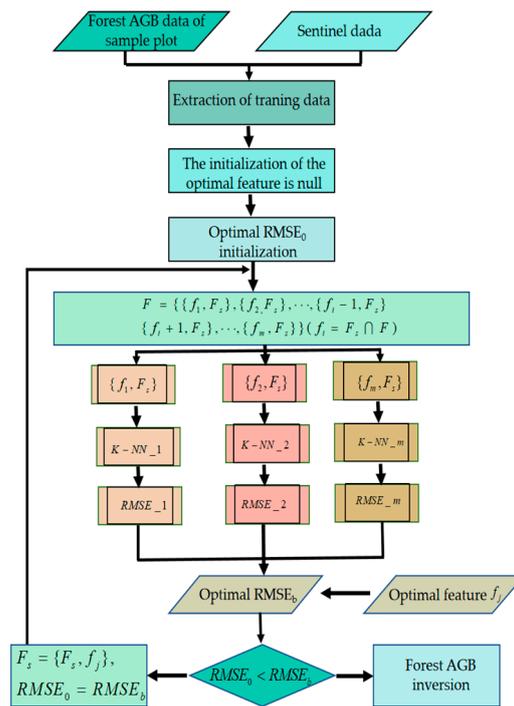


Figure 4. Flow chart of the KNN-FIFS algorithm.

2.4.4. Accuracy Evaluation Method

In this study, leave-one-out cross-validation (LOOCV) was used to evaluate model-inversed results. The method is to take 1 sample from the total sample set sequentially without repetition as validation, and the rest of the samples are used to build the model and carry out the inversion of the forest AGB, and so on until all the sample plots are used as a validation sample at one time, and finally take the mean value for the accuracy evaluation of the inversion results. The validation indices used here include coefficient of determination (R^2 , Equation (5)), root mean square error (RMSE, Equation (6)), relative root mean square error (rRMSE, Equation (7)), and mean absolute percentage error (MAPE, Equation (8)). The value of R^2 ranges from 0 to 1, and the closer it is to 1, the higher the accuracy of the inversion results, and vice versa, the lower the accuracy [34]. RMSE and rRMSE indicate the difference between the estimated and measured values; the smaller values indicate higher inversion accuracy.

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

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - y_i)^2}{n}} \tag{6}$$

$$rRMSE = \frac{RMSR}{\bar{y}} \times 100\% \tag{7}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - y_i}{Y_i} \right| \times 100\% \tag{8}$$

where Y_i is the measured value of forest AGB in the sample site, y_i is the inverted value of forest AGB in the sample site, n is the total number of sample sites, and \bar{y} is the average value of measured forest AGB.

3. Results and Analysis

3.1. Inversion Modeling of Forest AGB

RS features and ground-truth sample plots under three combinations of single Sentinel-1, single Sentinel-2, and joint Sentinel-1 and Sentinel-2 at Puer and Genhe were inputted into the MLSR, KNN-FIFS, and RF models as the independent variables and dependent variables, respectively, to estimate forest AGBs. The inversion results of the models were validated by four statistical parameters, and the estimation accuracies of the models were compared at two different test sites.

3.1.1. Construction of the MLSR

When the multi-source RS data are three combinations of single Sentinel-1, single Sentinel-2, and joint Sentinel-1&Sentinel-2, different types of RS features from different regions are inputted into the MLSR model, and optimizing is performed to derive the following features and form the optimal model. The optimized RS feature and related constructed models are shown in Table 3. The MTCI parameter was selected in both test sites and suggested that this parameter is more relevant and sensitive to forest AGB. This is because the red and red-edge bands were strongly absorbed by vegetation and are highly sensitive to the health status and physicochemical parameters of vegetation. In the other feature optimized in the test site of Genhe, the *vh_me3* parameter was selected both in the single Sentinel-1 and combination of Sentinel-1 and Sentinel-2 data. The results suggested that the mean value matches the forest-reflecting pattern within the image element. It is because the mean value reflects the pixel texture pattern, and larger values indicate that the texture pattern is more obvious and easier to describe.

Table 3. MLSR model equations and feature optimization results.

	Data Source	Model	Parameter
Puer	Sentinel-1	$y = 362.954 - 70.900x_1 - 213.408x_2$	en_vv7, JHL
	Sentinel-2	$y = 81.461 + 35.335x_1 + 0.035x_2$	MTCI, DVI
	Sentinel-1&Sentinel-2	$y = 161.639 + 763.706x_1 - 0.641x_2$	sm_vv9, B2
Genhe	Sentinel-1	$y = 2591.421 - 48.832x_1 - 39.001x_2$	vh_me3, vv_cor3
	Sentinel-2	$y = -121.999 + 80.406x_1$	MTCI
	Sentinel-1&Sentinel-2	$y = 1735.196 - 34.738x_1 + 15.221x_2 - 47.440x_3$	vh_me3, S2REP, vh_en9

3.1.2. Construction of KNN-FIFS

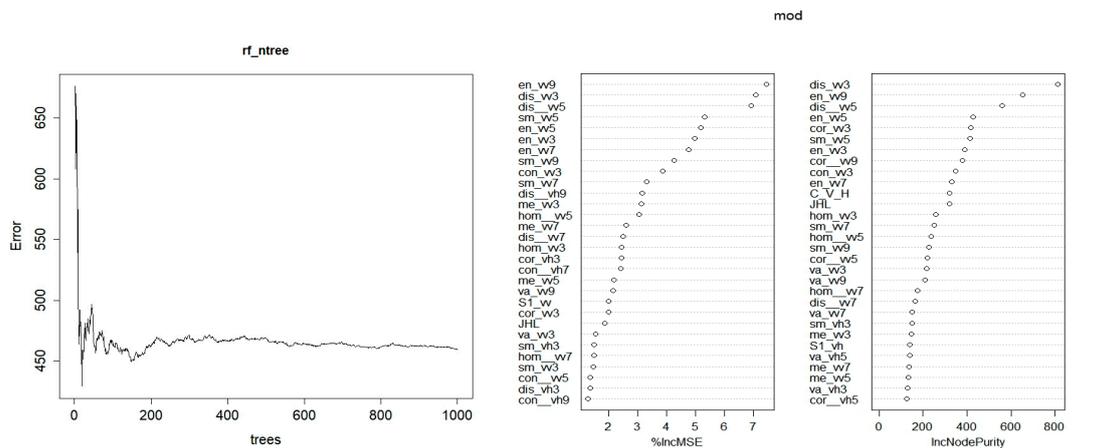
The optimal model parameters for the KNN-FIFS algorithm are shown in Table 4. In Puer, B6 and S2REP parameters were selected from a single Sentinel-2 and combinations of Sentinel-1 and Sentinel-2. The results indicated that the red-edge bands and short-wave infrared bands of Sentinel-2 are more sensitive to forest AGB and more significant than other RS features in forest AGB estimation. Among the many vegetation factors, the factors involved in the calculation of the red-edge band have the strongest correlation with the forest AGBs. The reason is that they can effectively relieve the signal saturation problem corresponding to the high biomass value and the dense canopy. In addition, the selected features in the Puer and Genhe are mostly divided into features with windows of 7 and 9, indicating that the size of the window has a significant impact on the correlations between RS features and forest AGB.

Table 4. KNN-FIFS model parameters and feature optimization results.

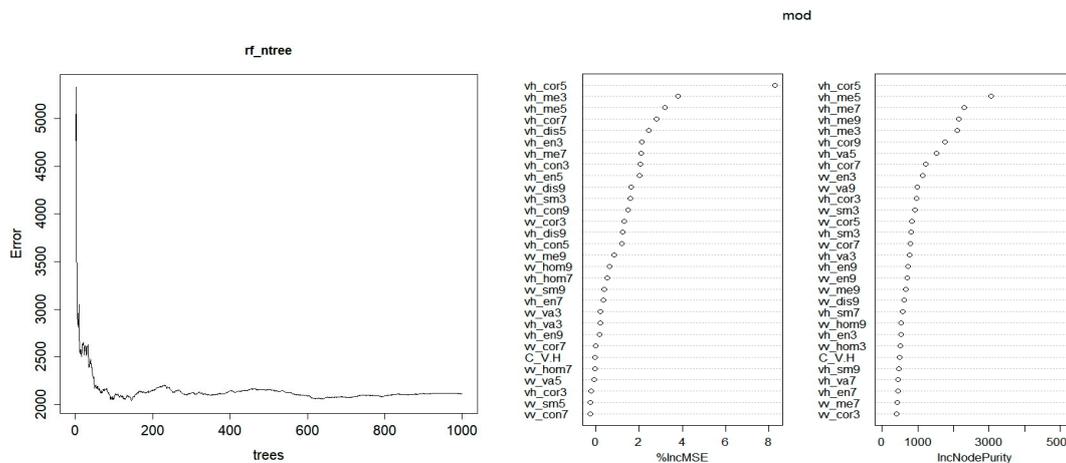
Test Site	Data Source	K	Size of the Window	Feature Selection
Puer	Sentinel-1	2	9 × 9	sm_vv3, cor_vv3, JHL, sm_vv9, hom_vh5, con_vv9, cor_vv7, en_vv9
	Sentinel-2	2	1 × 1	B6, S2REP, NDVI, IPVI
	Sentinel-1&Sentinel-2	1	1 × 1	B6, S2REP, B5, va_vv9, va_vv7
Genhe	Sentinel-1	3	3 × 3	con7_vv, va5_vh
	Sentinel-2	5	5 × 5	IPVI, B4, TNDVI, TSAVI
	Sentinel-1&Sentinel-2	1	5 × 5	sm7_vh, RVI, va7_vh, cor9_vv, sm5_vh, vh, dis9_vv

3.1.3. Construction of the RF

There are two important parameters, *mtry* and *ntree*, in the RF. *ntree* is the number of decision trees, *mtry* is the number of random features, and its size is usually set to one-third of the total number of variables by default in regression problems [35]. In order to determine the size of the parameter *ntree*, the plot function in the RF package is used to draw the trend graph of the regression error with the change of the number of decision trees, and it can be seen from Figure 5 that the regression error tends to be stable when the number of decision trees reaches 1000. In order to ensure the credibility of the results and the efficiency of the operation, the value of 1000 for *ntree* was selected.

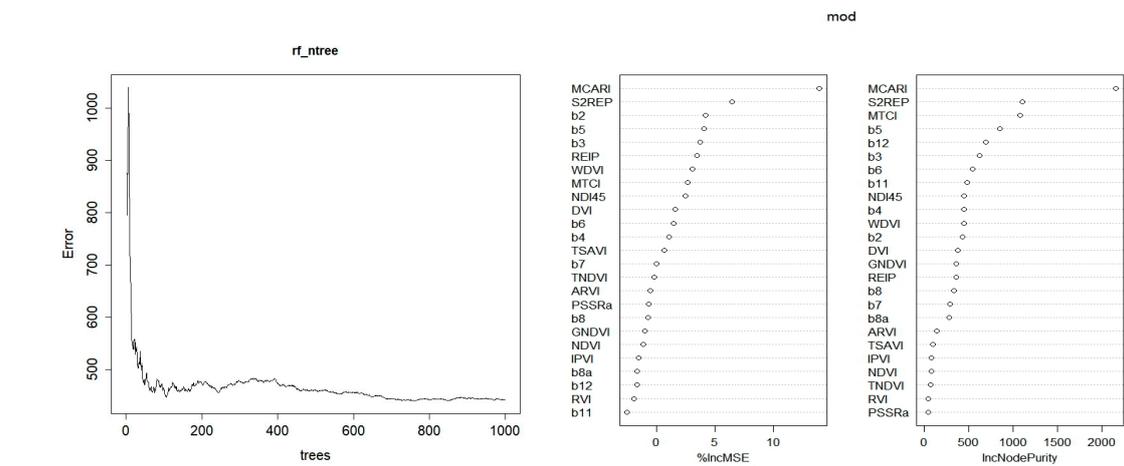


(a) Puer Sentinel-1

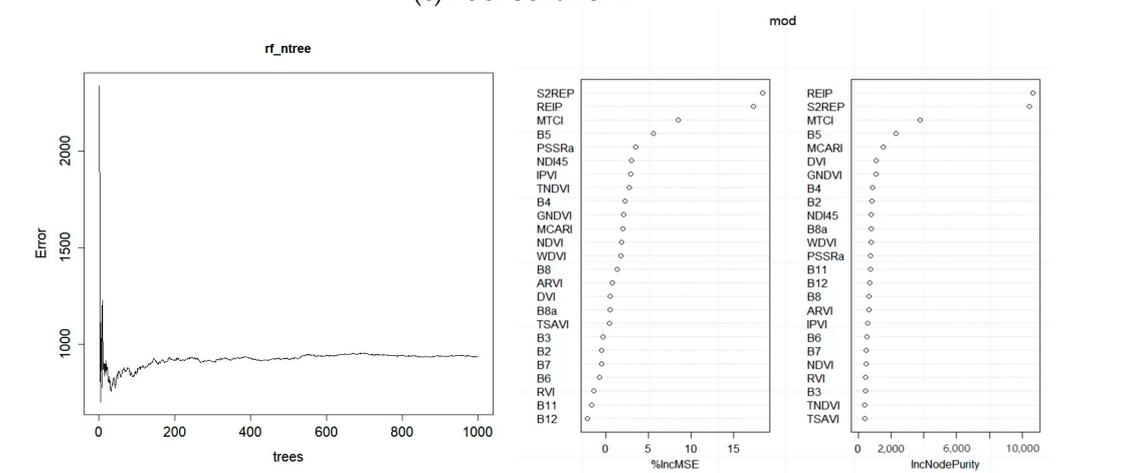


(b) Genhe Sentinel-1

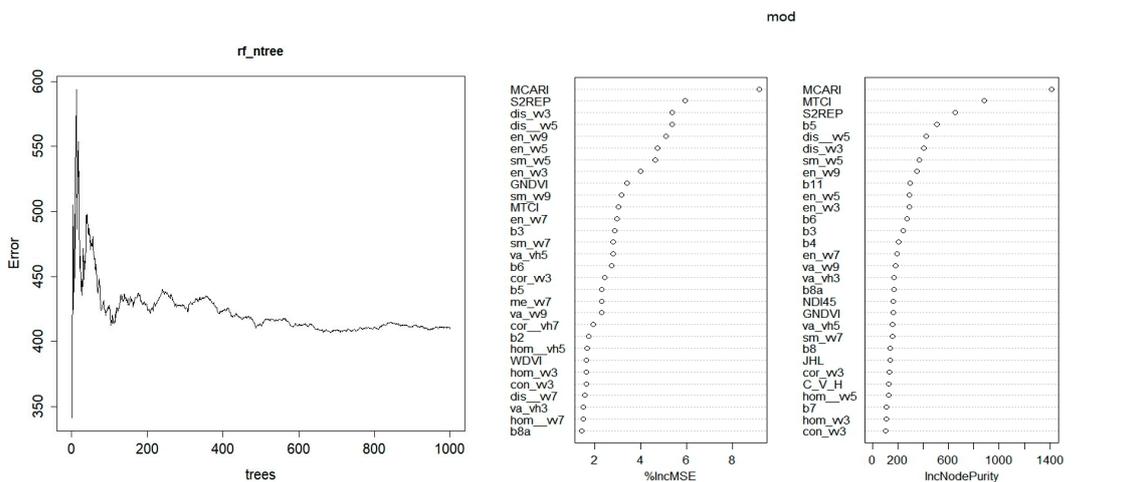
Figure 5. Cont.



(c) Puer Sentinel-2

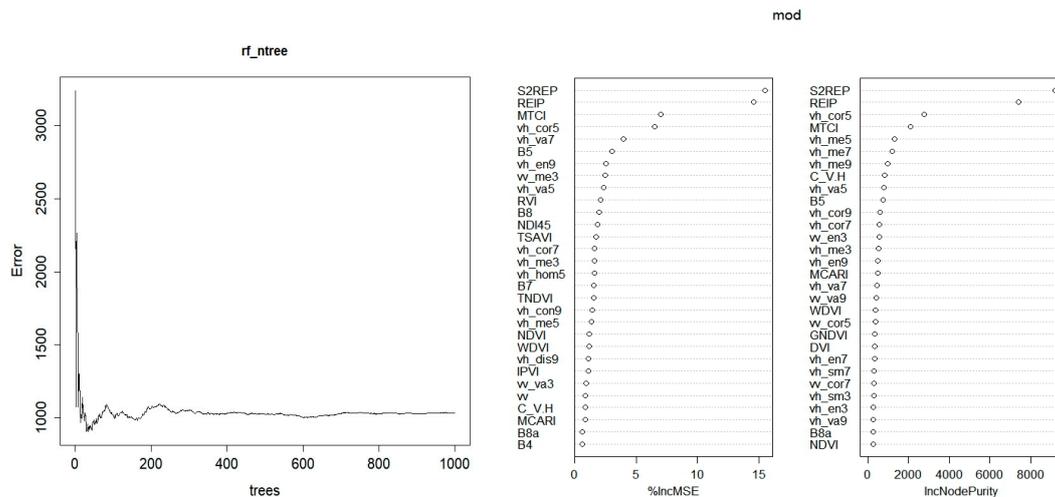


(d) Genhe Sentinel-2



(e) Puer Sentinel-1&Sentinel-2

Figure 5. Cont.



(f) Genhe Sentinel-1&Sentinel-2

Figure 5. Plot of RF error trend with the number of decision trees and importance ranking of remote sensing features.

RS features are determined by two metrics in the RF parameters, named the incremental model MSE (%IncMSE) and the node purity of the model tree (IncNodePurity). When the independent variable is used as out-of-bag data, the larger the parameter values of these two metrics, the more important the feature. The importance of the features was calculated by the important function in the RF package, and each RS feature was ranked in descending order according to the model MSE increment. According to the ordering of RF feature importance, variables with variable ordering in the top 10 were selected for inversion model training and validation [31,36]. The importance ranks of the RS variables used in the two test sites are shown in Figure 5.

The results of feature optimization for RF algorithms are summarized in Table 5. For using a single Sentinel-1 data source, entropy, non-similarity, and contrast are selected as the optimized features. Among them, entropy indicates the degree of complexity of the reflected RS characterization; the bigger the entropy value, the higher the complexity, and the larger the amount of information it contains. The non-similarity reflects the fact that the greater the gray scale difference within the image element, the clearer the visual effect of the image is. The results indicated that the vegetation cover portion of the image element has a high degree of fractional anisotropy, which can better improve the model inversion accuracy. As for the single Sentinel-2 data, S2REP, B5, and NDI45 were selected to participate in the modeling of the combination of features at the two test sites. Red-edge band information, including S2REP, B5, and NDI45, responds to the small changes in the structure of vegetation canopies. For the combination of Sentinel-1 and Sentinel-2, entropy and S2REP were selected as optimized RS features in both of the two test sites. From the perspective of window size, the texture features extracted at window sizes of 9×9 and 3×3 were ranked relatively high in the single data and the combination of Sentinel-1 and Sentinel-2. The results indicated that the small window texture has a strong ability to explain the changes of the forest AGBs and has certain advantages in their inversion.

3.2. Forest AGB Inversion Results

Tables 6 and 7 show the results of the three forest AGB estimation models in the two test sites, and the results of the estimation accuracy are as follows: In subtropical coniferous forests in Puer, the KNN-FIFS model for single Sentinel-1 data performed better than the MLSR and RF models. The R^2 differences between KNN-FIFS and MLSR are 0.10, the value between KNN-FIFS and RF is 0.40, and the differences in RMSE values are 2.44 t/hm^2 and 5.25 t/hm^2 , respectively. In the cold-temperate coniferous forests at Genhe, the KNN-FIFS model performed best for single Sentinel-1 data. However, here RF performed better than

the MLSR model, with R^2 differences of 0.09 and 0.20 and RMSE differences of 2.84 t/hm² and 1.61 t/hm², respectively. In subtropical coniferous forests in Puer, the KNN-FIFS model with single Sentinel-2 data performed better than the MLSR; the worst performance was RF, with a difference in R^2 of 0.14 and 0.23 and a difference in RMSE of 3.01 t/hm² and 3.12 t/hm². In Genhe, the KNN-FIFS model outperformed the MLSR, while the MLSR outperformed the RF with a difference in R^2 of 0.07 and 0.20 and a difference in RMSE of 2.79 t/hm² and 4.65 t/hm². In both test sites, the inversion accuracies of the three models using a combination of Sentinel-1 and Sentinel-2 are higher than using single data sources. A higher inversion accuracy was obtained in Puer. The inversion accuracy of the three models in the two test sites had small differences, with R^2 differences of 0.11 and 0.22 and RMSE differences of 2.0 t/hm² and 19.11 t/hm² in Puer and R^2 differences of 0.06 and 0.24 and RMSE differences of 1.14 t/hm² and 3.69 t/hm² in Genhe.

Table 5. Optimal feature selection results under the RF algorithm.

Test Site	Data Source	Major Parameter
Puer	Sentinel-1	en_vv9, dis_vv3, dis_vv5, sm_vv5 en_vv5, en_vv3, en_vv7, sm_vv9 con_vv3, sm_vv7
	Sentinel-2	MCARI, S2REP, B2, B5, B3, REIP, WDV, MT, NDI45, DVI
	Sentinel-1&Sentinel-2	MCARI, S2REP, dis_vv3, dis_vv5 en_vv9, en_vv5, sm_vv5, en_vv3 GNDVI, sm_vv9
Genhe	Sentinel-1	cor_vh5, me_vh3, me_vh5, cor_vh7, dis_vh5, en_vh3, me_vh7, con_vh3, en_vh5, dis_vv9
	Sentinel-2	S2REP, REIP, MT, B5, PSSRa, NDI45, IPVI, TNDVI, B4, GNDVI
	Sentinel-1&Sentinel-2	S2REP, REIP, MT, vh_cor5, vh_va7, B5, vh_en9, vv_me3, vh_va5, RVI

Table 6. The accuracy indicators of forest AGB inversion using three models in Puer.

Test Site	Inversion Model	Data Source	Fitting Result	
			R^2	RMSE (t/hm ²)
Puer	MLSR	Sentinel-1	0.414	17.78
		Sentinel-2	0.345	18.79
		Sentinel-1&Sentinel-2	0.461	17.04
	KNN-FIFS	Sentinel-1	0.512	15.34
		Sentinel-2	0.480	15.78
		Sentinel-1&Sentinel-2	0.568	15.04
	RF	Sentinel-1	0.116	20.59
		Sentinel-2	0.255	18.90
		Sentinel-1&Sentinel-2	0.348	18.06

Table 7. The accuracy indicators of forest AGB inversion using three models in Genhe.

Test Site	Inversion Model	Data Source	Fitting Result	
			R^2	RMSE (t/hm ²)
Genhe	MLSR	Sentinel-1	0.380	35.38
		Sentinel-2	0.320	36.20
		Sentinel-1&Sentinel-2	0.446	33.43
	KNN-FIFS	Sentinel-1	0.470	32.54
		Sentinel-2	0.390	33.41
		Sentinel-1&Sentinel-2	0.511	32.29
	RF	Sentinel-1	0.274	34.15
		Sentinel-2	0.188	38.06
		Sentinel-1&Sentinel-2	0.345	35.98

The analysis of the above inversion results shows that MLSR, RF, and KNN-FIFS models constructed by combining Sentinel-1 and Sentinel-2 data sources have achieved the highest accuracy in forest AGB inversion, which indicates that the combination of SAR data and multi-spectral RS data can play its advantages in characterizing different levels of forest structure, thus improving the accuracy of forest AGB inversion and its accuracy in the two test sites. The three models can effectively remove RS features with poor sensitivity and low correlation with the forest AGB by performing feature optimization with different algorithms during the establishment of the model, which can effectively improve forest AGB inversion accuracy. KNN-FIFS had the most robust forest AGB estimation among the three feature-optimizing methods and was also the most effective method. Under three different RS data sources, the MLSR and the KNN-FIFS models showed similar performance both in subtropical Simao Pine forests and cold-temperate Xing'an Larch forests. While MLSR performed slightly better than RF, the overall inversion accuracy was higher in subtropical coniferous forests than in cold-temperate coniferous forests.

From the scatter plot (Figure 6) analysis, it seems that the MLSR, RF, and KNN-FIFS models show different degrees of overestimation of low values and underestimation of high values under the three combinations of the two single data sources. When a single Sentinel-1 data source is used in Puer, the inversion results of the model show the phenomena of low-value overestimation and high-value underestimation, while when the two data sources are combined, the inversion results are closer to the 1:1 line, and the phenomena of high-value underestimation and low-value overestimation are not obvious. This may result from the fact that the penetration ability of Sentinel-1 SAR data are better than that of multi-spectral data on forests, and it can represent the canopy information of forests in a vertical structure in a more visual way, while the spectral signals of multi-spectral images can only reflect the canopy information of forests in a horizontal direction. The main reason is that Sentinel-1 SAR data can penetrate the forest better than multi-spectral data and can more intuitively represent the information on the vertical structure of the forest canopy. When a single Sentinel-2 data source is used, then the overestimation phenomenon is more obvious in the RF model, while the overestimation phenomenon is significantly reduced in the KNN-FIFS model, probably because KNN-FIFS can make better use of the correlation and select the feature factor that has a higher correlation with the forest AGB. When the two data sources are combined, the correlation between measured and predicted values performs well with the three models. Probably because there is a complementary nature to the data sources when multi-spectral and SAR data are combined, and they can better represent both the horizontal and vertical structure of the forest.

Analysis of the scatterplot (Figure 7) shows that the three models exhibit different degrees of overestimation of low values and underestimation of high values for the three combinations of the two data sources. Using single Sentinel-1 data source in the MLSR, RF, and KNN-FIFS models shows a more discrete state. The results from the RF model (Figure 7a,d,g) show a more obvious overestimation of the low values and an underestimation of the high values, and there is a certain degree of saturation phenomenon, which may be due to the fact that the Sentinel-1 C-band SAR features are more sensitive to the changes in scattering from the forest canopy. Forest AGB estimated results showed a more obvious saturation at around 85 t/hm² using single Sentinel-2 data (Figure 7e) and the RF algorithm. When combined Sentinel-1 and Sentinel-2 data sources were used, the overall fit was higher than the results when only a single data source was used, especially in the inversion of the KNN-FIFS model. The best performance of KNN-FIFS may result from the fact that the estimated parameters are only related to the neighboring K samples in the KNN-FIFS fitting method, and the weighted summation of the number of measured sample points reduces the problem of the imbalance of the sample points [37].

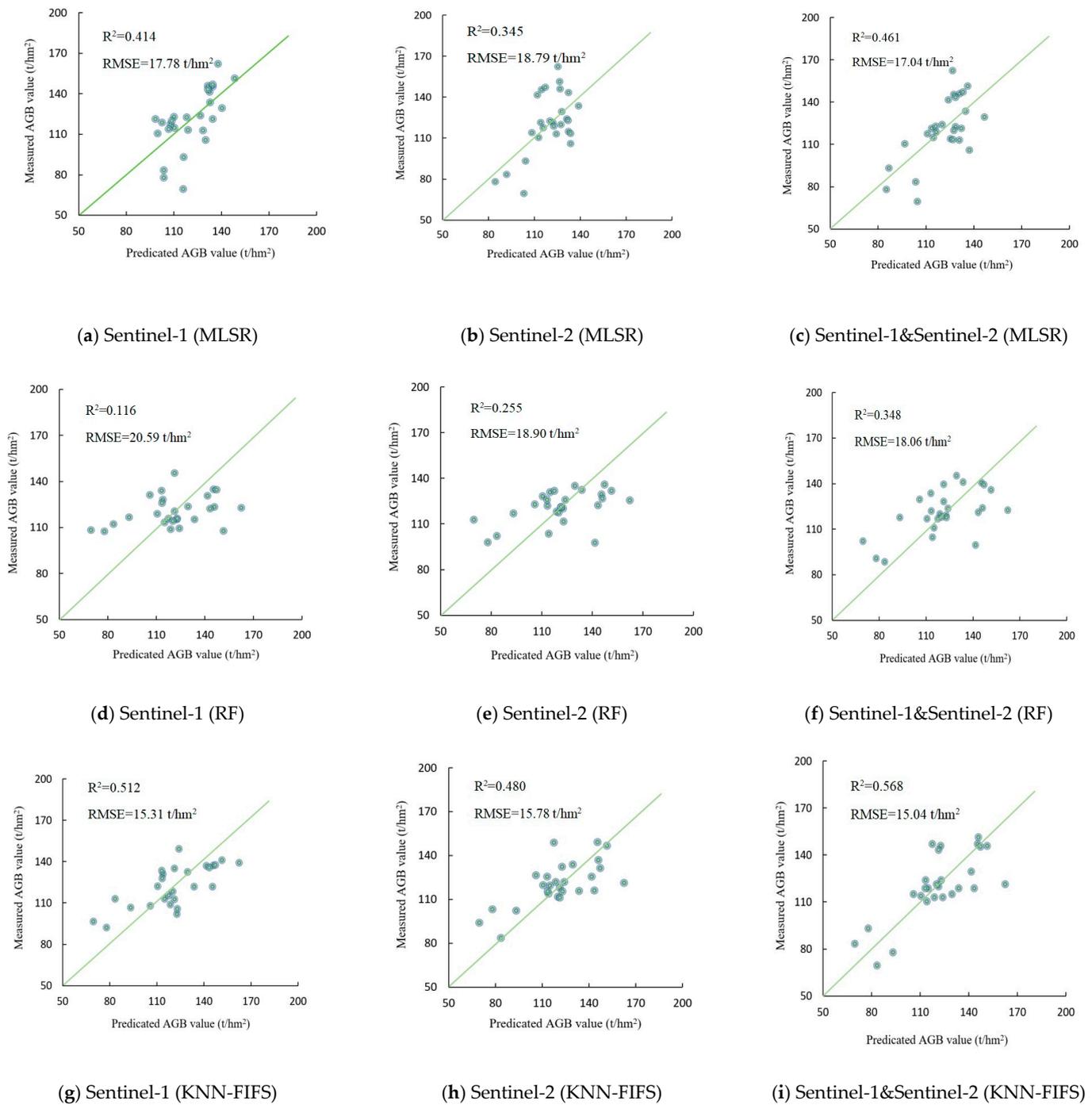


Figure 6. Forest AGB cross-validation results of MLSR, RF, and KNN-FIFS are inverted in Puer.

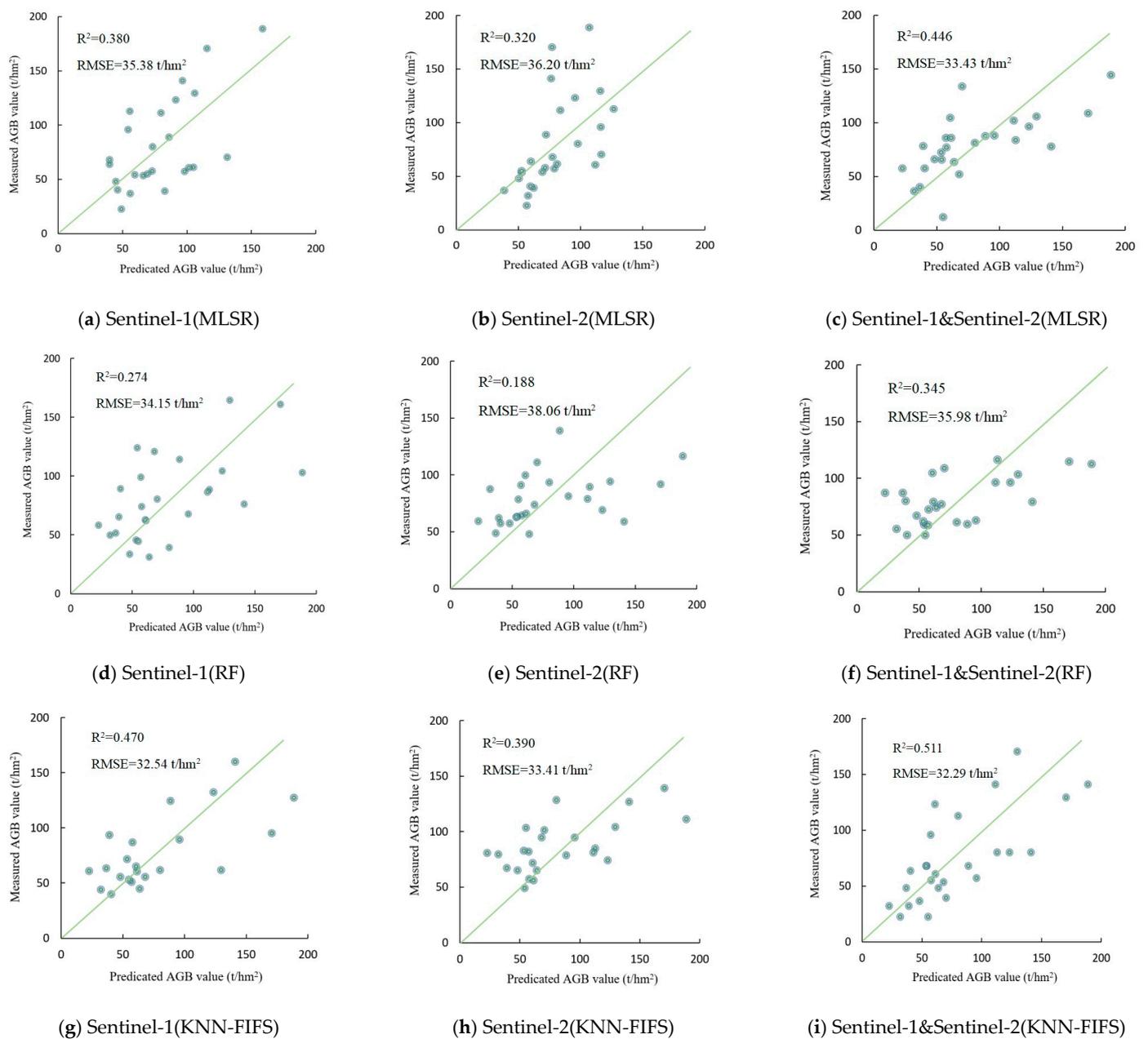


Figure 7. Forest AGB cross-validation results of MLSR, RF, and KNN-FIFS are inverted in Genhe.

3.3. Comparative Analysis of Three Data Sources with Different Features Optimization Algorithms

Using the four indicators of R^2 , RMSE, rRMSE, and MAPE, the performance of the MLSR, KNN-FIFS, and RF models was compared and analyzed in Figure 8. According to Figure 8, we found that under the data combination forms in both study areas, the R^2 values for the 3 models range from 0.116 to 0.568. Among them, the R value for KNN-FIFS is the highest, and the R value for RF is the lowest. In Puer, under the combination of Sentinel-1 and Sentinel-2, the RMSE from KNN-FIFS is the lowest (15.04 t/hm^2). The estimation results from KNN-FIFS have the lowest RMSE values, and the estimation results from RF have the highest RMSE values. But in the case of using single Sentinel-1 data in Genhe, the RMSE value from MLSR is the highest (35.38 t/hm^2). The values of rRMSE also confirmed the best performance of KNN-FIFS, with the lowest value of 12.4%. The analysis from MAPE showed the lowest value obtained from KNN-FIFS and the highest accuracy in Puer. However, the performance of KNN-FIFS in Genhe is not stable, with large errors, although the error of MLSR is relatively small. Overall, in two study areas and three data

combinations, KNN-FIFS generally showed the highest accuracy, indicating the robustness of feature optimization in this algorithm.

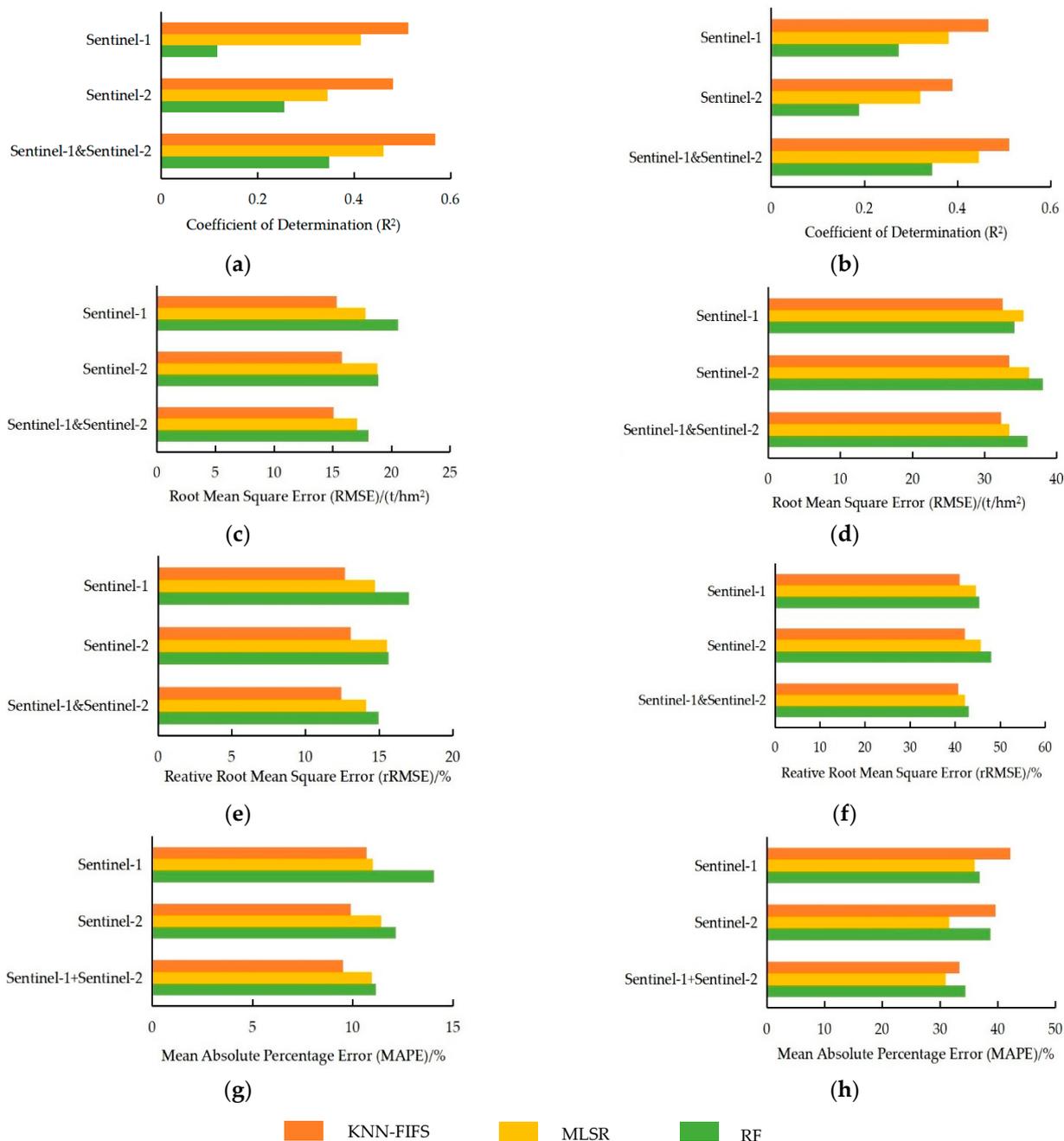


Figure 8. Accuracy evaluation and comparison. (a,c,e,g) represent the Puer test site; (b,d,f,h) represent the Genhe test site.

4. Discussion

(1) Estimation of forest AGB with single Sentinel-1 and Sentinel-2 data under different feature optimizing methods.

Single Sentinel-1 and single Sentinel-2 data have different estimated potentials in coniferous forest AGB estimation under different feature-optimizing approaches. The R^2 of the inversion results of the MLSR, RF, and KNN-FIFS models using single Sentinel-1 data were 0.414, 0.116, and 0.512 in Puer, whereas those of the inversion results from Genhe were 0.380, 0.274, and 0.470. Pan et al. used Sentinel-1 as the RS data source and took the

fir forest in Jiangle County as the test sample, and the highest precision R^2 of the inversion results was 0.636, which was slightly higher than the results of this study. Probably due to differences in texture factors across different study areas, the accuracy of this study area is slightly lower [15]. In the study of Shi et al., GF-1 and GF-3 were selected as the data sources for the inversion study of the total biomass of forest AGB and component AGB, in which the results of the Yunnan pine forest inversion study of the total AGB of a single GF-3 data were slightly lower than those of Puer in this study and, at the same time, slightly higher than those of Genhe in this study, with an R^2 of 0.500. The better performance may result from the quad-polarization information from GF-3. Quad-polarization information can better characterize forest structure than the dual-polarization sentinel data in this study [38]. In addition, Xu et al. used Sentinel-1, Landsat-8 OLI, and continuous forest inventory data as the main data sources and took subtropical evergreen broadleaf forests in Guidong to explore the performance of forest AGB inversion using single RS data sources. In this study, the R^2 for the inversion result using single Sentinel-1 data were 0.49, which was slightly lower than our results. The reason may be that SAR data has stronger penetration capability in coniferous forests than in broadleaf forests [39].

The R^2 of the inversion results of the MLSR, RF, and KNN-FIFS models using single Sentinel-2 data in Puer were 0.345, 0.255, and 0.480, while those values of the inversion results in Genhe were 0.320, 0.188, and 0.390. Wang et al. selected Sentinel-2, Landsat-8 OLI, and GF-1 as data sources to estimate forest AGB and found Sentinel-2 data had the best results with an $R^2 = 0.60$ and an RMSE = 21.40 t/hm². The results also indicated that the red-edge band has a strong sensitivity to forest AGB and has great potential for quantitative inversion studies [40]. Guo et al. used Sentinel-2 data as a data source and constructed an MLSR model to invert coniferous forest AGB in Inner Mongolia. The R^2 of the inversion results was 0.765, and the RMSE was 39.49 t/hm² [16]. The results are slightly higher than those in our study. It may be due to the influence of topographic factors, which affect forest diversity to a large extent, distribution, and local heterogeneity, thus directly affecting forest productivity and structure. Lopez-Serrano et al. used Landsat-8 to study temperate coniferous forests in Mexico. Band indices, vegetation indices, texture indices, and topographic and climatic variables were extracted and used for forest AGB estimation through the RF and support vector regression (SVR) algorithms. They found that SVR performed better in coniferous forest AGB inversion with $R^2 = 0.8$ and RMSE = 8.20 t/hm². The results revealed that considering climatic variables helps to improve the inversion accuracy of forest AGB with RS features since climatic characteristics are one of the most important factors affecting the distribution and growth of forest species [41].

- (2) Estimation of forest AGB for a combination of Sentinel-1 and Sentinel-2 data under different feature optimizing methods.

The R^2 of the inversion results of the MLSR, RF, and KNN-FIFS models using a combination of Sentinel-1 and Sentinel-2 data in Puer were 0.461, 0.348, and 0.568, while those values in Genhe were 0.446, 0.345, and 0.511. Pan et al. used Sentinel-1 and Sentinel-2 to estimate coniferous forest AGB in subtropical regions, and the inversion result of R^2 was 0.575, which was similar to that of the subtropical coniferous forests in Puer in this study. However, the values were slightly higher than those obtained in the cold-temperate coniferous forests in Genhe in our study [14]. In the study of Liu et al., the combination of Sentinel-1 and Sentinel-2 was used as the data source for forest AGB estimation as well, but in this study, the effects of topographic factors on forest AGB were considered. In this study, with the comparison of the performance of the MLSR, BP neural network, and SVR models, the best performance for forest AGB estimation was acquired by the BP neural network model with an R^2 of 0.821 and an RMSE of 32.39 t/hm². The results of the study are higher than the results of our study. The results of the study showed the potential of a combination of optical data and SAR data for improving the inversion accuracy of forest AGB [42]. Forkuor et al. used Sentinel-1, Sentinel-2, and RF for forest AGB estimation in West African dryland forests and showed that the joint Sentinel-1 and Sentinel data could be better for inversion of forest AGB with $R^2 = 0.90$ and RMSE = 54.5 t/hm² [43].

(3) The advantages of Forest AGB estimation using the KNN-FIFS algorithm.

KNN-FIFS shows the most robustness in forest AGB inversion in both test sites; moreover, KNN-FIFS achieved superior inversion results in forest AGB inversion both using different data sources and at different test sites. Shi et al. used the KNN-FIFS model and GF-3 data for forest AGB estimation in Yiliang and also confirmed its better performance with $R^2 = 0.500$ and $RMSE = 14.11 \text{ t/hm}^2$, which is slightly lower than the inversion results of our study accuracy with $R^2 = 0.512$ and $RMSE = 15.34 \text{ t/hm}^2$ [38]. Han et al. also used KNN-FIFS, Landsat-8 OLI, and the airborne SAR P-band for forest AGB inversion. The results from them were slightly higher than the results of our study, with $R^2 = 0.770$ and $RMSE = 22.74 \text{ t/hm}^2$. The higher R^2 value and lower RMSE value may result from the involved P-band SAR features, which have better penetration capability in forests and can interact with the large branches and trunks that account for the main AGB in forests [30]. Ji et al. used ALOS-1 PALSAR-1 and ALOS-2 PALSAR as data sources and KNN-FIFS as an inversion model to estimate forest AGB in Genhe. The inversion accuracy of KNN-FIFS was outperformed by RF and KNN, and its R^2 was 0.37, which was lower than the value obtained in this study. The lower R^2 values in their study may be due to the incomplete coverage of RS images in the study area [44]. The study confirmed the robustness of KNN-FIFS and the best performance of combined Sentinel-1 and Sentinel-2 in forest AGB estimations.

5. Conclusions

In this study, we investigated the potential of a feature optimization inversion model for forest AGB estimation using Sentinel-1 SAR data, Sentinel-2 multispectral data, and a combination of them. Larch pure forests in northeastern China and Simao pine pure forests in southwestern China are investigated in this study. Through this study, we concluded that: (1) Combining Sentinel-1 and Sentinel-2 data allows a certain degree of information complementarity and higher estimation accuracy than only using Sentinel-1 and Sentinel-2 alone. (2) Comparative analyses of the two test sites in two typical coniferous forests with three combinations of two data sources showed that tree species or AGB levels may result in the effect of forest AGB estimation accuracy. (3) All of MLR, RF, and KNN-FIFS were subjected to feature optimization during the forest AGB modeling; among them, KNN-FIFS showed the best suitability and promotability in different forest scenes and different test sites. However, due to the limitations of field data collection, only 26 larch pure forest sample plots and 27 Simao pine pure forest sample plots were used for modeling analysis in this study. Although the results are convincing, more different observations, such as LiDAR data, UAV data, and L- and P-band SAR data with longer wavelengths, are needed to be explored and validated in the future [7,45,46]. In addition, only one forest type, like pure coniferous forest, was used in this study, and other different forest types, such as broadleaf forest and mixed coniferous forest, need to be further studied in the future [7,11,45].

Author Contributions: Conceptualization, Y.J.; methodology, F.Z.; project administration, Y.J.; resources, Y.J.; software, F.Z. and Y.J.; supervision, Y.J. and W.Z.; validation, F.Z.; writing—original draft preparation, F.Z. and Y.J.; writing—review and editing, Y.J., A.M. and W.Z.; visualization, F.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China with grant numbers of 32160365, 32371869, and 42161059 and the Yunnan Province agriculture joint special project with grant numbers of 202301BD070001-058.

Data Availability Statement: Remote sensing image data for this study were obtained from the European Space Agency (ESA, <http://www.esa.int/ESA>). Anyone needing to use the ground data in this paper should contact the corresponding authors.

Acknowledgments: The authors would like to thank ESA for providing remote sensing image data. The authors are grateful to Zongtao Han of Weihai Wuzhou Navi-Tech for providing the KNNFIFS model.

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

References

- Zhu, J.; Huang, Z.; Sun, H.; Wang, G. Mapping forest ecosystem biomass density for Xiangjiang River Basin by combining plot and remote sensing data and comparing spatial extrapolation methods. *Remote Sens.* **2017**, *9*, 241. [[CrossRef](#)]
- Houghton, R.A. Above-ground forest biomass and the global carbon balance. *Glob. Chang. Biol.* **2010**, *11*, 945–958. [[CrossRef](#)]
- Wang, M.; Zhang, W.; Ji, Y.; Marino, A.; Xu, K.; Zhao, L.; Shi, J.; Zhao, H. Aboveground Biomass Retrieval in Tropical and Boreal Forests Using L-Band Airborne Polarimetric Observations. *Forests* **2023**, *14*, 887. [[CrossRef](#)]
- Gao, Y.L.; Wu, D.S.; Fang, L.M. Estimation of forest aboveground biomass in Linhai based on sentinel data. *Zhejiang For. Sci. Technol.* **2022**, *42*, 40–49. [[CrossRef](#)]
- Zhao, P.P.; Lu, D.S.; Wang, G.G.; Wu, C.P.; Huang, Y.J.; Yu, S.Q. Examining spectral reflectance saturation in Landsat imagery and corresponding solutions to improve forest aboveground biomass estimation. *Remote Sens.* **2016**, *8*, 469. [[CrossRef](#)]
- Liu, Q.; Yang, L.; Liu, Q.H.; Li, J. Review of forest above-ground biomass inversion methods based on remote sensing technology. *J. Remote Sens.* **2015**, *19*, 62–74. [[CrossRef](#)]
- Tian, X.M.; Zhang, X.L. Estimation of forest aboveground biomass by remote sensing. *J. Beijing For. Univ.* **2021**, *43*, 137–148. [[CrossRef](#)]
- Hosseini, M.; McNairn, H.; Mitchell, S.; Robertson, L.D.; Davidson, A.; Ahmadian, N.; Bhattacharya, A.; Borg, E.; Conrad, C.; Dabrowska-Zielinska, K.; et al. A Comparison between Support Vector Machine and Water Cloud Model for Estimating Crop Leaf Area Index. *Remote Sens.* **2021**, *13*, 1348. [[CrossRef](#)]
- Zhao, P.P.; Lu, D.S.; Wang, G.X.; Liu, L.J.; Li, D.Q.; Zhu, J.R.; Yu, S.Q. Forest aboveground biomass estimation in Zhejiang Province using the integration of Landsat TM and ALOS PALSAR data. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *53*, 1–15. [[CrossRef](#)]
- Shao, Z.F.; Zhang, L.J.; Wang, L. Stacked sparse autoencoder modeling using the synergy of airborne LiDAR and satellite optical and SAR data to map forest above-ground biomass. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 5569–5582. [[CrossRef](#)]
- Navarro, J.A.; Algeet, N.; Fernández-Landa, A.; Esteban, J.; Guillen-Climent, M.L. Integration of UAV, Sentinel-1 and Sentinel-2 Data for Mangrove Plantations Aboveground Biomass Monitoring in Senegal. *Remote Sens.* **2019**, *11*, 77. [[CrossRef](#)]
- Shao, Z.F.; Zhang, L.J. Estimating Forest Aboveground Biomass by Combining Optical and SAR Data: A Case Study in Genhe, Inner Mongolia, China. *Sensors* **2016**, *16*, 834. [[CrossRef](#)] [[PubMed](#)]
- Li, M.Z.; Yu, X.T.; Gao, Y.K.; Fan, W.Y. Remote sensing quantification on forest biomass based on SAR polarization decomposition and Landsat data. *J. Beijing For. Univ.* **2018**, *40*, 1–10. [[CrossRef](#)]
- Pan, J.L.; Xin, Y.Q.; Huang, J.P.; Wang, X.Y. Estimation of forest above-ground biomass based on GF-3 PolSAR data and Landsat-8 OLI data. *J. Cent. South Univ. For. Technol.* **2020**, *40*, 83–90. [[CrossRef](#)]
- Pan, L.; Sun, Y.J. Estimation of cunninghamia lanceolate forest biomass based on Sentinel-1 image texture information. *J. Northeast. For. Univ.* **2018**, *16*, 58–62. [[CrossRef](#)]
- Guo, Z.Q.; Zhang, X.L.; Wang, Y.T. Ability evaluation of coniferous forest aboveground biomass inversion using Sentinel-2A multiple characteristic variables. *J. Beijing For. Univ.* **2020**, *42*, 27–38. [[CrossRef](#)]
- 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. Interdiscip. J.* **2022**, *282*, 113232. [[CrossRef](#)]
- Cheng, S.; Xu, Z.; Su, Y.; Zhen, L. Spatial and Temporal Flows of China's Forest Resources: Development of a Framework for Evaluating Resource Efficiency. *Ecol. Econ.* **2010**, *69*, 1405–1415. [[CrossRef](#)]
- Li, W.Y.; Zhang, Z.R.; Xin, Y.X.; Wang, F.; Xin, P.Y.; Gao, J. Needle phenotype variation among natural populations of *Pinus yunnanensis*, *P. kesiya* var. *langbianensis*, and *P. kesiya*. *Chin. J. Plant Ecol.* **2023**, *47*, 833–846. [[CrossRef](#)]
- Hu, T.; Hu, H.; Li, F. Long-term effects of post-fire restoration types on nitrogen mineralisation in a *Dahurian larch* (*Larix gmelinii*) forest in boreal China. *Sci. Total Environ.* **2019**, *679*, 237–247. [[CrossRef](#)]
- Xu, H.; Zhang, Z.Y.; Ou, G.L. *Estimation and Distribution of Forest Biomass and Carbon Stocks in Yunnan Province*; Yunnan Science and Technology Press: Kunming, China, 2019.
- Chen, C.G.; Zhu, J.F. *Biomass Tables for Main Tree Species in Northeast China*; China Forestry Publishing House: Beijing, China, 1989.
- Zhao, Y.S. *Principles and Methods of Remote Sensing Applications*; Science Press: Beijing, China, 2011.
- Sarker, M.L.R.; Nichol, J.; Iz, H.B.; Ahmad, B.B.; Rahman, A.A. Forest Biomass Estimation Using Texture Measurements of High-Resolution Dual-Polarization C-Band SAR Data. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 3371–3384. [[CrossRef](#)]
- Chen, L.; Hao, W.Q.; Gao, D.L. The latest applications of optical image texture in forestry. *J. Beijing For. Univ.* **2015**, *37*, 1–12.
- Hao, L.; Liu, H.; Chen, Y.F.; Wu, Y.H. Remote sensing estimation of forest growing stock volume based on spectral and texture information. *Mt. Res.* **2017**, *35*, 246–254. [[CrossRef](#)]
- Haralick, R.M. Statistical and structural approaches to texture. *Proc. IEEE* **2005**, *67*, 786–804. [[CrossRef](#)]
- Chen, L.; Wang, Y.Q.; Ren, C.Y.; Zhang, B.; Wang, Z.M. Assessment of multi-wavelength SAR and multispectral instrument data for forest aboveground biomass mapping using random forest kriging. *For. Ecol. Manag.* **2019**, *447*, 12–25. [[CrossRef](#)]
- Dang, A.T.N.; Nandy, S.; Srinet, R.; Luong, N.V.; Ghosh, S.; Kumar, A.S. Forest aboveground biomass estimation using machine learning regression algorithm in Yok Don National Park, Vietnam. *Ecol. Inform.* **2018**, *50*, 24–32. [[CrossRef](#)]
- Han, Z.T.; Jiang, H.; Wang, W.; Li, Z.Y.; Chen, E.X.; Yan, M.; Tian, X. Forest above-ground biomass estimation using KK-FIFS method based on multi-source remote sensing data. *Sci. Silvae Sin.* **2018**, *54*, 70–79.

31. Han, Z.T. *Forest Above-Ground Biomass Estimation Using Feature Selection Based on Remote Sensing Data*; Fuzhou University: Fuzhou, China, 2017.
32. Breiman, L. *Random Forests*; University of California: Berkeley, CA, USA, 2001; pp. 5–26.
33. Li, X.H. Using random forest for classification and regression. *Chin. J. Appl. Entomol.* **2013**, *50*, 1190–1197.
34. Shi, J.L. *Estimation of Forest Fraction Biomass Supported by Remote Sensing Feature Extraction and Preference of Each Gaoen (GF) Series Satellite*; Southwest Forestry University: Kunming, China, 2023.
35. Pham, L.T.H.; Brabyn, L. Monitoring mangrove biomass change in Vietnam using SPOT images and an object-based approach combined with machine learning algorithms. *Isprs J. Photogramm. Remote Sens.* **2017**, *128*, 86–97. [[CrossRef](#)]
36. Li, C.M. *A Study of the Retrieval Methods of Forest Above-Ground Biomass Based on Multi-Source Data*; Southwest Forestry University: Kunming, China, 2015.
37. Ju, Y.L.; Ji, Y.J.; Huang, J.M.; Zhang, W.F. Inversion of forest aboveground biomass using a combination of LiDAR and multispectral data. *J. Nanjing For. Univ. (Nat. Sci. Ed.)* **2022**, *46*, 58–68. [[CrossRef](#)]
38. Shi, J.M.; Zhang, W.F.; Zeng, P.; Zhao, L.X.; Wang, M.J. Inversion of forest aboveground biomass from combined images of GF-1 and GF-3. *J. Beijing For. Univ.* **2022**, *44*, 70–81. [[CrossRef](#)]
39. Xu, Z.Y.; Li, Y.C.; Li, M.Y.; Li, C.; Wang, L. Forest biomass retrieval based on Sentinel-1A and Landsat 8 image. *J. Cent. South Univ. For. Technol.* **2020**, *40*, 147–155. [[CrossRef](#)]
40. Wang, X.Y.; Zhang, W.F.; Li, Y.; Yang, X.B. Estimation of forest above-ground biomass from optical remote sensing data by optimizing features. *J. Northeast. For. Univ.* **2022**, *50*, 47–54. [[CrossRef](#)]
41. López-Serrano, P.M.; Domínguez, J.L.C.; 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](#)]
42. Liu, S. *Forest Biomass Estimation in Nanchuan District of Chongqing CITY Using a Combination of Sentinel-1 and Sentinel-2 Data*; Chengdu University of Technology: Chengdu, China, 2020. [[CrossRef](#)]
43. Forkuor, G.; Zoungrana, J.B.; Dimobe, K.; Ouattara, B.; Vadrevu, K.P.; Tondoh, J.P. Above-ground biomass mapping in West African dryland forest using Sentinel-1 and 2 datasets—A case study. *Remote Sensing Environ.* **2020**, *236*, 111496. [[CrossRef](#)]
44. Ji, Y.J.; Zeng, P.; Zhang, W.F.; Zhao, L. Forest Biomass Inversion Based on KNN-FIFS with Different ALOS Data. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021. [[CrossRef](#)]
45. Lu, D.; Chen, Q.; Wang, G.; Liu, L.; Li, G.; Moran, E. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *Int. J. Digit. Earth.* **2014**, *9*, 63–105. [[CrossRef](#)]
46. Santoro, M.; Cartus, O.; Fransson, J.E.S.; Wegmüller, U. Complementarity of X-, C-, and L-Band SAR Backscatter Observations to Retrieve Forest Stem Volume in Boreal Forest. *Remote Sens.* **2019**, *11*, 1563. [[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.