Estimation of Broadleaf Tree Canopy Height of Wolong Nature Reserve Based on InSAR and Machine Learning Methods

Xinyi Liu 1,2, Li He 1,3,*, Zhengwei He 1,2 and Yun Wei 1,2

1 State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu 610059, China
2 College of Earth Sciences, Chengdu University of Technology, Chengdu 610059, China
3 College of Tourism and Urban-Rural Planning, Chengdu University of Technology, Chengdu 610059, China
* Correspondence: heli2020@cdut.edu.cn

Abstract: Tree height is an important parameter for calculating forest carbon sink and assessing forest carbon cycle. In order to obtain forest tree height over a large area both efficiently and at a low cost, this study proposed an Interferometric Synthetic Aperture Radar (InSAR) combined with a machine learning method to estimate the tree canopy height. The forest height in the study area was obtained using Unmanned Aerial Vehicle (UAV) photogrammetry, which was considered to be the true canopy height. Two machine learning methods (Random Forest, Multi-layer perceptron) were used to establish the relationship between phase center height calculated by InSAR DEM differential interference method and coherent amplitude method with true canopy height. The topographic factor, backward scattering coefficient and coherence coefficient were introduced into the relationship model. It was found that the accuracy of tree height estimation using random forest and two InSAR methods can reach 0.95 and 0.94. The root-mean-square error was 1.76 m, 1.86 m, respectively. The accuracy of tree height estimation using multi-layer perceptron and two InSAR methods was 0.25 and 0.2. The root-mean-square error was 3.96 m and 4.13 m. The results indicated that the combination of InSAR and machine learning can estimate canopy height efficiently and at a low cost. Moreover, the integrated learning algorithm random forest demonstrated better stability and higher accuracy than the single learning algorithm multi-layer perceptron.

Keywords: InSAR; canopy height; machine learning; UAV; broadleaf forest; Wolong Nature Reserve

1. Introduction

Forest height is one of the most important vertical structure parameters, which is not only an essential element for studies on species composition, stand quality and climate effects [1], but also a key parameter for simulating ecosystem service functions, forest fires and forest biomass monitoring [2–4]. Accurate estimation of forest vertical structure parameters is of great significance for global carbon cycle research and forest resource monitoring and management [5]. In the process of estimating forest height, the traditional method based on the ground sample standard survey is not only time-consuming and laborious, but it is also difficult to obtain spatially continuous measurements of forest parameters, which cannot satisfy the requirements of modern forest resource management and the ecological environment scientific research. Remote sensing technology provides an effective approach for estimating forest structure parameters at regional and global scales. Interferometric synthetic aperture radar (InSAR) has a series of characteristics such as high temporal resolution, all-weather, no cloud interference and strong penetration ability. It is one of the effective remote sensing tools for estimating forest parameters over large areas. Forest height estimation using InSAR is based on InSAR’s acquisition of the scattering phase center height of vegetation, which is affected by vegetation structure, scattering mechanisms and sensor characteristics (wavelength, polarization, viewing angle,
etc.) [6–10]. InSAR cannot be easily restricted by various climatic conditions and hence, can interact with the interior of vegetation canopy (e.g., X-band: frequency range 8–12.5 GHz, wavelength range 2.4–3.75 cm, C-band: frequency range 4–8 GHz, wavelength range 3.75–7.5 cm), and provide new information that is not available from visible and infrared remote sensing. Moreover, forest height estimation using InSAR only requires processing image information and does not require the establishment of relationships with other forest parameters. Now InSAR has been widely used in forest height studies [6,11,12].

There are two main ideas for estimating forest height using InSAR techniques: (1) based on phase information, forest height is estimated by interferometric phase difference representing two scattering mechanisms, top of canopy and understory structure, also known as the DEM differencing method; (2) based on coherence magnitude information, forest height is inferred from coherence information by establishing the relationship between body scattering decoherence and forest height (random body scattering model) [6]. In vegetation-covered areas, the information obtained from InSAR includes the height of the vegetation scattering phase center, which is the basis for forest height estimation. In the study of forest height estimation using the DEM difference method, Cloude et al. [13] first used Digital Elevation Model (DEM) differential inversion to obtain vegetation height from polarimetric interferometric SAR data in 1998, then various forest tree height inversion algorithms started to emerge. Izzawati et al. [14] used the DSM obtained by X-band InSAR to distinguish the height of plantation from Ordnance Survey DEM(OSDEM) data. Kenyi et al. [15] obtained forest height by differentiated processing of the C-band Space Shuttle Radar Terrain Mission (SKTM), DEM, and National Elevation Dataset (NED). For satellite-based SAR, Soja and Ulander [16] used TanDEM-X data combined with LiDAR-derived high-precision DEM to obtain forest height estimation results. Garestier et al. [17] found that the height difference between the phase centers of HV (cross polarization mode in which radar transmits electromagnetic waves horizontally and receives them vertically) and HH polarization (copolization mode in which radar transmits and receives electromagnetic waves horizontally) channels reflected the canopy height of pine forests. The multi-frequency interferometry technology uses the difference of penetration ability of forest wavelength and the difference of high and low frequency SAR to obtain the height of scattering phase center of vegetation body and woodland. Subsequently, it utilizes the phase difference to estimate the forest height. For example, the dual-frequency InSAR method uses the long-wavelength (e.g., L-band: frequency range 1–2 GHz, wavelength range 15–30 cm, P-band: frequency range 280–390 MHz, wavelength range 77–107 cm) InSAR data to penetrate the tree canopy, and then subtracts DTM from the interference surface height of single-channel short-wavelength band (such as X-band) to obtain the estimation of forest canopy height [18]. Pang et al. [19] used the C-band and L-band of SIR-SAR for interferometry to estimate the forest stand height. Santos et al. [20] used X-band and P-band InSAR phase center to obtain forest height. Balzter et al. [8] estimated forest canopy height while utilizing airborne X-band VV-polarized single-pass and L-band HH-polarized repeat-pass SAR interferometry. Sexton et al. [21] obtained forest height by differencing the heights obtained from X-band and P-band InSAR data. Moreover, in the study of forest height estimation the coherent magnitude method was used. Cloude proposed the coherent phase-amplitude inversion method in 2005, also known as SINC model, which demonstrated high stability and can ignore the influence of topographic factors [22]. The feasibility of SINC model was verified using TanDEM-X bipolarization data in high latitudes and temperate rainforests regions [23]. Qi et al. [24] used CAMSAR’s X-band to verify the SINC model and proved its higher practical value. Olesk et al. [25] utilized the SINC model to retrieve forest height from Tandem-X InSAR data and found that the estimation accuracy of deciduous season in broadleaf forest was higher than that of non-deciduous season.

However, when using the above two methods for forest height estimation, it is necessary to collect field data or obtain high-precision Light detection and ranging (LiDAR) data as the real tree height value for verification. These data can also be costly in terms of
labor and materials. In practical applications, we can establish the relationship between phase center height and true height depending on the vegetation type and wavelength, and then derive the true vegetation height [26]. Both classical statistical regression methods and machine learning methods can be used to establish relationships between variables. Compared with traditional statistical methods, machine learning methods have higher accuracy and are suitable for forest areas affected by multiple variables. At present, machine learning is widely used for the inversion of forest height. KS Džeroski [27] modelled vegetation height and canopy cover in Slovenia by combining LiDAR, Landsat satellite data and machine learning. WANG et al. [28] used spatially weighted geographic random forest (GRF) and traditional random forest (TRF) methods to predict the canopy height of mixed dry woodland in complex mountainous areas. Maryam et al. [29] used the machine learning method to combine multi-baseline polarization interferometric SAR (PolInSAR) data with LiDAR to improve the forest canopy height estimation.

However, there are few studies in the past that have used machine learning to establish the relationship between the phase center height acquired by InSAR and the real tree height acquired by UAV. Additionally, no matter which method was used in the previous study, both the image data used for tree height estimation and the measured data used for validation have high precision and large acquisition cost. The research objectives of this study are: 1) Establish the relationship between InSAR phase center height and real tree height using machine learning methods; 2) Verify the applicability of open-source InSAR data and UAV data in forest tree height estimation, then estimate the forest height efficiently and at a low cost. In this study, open source and easily accessible Sentinel-1A C-band data was selected, and two common machine learning methods, random forest and multi-layer perceptron (also known as Deep Learning network), were selected. They are also typical algorithms of integrated learning and single learner, respectively. In this study, the UAV images and InSAR data were utilized to obtain tree height data. The tree height extracted by the UAV demonstrated a small deviation from the actual tree height [30,31]. So, the tree height obtained by the UAV was considered the real tree height in this study. The height obtained by InSAR was considered as phase center height [26]. This study establishes the relationship between the phase center height which calculated by DEM differential interference method and coherent magnitude method with the real tree height, respectively. Finally, two machine learning methods were utilized to evaluate the accuracy of tree height estimation.

2. Materials and Methods
2.1. Study Area Overview

The Wolong Nature Reserve in Wenchuan County and Sichuan Province, with latitude 30°45’37” N~31°19’23” N and longitude 102°51’44” E~103°29’11” E, is the research object of this study while covering an area of 1736.81 km². Wolong Nature Reserve is located in the junction of the mountainous area and plain area, with the highest elevation of 6132 m in the west adjacent to Mountain Siguniang and the lowest elevation of 898 m in the east to Yingxiu Town of Wenchuan City, which is the end part of Zipingpu Reservoir, with an overall average elevation of 3544 m a.s.l. (Figure 1). The DEM data was obtained from ESA’s PALSAR 12.5 m resolution. The topography in the area was undulating, with an average slope of 31.4°. The slope of the area above 25° reached 74.7%, among which the slope of the Zhenghe River and the north, as well as the south sides of the Yuzi River, demonstrated the most obvious undulation (Figure 2).
The reserve belongs to the eastern edge of the climate zone of the Tibetan Plateau. Considering its geographical location and topography, it has formed a typical subtropical inland mountain climate. The area is dry in winter and has a lot of precipitation in summer. Three-fifths of the study area has more than 80% vegetation cover.

2.2. Sample Sites Data Processing

A total of 64 sample sites (Figure 2) were recorded from the field in October 2020, with elevations ranging from 940 m a.s.l. to 4193 m a.s.l. According to the community species composition, the 64 sample sites included 18 shrub forests, 9 tree forests, and 36 mixed tree and shrub forests. According to the nature of trees, there were 32 broad-leaved forests, 31 coniferous forests, and 1 bamboo forest. The distribution of vegetation in the sample plots demonstrated that the study area was rich in species diversity. As a large proportion of
trees and mixed trees occupy the study area, it is particularly important to estimate the height of trees.

The field measurements included geodetic values, elevation, slope, aspect, disturbance level, disturbance type, and community depression of each sample site. For trees greater than 1.5 m in height, the diameter at breast height, root diameter, tree height, crown width and LAI were measured. Considering the trees less than 1.5 m in height, the basal diameter, root diameter, tree height and crown width were measured. For shrubs, the clump diameter, root diameter, crown diameter and height were measured. Among them, the coordinates and elevation were measured using a handheld GPS locator (Zhitu H50). The locator utilized BeiDou, GPS, GLONASS multi-star high-precision positioning and coordinate systems including WGS84, Xian80, Beijing54 and CGS2000. The terrain was decided by a multifunctional slope measuring instrument (model JZC-B2), which had a high-strength level tube (double emphasis on the ring), a large dial and a slope angle comparison table. The LAI was measured using vegetation canopy analyzer LAI2000 and a portable app “Smart LAI”. Moreover, tape measurement was also utilized to determine the diameter at the breast height of trees.

2.3. UAV Image Processing and Extraction of Tree Height at the Sampling Site

The UAV equipment is DJI Phantom 4 Pro, which utilizes Flight Autonomy technology, quadrotor UAV as the flight platform with 6 vision sensors, main camera, 2 groups of IR sensors, 1 group of ultrasonic sensors, GPS/GLONASS dual-mode satellite positioning system, IMU and compass dual redundant sensors. DJIGS Pro was used as the ground station system for field image data acquisition. The effective pixels were 20 million for the single-lens visible sensor and the single-frame resolution was 5472 × 3648 pixels. The UAV images were acquired in October 2020. The flight height of the UAV was determined to be 120 m. The latitude cover was between 31°1′33″ N–31°1′40″ N, whereas the longitude cover was between 103°10′1″ E–103°10′8.4″ E. The total covered area was 0.578 km², with an average ground sampling distance (GSD) of 3.79 cm. The UAV flew in the horizontal direction according to “s” trajectory. The flight trajectory is shown in Figure 3a, and the map of the photographed features is shown in Figure 3b.

A total of five orthophoto maps corresponding to the sample areas were considered for field sampling. There are a total of 56 photos in the sample plot studied in this paper. The photogrammetric processing software Pix4D Mapper was used in the later stage to pre-process the acquired raw images for aerial photography, map formation, feature point extraction and matching, automatic aerial triangulation and digital orthophoto generation.
The pre-processing process mainly included: (1) acquiring raw data, i.e., confirming the integrity of raw image data, viewing POS data files, camera parameters and control point files; (2) establishing the project and importing data, i.e., creating the project in Pix4D Mapper, including initialization settings, point cloud encryption, regional 3D reconstruction and orthophoto generation.

In this study, as the tree canopy height was estimated for the study area, the sample area with high forest depression was selected. The field sample site photo and the 3D point cloud images generated by the UAV are shown in Figure 4. The Digital Surface Model (DSM) and Digital Terrain Model (DTM) generated by the UAV with a resolution of 3.79 cm are shown in Figure 5. The GSD criterion was used to define the spatial resolution of DSM/DTM. The threshold segmentation used for point cloud segmentation is the most common method to determine the maximum variance between classes, namely the OTSU method [26].

![Field sample site photo (a) and 3D point cloud map (b)](image)

**Figure 4.** Field sample site photo (a) and 3D point cloud map (b).

![DSM (a) and DTM (b)](image)

**Figure 5.** DSM (a) and DTM (b).

The mathematical expression of OTSU is mentioned as follows:

Assuming the threshold value as "t", it falls between the maximum value $Z_{\text{max}}$ and the minimum value $Z_{\text{min}}$. The point cloud height value $Z_i$ is not bigger than t for $\{Z_1, Z_2, \ldots, Z_l\}$, and bigger than t for $\{Z_{l+1}, Z_{l+2}, \ldots, Z_n\}$. The threshold value "t" classifies the point clouds into two classes as ground point clouds G (ground) and tree point clouds T (tree). Therefore, the probability of one of the point clouds can be classified into two classes as:

\[
P_G(t) = \frac{1}{n} P_G(t) = \frac{n-1}{n}
\]

(1)
where, \( P_G(t) \) is the probability of the point cloud assigned to the ground point and \( P_T(t) \) is the probability of the point cloud assigned to the tree point.

The mean height of tree point clouds was mentioned as “\( \mu_T(t) \)”, whereas the mean height of ground point clouds was noted as “\( \mu_G(t) \)”. Moreover, the mean height of all point clouds was noted as “\( \mu \)”. The mean squared differences between the two types of point clouds were termed as \( \sigma^2_T(t) \) and \( \sigma^2_G(t) \). The mean-variance of all point cloud heights was “\( \sigma^2 \)”. For the point cloud of the same tree, the mean height “\( \mu \)” and the mean square height “\( \sigma^2 \)” were fixed. In addition, there was no relationship with the threshold “\( t \)”. The intra-class variance of the two classes of point cloud heights segmented by the threshold “\( t \)” was noted as “\( \sigma^2_B(t) \)”. The inter-class variance of the two classes of point cloud heights segmented by the threshold “\( t \)” was termed as “\( \sigma^2_{GT}(t) \)”. According to the principle of selecting the optimal threshold, once “\( \sigma^2_{GT}(t) \)” is largest, the value of “\( t \)” will be optimal. After finding the optimal threshold \( t \), the point cloud was divided into two parts as ground point cloud and tree point cloud, followed by the generation of DSM and DTM for sample areas.

As shown in Figure 4a and based on the field sample survey, the canopy closure in the area was between 0.7–1. There were 60 trees in the 10 m × 10 m sample square with a density of 1, while most of the trees are in the growing stage. The tree crown width was between 1–2 m, and the average tree height was about 9–11 m. The tree forest occupied about 40% of the UAV area (Figure 3). While using the systematic sampling method, 420 sample points were taken with a sampling spacing of 20 m. After eliminating the non-tree areas and some sample points located in the clearing, 152 sampling points were finally left (Figure 6).

Figure 6. Sample area sampling points.

Since the resolution of UAV data and InSAR data were 0.038 m and 20 m, respectively, it was necessary to resample the UAV image to 20 m. By using the DSM generated by UAV and DTM, the Canopy Height Model (CHM) was calculated, CHM value, slope value, the aspect value generated by DTM were extracted, the height generated by InSAR was also calculated.
2.4. InSAR Data Processing

The Sentinel-1A used in this study was a Level-1 data product for SLC images in IW mode, which was imaged on 1 October 2020 and 25 October 2020, with VH and VV polarization and 5 m × 20 m ground resolution. The Sentinel-1A data was from European Space Agency (ESA)’s Copernicus program.

The process of generating InSAR was performed by SNAP software. The first step was to register the image with parameters set that included (1) setting the master and slave images; (2) setting the subswath; (3) setting the satellite orbit file; and (4) setting the auxiliary DEM data. The DEM data was obtained from ALOS PALSAR 12.5 m DEM and the orbit file was from ESA’s official track information website. The second step was to form the interferogram. The corresponding interference results were obtained based on Orb_Stack generated in the previous step. Subsequently, the topographic phase was removed, and the terrain was corrected using 12.5 m DEM. To reduce noise for multi-look processing, Gaussian filtering was utilized, followed by Goldstein phase filtering. In the final stage, the geocoding of the results was conducted.

According to the InSAR inversion principle of forest structure parameters (Figure 7) [8,10], the change in scattering characteristics of target ground objects and the elevation of corresponding points can be obtained by the interference phase difference of radar images in different periods when ignoring the influence of different factors in the electromagnetic wave transmission process. For different polarizations at the same wavelength, the ability of cross-polarization to receive echo signals was weaker than that of co-polarization. It is generally believed that the echoes of cross-polarization come from the canopy, while co-polarization comes from the ground. This causes different polarization methods to obtain different radar backscattering coefficients. In this study, cross-polarization VH and co-polarization VV of Sentinel-1A radar was used for calculation. The main image was acquired on 1 October 2020, while the second image was acquired on October 15, 2020. The VH and VV radar images were interferometrically processed based on different interferometric phases corresponding to different ground elevations. The PALSAR DEM data of 12.5 m resolution was subsequently utilized to simulate and remove the topographic phases from the interferogram of the study area. In the calculation process, the body scattering dominated the echo signal of the VH cross-polarization mode, whereas the surface scattering dominated the echo signal of the VV co-polarization mode.

![Figure 7. Geometric relations for SAR interferometry.](image-url)

A1 and A2 represent two antenna positions, the baseline B is the distance between A1 and A2, r1 and r2 are the distance of the antenna to ground point P, h is the elevation of P, P0 is the point on the reference plane equidistant from A1 to P, θ and θ0 are the point view of A to P and P0, φθ is the amount of change in the view of A1 from P0 to P. The angle...
between the baseline and the horizontal direction is $\alpha$, and $B_\perp$ is the component of $B$ in the vertical SAR line of sight direction.

2.4.1. DEM Differential Interference Method

The basic idea of this method is to calculate the canopy phase and ground phase separately, while using the vertical effective wave number $k_z$ and the canopy-ground phase difference to obtain the vegetation height. In the polarization system, the representative scattering becomes closer to the plane scattering once the effective ground-to-body amplitude ratio of a certain polarization channel becomes larger. In this case, the scattering center phase can be regarded as ground phase. If the VH polarization is dominated by body scattering and VV polarization is dominated by ground scattering, then the calculation formula can be presented as Equation (2).

$$\phi_s = \arg(\gamma_{VV}), \phi_v = \arg(\gamma_{VH})$$

(2)

$\gamma_{VV}$ represents the phase after interference processing in the co-polarization mode; $\gamma_{VH}$ represents the phase after interference processing in the cross-polarization mode; $\phi_v$ represents the scattering phase center of the vegetation canopy, and $\phi_s$ represents the scattering phase center of the ground surface.

The height of the vegetation $h_v$ can be represented by Equation (3).

$$h_v = \frac{\phi_v - \phi_s}{k_z}$$

(3)

Out of which, the “$k_z$” can be expressed as:

$$k_z = \frac{2p\pi B_\perp}{\lambda R \sin \theta}$$

(4)

where $k_z$ is vertical effective wave number, $p$ denotes the InSAR data acquisition mode, $B_\perp$ is the vertical baseline, $\lambda$ is wavelength, $\theta$ is the radar view, and $R$ is the slope distance [10]. If the vegetation canopy is high and sparsely distributed, then the extinction coefficient will be small and the phase center will be high due to the structure. However, if the vegetation canopy is densely distributed, the phase center may be located anywhere between half of the tree height and the canopy, depending on the vertical structural variation of the vegetation and the average extinction coefficient.

2.4.2. Coherent Magnitude Method

The coherence magnitude method is a model based on the theory of interference decoherence. Due to the short time interval of the C-band InSAR data used in this study, the problem of temporal decoherence was not considered. The C-band InSAR data is dominated by body scattering. Using body scattering decoherence model, also known as SINC model, can estimate the forest height.

If the scattering model is considered as an independent vegetation body, the amplitude ratio of the ground to body is considered to be reduced to 0. If the extinction coefficient was 0, the coherence coefficient is only related to the body scattering coherence. If the ground reflection and extinction coefficients were ignored, the coherence coefficient amplitude can only relate to the tree height. Once the backscattered energy of the vegetation layer is evenly distributed, the coherent tree height $z$ will be consistent with the SINC model. The body scattering decoherence model is expressed as:

$$\gamma = \exp\left(\frac{k_z h_v}{2}\right) \sin \left(\frac{1}{2} k_z h_v\right)$$

(5)
Body scattering decoherence $\gamma$ and forest height $h_0$ are satisfied as a function of SINC [10] and the simple inversion model between the coherence coefficient amplitude and the tree height can be established using the SINC function as follows:

$$h_0 \approx \frac{2\pi}{k} \left( \pi - 2 \sin^{-1}\left(\frac{|\gamma|^{0.8}}{\pi}\right) \right)$$

(6)

This method only considers the amplitude information of the complex coherence coefficients and does not consider its phase information. Therefore, this method is also called the coherent amplitude method.

2.5. Machine Learning Methods Used for Tree Height Estimation

Two machine learning methods-random forest (RF) and multilayer perceptron (MLP) were used to combine the phase center height generated by InSAR technology with canopy height generated by UAV. Both of them were run in anaconda using python. Each method included three steps. Firstly, the input variables of two methods were phase center height, slope, aspect, backscattering coefficient and coherence coefficient, in which the topography factor is removed from the input variables when using the coherent magnitude method, and the output variables was canopy height. Secondly, the dataset was divided into two sets: 70% of dataset was taken as the training set, and 30% of them was considered as the testing set. Finally, the testing set was used to evaluate the performance of each method in estimating tree height.

The main idea of modeling using random forests is as follows.

1) The number of training cases (samples) $N$ was 152 and the number of features $M$ was 5;
2) The number of input features, $m$, was used to determine the decision outcome of a node in the decision tree, where $m$ was much smaller than $M$;
3) From the 152 training cases (samples), a training set was formed by put-back sampling (i.e., bootstrap sampling) 106 times, and the unsampled cases (samples) were used as predictions to evaluate their errors;
4) For each node, “$m$” features were randomly selected and the decision of each node in the decision tree was determined based on these features. According on these $m$ features, its optimal split was calculated;
5) Each tree grew intact without pruning, which was likely to be adopted after building a normal tree classifier.

For the multilayer perceptron approach, we complied with the main principle of multilayer perceptron. This experiment fixed the number of hidden layers to 1, which can fit any function “containing a continuous mapping from one finite space to another finite space” [32]. The number of neurons in the input layer was 5 and the number of neurons in the output layer was 1.

3. Results

3.1. Visualization and Comparative Analysis of Model Parameters

Two machine learning methods were used to build the regression models in this study. Figures 8 and 9 show the input and output tree height data of the model, and other model input parameters, respectively. Figure 9 shows the box plot and line plot of the three-tree height distribution of some sample points, namely, the actual tree height, the tree height of DEM differential interference method and the tree height of coherence amplitude method.

As shown in Figure 8a, there was no obvious correlation between the phase center height values obtained by the DEM differential interference method and the actual height. The correlation between the two was calculated and the correlation coefficient was determined to be 0.032. Its histogram distribution also differed significantly from that of the real tree height and the overall estimated tree height was slightly higher than the actual tree height. The relationship between them was complex rather than a simple linear relationship.
The mean value of the actual tree height obtained from the UAV photogrammetry was 9 m, while the mean value of the phase center height obtained from the DEM differential interference method was 16 m. As shown in Figure 8b, there was a large gap between them, which was caused by the uncertainty of the phase center of the HV polarized canopy. It was influenced by the extinction coefficient of the forest and the spatial variation of the vertical structure of the canopy. Even if the extinction coefficient was small, the phase center might change due to the structural influence of the trees themselves. On the other hand, the vegetation scattering information is the backscattered echoes received by the HV polarization channel. Moreover, the intermixed two-way scattering can also affect the effective phase center estimation of canopy scattering. In addition, the resolution of the SAR image was 20 m and the UAV image resolution was 3.79 cm. The InSAR data used C-band in this study, which was weak during penetration while being easily affected by terrain. Therefore, a large gap between the two results resulted.

![Figure 8. Comparison line plots (a) and box plots (b).](image)

![Figure 9. The slope map (a) aspect map (b) and coherence coefficient map (c) of study area.](image)

However, the results of the phase center height established by the coherence amplitude method have the same distribution trend as the actual tree height line graph distribution. The correlation between the two was calculated and the correlation coefficient was determined to be 0.379. Although the inversion mean value of SINC model was around 20 m, the trend of its inversion value was closer to the actual tree height value. It was slightly higher than the actual value as this algorithm simplified the interference factor as much as
possible. The obtained echo information cannot be pure body scattering information and the extinction coefficient cannot completely converge to 0. As the topography and stand structure of each sample site were different, it caused different degrees of influence on the inversion results. The use of SINC function for tree height inversion mainly utilized the coherence coefficient in the image, which was not influenced by the topography. Therefore, the extracted height values were closer to the actual values.

The topography of the study area was found to be complex. The tree height distribution obtained by the DEM differential interference method indicated that the tree height obtained by this method was significantly affected by the topography. The slope orientation data of the study area were extracted using 12.5 m DEM data and the backscatter coefficients of the sample sites were extracted from the SAR images. Moreover, the coherence coefficient maps were generated by InSAR (Figure 9).

### 3.2. Tree Height Estimation Based on InSAR and Random Forest

The two-phase center heights obtained by DEM difference method and coherent magnitude method based on InSAR technology were modeled with the real tree heights obtained by the UAV for random forest with $R^2$ values of 0.95 and 0.94, respectively. Figure 10 shows the difference between the model predictions and the true values when estimating the forest height using random forest. Figure 11 shows the accuracy and fitted equations of the random forest model for estimating tree height. Figures 12 and 13 show the prediction residuals of the DEM difference method, coherent magnitude method with random forest for modeling, respectively.

**Figure 10.** Fitted results of two InSAR methods with random forest. The left (a) represents a comparison between interference difference tree heights fitted using random forest and the right (b) represents a comparison of coherent amplitude tree heights fitted using random forest.

**Figure 11.** DEM differential interference method with random forest estimation of tree height precision and fitting equation (a). Coherent amplitude method with random forest estimation of tree height precision and fitting equation (b).
As shown in Figure 11a, the regression model established between the phase center height measured by the DEM differential interference method and actual tree height using random forest had a root mean square error of 1.76 m. The linear fit between the predicted and true values was 0.95, which indicated that the forest tree height could be estimated relatively well using the random forest algorithm. It also indicated that the tree height values measured using differential interferometry could be interconverted with the actual tree height after correction for topographic factors. The use of the random forest algorithm can solve the problem associated with the scattered phase center height and canopy height of C-band SAR data having different resolutions and complex terrain factors.

As shown in Figure 11b, the regression model between the phase center height of the coherent magnitude method and actual tree height using random forest had a root mean square error of 1.86 m. The linear fit between the predicted and true values was 0.94. Random forest is still valid for the height measurement of the coherent magnitude method. This model eliminates the terrain factor from the input parameters. It further illustrates that the SINC model was not affected by the terrain factor in the inversion of tree height.

As shown in Figure 12, the residual distribution was consistent with the model assumptions. The variables fluctuated around the zero value while indicating the use of the DEM differential method combined with random forest can effectively estimate tree height and resolve the canopy phase center uncertainty caused by topographic problems or single-baseline radar systems.

As shown in Figure 13, the residual distribution of the estimated tree height using the coherent magnitude method was consistent with the model assumptions. As the variables fluctuated around the zero value, it indicated that the use of the SINC model in combination with random forest can effectively estimate the tree height without considering the influence of topographic factors.

3.3. Tree Height Estimation Based on InSAR and Multilayer Perceptron

The two-phase center heights obtained using the DEM difference method and the coherent magnitude method based on InSAR technology were modeled with the real tree heights obtained by the UAV for multilayer perceptron with $R^2$ values of 0.25 and 0.2, respectively. Figure 14 shows the difference between the model predictions and the true values when estimating the forest height using multilayer perceptron. Figure 15 shows the accuracy and fitted equations of the multilayer perceptron model for estimating tree height. Figures 16 and 17 show the prediction residuals of the DEM difference method, coherent magnitude method with multilayer perceptron for modeling, respectively.
random forest had a root mean square error of 1.76 m. The linear fit between the predicted and true values was 0.95, which indicated that the forest tree height could be estimated relatively well using the random forest algorithm. It also indicated that the tree height values measured using differential interferometry could be interconverted with the actual tree height after correction for topographic factors. The use of the random forest algorithm can solve the problem associated with the scattered phase center height and canopy height of C-band SAR data having different resolutions and complex terrain factors.

As shown in Figure 11b, the regression model between the phase center height of coherent magnitude method and actual tree height using random forest had a root mean square error of 1.86 m. The linear fit between the predicted and true values was 0.94. Random forest is still valid for the height measurement of the coherent magnitude method. This model eliminates the terrain factor from the input parameters. It further illustrates that the SINC model was not affected by the terrain factor in the inversion of tree height.

As shown in Figure 12, the residual distribution was consistent with the model assumptions. The variables fluctuated around the zero value while indicating the use of the DEM differential interference method combined with random forest can effectively estimate tree height and resolve the canopy phase center uncertainty caused by topographic problems or single-baseline radar systems.

As shown in Figure 13, the residual distribution of the estimated tree height using the coherent magnitude method was consistent with the model assumptions. As the variables fluctuated around the zero value, it indicated that the use of the SINC model in combination with random forest can effectively estimate the tree height without considering the influence of topographic factors.

3.3. Tree Height Estimation Based on InSAR and Multilayer Perceptron

The two-phase center heights obtained using the DEM difference method and the coherent magnitude method based on InSAR technology were modeled with the real tree heights obtained by the UAV for multilayer perceptron with R² values of 0.25 and 0.2, respectively. Figure 14 shows the difference between the model predictions and the true values when estimating the forest height using multilayer perceptron. Figure 15 shows the accuracy and fitted equations of the multilayer perceptron model for estimating tree height. Figures 16 and 17 show the prediction residuals of the DEM difference method, coherent magnitude method with multilayer perceptron for modeling, respectively.

Figure 14. The fitted results of two InSAR methods with multilayer perceptron. The left one (a) represents the comparison of interference difference tree heights fitted using multilayer perceptron, and the right one (b) shows the comparison of coherent amplitude tree heights fitted using multi-layer perceptron.
Figure 15. DEM differential interference method with multilayer perceptron estimation of tree height precision and fitting equation (a). Coherent amplitude method with multilayer perceptron estimation of tree height precision and fitting equations (b).

As shown in Figure 15a, the root mean square error of tree height prediction using multilayer perceptron was found to be 3.96 m, and the linear fit between the predicted and true values was 0.25. Its accuracy is lower than the accuracy of 0.95 by using random forest. The reason for this result could be attributed to the random forest which is an ensemble learning algorithm with strong model generalization ability. Even though the multilayer perceptron can be used to infer unknown relationships between data, its arithmetic power is still limited. Moreover, a multilayer perceptron is a single learner with the characteristic of a “black box”, while having slightly lower stability than that of the random forest.

As shown in Figure 15b, the root mean square error of tree height prediction using multi-layer perceptron was 4.13 m, and the linear fit between predicted and true values was...
0.2. In addition, there was a large error between model prediction results and the actual values. The multi-layer perceptron is not suitable for simulating forest tree height. The simulation accuracy is much less than using the random forest method.

As shown in Figures 16 and 17, the residuals of predicted tree height using multilayer perceptron fluctuated widely. It could be assumed that the error with equal variance did not hold, the performance of multilayer perceptron in tree height estimation is worse than random forest.

4. Discussion

This study successfully established the model relationship between the phase center height and the real tree height and proved that the forest tree height can be effectively estimated by the combination of InSAR technique and machine learning methods. The method proposed in this paper allows us to obtain the forest height over a large area at low cost and effectively, where the combination of DEM difference method and random forest can achieve higher estimation accuracy, while the combination of coherent magnitude method and random forest can ignore the influence of topographic factors.

4.1. Analysis of Tree Height Estimation Results

The determination of forest tree height is a key element in studies on the estimating forest parameters such as forest biomass. In previous studies, two main methods using UAV airborne LiDAR and photogrammetry were used to extract tree height. The airborne LiDAR can provide detailed estimation of forest vertical scale information [33] and LiDAR-generated 3D point clouds can segment tree structures to improve the estimation accuracy. However, using airborne LiDAR for multi-time and large-scale forest monitoring is usually difficult in practical applications. While using UAV photogrammetry for tree height extraction can also achieve higher accuracy [34,35]. Therefore, the UAV tree height in this study can be regarded as the real tree height. However, while estimating forest tree
height over a large area, InSAR data is still considered the mainstream. The difficulties in estimating forest tree height using InSAR data currently lie in the acquisition of high-precision data and mechanistically complex parameter models. While using InSAR data to invert forest tree heights, Feng et al. [36] used an improved three-stage algorithm that considered topographic factors to extract tree heights in sample areas along transmission line corridors based on ALOS PALSAR dual-polarization satellite data. They verified the accuracy of extracted tree heights using field-collected airborne LiDAR data. Liao [12] used P-band SAR data, PolInSAR and TomoSAR techniques to conduct an in-depth study on the influence of complex subsurface, topographic relief, and saturation in the estimation of forest above-ground forest biomass. It is closely related to the forest structure parameters. In addition, the use of InSAR data for canopy height inversion can also be affected by surface diagonality. Although L-band and P-band data have strong penetration, they can also have large acquisition costs, in comparison to C-band data, which are easier to obtain and demonstrate better results in forests with low biomass. The analysis of C-band interferometric height discontinuity at forest edges reported by Hagberg et al. [37] and the effective interferometric C-band height from ERS-1 reported by Askne et al. [38] proved its validity.

This study proved that combining C-band InSAR data and machine learning methods can establish the relationship between phase center height and real tree height. It can efficiently and cost-effectively estimate the forest height values over a large area, and also solve the problems of high cost of data acquisition and complicated inversion models in previous studies. Nasiri et al. [39] proved the suitability of integrating UAV-obtained RGB images, Sentinel-2 data, and ML models for the estimation of FCC, Yu et al. [40] found that the forest structure classification achieved from two-seasonal UAV optic images and LiDAR-derived DSMs can reach 0.9 with the application of an optimal machine learning approach. A combination of InSAR techniques, aerial photography images, and machine learning has been shown to be useful in the study of forest structure parameters. For this study, our results showed that the use of DEM difference method combined random forest was influenced by topography with the highest accuracy of 0.95 and a root-mean-square error of 1.76 m, the accuracy is significantly improved compared to the root-mean-square error of 3.5 m and 7 m in the height extraction studies of Balzter et al. [8] and Kenyi et al. [41] while using SRTM C-band and USGS DTM data. The estimation accuracy of the coherent amplitude method combined random forest regression model can reach 0.94 and the root-mean-square error was 1.86 m without considering phase information, it also indicated that using coherent amplitude method to estimate tree height is less affected by topographic factors. However, the performance of multilayer perceptron method was poor when compared with RF. The accuracy of two InSAR methods combined with multilayer perceptron was 0.25 and 0.2, and the root-mean-square error was 3.96 m and 4.13 m. However, the root mean square error of InSAR combined MLP is still better than the results of the previous study. The two machine learning methods used in this study have their own advantages. For example, RF has characteristics such as fast training, high accuracy and resistance to overfitting [42], MLP has strong nonlinear fitting ability with multiple layers to learn the representation in data [43]. However, this experiment showed that, despite the advantages of each of the two machine learning methods, RF performs better than MLP in estimating the forest height, the stability and accuracy of MLP which is a single learner is slightly inferior compared RF.

4.2. Analysis of Uncertainty of This Study

In this study, the tree height obtained by UAV photogrammetry was used as the real tree height. Although it has been proved that the tree height value obtained by UAV can be used as the real tree height value in the actual study, there are still errors using this method, which come from the influence of terrain in field measurement, human operation error, sensor error, etc.
The InSAR data used in the study is the open-source Sentinel-1A data, which is one of the main features of this study. The choice of data source does not require an excessive cost, but it also brings some uncertainties. The radar scattering mechanism in forest areas is complex and difficult to control, radar data with longer wavelengths may have a better performance. In addition, InSAR and UAV have different resolutions and have some computational complexity in scale conversion.

In the selection of machine learning methods, we chose two machine learning algorithms that are more commonly used and have certain representativeness, then compared the estimation accuracy of both. However, the current machine learning models are diverse, exploring machine learning tree-height estimation models with higher accuracy is a development direction for future research.

Besides, there are some limitations in this study; the object is broadleaf forest in the growing stage, and the applicability of this approach to coniferous forest and broadleaf forest in the mature stage is unclear. The complexity of the method lies in the inconsistent resolution of the InSAR and UAV images, the selection of the scale conversion method needs to be further explored. In future studies, the research object can be extended to other natural forests or plantations, the InSAR data used could be converted from repeat-pass InSAR to single-pass InSAR, and machine learning methods with higher accuracy can be developed.

5. Conclusions

This study used two InSAR technologies (DEM difference method and coherent magnitude method) to calculate forest tree height and relate it to real tree height through two machine learning methods (RF, MLP). The experimental results demonstrated that the accuracy of forest tree height estimation used DEM difference method combined with machine learning can be as high as 0.95 when topographic factors were added, while the accuracy of forest tree height estimation using the coherent magnitude method combined with machine learning is not affected by topographic factors and can be as high as 0.94. Among them, RF had a better prediction ability than MLP, which indicated that the use of C-band InSAR data and machine learning method can effectively save costs and greatly reduce the difficulty of acquiring actual field data and remote sensing high-precision data. InSAR technology combined random forest method had gained efficient results in estimating forest height. This study used sentinel-1 repeat-pass InSAR and two machine learning methods to estimate broadleaf forest canopy height. In future studies, we suggest to further explore the applicability of single-pass InSAR in tree height estimation and extended this approach to coniferous forest and plantations and developed machine learning models with higher inversion accuracy.

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

Funding: This research was supported by Natural Science Foundation of Sichuan Province (2022NSFSC1040) and supported by the Independent Research Project of the State Key Laboratory of Geohazard Prevention and Geoenvironment Protection Independent Research Project (SKLG2021Z003).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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


