

Article

Mapping the Forest Height by Fusion of ICESat-2 and Multi-Source Remote Sensing Imagery and Topographic Information: A Case Study in Jiangxi Province, China

Yichen Luo ^{1,2,†}, Shuhua Qi ^{1,*,†}, Kaitao Liao ^{1,2} , Shaoyu Zhang ¹, Bisong Hu ¹  and Ye Tian ¹

¹ Key Laboratory of Poyang Lake Wetland and Watershed Research (Ministry of Education), School of Geography and Environment, Jiangxi Normal University, Nanchang 330022, China

² Jiangxi Academy of Water Science and Engineering, Nanchang 330029, China

* Correspondence: qishuhua11@jxnu.edu.cn

† These authors contributed equally to this work.

Abstract: Forest canopy height is defined as the distance between the highest point of the tree canopy and the ground, which is considered to be a key factor in calculating above-ground biomass, leaf area index, and carbon stock. Large-scale forest canopy height monitoring can provide scientific information on deforestation and forest degradation to policymakers. The Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) was launched in 2018, with the Advanced Topographic Laser Altimeter System (ATLAS) instrument taking on the task of mapping and transmitting data as a photon-counting LiDAR, which offers an opportunity to obtain global forest canopy height. To generate a high-resolution forest canopy height map of Jiangxi Province, we integrated ICESat-2 and multi-source remote sensing imagery, including Sentinel-1, Sentinel-2, the Shuttle Radar Topography Mission, and forest age data of Jiangxi Province. Meanwhile, we develop four canopy height extrapolation models by random forest (RF), Support Vector Machine (SVM), K-nearest neighbor (KNN), Gradient Boosting Decision Tree (GBDT) to link canopy height in ICESat-2, and spatial feature information in multi-source remote sensing imagery. The results show that: (1) Forest canopy height is moderately correlated with forest age, making it a potential predictor for forest canopy height mapping. (2) Compared with GBDT, SVM, and KNN, RF showed the best predictive performance with a coefficient of determination (R^2) of 0.61 and a root mean square error (RMSE) of 5.29 m. (3) Elevation, slope, and the red-edge band (band 5) derived from Sentinel-2 were significantly dependent variables in the canopy height extrapolation model. Apart from that, Forest age was one of the variables that the RF moderately relied on. In contrast, backscatter coefficients and texture features derived from Sentinel-1 were not sensitive to canopy height. (4) There is a significant correlation between forest canopy height predicted by RF and forest canopy height measured by field measurements ($R^2 = 0.69$, RMSE = 4.02 m). In a nutshell, the results indicate that the method utilized in this work can reliably map the spatial distribution of forest canopy height at high resolution.

Keywords: forest canopy height; ICESat-2; Sentinel-1; Sentinel-2; topographic information; forest age; machine learning



Citation: Luo, Y.; Qi, S.; Liao, K.; Zhang, S.; Hu, B.; Tian, Y. Mapping the Forest Height by Fusion of ICESat-2 and Multi-Source Remote Sensing Imagery and Topographic Information: A Case Study in Jiangxi Province, China. *Forests* **2023**, *14*, 454. <https://doi.org/10.3390/f14030454>

Academic Editor: Helmi Zulhaidi Mohd Shafri

Received: 9 December 2022

Revised: 18 February 2023

Accepted: 20 February 2023

Published: 22 February 2023



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

1. Introduction

Forest ecosystems, as an integral component of terrestrial ecosystems, play a vital role in global climate change and carbon sinks [1]. Forest canopy height is defined as the distance between the highest point of the tree canopy and the ground, which is used to calculate above-ground biomass, leaf area index, and carbon stocks [2–4]. In addition, large-scale forest canopy height monitoring can provide scientific information regarding deforestation and forest degradation to policymakers [5,6].

Remote sensing technology is regarded as a useful technique for global and regional forest canopy height mapping [7,8]. Traditional optical remote sensing for forest canopy

height mapping is based on reflectance information, which has limitations such as quick signal saturation and sensitivity to meteorological conditions, such as clouds and rain [9,10]. In contrast, a microwave radar can operate in any weather condition, regardless of clouds or rain. However, in areas with varied terrain, microwave radar emissions can be influenced by signal saturation, which dramatically reduced the accuracy of canopy height estimates [11]. Light Detection and Ranging (LiDAR), in contrast to the preceding two types of technology, measures targets by producing pulsed lasers that can penetrate dense forest surfaces to acquire information regarding both the understory and surface. As a result, LiDAR is regarded as one of the most precise methods for the measurement of forest structure parameters [6,12].

The Geoscience Laser Altimeter System (GLAS) onboard the Ice, Cloud, and Land Elevation Satellite (ICESat) was operated by NASA in 2010, which was the first satellite-based LiDAR to be used to observe ice cap and forest structure parameters [13]. A series of researchers have undertaken research regarding the integration of the GLAS and other optical data for large-scale forest canopy height mapping [14–16]. For example, in 2010, Lefsky [15] mapped the first global 500 m resolution forest canopy height map by combining GLAS and MODIS (NASA). Simard et al. [16] used random forest (RF) to map 1 km of global resolution forest canopy height based on the GLAS and seven remote sensing images. In addition, in 2016, Wang et al. [17] used the GLAS and 13 auxiliary variables to globally map forest canopy height at a 500 m resolution using the balanced random forest algorithm. However, the GLAS has a low spot density and a big spot size, resulting in low accuracy and resolution in forest canopy height mapping products [18]. In 2018, the Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) was launched by NASA at Vandenberg Air Force Base, which carried the Advanced Topographic Laser Altimeter System (ATLAS) [19,20]. The ATLAS instrument employs multi-beam, micropulse, and photon-counting LiDAR technology. Multiple beams can map more information at the same time compared with a single beam. Meanwhile, the micropulse increases the emission frequency and thus the spot density, which can greatly improve measurement accuracy [19]. The ATLAS instrument is not primarily charged with studying vegetation, but the ATL08 data packages for both land and vegetation provide more opportunities for large-scale forest investigations. On the other hand, ATL08 data can only offer a footprint rather than continuous forest canopy height information at regional and global scales [20–22]. Furthermore, the Global Ecosystem Dynamics Investigation (GEDI) is another high-resolution laser ranging system operated by NASA [23]. The GEDI is designed to observe the forest areas between 51.6° N and S. Meanwhile, the basic metric of the GEDI are wave-forms related to the density profile of canopy height for 25 m footprints [24]. To date, the GEDI has been widely used for global and regional forest canopy height mapping. In 2019, Potapov et al. [25] integrated the GEDI and Landsat data to create a global, 30 m resolution map of forest canopy height.

To overcome the spatial discontinuity of LiDAR measurements [21,26], several studies have shown that integrating multi-source continuous remote sensing data with LiDAR data could extrapolate forest canopy height from the footprint scale to the regional scale, resulting in wall-to-wall forest canopy height maps [14,16,18,27]. Furthermore, it was found that the large-scale spatial feature information provided by optical remote sensing data such as Landsat (NASA), Sentinel (ESA), and MODIS correlated well with the vegetation structure information acquired from LiDAR [6,7,16,21]. Wu et al. [28] constructed an ecological zoning random forest algorithm based on ATLAS and Landsat for 30 m resolution forest height mapping in China. Compared with the optical satellite data, Sentinel-2's red-edge bands were shown to provide more accurate information regarding vegetation growth [29]. Zhang et al. [30] fused ATLAS, Landsat, and Sentinel to map boreal forest canopy height via RF. In contrast, this study demonstrated that the red-edge band (band 5) and the NDRE derived from band 5 in Sentinel-2 are significantly dependent features in RF. In addition, Sentinel-1 SAR data, topographic information, textural features, and climatic data were used to map global and regional forest canopy height maps [1,7,18]. In the Himalayan Foothills of India, Nandy et al. [1] mapped forest canopy height by inte-

grating ICESat-2 and Sentinel-1 data using the RF algorithm. Meanwhile, Sothe et al. [31] mapped forest canopy height in Canada by combining GEDI and ICESat-2 with PALSAR and Sentinel in 2020, which demonstrated the potential of the L-band for forest canopy height prediction.

Machine learning models, as statistical predictive models, have many advantages, including lower computational complexity, less need for parameter tuning, stronger classification and regression capabilities, and higher performance in integrating multi-source data [5,32]. Machine learning models such as RF, the Support Vector Machine (SVM), K-nearest neighbor (KNN), Gradient Boosting Decision Tree (GBDT), and Mixed Logistic Regression (MLR) have been frequently employed for forest canopy height extrapolation [5,6,21,27,33]. For example, Zhu et al. [14] generated a forest canopy height map with a spatial resolution of 30 m, utilizing Multiple Altimeter Beam Experimental Lidar (MABEL) and Landsat 8 OLI. Jiang et al. [6] successfully developed a stacking algorithm consisting of MLR, RF, SVM, and KNN to map the forest canopy height in northern China. Compared with the single algorithm, it was shown that the stacking algorithm effectively improved the extrapolation accuracy. Xi et al. [5] developed machine learning models for different forest types by integrating ICESat-2, Sentinel, and topographic information. This study demonstrated the better performance of machine learning models developed based on different forest types relative to the whole forest.

Although numerous studies have successfully mapped continuous forest canopy height by fusing satellite-based LiDAR data and remotely sensed images, few have considered the relationship between forest-related indicators and canopy height when conducting forest canopy height mapping [34]. Previous research has demonstrated that forest age and forest canopy height are closely related [35]. Several methods for combining forest canopy height information to map forest age have been proposed [36–38]. Furthermore, no relevant studies have explored the potential of forest age in forest canopy height mapping. In addition, the spatial and temporal inconsistency of remote sensing and validation datasets is another challenge in mapping regional forest canopy height [39].

The goal of this research was to create a high-resolution forest canopy height map for Jiangxi Province in southeastern China by utilizing machine learning models that incorporate ICESat-2, Sentinel, and forest age, and then to assess the forest canopy height map using time-consistent validation data. To achieve this goal, four specific objectives were proposed: (1) investigating the relationship between forest age and forest canopy height at the sample scale and providing insight into the key drivers of canopy height extrapolation models in the region, (2) developing four forest canopy height extrapolation models based on machine learning models to explore significant predictors that relate to forest canopy height, (3) generating a Jiangxi Province forest canopy height map based on the best-performing model, and (4) validating the forest canopy height map using field measurements and comparing it with existing forest canopy height maps.

2. Materials and Methods

2.1. Study Area

The study was conducted in Jiangxi Province, which is located in southeastern China and has an area of about 166,900 km² (Figure 1). The research area's topography is complex and diversified, dominated by hills and mountains, and it features wide basins and valleys with an average elevation above sea level of 245.6 m. Jiangxi Province is abundant in water resources, with a thick network of rivers connecting to the Yangtze River via Poyang Lake, China's largest freshwater lake. The climate is a humid subtropical monsoon climate with an annual temperature range of 16.3 to 19.5 °C. Furthermore, the region has a frost-free period of 240–307 days and an annual precipitation average of 1341–1943 mm.

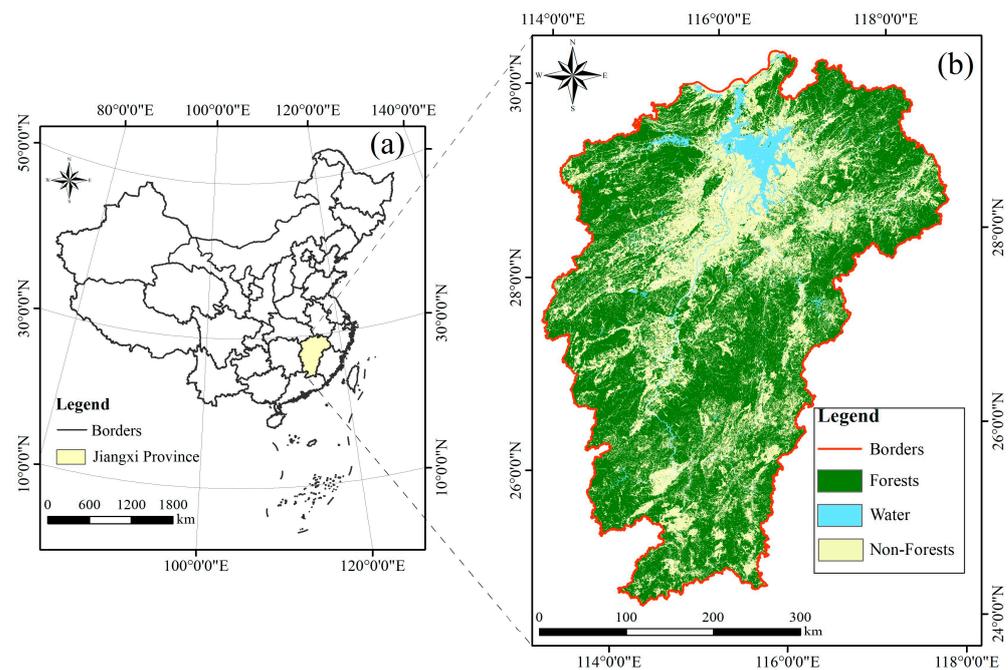


Figure 1. (a) Location of the study area in China. (b) Land cover, including forest, water, and non-forest, in the study area.

Jiangxi Province is rich in forest resources, with a forest coverage rate of 63.35%, ranking second in China. Broad-leaved evergreen forests are the zonal vegetation of Jiangxi Province. Due to the region's broad geographical span from north to south, the southern section is influenced by the southern subtropical climate and flora components, and the vegetation features exhibit some southern subtropical characteristics. In the north, there are more tropical flora components, and some warm temperate flora components are increasingly mixed. The dominant tree species in the region are Masson's pine (*Pinus massoniana* Lamb), Chinese fir (*Cunninghamia lanceolata*), and slash pine (*Pinus elliottii*).

2.2. Data Acquisition and Processing

In this study, the following data were used to accomplish the objectives: (1) ATL08 data collected using ICESat-2, (2) SAR data collected using Sentinel-1, (3) optical image data collected using Sentinel-2, (4) topographic data collected from the Shuttle Radar Topography Mission (SRTM), (5) forest age data in Jiangxi Province calculated using a time-series change monitoring algorithm, (6) the Land Use Cover and Change product collected using GlobeLand30, and (7) validation data provided by the 7th Jiangxi Province Forest Resources Second Class Survey. Canopy height information from the ICESat-2 point and 33 predictors (Table 1) corresponding to the geographical location were obtained from the data sources listed above and were used in subsequent machine learning modeling.

2.2.1. ATL08 Data

The ATLAS instrument uses green (532 nm) laser light and single-photon sensitive detection to measure the time of flight and, subsequently, surface height along each of its six beams [19]. ATLAS data are distributed as Hierarchical Data Format Version 5 (HDF-5) through the National Snow and Ice Data Center (NSIDC, <https://nsidc.org/data/icesat-2> (accessed on 10 January 2022)), which is divided into 4 levels (ATL01-ATL21, without ATL05) [40]. ATL08 is a product that was developed in response to one of ICESat-2's scientific tasks: measuring vegetation canopy height as a basis for estimating large-scale biomass and biomass change, and it provides us with several parameters such as surface elevation, absolute canopy height, relative canopy height, and so on. Each metric is derived from photons classified as ground or canopy within a 100-m segment. The primary

canopy height parameter in the data product, h_{canopy} , is defined as 98% canopy height in 100 m segments [20]. In addition, as a parameter used to describe the terrain elevation, $h_{te_best_fit}$ is defined as the best-fit terrain elevation at the mid-point location of each 100 m segment [20].

The ATL08 data (580876 ICESat-2 points) from 2020 were utilized in this study to ensure the chronological validity of the verified data and the coverage of the research area. As noise points are always present, we used several methods to filter out potentially erroneous ICESat-2 points. ICESat-2 points with h_{canopy} less than 2 m or greater than 50 m were filtered out based on forest structure information in the research area. Furthermore, SRTM was utilized as terrain calibration data, and if the discrepancy between the $h_{te_best_fit}$ and SRTM was greater than 20 m, it was assumed that the point was impacted by clouds or the atmosphere [18]. At the same time, strong beam data recorded at night was used in the study as it was shown to have the highest performance in canopy height retrieval [41]. In summary, 5777 filtered ICESat-2 points (Figure 2a) based on the method described above were used for subsequent machine modeling.

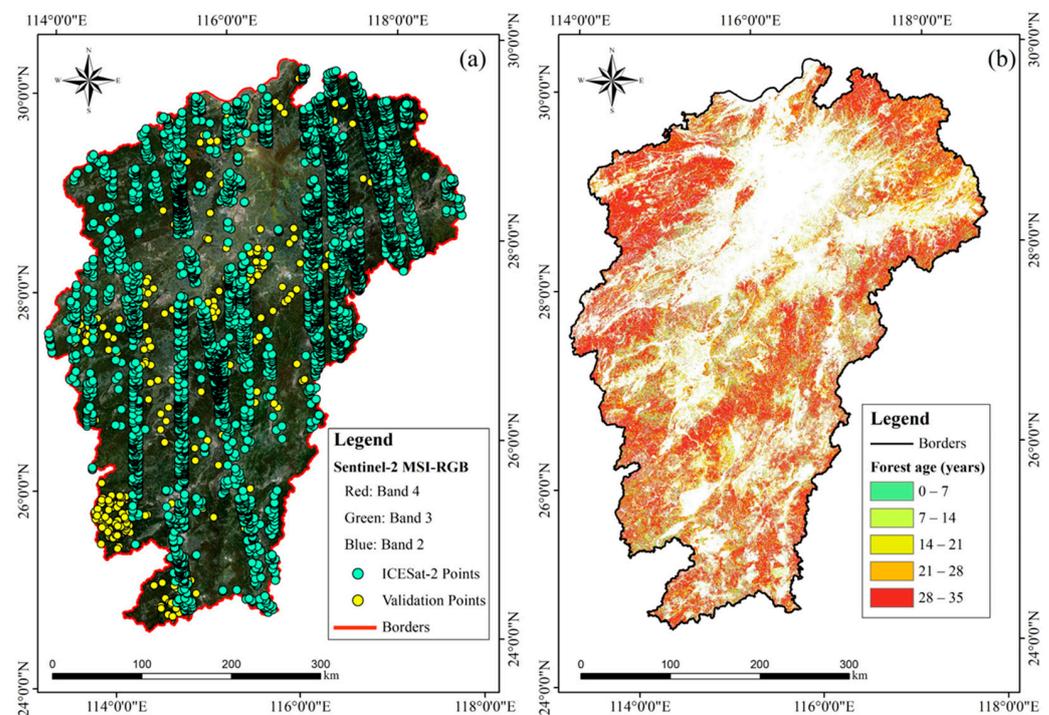


Figure 2. (a) Distribution of ICESat-2 points and field measurements in Sentinel-2 RGB composite images. (b) Forest age of Jiangxi Province.

2.2.2. Forest Age

Forest age data were used in this study as key input data in models to estimate forest canopy height, which were created by the contributing author last year. The data showed that the secondary forest age ranged from 0 to 34 years annually, and the stable forest with no changes in the time series was defined as having an age of 35. It was mapped using multiple change detection algorithms with dense Landsat time series and inventory data, which presents the first map of forest age (Figure 2b) in 2020 at a 30 m resolution in Jiangxi Province. Meanwhile, we used 160 field measurements to validate the forest age map for Jiangxi Province. The variable from the forest age product was reliable to establish the relationship with forest canopy height due to the high accuracy of the coefficient of determination (R^2) = 0.87; root mean square error (RMSE) = 3.17 years.

2.2.3. Sentinel Data

Sentinel-1 is a polar-orbiting binary constellation with the ability to provide daily, day/night, and all-weather medium-resolution observations for the continuation and improvement of SAR operational services and applications, which can provide single (HH and VV) or dual polarization modes (HH + HV and VV + VH) [42]. Sentinel-2 provides high-resolution multispectral image acquisition on a global scale through the Multi-Spectral Instrument (MSI), which carries 13 spectral bands from the visible and near-infrared to the short-wave infrared [43]. In addition, the red-edge band obtained using the MSI is widely used to monitor vegetation because it is sensitive to vegetation growth [44].

In this study, Sentinel data regarding the vegetation growing season (2020) were acquired and processed in Google Earth Engine (GEE, <https://earthengine.google.com/> (accessed on 28 April 2022)). The backscatter coefficients, VV and VH, of Sentinel-1 were obtained from Sentinel-1 SAR GRD products. Sentinel-2 images from “Level-2A” data were cloud-masked using the s2cloudless algorithm [45]. Considering the spatial consistency and availability of data, both Sentinel-1 and Sentinel-2 data (Figure 2a) were resampled to a 30 m resolution and processed using a median image composition. Additionally, composite images were used to generate eighteen texture features and three vegetation indices as the base image. In general, features contained in the gray-level co-occurrence matrix (GLCM) in Table 1 are used to characterize the texture; the features include Mean, Variance, Homogeneity, Contrast, Dissimilarity, Entropy, Angular Second Moment, and Correlation [46]. Texture features are calculated using a GLCM with third-order kernels based on the backward scattering coefficient. The three normalized (NDRE_BAND5, NDRE_BAND6, and NDRE_BAND7) red-edge indices are calculated based on the three red-edge bands and the infrared band [47]. The formula for NDRE is as follows:

$$\text{NDRE} = \frac{\text{NIR} - \text{RedEdge}}{\text{NIR} + \text{RedEdge}} \quad (1)$$

where *NIR* represents the near-infrared band (band 8), and *RedEdge* represents the three red-edge bands of band 5, 6, and 7 in Sentinel-2.

Table 1. Predictors obtained by multi-source data.

Data Source	Predictors	Description	References
Forest Age	FOREST_AGE	30 m-resolution forest age map in Jiangxi Province	-
	VV	VV band extracted from Sentinel-1	[42]
	VH	VH band extracted from Sentinel-1	[42]
Sentinel-1	VV_MEAN	Mean Value calculated by GLCM based on VV	[46]
	VV_VAR	Variance calculated by GLCM based on VV	
	VV_HOM	Homogeneity calculated by GLCM based on VV	
	VV_CON	Contrast calculated by GLCM based on VV	
	VV_DISS	Dissimilarity calculated by GLCM based on VV	
	VV_ENT	Entropy calculated by GLCM based on VV	
	VV_SEC	Angular Second Moment calculated by GLCM based on VV	
	VV_COR	Correlation calculated by GLCM based on VV	
	VH_MEAN	Mean Value calculated by GLCM based on VH	
	VH_VAR	Variance calculated by GLCM based on VH	
	VH_HOM	Homogeneity calculated by GLCM based on VH	
	VH_CON	Contrast calculated by GLCM based on VH	
	VH_DISS	Dissimilarity calculated by GLCM based on VH	
	VH_ENT	Entropy calculated by GLCM based on VH	
	VH_SEC	Angular Second Moment calculated by GLCM based on VH	
VH_COR	Correlation calculated by GLCM based on VH		

Table 1. Cont.

Data Source	Predictors	Description	References
Sentinel-2	S2_BAND2	Blue band extracted from Sentinel-2	[43]
	S2_BAND3	Green band extracted from Sentinel-2	
	S2_BAND4	Red band extracted from Sentinel-2	
	S2_BAND5	Vegetation Red Edge band extracted from Sentinel-2 (705 nm)	
	S2_BAND6	Vegetation Red Edge band extracted from Sentinel-2 (740 nm)	
	S2_BAND7	Vegetation Red Edge band extracted from Sentinel-2 (782 nm)	
	S2_BAND8	NIR band extracted from Sentinel-2	
	S2_BANDA	Narrow NIR band extracted from Sentinel-2	
	NDRE_BAND5	Normalized difference red-edge vegetation index based on S2_BAND5 and S2_BAND8	[47]
	NDRE_BAND6	Normalized difference red-edge vegetation index based on S2_BAND6 and S2_BAND8	
	NDRE_BAND7	Normalized difference red-edge vegetation index based on S2_BAND7 and S2_BAND8	
SRTM	ELEVATION	Elevation extracted from SRTM	[48]
	SLOPE	Slope extracted from SRTM	
	ASPECT	Aspect extracted from SRTM	

2.2.4. Auxiliary Data

The SRTM uses one-way interferometry to acquire topographic data. The SRTM collects radar data for 80% of the Earth's surface between 60° N and 56° S, which is used to construct a global digital elevation model. In this study, the SRTM was called and processed using GEE to obtain the elevation, slope, and aspect of the study area [48].

GlobeLand30 2020, a 30 m resolution global land cover data product, is the latest version developed in China. It includes 10 land cover classifications, including forests. The product has an overall accuracy of 85.72% [49]. GlobeLand30 2020 was used for the study area in this study, which effectively helped us to distinguish between forested and non-forested areas.

The 7th Jiangxi Province Forest Resources Second Class Survey was completed in Jiangxi Province in 2019–2020, which established a database of forest resources in Jiangxi Province consisting of irregular blocks [6]. Each block consists of the same forest species and contains geographic coordinates, forest type, forest canopy height, forest age, etc. Meanwhile, the forest canopy height represents the average height of the forest canopy within the whole block. In this study, 380 field measurement points (Figure 2a) converted from the center of gravity of blocks were used to validate the forest canopy height maps generated by the machine learning model. The coordinates of the field measurement points were used as a reference, and cell values at locations were extracted from one raster and recorded in the attribute table of the point.

Previous research has generated forest canopy height maps of China at 30 m resolutions using various methodologies [18,50]. Zhu [18] produced a 30 m resolution forest altitude map of China ($R^2 = 0.38$; RMSE = 2.67 m) using a random forest model in collaboration with data from satellite-based LiDAR (GEDI and ICESat-2) and optical imagery (Landsat-8 and Sentinel). Meanwhile, Liu et al. [50] developed a deep learning interpolation model by integrating the GEDI and ICESat-2 to map another 30 m resolution forest canopy height map of the Chinese region ($R^2 = 0.6$; RMSE = 4.88 m). In addition, the Global Forest Canopy Height (GFCH) mapped based on the GEDI and Landsat by Potapov et al. [25] in 2019 is also open access (<https://glad.umd.edu/dataset/gedi/> (accessed on 11 February 2023)). We gathered the forest canopy height in Jiangxi Province from three data [18,25,50] using the Jiangxi Province boundary mask to assess the correctness of the forest canopy height map in this study.

2.3. Methods

In this study, correlation analysis was used to understand the relationship between canopy height and predictors, and in addition, machine learning models were used to link canopy height and predictors and perform continuous forest canopy height mapping. Meanwhile, SPSS Statistics software (version 25) was used for correlation analysis, and R (version 4.2.1) was used for machine learning modeling. A workflow (Figure 3) was created to summarize the main steps of the implemented methodology in the study.

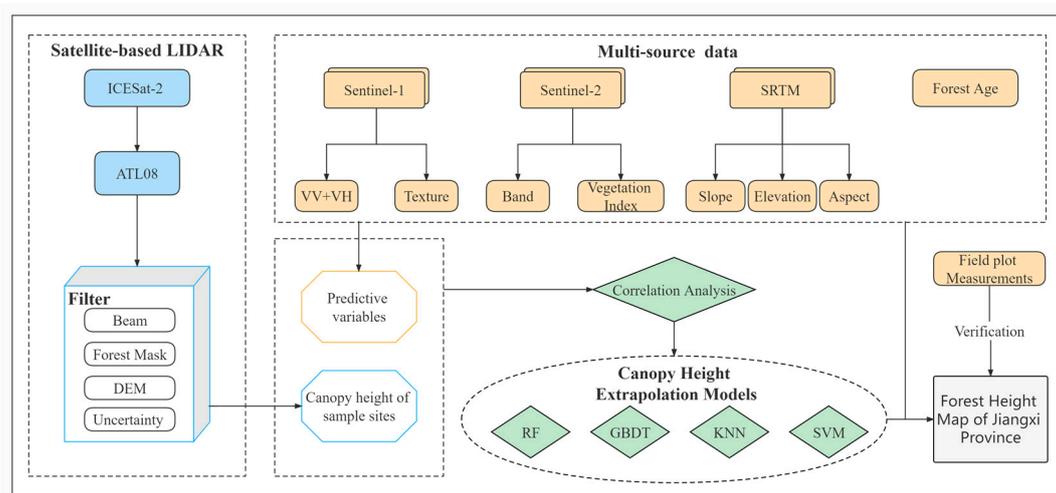


Figure 3. Workflow of integration of ICESat-2 and multi-source data for forest canopy height mapping in Jiangxi Province.

2.3.1. Correlation Analysis

Correlation analysis is the examination of two or more variables that may be correlated, and it often uses the correlation coefficient to determine the degree of correlation between two variables. The correlation coefficient is between -1 and 1 . The more connected the variables are, the closer it gets to the extreme value [51]. In this study, the Spearman correlation coefficient (r_s) was utilized to link the forest canopy height and the relevant predictors at ICESat-2 sites based on the distribution features of the data. To reduce the dimensionality of the predictors, a p -value < 0.01 and an $|r_s| > 0.2$ were used for model development. The formula of r_s is as follows:

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (2)$$

where r_s represents the Spearman correlation coefficient; d_i represents the difference in the place value of the data pair; and n represents the number of observation samples.

2.3.2. Model Development and Hyperparameter Optimization

Four machine learning models (RF, GBDT, SVM, and KNN) were utilized to construct canopy height extrapolation models for the proper estimation of forest canopy height in Jiangxi Province. RF is a bagging integration algorithm that builds a robust model by integrating multiple decision trees and has better generalization performance and less possible overfitting compared with a single decision tree [52]. GBDT, similar to RF, is an integrated model based on decision trees that uses a boosting algorithm to concatenate the decision trees and eventually approximate the true value by minimizing the model's bias [53]. The basic principle of an SVM is to discover a hyperplane that minimizes the distance between it and all of the data. Although the SVM technique is best suited for classification issues, it can also perform well in regression [6]. The final method, KNN,

solves the regression problem by looking for comparable K samples based on spectral distance and making predictions using the mean of the samples [54].

In this study, 80% of the samples ($n = 4621$) served as a training set, while 20% of the data ($n = 1156$) served as a testing set. The RandomForest (version 4.7.1) package was used for the RF algorithm [52], the gbm (version 2.1.8.1) package was used for the GBDT algorithm [53], the e1071 (version 1.7.11) package was used for the SVM algorithm [55], and the Caret (version 6.0.92) package was used for the KNN algorithm [56]. To improve the model's performance, the parameters were modified by running the "train" function in the Caret package several times. The hyperparameter optimization approach was developed using a grid search and 10-fold cross-validation [56]. The Ntree and Mtry parameters had to be input to the RF algorithm. The "randomforest" function was also used to test the value of the Ntree. After determining the optimal Ntree, a hyperparameter optimization approach was utilized to determine the best Mtry. The GBDT algorithm required the input of four parameters: Trees, Depth, Minobsinnode, and Shrinkage, which were divided into two groups for the implementation of the hyperparameter optimization algorithm. The linear function was used as the kernel function in the SVM algorithm, and the only parameter that needed to be changed was the cost value (C). A range was established based on various changes to determine the best-performing cost value. Simultaneously, for the KNN algorithm, the Euclidean distance was set to the spectral distance, and the optimal number of neighborhood points (K) was found using a hyperparametric optimization algorithm. All the parameters are shown in Table 2, which presents the parameters and descriptions of them in the machine learning models. The RMSE was used as an evaluation tool for hyperparametric optimization algorithms, and the parameters with the lowest RMSE were used to develop the canopy height extrapolation model.

Table 2. Parameters and descriptions in machine learning models.

Model	Parameter	Description
RF	Ntree	Number of decision trees
	Mtry	Number of features in each decision tree
SVM	Cost	Penalty factor
KNN	K	Number of neighboring points
GBDT	Trees	Number of iterative regression trees
	Depth	Maximum depth of decision tree
	Minobsinnode	The minimum number of observations in the terminal nodes of the decision tree
	Shrinkage	Learning rate or step reduction

2.3.3. Continuous Forest Canopy Height Mapping and Validation

Once all of the parameters were optimized, four canopy height extrapolation models were created, with the canopy height in the training set serving as the dependent variable and other remote sensing data serving as the predictors. The canopy heights of the testing set were assessed using R^2 and RMSE. Furthermore, the best-performing model was used to map the forest canopy height in Jiangxi Province. Finally, all the forest canopy height maps were evaluated using field measurements. The formulas for R^2 and RMSE are as follows:

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

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

where n represents the number of observations; x_i represents the i -th observation; y_i represents the i -th predicted value; \bar{y} represents the average of the predicted values.

3. Results

3.1. Correlation between Canopy Height and Predictors

For each ICESat-2 point, the Spearman correlation coefficient was used to investigate the relationship between canopy height and 33 predictors. The statistical results (Table 3) showed that 26 variables were significantly correlated with canopy height, and 7 variables were not significantly correlated. Among the 26 significantly correlated variables, the r_s ranged from 0.08 to 0.62.

Table 3. r_s of predictors and canopy height at sample scale.

Predictors	r_s	Predictors	r_s	Predictors	r_s
FOREST_AGE	0.415 **	VV_COR	0.007	NDRE_BAND5	0.454 **
VH_DISS	0.313 **	VV_DISS	0.326 **	NDRE_BAND6	0.361 **
VH_ENT	0.316 **	VV_HOM	−0.316 **	NDRE_BAND7	0.007
VH_CON	0.310 **	VV_MEAN	0.147 **	S2_BAND2	−0.278 **
VH_COR	0.011	VV_SEC	−0.327 **	S2_BAND3	−0.409 **
VH_HOM	−0.302 **	VV_VAR	0.326 **	S2_BAND4	−0.411 **
VH_MEAM	0.191 **	SLOPE	0.548 **	S2_BAND5	−0.436 **
VH_SEC	−0.315 **	ELEVATION	0.620 **	S2_BAND6	−0.078 **
VH_VAR	0.309 **	ASPECT	0.031 *	S2_BAND7	0.018
VV_ENT	0.327 **	VV	0.163 **	S2_BAND8	0.016
VV_CON	0.322 **	VH	0.128 **	S2_BAND8A	0.034 *

– represents negative correlation, * represents p -value < 0.05, ** represents p -value < 0.01.

It is worth noting that elevation ($r_s = 0.62$) and slope ($r_s = 0.55$) were strongly connected with canopy height among all the statistically correlated variables. Sentinel-2 characteristics such as NDRE_BAND5 ($r_s = 0.45$) and S2_BAND5 ($r_s = 0.44$) were also moderately linked with canopy height. Sentinel-1-derived information was less correlated (r_s from 0.13 to 0.32), reflecting the limited penetration of Sentinel-1's C-band into the forest [27]. Furthermore, forest age showed a link with canopy height ($r_s = 0.42$), implying that it may contribute to canopy height extrapolation modeling. The correlation analysis and variable selection method were utilized to choose 21 variables (Figure 4) for models development.

Correlation Analysis

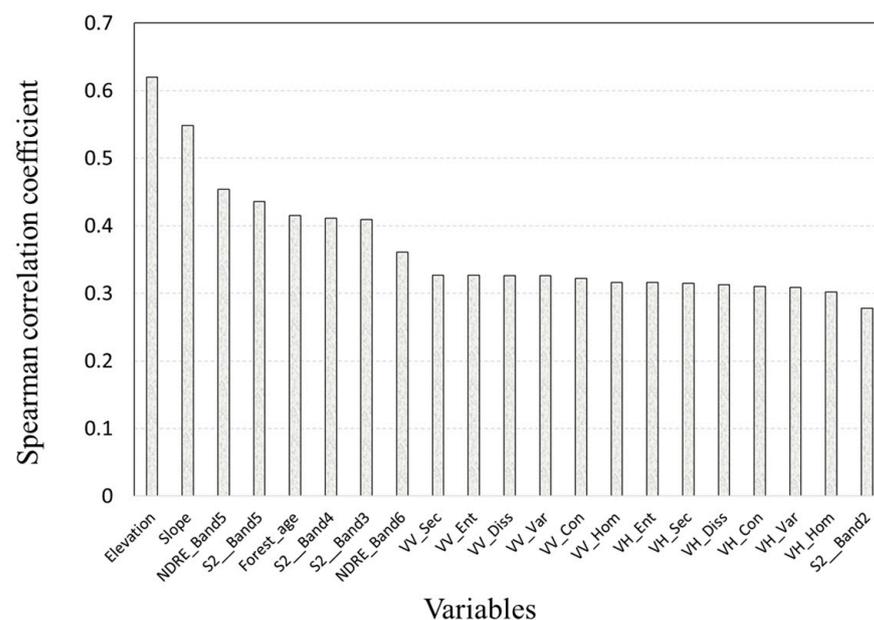


Figure 4. Correlation analysis of the 21 characteristic variables used in the model with canopy height.

3.2. Hyperparameter Optimization for Canopy Height Extrapolation Models

Ntree was initially set to 300 for the RF model, and the OOB error was steady after Ntree exceeded 150, whereupon it was set to 150 (Figure 5a). Furthermore, in the hyperparametric method, the ideal Ntree was used to obtain the Mtry, and Figure 5b shows that the RMSE was lowest when Mtry was 9. In contrast to RF, in this study, the KNN and SVM models required the development of a parameter. When K was set to 16 (Figure 5c) in KNN and cost was set to 0.3 (Figure 5d) in the SVM, the smallest RMSE was attained. Searching the mesh for all parameters at the same time in a model with several parameters may have resulted in computational complexity. Hence, for a four-parameter GBDT model, as shown in Figure 5e,f, the four parameters were divided into two groups and searched in the grid, and the best parameters were shown to be a combination of the two searches.

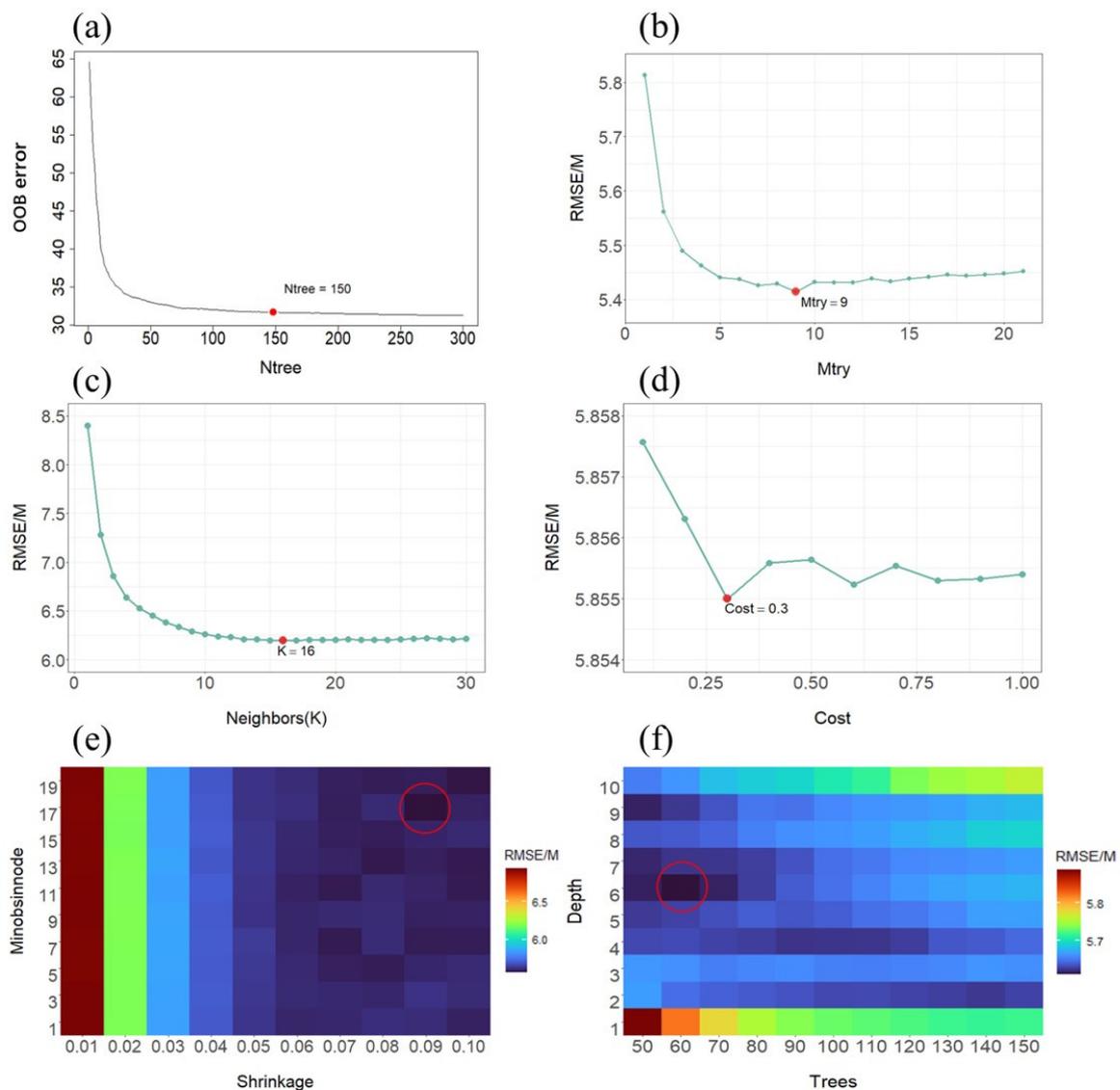


Figure 5. Parameters in the machine learning model selected based on cross-validation. (a) Ntree in the RF, (b) Mtry in the RF, (c) K in the KNN, (d) cost in the SVM, (e) Minobsinnode and Shrinkage in the GBDT, (f) Depth and Trees in the GBDT.

3.3. Comparison of Canopy Height Extrapolation Models

Once the optimal parameters in the model were determined, four machine learning models were utilized to establish the relationship between canopy height and predictors

for the ICESat-2 point. The models were validated using the testing set to further verify their correctness and compare their performance using R^2 and RMSE metrics. In Figure 6, the R^2 for all models ranged from 0.47 to 0.61, and the RMSE ranged from 5.29 to 6.18 m. The best-fitting model was determined to be RF ($R^2 = 0.61$ and RMSE = 5.29 m). Figure 6a,b indicate that the two integrated models obtained from the decision tree, GBDT and RF, produced relatively similar results. Additionally, the mean coefficient of determination ($R^2 = 0.55$) and mean root mean square error (RMSE = 5.64 m) were calculated, indicating that all four models could extrapolate the sample plot scale to the regional scale. Lastly, Figure 6 shows all the model-predicted and actual canopy height values were significantly correlated (p -value < 0.01).

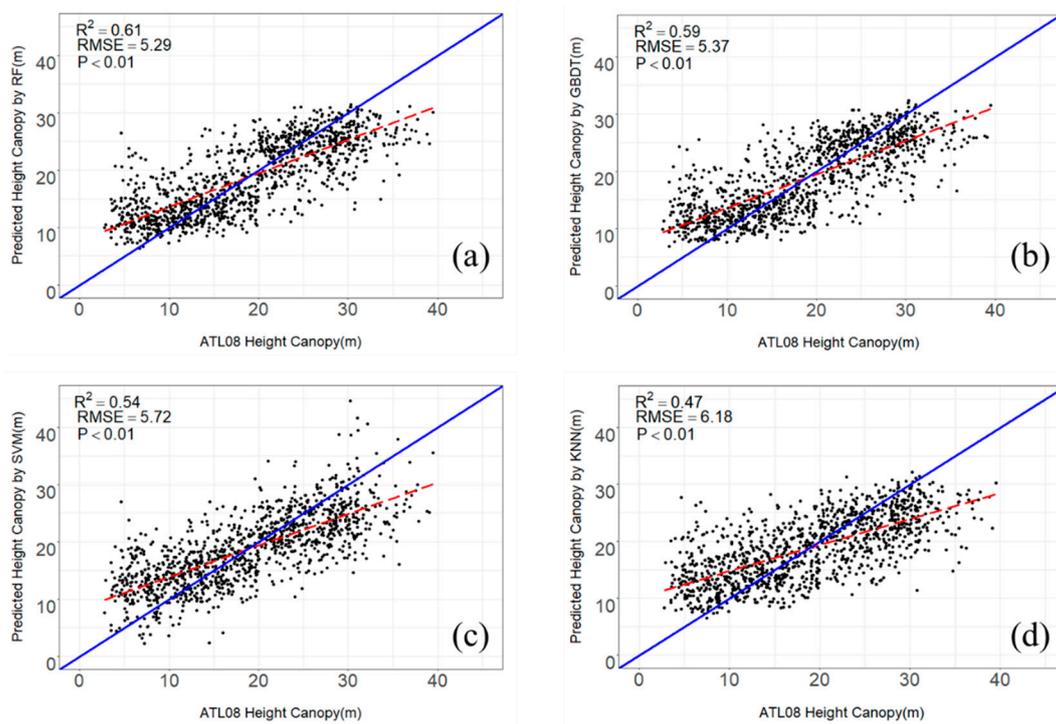


Figure 6. Comparison of accuracy of four machine learning models: (a) RF, (b) GBDT, (c) SVM, (d) KNN.

In Table 3, the Spearman coefficient was used to measure the linear relationship between the canopy height and predictors. However, the canopy height and predictors are not generally linearly related in environmentally complex forests [6]. To further explore the key drivers in the extrapolation process of forest canopy height, the importance evaluation in the random forest model (Figure 7) was used to analyze the degree of contribution of the variables in the canopy height extrapolation model. The results of Figure 7 show that the random forest model significantly depended on elevation, slope, S2_BAND5, and NDRE_BAND5, with elevation being the variable that contributed the most to the model. Meanwhile, Figure 7 suggests the backward scattering coefficient and texture features derived from Sentinel-1 were shown to have relatively low contributions to the model.

3.4. Mapping Wall-to-Wall Map of Forest Canopy Height

Continuous spatial feature information and canopy height extrapolation models provide the possibility of mapping forest canopy height in wall-to-wall areas. In Figure 8a, a 30-m resolution forest canopy height map for Jiangxi Province was generated based on the RF model and 21 consecutive remote sensing images. The heights of the predictions (Figure 8a) in the area ranged from 4.88 to 35.55 m, with an average height of 23.8 m. In

general, the forest canopy height was higher in the mountains and lower in the plains and hills. Figure 8b shows that forest heights recorded in the field ranged from 2.7 m to 31.1 m, while samples were utilized to validate the forest canopy height map. The validation (Figure 8c) reveals that the canopy height extrapolation model we created could effectively map the forest canopy height in the research area ($R^2 = 0.69$; RMSE = 4.02 m).

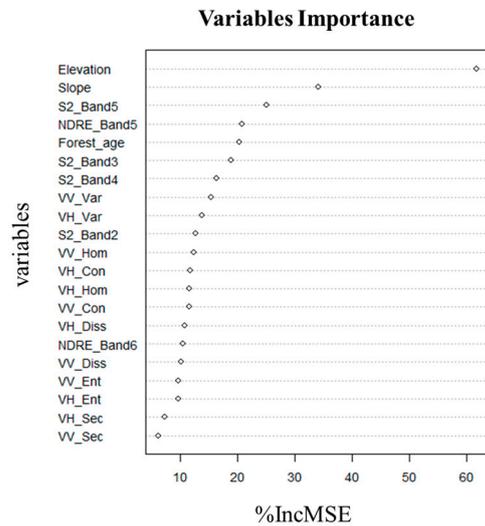


Figure 7. Importance assessment based on %IncMse in the RF model.

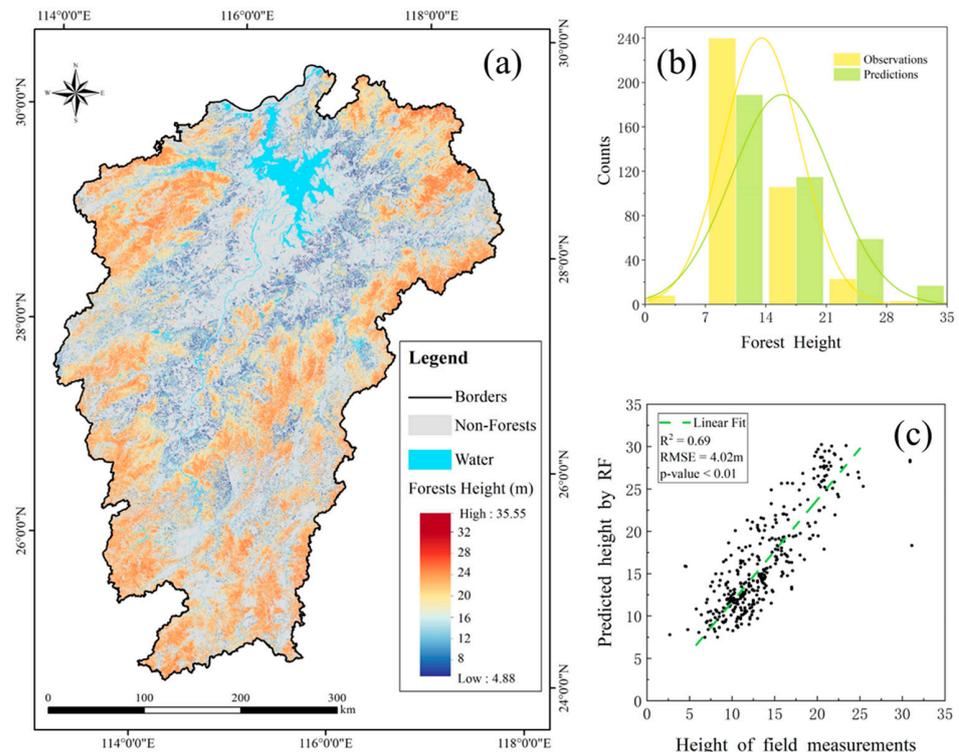


Figure 8. (a) A 30 m resolution forest canopy height map of Jiangxi Province based on the RF model. (b) The counts of field-measured height and predicted tree height. (c) Validation of forest canopy height map produced using RF model.

3.5. Comparison of Forest Canopy Height Maps

In Figure 8a, the validation shows that our method can produce reliable forest canopy height maps. Nonetheless, for a more comprehensive analysis of the accuracy of the forest

canopy height maps, one global [25] and two Chinese [18,50] forest canopy height maps were processed using Jiangxi masks and validated using the same field measurements. Figure 9 shows that the R^2 of the validation ranged from 0.24 to 0.39, while the RMSE ranged from 4.11 to 5.48 m. The results show that the forest canopy height map of the Jiangxi Province produced in this study had higher accuracy.

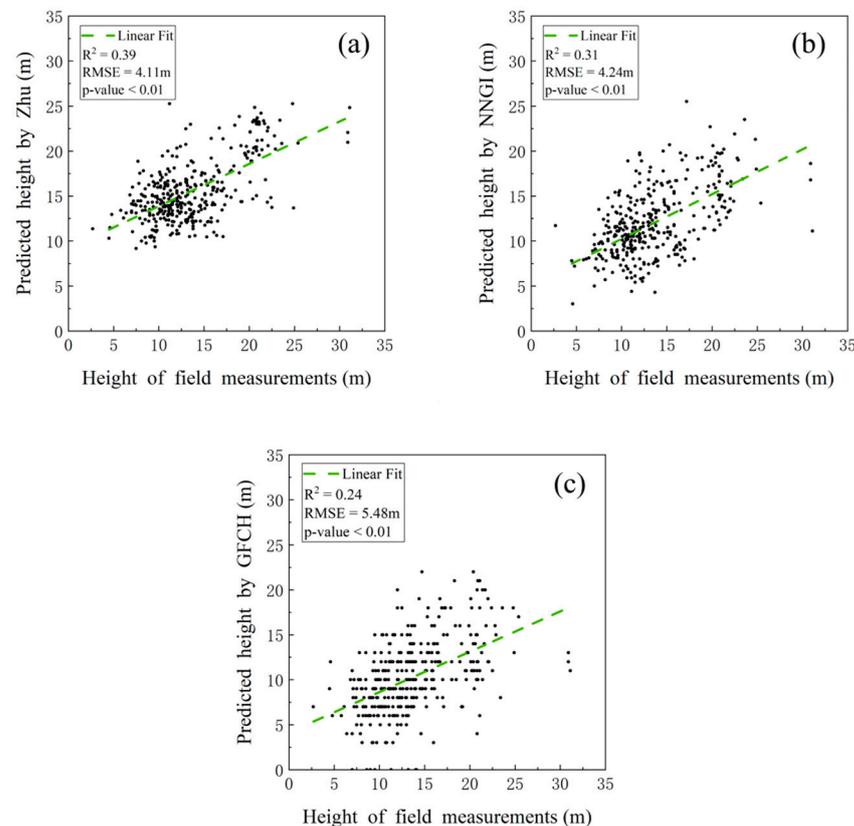


Figure 9. Validation of the forest height of Jiangxi Province generated by (a) Zhu [18], (b) NNGI [50], (c) GFCH [25].

4. Discussion

4.1. Key Drivers in the Extrapolation of Forest Canopy Height

Predictors derived from remote sensing data such as optical images can be efficiently integrated with ICESat-2 or MABEL modeling data to produce continuous forest canopy height maps [1,5–7,14]. However, few studies have discussed the role of other forest structures in forest canopy height mapping. Several studies have already shown a potential correlation between forest age and forest canopy height [36–38]. In this study, we explored continuous forest canopy height mapping using high-resolution forest age maps generated using a continuous change detection technique as predictors, along with other optical remote sensing data. The results in Table 3 indicate that there is a moderate correlation between forest age and forest canopy height at the sample scale.

Meanwhile, the importance ratings of the predictors (Figure 7) generated by the RF showed that forest age was one of the variables that the RF moderately relied on. These facts stated above imply that forest age has the potential to predict forest canopy height. From another perspective, the results in Figure 7 show that S2_BAND5 and NDRE_BAND5 are significantly dependent variables for forest canopy height prediction. Notably, this result has similarity with the results of Zhang et al. [30] and implies that the red-edge band can characterize vegetation growth to some extent, as demonstrated in previous studies [44,47]. In addition, Figure 7 shows that forest canopy height is also significantly affected by elevation and slope. One obvious reason for this is that changes in elevation

and slope impact climate, temperature, and sunlight, all of which have an indirect effect on forest growth [57,58]. On the other hand, the impact of human activities on a forest will vary with elevation and slope.

4.2. Accuracy of Canopy Height Extrapolation Model and Forest Canopy Height

In Figure 6, all models overestimate low forests and underestimate forests with high heights. In fact, this bias appeared in previous studies that integrated ICESat-2 and different predictor variables using RF [5–7,18,28,30,31], KNN [6], SVM [6], GBDT [5], MLR [6], and a deep learning model [7] for continuous forest height mapping. However, the reason for this bias in the prediction of forest height maps via machine learning models is not clear. In addition, RMSE has been frequently used in the assessment and comparison of canopy height extrapolation models [5,6,31], and other metrics such as mean absolute error (MAE) and RMSE% should also be considered in subsequent studies to reduce uncertainty.

The results of Figures 8c and 9 reveal that the forest canopy height map generated in this study was more accurate. It is worth noting that the difference in accuracy between the current study and previous results may be attributable to a temporal mismatch between the validation data and ICESat-2 [28]. From another perspective, for global and national forest canopy height mapping, there are more complex elements to consider compared with regional studies. Both global and national data have achieved good accuracy [18,25,50,59]; however, part of the data may have lower accuracy due to geographical heterogeneity. Hence, forest canopy height mapping studies are important for forest resource management in the region.

4.3. Limitations and Prospects

In this study, 33 predictors were used to build a machine learning model to map wall-to-wall forest canopy height, including optical images and terrain features, as well as forest age, among others. Furthermore, a high-resolution forest canopy height map of the study area was mapped with a canopy height extrapolation model, which was developed using machine learning models. To investigate the potential of these variables in the estimation of forest canopy height, as seen in Figure 7, the feature variables in this study could be effectively linked to the canopy height in ICESat-2 points, with forest age performing particularly well as a new variable. Therefore, continuous forest canopy height maps can be used as predictors to map continuous forest structure parameters such as above-ground biomass and leaf area index in the future. Furthermore, due to latitudinal and longitudinal variances, environmental parameters such as precipitation and temperature should be incorporated as predictors for national or global forest canopy height mapping [50,58]. In terms of data, multicollinearity may have an impact on the ranking of importance in RF, and this issue should be taken into account in subsequent studies. In addition, several trials require additional consideration. ICESat-2 provides us with strip samples [19], which will leave many gaps that the GEDI [41] can fill in the study area. In terms of data, the number of vegetation signal photons in the ATL08 product is very low, and the vertical sampling error may underestimate the canopy height at the ICESat-2 point [20]. From another point of view, distinct forest species may have different canopy height extrapolation models [5]. Given the diverse biodiversity in the study area, canopy height extrapolation models for different tree species should be built once an appropriate categorization of the tree species is obtained.

5. Conclusions

In this study, we integrated ICESat-2, Sentinel, SRTM, and forest age data from Jiangxi Province, proposed a new method for variable selection, and developed four canopy height extrapolation models via RF, GBDT, KNN, and the SVM to link canopy height and spatial feature information. The model (RF) with the best-performance was used to map the forest canopy height in Jiangxi Province. Based on the results of this study, the following conclusions can be drawn: (1) Forest canopy height is moderately correlated with for-

est age, meaning it has the potential to be a predictor for forest canopy height mapping. (2) RF has the best-performance of the four canopy height extrapolation models, although it still underestimates higher forests and overestimates lower forests. Nonetheless, the use of multiple models can produce more robust results than the use of a single model. (3) Elevation, slope, and the red-edge band (band 5) derived from Sentinel-2 were significantly dependent variables in the canopy height extrapolation model. Apart from that, forest age was one of the variables that the RF moderately relied on. In contrast, backscatter coefficients and texture features derived from Sentinel-1 were not sensitive to canopy height. (4) There was a significant correlation between forest canopy height predicted using RF and forest canopy height measured using field measurements ($R^2 = 0.69$; RMSE = 4.02 m). In a nutshell, the results show that the method utilized in this work can reliably map the spatial distribution of forest canopy height.

Author Contributions: Conceptualization, Y.L., S.Q., K.L., B.H. and Y.T.; Methodology, Y.L. and B.H.; Formal analysis, Y.L. and K.L.; Software, Y.L.; Investigation, Y.L., S.Q. and K.L.; Resources, K.L. and S.Z.; Data curation, Y.L., S.Q. and S.Z.; Writing—original draft, Y.L.; Writing—review & editing, S.Q., S.Z., B.H. and Y.T.; Visualization, Y.L.; Project administration, S.Q.; Supervision, S.Q.; Funding acquisition, S.Q. and K.L. All authors have read and agreed to the published version of the manuscript.

Funding: The research was funded by the Water Conservancy Science and Technology Project of Jiangxi Province, China (No. 202124ZDKT25) and the National Natural Science Foundation of China (No. 41867012).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available from the corresponding author by reasonable request.

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

References

- Nandy, S.; Srinet, R.; Padalia, H. Mapping Forest Height and Aboveground Biomass by Integrating ICESat-2, Sentinel-1 and Sentinel-2 Data Using Random Forest Algorithm in Northwest Himalayan Foothills of India. *Geophys. Res. Lett.* **2021**, *48*, e2021GL093799. [\[CrossRef\]](#)
- Asner, G.P.; Hughes, R.F.; Mascaró, J.; Uowolo, A.L.; Knapp, D.E.; Jacobson, J.; Kennedy-Bowdoin, T.; Clark, J.K. High-Resolution Carbon Mapping on the Million-Hectare Island of Hawaii. *Front. Ecol. Environ.* **2011**, *9*, 434–439. [\[CrossRef\]](#)
- Asner, G.P.; Mascaró, J.; Muller-Landau, H.C.; Vieilledent, G.; Vaudry, R.; Rasamoelina, M.; Hall, J.S.; van Breugel, M. A Universal Airborne LiDAR Approach for Tropical Forest Carbon Mapping. *Oecologia* **2012**, *168*, 1147–1160. [\[CrossRef\]](#)
- Zeng, Z.; Piao, S.; Li, L.Z.X.; Zhou, L.; Ciais, P.; Wang, T.; Li, Y.; Lian, X.; Wood, E.F.; Friedlingstein, P.; et al. Climate Mitigation from Vegetation Biophysical Feedbacks during the Past Three Decades. *Nat. Clim. Chang.* **2017**, *7*, 432–436. [\[CrossRef\]](#)
- Xi, Z.; Xu, H.; Xing, Y.; Gong, W.; Chen, G.; Yang, S. Forest Canopy Height Mapping by Synergizing ICESat-2, Sentinel-1, Sentinel-2 and Topographic Information Based on Machine Learning Methods. *Remote Sens.* **2022**, *14*, 364. [\[CrossRef\]](#)
- Jiang, F.; Zhao, F.; Ma, K.; Li, D.; Sun, H. Mapping the Forest Canopy Height in Northern China by Synergizing ICESat-2 with Sentinel-2 Using a Stacking Algorithm. *Remote Sens.* **2021**, *13*, 1535. [\[CrossRef\]](#)
- Li, W.; Niu, Z.; Shang, R.; Qin, Y.; Wang, L.; Chen, H. High-Resolution Mapping of Forest Canopy Height Using Machine Learning by Coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 Data. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *92*, 102163. [\[CrossRef\]](#)
- Lu, D.; Chen, Q.; Wang, G.; Liu, L.; Li, G.; Moran, E. A Survey of Remote Sensing-Based Aboveground Biomass Estimation Methods in Forest Ecosystems. *Int. J. Digit. Earth* **2016**, *9*, 63–105. [\[CrossRef\]](#)
- Chopping, M.; Moisen, G.G.; Su, L.; Laliberte, A.; Rango, A.; Martonchik, J.V.; Peters, D.P.C. Large Area Mapping of Southwestern Forest Crown Cover, Canopy Height, and Biomass Using the NASA Multiangle Imaging Spectro-Radiometer. *Remote Sens. Environ.* **2008**, *112*, 2051–2063. [\[CrossRef\]](#)
- Balzter, H.; Luckman, A.; Skinner, L.; Rowland, C.; Dawson, T. Observations of Forest Stand Top Height and Mean Height from Interferometric SAR and LiDAR over a Conifer Plantation at Thetford Forest, UK. *Int. J. Remote Sens.* **2007**, *28*, 1173–1197. [\[CrossRef\]](#)
- Wagner, W. Large-Scale Mapping of Boreal Forest in SIBERIA Using ERS Tandem Coherence and JERS Backscatter Data. *Remote Sens. Environ.* **2003**, *85*, 125–144. [\[CrossRef\]](#)

12. Narine, L.L.; Popescu, S.C.; Malambo, L. Synergy of ICESat-2 and Landsat for Mapping Forest Aboveground Biomass with Deep Learning. *Remote Sens.* **2019**, *11*, 1503. [[CrossRef](#)]
13. Zwally, H.J.; Schutz, B.; Abdalati, W.; Abshire, J.; Bentley, C.; Brenner, A.; Bufton, J.; Dezio, J.; Hancock, D.; Harding, D.; et al. ICESat's Laser Measurements of Polar Ice, Atmosphere, Ocean, and Land. *J. Geodyn.* **2002**, *34*, 405–445. [[CrossRef](#)]
14. Zhu, X.; Wang, C.; Nie, S.; Pan, F.; Xi, X.; Hu, Z. Mapping Forest Height Using Photon-Counting LiDAR Data and Landsat 8 OLI Data: A Case Study in Virginia and North Carolina, USA. *Ecol. Indic.* **2020**, *114*, 106287. [[CrossRef](#)]
15. Lefsky, M.A. A Global Forest Canopy Height Map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System: A GLOBAL FOREST CANOPY HEIGHT MAP. *Geophys. Res. Lett.* **2010**, *37*. [[CrossRef](#)]
16. Simard, M.; Pinto, N.; Fisher, J.B.; Baccini, A. Mapping Forest Canopy Height Globally with Spaceborne Lidar. *J. Geophys. Res.* **2011**, *116*, G04021. [[CrossRef](#)]
17. Wang, Y.; Li, G.; Ding, J.; Guo, Z.; Tang, S.; Wang, C.; Huang, Q.; Liu, R.; Chen, J.M. A Combined GLAS and MODIS Estimation of the Global Distribution of Mean Forest Canopy Height. *Remote Sens. Environ.* **2016**, *174*, 24–43. [[CrossRef](#)]
18. Zhu, X. Forest Height Retrieval of China with a Resolution of 30m Using ICESat-2 and GEDI Data. Ph.D. Thesis, Aerospace Information Research Institute, Chinese Academy of Sciences (CAS), Beijing, China, 2021.
19. Neumann, T.A.; Martino, A.J.; Markus, T.; Bae, S.; Bock, M.R.; Brenner, A.C.; Brunt, K.M.; Cavanaugh, J.; Fernandes, S.T.; Hancock, D.W.; et al. The Ice, Cloud, and Land Elevation Satellite—2 Mission: A Global Geolocated Photon Product Derived from the Advanced Topographic Laser Altimeter System. *Remote Sens. Environ.* **2019**, *233*, 111325. [[CrossRef](#)]
20. Neuenschwander, A.; Pitts, K. The ATL08 Land and Vegetation Product for the ICESat-2 Mission. *Remote Sens. Environ.* **2019**, *221*, 247–259. [[CrossRef](#)]
21. Narine, L.L.; Popescu, S.; Zhou, T.; Srinivasan, S.; Harbeck, K. Mapping Forest Aboveground Biomass with a Simulated ICESat-2 Vegetation Canopy Product and Landsat Data. *Ann. For. Res.* **2019**, *62*, 69–86. [[CrossRef](#)]
22. Narine, L.L.; Popescu, S.; Neuenschwander, A.; Zhou, T.; Srinivasan, S.; Harbeck, K. Estimating Aboveground Biomass and Forest Canopy Cover with Simulated ICESat-2 Data. *Remote Sens. Environ.* **2019**, *224*, 1–11. [[CrossRef](#)]
23. Dubayah, R.; Blair, J.B.; Goetz, S.; Fatoyinbo, L.; Hansen, M.; Healey, S.; Hofton, M.; Hurtt, G.; Kellner, J.; Luthcke, S.; et al. The Global Ecosystem Dynamics Investigation: High-Resolution Laser Ranging of the Earth's Forests and Topography. *Sci. Remote Sens.* **2020**, *1*, 100002. [[CrossRef](#)]
24. Patterson, P.L.; Healey, S.P.; Ståhl, G.; Saarela, S.; Holm, S.; Andersen, H.-E.; Dubayah, R.O.; Duncanson, L.; Hancock, S.; Armston, J.; et al. Statistical Properties of Hybrid Estimators Proposed for GEDI—NASA's Global Ecosystem Dynamics Investigation. *Environ. Res. Lett.* **2019**, *14*, 065007. [[CrossRef](#)]
25. Potapov, P.; Li, X.; Hernandez-Serna, A.; Tyukavina, A.; Hansen, M.C.; Kommareddy, A.; Pickens, A.; Turubanova, S.; Tang, H.; Silva, C.E.; et al. Mapping Global Forest Canopy Height through Integration of GEDI and Landsat Data. *Remote Sens. Environ.* **2021**, *253*, 112165. [[CrossRef](#)]
26. Hansen, M.C.; Potapov, P.V.; Goetz, S.J.; Turubanova, S.; Tyukavina, A.; Krylov, A.; Kommareddy, A.; Egorov, A. Mapping Tree Height Distributions in Sub-Saharan Africa Using Landsat 7 and 8 Data. *Remote Sens. Environ.* **2016**, *185*, 221–232. [[CrossRef](#)]
27. Huang, W.; Min, W.; Ding, J.; Liu, Y.; Hu, Y.; Ni, W.; Shen, H. Forest Height Mapping Using Inventory and Multi-Source Satellite Data over Hunan Province in Southern China. *For. Ecosyst.* **2022**, *9*, 100006. [[CrossRef](#)]
28. Wu, Z.; Shi, F. Mapping Forest Canopy Height at Large Scales Using ICESat-2 and Landsat: An Ecological Zoning Random Forest Approach. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–16. [[CrossRef](#)]
29. Vuolo, F.; Zóltak, M.; Pipitone, C.; Zappa, L.; Wennig, H.; Immitzer, M.; Weiss, M.; Baret, F.; Atzberger, C. Data Service Platform for Sentinel-2 Surface Reflectance and Value-Added Products: System Use and Examples. *Remote Sens.* **2016**, *8*, 938. [[CrossRef](#)]
30. Zhang, T.; Liu, D. Mapping 30 m Boreal Forest Heights Using Landsat and Sentinel Data Calibrated by ICESat-2. *Authorea*, 6 December 2021.
31. Sothe, C.; Gonsamo, A.; Lourenço, R.B.; Kurz, W.A.; Snider, J. Spatially Continuous Mapping of Forest Canopy Height in Canada by Combining GEDI and ICESat-2 with PALSAR and Sentinel. *Remote Sens.* **2022**, *14*, 5158. [[CrossRef](#)]
32. Liu, T.; Abd-Elrahman, A.; Morton, J.; Wilhelm, V.L. Comparing Fully Convolutional Networks, Random Forest, Support Vector Machine, and Patch-Based Deep Convolutional Neural Networks for Object-Based Wetland Mapping Using Images from Small Unmanned Aircraft System. *GIScience Remote Sens.* **2018**, *55*, 243–264. [[CrossRef](#)]
33. Neuenschwander, A.; Guenther, E.; White, J.C.; Duncanson, L.; Montesano, P. Validation of ICESat-2 Terrain and Canopy Heights in Boreal Forests. *Remote Sens. Environ.* **2020**, *251*, 112110. [[CrossRef](#)]
34. Zhu, X.; Wang, C.; Xi, X.; Nie, S.; Yang, X.; Li, D. Research progress of ICESat-2/ATLAS data processing and applications. *Infrared Laser Eng.* **2020**, *49*, 20200259. [[CrossRef](#)]
35. Zhang, C.; Ju, W.; Chen, J.M.; Li, D.; Wang, X.; Fan, W.; Li, M.; Zan, M. Mapping Forest Stand Age in China Using Remotely Sensed Forest Height and Observation Data: CHINA'S FOREST STAND AGE MAPPING. *J. Geophys. Res. Biogeosciences* **2014**, *119*, 1163–1179. [[CrossRef](#)]
36. Yu, Z.; Zhao, H.; Liu, S.; Zhou, G.; Fang, J.; Yu, G.; Tang, X.; Wang, W.; Yan, J.; Wang, G.; et al. Mapping Forest Type and Age in China's Plantations. *Sci. Total Environ.* **2020**, *744*, 140790. [[CrossRef](#)]
37. Zhang, Y.; Yao, Y.; Wang, X.; Liu, Y.; Piao, S. Mapping Spatial Distribution of Forest Age in China. *Earth Space Sci.* **2017**, *4*, 108–116. [[CrossRef](#)]

38. Racine, E.B.; Coops, N.C.; St-Onge, B.; Bégin, J. Estimating Forest Stand Age from LiDAR-Derived Predictors and Nearest Neighbor Imputation. *For. Sci.* **2014**, *60*, 128–136. [[CrossRef](#)]
39. García, M.; Saatchi, S.; Ustin, S.; Balzter, H. Modelling Forest Canopy Height by Integrating Airborne LiDAR Samples with Satellite Radar and Multispectral Imagery. *Int. J. Appl. Earth Obs. Geoinformation* **2018**, *66*, 159–173. [[CrossRef](#)]
40. Markus, T.; Neumann, T.; Martino, A.; Abdalati, W.; Brunt, K.; Csatho, B.; Farrell, S.; Fricker, H.; Gardner, A.; Harding, D.; et al. The Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2): Science Requirements, Concept, and Implementation. *Remote Sens. Environ.* **2017**, *190*, 260–273. [[CrossRef](#)]
41. Liu, A.; Cheng, X.; Chen, Z. Performance Evaluation of GEDI and ICESat-2 Laser Altimeter Data for Terrain and Canopy Height Retrievals. *Remote Sens. Environ.* **2021**, *264*, 112571. [[CrossRef](#)]
42. Torres, R.; Snoeij, P.; Geudtner, D.; Bibby, D.; Davidson, M.; Attema, E.; Potin, P.; Rommen, B.; Floury, N.; Brown, M.; et al. GMES Sentinel-1 Mission. *Remote Sens. Environ.* **2012**, *120*, 9–24. [[CrossRef](#)]
43. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; . Sentinel-2: ESA’s Optical High-Resolution Mission for GMES Operational Services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [[CrossRef](#)]
44. Liu, Y.; Gong, W.; Xing, Y.; Hu, X.; Gong, J. Estimation of the Forest Stand Mean Height and Aboveground Biomass in Northeast China Using SAR Sentinel-1B, Multispectral Sentinel-2A, and DEM Imagery. *ISPRS J. Photogramm. Remote Sens.* **2019**, *151*, 277–289. [[CrossRef](#)]
45. Sanchez, A.H.; Picoli, M.C.A.; Camara, G.; Andrade, P.R.; Chaves, M.E.D.; Lechler, S.; Soares, A.R.; Marujo, R.F.B.; Simões, R.E.O.; Ferreira, K.R.; et al. Comparison of Cloud Cover Detection Algorithms on Sentinel-2 Images of the Amazon Tropical Forest. *Remote Sens.* **2020**, *12*, 1284. [[CrossRef](#)]
46. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *SMC-3*, 610–621. [[CrossRef](#)]
47. Jorge, J.; Vallbé, M.; Soler, J.A. Detection of Irrigation Inhomogeneities in an Olive Grove Using the NDRE Vegetation Index Obtained from UAV Images. *Eur. J. Remote Sens.* **2019**, *52*, 169–177. [[CrossRef](#)]
48. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The Shuttle Radar Topography Mission. *Rev. Geophys.* **2007**, *45*, RG2004. [[CrossRef](#)]
49. Chen, J.; Chen, J. GlobeLand30: Operational Global Land Cover Mapping and Big-Data Analysis. *Sci. China Earth Sci.* **2018**, *61*, 1533–1534. [[CrossRef](#)]
50. Liu, X.; Su, Y.; Hu, T.; Yang, Q.; Liu, B.; Deng, Y.; Tang, H.; Tang, Z.; Fang, J.; Guo, Q. Neural Network Guided Interpolation for Mapping Canopy Height of China’s Forests by Integrating GEDI and ICESat-2 Data. *Remote Sens. Environ.* **2022**, *269*, 112844. [[CrossRef](#)]
51. Wessler, C. The Spearman Correlation Formula. *Science* **1905**, *22*, 309–311. [[CrossRef](#)]
52. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
53. Friedman, J.H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
54. Guo, G.; Wang, H.; Bell, D.; Bi, Y.; Greer, K. KNN Model-Based Approach in Classification. In *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE.*; Meersman, R., Tari, Z., Schmidt, D.C., Eds.; Springer: Berlin/Heidelberg, Germany, 2003; Volume 2888, pp. 986–996. ISBN 978-3-540-20498-5.
55. Dimitriadou, E.; Hornik, K.; Leisch, F.; Meyer, D.; Weingessel, A. Misc Functions of the Department of Statistics (E1071), TU Wien. *R Package* **2008**, *1*, 5–24.
56. Kuhn, M. Building Predictive Models in R Using the Caret Package. *J. Stat. Softw.* **2008**, *28*, 1–26. [[CrossRef](#)]
57. Yang, Y.; Watanabe, M.; Li, F.; Zhang, J.; Zhang, W.; Zhai, J. Factors Affecting Forest Growth and Possible Effects of Climate Change in the Taihang Mountains, Northern China. *For. Int. J. For. Res.* **2006**, *79*, 135–147. [[CrossRef](#)]
58. Tao, S.; Guo, Q.; Li, C.; Wang, Z.; Fang, J. Global Patterns and Determinants of Forest Canopy Height. *Ecology* **2016**, *97*, 3265–3270. [[CrossRef](#)]
59. Lang, N.; Jetz, W.; Schindler, K.; Wegner, J.D. A High-Resolution Canopy Height Model of the Earth. *arXiv* **2022**, arXiv:2204.08322.

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.