forests

Special Issue Reprint

# LiDAR Remote Sensing for Forestry

Edited by
Zhenyang Hui, Penggen Cheng, Bo Liu and Mark Vanderwel

mdpi.com/journal/forests

MDPI

# LiDAR Remote Sensing for Forestry

# LiDAR Remote Sensing for Forestry

Guest Editors

**Zhenyang Hui**
**Penggen Cheng**
**Bo Liu**
**Mark Vanderwel**

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

*Guest Editors*

Zhenyang Hui
School of Surveying and Geoinformation Engineering
East China University of Technology
Nanchang
China

Penggen Cheng
School of Surveying and Geoinformation Engineering
East China University of Technology
Nanchang
China

Bo Liu
School of Surveying and Geoinformation Engineering
East China University of Technology
Nanchang
China

Mark Vanderwel
Department of Biology
University of Regina
Regina, SK
Canada

*Editorial Office*

MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Forests* (ISSN 1999-4907), freely accessible at: https://www.mdpi.com/journal/forests/special_issues/0WY4ROM2E4.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

**ISBN 978-3-7258-8807-8 (Hbk)**
**ISBN 978-3-7258-8808-5 (PDF)**
**https://doi.org/10.3390/books978-3-7258-8808-5**

© 2026 by the authors. Articles in this reprint are Open Access and distributed under the Creative Commons Attribution (CC BY) license. The reprint as a whole is distributed by MDPI under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

# Contents

**About the Editors** . . . . . **vii**

**Preface** . . . . . **ix**

**Zhenyang Hui, Lei Lin, Shuanggen Jin, Yuanping Xia and Yao Yevenyo Ziggah**
A Reliable DBH Estimation Method Using Terrestrial LiDAR Points through Polar Coordinate Transformation and Progressive Outlier Removal
Reprinted from: *Forests* **2024**, *15*, 1031, https://doi.org/10.3390/f15061031 . . . . . **1**

**Longwei Li, Linjia Wei, Nan Li, Shijun Zhang, Zhicheng Wu, Miaofei Dong and Yuyun Chen**
Extracting the DBH of Moso Bamboo Forests Using LiDAR: Parameter Optimization and Accuracy Evaluation
Reprinted from: *Forests* **2024**, *15*, 804, https://doi.org/10.3390/f15050804 . . . . . **18**

**Yongkyu Lee and Jungsoo Lee**
Evaluation of Accuracy in Estimating Diameter at Breast Height Based on the Scanning Conditions of Terrestrial Laser Scanning and Circular Fitting Algorithm
Reprinted from: *Forests* **2024**, *15*, 313, https://doi.org/10.3390/f15020313 . . . . . **34**

**Iván Raigosa-García, Leah C. Rathbun, Rachel L. Cook, Justin S. Baker, Mark V. Corrao and Matthew J. Sumnall**
Rethinking Productivity Evaluation in Precision Forestry through Dominant Height and Site Index Measurements Using Aerial Laser Scanning LiDAR Data
Reprinted from: *Forests* **2024**, *15*, 1002, https://doi.org/10.3390/f15061002 . . . . . **50**

**Yuncheng Deng, Jinliang Wang, Pinliang Dong, Qianwei Liu, Weifeng Ma, Jianpeng Zhang, et al.**
Registration of TLS and ULS Point Cloud Data in Natural Forest Based on Similar Distance Search
Reprinted from: *Forests* **2024**, *15*, 1569, https://doi.org/10.3390/f15091569 . . . . . **71**

**Shengbo Yu, Dajun Li, Xiaowei Xie, Zhenyang Hui, Xiaolong Cheng, Faming Huang, et al.**
Multi-Constraint and Shortest Path Optimization Method for Individual Urban Street Tree Segmentation from Point Clouds
Reprinted from: *Forests* **2026**, *17*, 27, https://doi.org/10.3390/f17010027 . . . . . **94**

**Song Chen, Ming Gong, Hua Sun, Ming Chen and Binbin Wang**
Estimation of Forest Canopy Height from Spaceborne Full-Waveform LiDAR Data Using a Bisection Approximation Decomposition Method
Reprinted from: *Forests* **2025**, *16*, 145, https://doi.org/10.3390/f16010145 . . . . . **118**

**Jiapeng Huang, Yunqiu Wang and Yang Yu**
Multi-Criteria Filtration and Extraction Strategy for Understory Elevation Control Points Using ICESat-2 ATL08 Product
Reprinted from: *Forests* **2024**, *15*, 2064, https://doi.org/10.3390/f15122064 . . . . . **136**

# About the Editors

**Zhenyang Hui**

Zhenyang Hui is a Professor and Deputy Director of the Institute of Resources and Environment at East China University of Technology, China. Zhenyang is interested in intelligent processing theories of LiDAR point clouds and intelligent remote sensing interpretation, and his current areas of research mainly focus on multi-platform forest LiDAR application, individual tree segmentation, forest structural parameter inversion and above-ground biomass estimation. The main goals of his current research are to develop robust point cloud filtering, registration and tree parameter fitting algorithms that adapt to complex mountain forests, bamboo stands and urban street trees, and the additional ecological benefits that high-precision LiDAR survey brings to forest ecosystem assessment, such as reliable carbon stock quantification, fine-scale biodiversity monitoring and scientific support for precise forest management and nature-based ecological restoration solutions.

**Penggen Cheng**

Penggen Cheng is a Professor and Doctoral Supervisor at the School of Surveying and Geoinformation Engineering, East China University of Technology, China. Penggen is interested in geospatial information analysis, three-dimensional spatial modeling and multi-source remote sensing data fusion, and his current areas of research mainly focus on LiDAR remote sensing application for forest resources, geological spatial information extraction and ecological environment monitoring. The main goals of his current research are to explore efficient integrated processing methods for multi-platform LiDAR and optical remote sensing data over mountainous forested regions, and the additional ecological and surveying benefits that integrated geospatial analysis delivers to forest management, such as precise inversion of forest carbon storage, fine mapping of forest stand structures, and solid technical foundations for regional ecological protection and sustainable forest resource utilization.

**Bo Liu**

Bo Liu is a Professor and Doctoral Supervisor at the School of Surveying and Geoinformation Engineering, East China University of Technology. Bo is interested in fuzzy topology theories, spatial database construction, and 3D geospatial data processing, and his current areas of research mainly focus on LiDAR point cloud feature extraction, spatial data fusion, and quantitative structural modeling for forest remote sensing. The main goals of his current research are to develop robust spatial analysis algorithms and adaptive data processing frameworks for complex forest environments, and the additional technical benefits that optimized 3D data processing brings to forest surveys, such as accurate forest structural parameter retrieval, reliable spatial data integration, and solid theoretical and technical support for intelligent forest resource monitoring and ecological spatial governance.

**Mark Vanderwel**

Mark Vanderwel is a Professor in the Department of Biology at the University of Regina. Mark is interested in the use of high-resolution remote sensing data to characterize the structure and demography of forest stands and their responses to environmental change. His research focuses on boreal stand dynamics, carbon sequestration, climate change, and fire. The main goals of his current research are to use new remote sensing technologies to better understand the climate sensitivity of regional forest carbon stores, and how climate-related changes in growth, mortality, and fire are shaped by tree- and stand-level factors.

# Preface

This collection of papers presents relevant results from scientific researchers regarding LiDAR remote sensing technologies applied to forest inventory and ecological monitoring across multi-platform LiDAR systems including terrestrial, unmanned aerial, airborne and spaceborne sensors, encompassing not only point cloud processing algorithms such as filtering, co-registration and individual tree segmentation, but also quantitative forest parameter inversion tasks like tree height, DBH, canopy structure and above-ground biomass estimation, as well as carbon cycle and carbon sequestration assessment. The challenges brought by complex terrain, dense canopy overlap and noisy photon data to the robustness and accuracy of LiDAR forest surveys are also investigated. Case studies covering urban street trees, natural coniferous forests, moso bamboo stands and pine plantations are discussed, as well as optimized processing pipelines, adaptive fitting algorithms and multi-criteria denoising strategies, which might deliver reliable technical support for precise forest management, ecological evaluation and smart urban green space governance.

**Zhenyang Hui, Penggen Cheng, Bo Liu, and Mark Vanderwel**
*Guest Editors*

forests

MDPI

Article

# A Reliable DBH Estimation Method Using Terrestrial LiDAR Points through Polar Coordinate Transformation and Progressive Outlier Removal

**Zhenyang Hui [1,2,3], Lei Lin [1,2,3], Shuanggen Jin [4,5], Yuanping Xia [1,2,3,*] and Yao Yevenyo Ziggah [6]**

1 Faculty of Geomatics, East China University of Technology, Nanchang 330013, China; huizhenyang2008@ecut.edu.cn (Z.H.); 2022120383@ecut.edu.cn (L.L.)
2 Jiangxi Key Laboratory of Watershed Ecological Process and Information, East China University of Technology, Nanchang 330013, China
3 Key Laboratory of Mine Environmental Monitoring and Improving around Poyang Lake of Ministry of Natural Resources, East China University of Technology, Nanchang 330013, China
4 School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China; sgjin@shao.ac.cn
5 Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China
6 Faculty of Mineral Resources Technology, University of Mines and Technology, Tarkwa 999064, Ghana; yyziggah@umat.edu.gh
* Correspondence: ypxia@ecut.edu.cn

**Abstract:** Diameter at breast height (DBH) is a crucial parameter for forest inventory. However, accurately estimating DBH remains challenging due to the noisy and incomplete cross-sectional points. To address this, this paper proposed a reliable DBH estimation method using terrestrial LiDAR points through polar coordinate transformation and progressive outlier removal. In this paper, the initial center was initially detected by rasterizing the convex hull, and then the Cartesian coordinates were transformed into polar coordinates. In the polar coordinate system, the outliers were classified as low and high outliers according to the distribution of polar radius difference. Both types of outliers were then removed using adaptive thresholds and the moving least squares algorithm. Finally, DBH was estimated by calculating the definite integral of arc length in the polar coordinate system. Twenty publicly available individual trees were adopted for the test. Experimental results indicated that the proposed method performs better than the other four classical DBH estimation methods. Furthermore, several extreme cases scanned using terrestrial LiDAR in practice, such as cross-sectional points with lots of outliers or larger data gaps, were also tested. Experimental results demonstrate that the proposed method accurately calculates DBH even in these challenging cases.

**Keywords:** diameter at breast height; terrestrial LiDAR; polar transformation; outlier; cross-sectional points

## 1. Introduction

Recent advancements in laser scanning technologies have made it possible to obtain accurate and detailed three-dimensional information about the structure of forests [1–3]. Terrestrial LiDAR (light detection and ranging), in particular, has emerged as a crucial method for monitoring forest resources due to its efficiency and precision [4,5]. The dense point clouds generated by this technology provide access to a wide range of forest attributes, including tree height [6], diameter at breast height (DBH, which is a measurement used in forestry to determine the diameter of a tree trunk at a standard height of 1.30 m above the ground) [7], crown width [8], tree volume [9], and above-ground biomass (AGB, which is an important metric used in forestry, ecology, and environmental studies to assess the carbon sequestration capacity of forests and vegetation) [10–12]. DBH, in particular, plays a vital role in forest inventory as it can be used to estimate various other tree metrics, such

as tree height [13,14], wood volume [15], and foliage area [16]. Furthermore, DBH is also a key parameter in many allometric models used for AGB estimation [17,18].

Traditional methods for measuring DBH typically rely on calipers or tapes, which are inefficient and labor-intensive [19]. Moreover, manual visual measurements are prone to errors [4]. In contrast, terrestrial LiDAR offers an accurate and non-destructive solution for estimating DBH [20]. The methods using terrestrial LiDAR can be classified into three categories: DBH estimation based on regular geometry, DBH estimation by simulating manual measurement, and DBH estimation based on the quantitative structure model (QSM). The first two categories generally calculate DBH directly from point clouds, while the last category relies on the built individual tree models to achieve DBH values [4].

In terms of the DBH estimation methods based on regular geometry, circle fitting [21–25] and cylinder fitting [26–28] are two typical forms for calculating the DBH. Circle fitting methods involve selecting a slice of points and estimating the corresponding circle center and radius using techniques such as least squares fitting [21], random sample consensus (RANSAC) [29–32], and Hough transform [4,33]. For example, Reddy et al. [32] used the RANSAC algorithm for circle fitting, while Liu et al. [30] used RANSAC to calculate DBHs at different tree heights and removed wrongly calculated values based on the PauTa criterion to obtain accurate estimates. Panagiotidis et al. [34] applied a least squares circle fit algorithm in combination with hierarchical cluster analysis for DBH extraction. Koren et al. [35] compared five different circle fitting methods for DBH estimation and found that all four methods, except for the Monte Carlo method, tended to underestimate the DBH. Koren et al. [24] found that varying the thickness from 10 cm to 100 cm resulted in significant changes in calculated DBHs. Other factors such as range and angular errors of terrestrial LiDAR also have a significant impact on circle fitting results, as demonstrated by Wang et al. [6] in their experiments.

Cylinder fitting is similar to circle fitting but avoids selecting a specific layer of points [36]. Most circle fitting techniques can be applied to cylinder fitting as well. For example, Panagiotidis and Abdollahnejad [37] employed the RANSAC algorithm with a cylinder fitting strategy using points from cross-sectional areas and used the estimated the DBH for stem volume calculation. Srinivasan et al. [28] applied cylinder fitting to different height bins and calculated the average diameter of different directions to retrieve the DBH.

In summary, DBH estimation based on regular geometry assumes that the cross-section points of the stem are distributed in a circular or cylindrical shape. These methods have simple principles and are easy to implement. However, both circle fitting and cylinder fitting methods are sensitive to outliers, which can lead to incorrect DBH estimation results when dealing with noisy points or cross-section points with data gaps.

The second type of DBH estimation method relies on simulating manual measurement. Instead of assuming a regular geometry for the stem profile, these methods simulate the process of manual measurement using a tape [38]. As a result, these methods can be applied to different shapes of stem profiles and are more applicable in real-world scenarios. In this kind of method, convex hull points are generally detected initially [39]. You et al. [40] applied B-spline curves to interpolate the convex hull points and simulate the path of the tape. Subsequently, the DBH could be calculated using the calculated path length along the interpolated smooth curve. However, the extraction of convex hull points is easily influenced by outliers at the outermost points. In order to obtain more accurate DBH estimation results, Stovall et al. [39] proposed a convex hull peeling method to gradually remove outlying points. You et al. [38] proposed a caliper simulation method, where two parallel tangent lines were simulated as the caliper arms towards the convex hull points. By calculating the average distance between these lines, the DBH can be estimated. However, experimental results showed that this method tends to overestimate the DBH. In summary, the main advantage of these methods lies in their ability to handle irregular shapes of cross-section points of the stem. However, implementing these methods generally requires dense point clouds of high quality. Additionally, the simulation process is typically time-consuming [38].

The last category is QSM-based DBH estimation methods. In these methods, the QSM must be constructed for each individual tree as the first step. Several well-known individual tree modeling methods exist, including SimpleTree [41], TreeQSM [42] and AdTree [43]. It is important to note that all three of these methods are open-source, which facilitates the tree modeling process. DBH can be obtained as a byproduct of the modeling result. Ravaglia et al. [4] successfully applied SimpleTree for modeling and achieved satisfactory results in terms of DBH estimation. The mean errors in their study ranged from 1.03 cm to 3.34 cm. However, Ravaglia et al. [4] also mentioned that filtering is generally necessary to remove outliers and improve accuracy. Ye et al. [44] calculated DBH by building a quantitative structure model (QSM) for trees. However, their findings show that there are obvious deviations between QSM-based DBHs and referenced DBHs. The main limitation of this kind of method is that the accuracy of DBH estimation heavily relies on the quality of the QSM construction result. If the modeling result is poor due to occluded tree points, obtaining accurate DBH estimates becomes challenging. As a result, QSM-based DBH estimation methods are generally not suitable for single-scanned tree points, which are highly influenced by occlusion.

To sum up, data gaps caused by occlusion and outliers are the two main factors affecting the accuracy of DBH estimation regardless of the chosen DBH estimation method. Although some researchers have attempted to address these issues, occlusion and noisy points still affect most methods. Therefore, there is an urgent need to develop more accurate and robust DBH estimation methods. In this study, the Cartesian coordinates of cross-section points are first transformed into polar coordinates. This transformation is performed because points in polar coordinates are distributed in a wave-like pattern, which facilitates outlier detection through curve fitting. The noisy points are then classified as high or low outliers and removed separately using adaptive threshold filtering and the moving least squares algorithm. Finally, DBH is estimated based on the definite integral of arc length in the polar coordinate system.

## 2. Materials and Methods

### *2.1. Datasets*

This study selected 20 individual trees from the terrestrial LiDAR datasets provided by Liang et al. [45]. These datasets were obtained in Evo, Finland (61.19° N, 25.11° E), using a Leica HDS6100 terrestrial laser scanner (Leica Geosystems AG, Heerbrugg, Switzerland). The scanner has a scanning view of 360° × 310° and a distance measurement accuracy of ±2 mm. Each individual tree was manually segmented from sample plots using an open-source visualization software called CloudCompare V2.13, and point clouds were scanned under different forest environments. Specifically, these 20 tree samples were chosen from three different forest plots measuring 32 by 32 m each. These plots contain a variety of tree species, including Scots pine, Norway spruce, silver birch, and downy birch. In each plot, both single-scan and multi-scan modes were used to capture the individual tree points. In the multi-scan mode, five scans were carried out within each plot: one at the center and four in the four quadrant directions.

The DBH values were measured manually using steel calipers at two perpendicular directions on-site. The mean value of these two measurements was considered the referenced DBH. In ideal conditions, the segmented points at breast height should be complete for each individual tree, as shown in Figure 1a,b. However, there are often outliers caused by the instrument itself or adjacent shrubs, which can introduce errors in DBH calculation, as shown in Figure 1c,d. Additionally, due to limitations in scanning view or occlusion, the segmented points at breast height may have data gaps, resulting in an incomplete circle-like shape, as shown in Figure 1e,f.

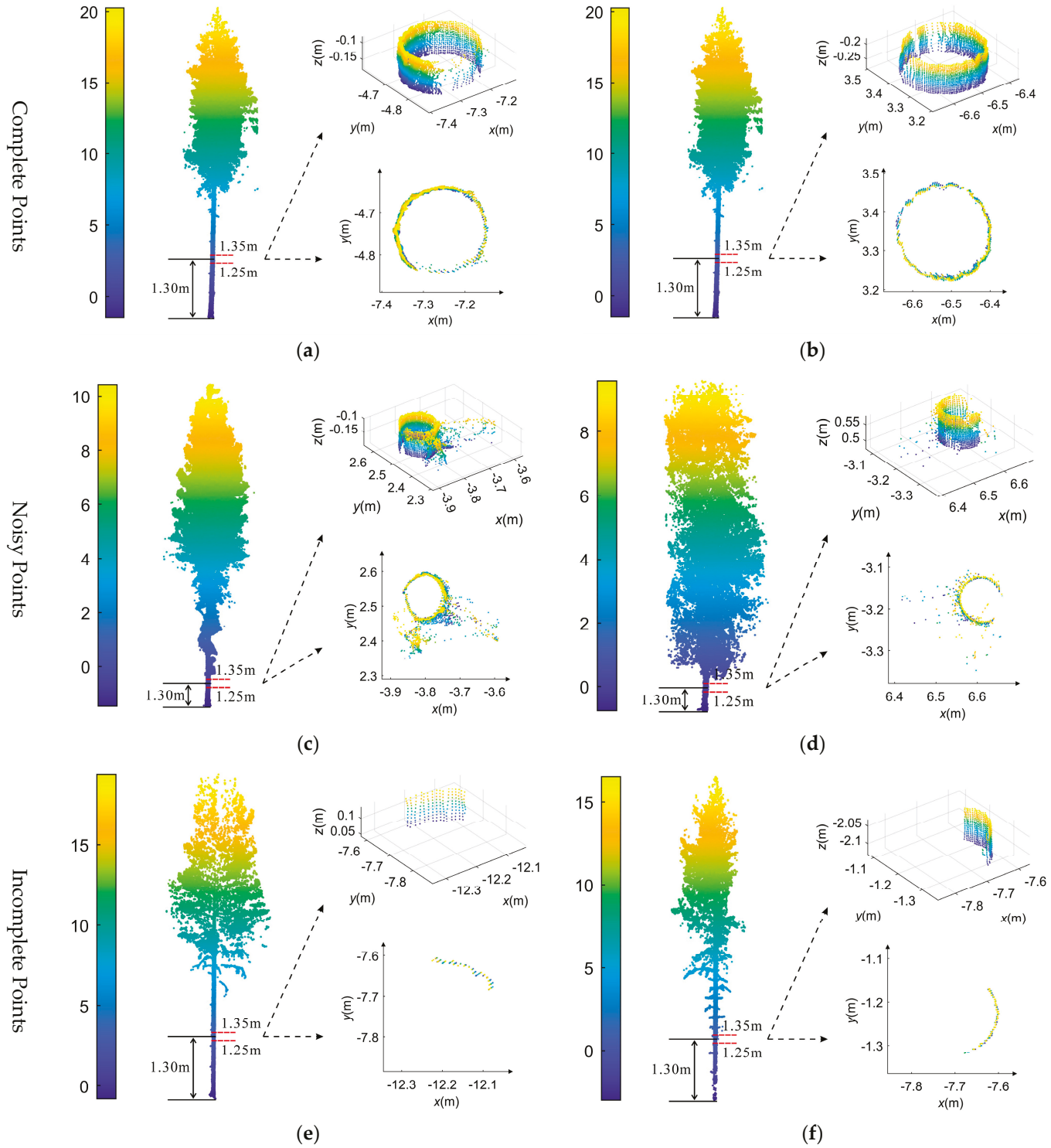



**Figure 1.** Different conditions of individual tree points: (**a**,**b**) complete cross-sectional points at breast height; (**c**,**d**) noisy cross-sectional points at breast height; (**e**,**f**) incomplete cross-sectional points at breast height.

Among the selected 20 individual trees, all three cases mentioned above (complete points, noisy points, and incomplete points) were present. Therefore, the proposed method's robustness against different scenarios can be tested.

### *2.2. Methodology*

The flowchart in Figure 2 illustrates the proposed method. DBH is commonly defined as the stem diameter at a distance of 1.30 m from the tree root. Therefore, a segment of point clouds between 1.25 m and 1.35 m from the tree root is initially segmented. These points are then projected onto a horizontal plane. To convert the points from Cartesian coordinates into polar coordinates, an initial center is estimated by rasterizing the convex hull formed by x–y coordinates. After transforming them into polar coordinates, it becomes apparent that there are outliers caused by the instrument itself or adjacent shrubs. These outliers significantly affect the accuracy of DBH estimation and are defined in this paper as high and low outliers. High outliers are removed using an adaptive threshold, while low outliers are identified through iterative use of the moving least squares algorithm. Once the outliers are removed, the polar coordinates are transformed back to Cartesian coordinates. Subsequently, a rough estimate of the circle center can be obtained using the least squares algorithm. Based on this newly calculated center, the Cartesian coordinates are further transformed into polar coordinates. Finally, DBH can be estimated by calculating the definite integral of arc length in the polar coordinate system. In summary, this paper includes three main steps: (i) fast initial center estimation by rasterizing convex hull, (ii) polar coordinate transformation and high/low outlier identification, and (iii) DBH estimation based on the definite integral of arc length.

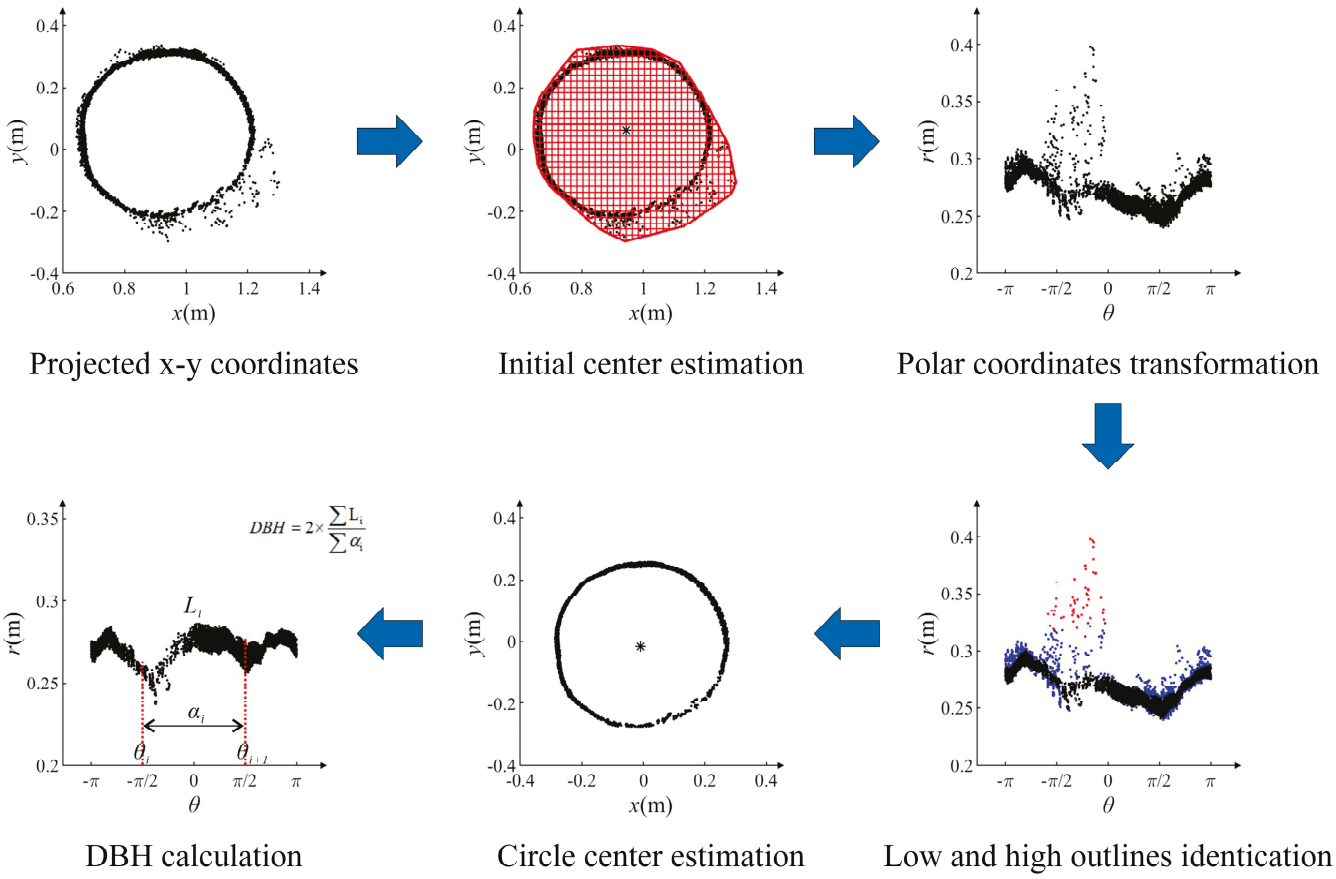



**Figure 2.** Flowchart of the proposed method. The asterisk is the initial estimated center.

#### 2.2.1. Fast Initial Center Estimation by Rasterizing Convex Hull

The cross-section of the stem usually resembles a circle shape, which allows the DBH to be estimated as the diameter of a fitting circle. Ideally, if the projected horizontal points used for DBH calculation were perfectly distributed along the circle (Figure 3a), their distances from the center would be equal. In other words, their polar radii should be equal when transforming Cartesian coordinates into polar coordinates with the circle center

selected as the pole (Figure 3b). However, in practice, the sliced points of the stem do not strictly form a circle shape, as shown in Figure 3c. As a result, the corresponding transformed polar coordinates exhibit a wave-like curve, as shown in Figure 3d.

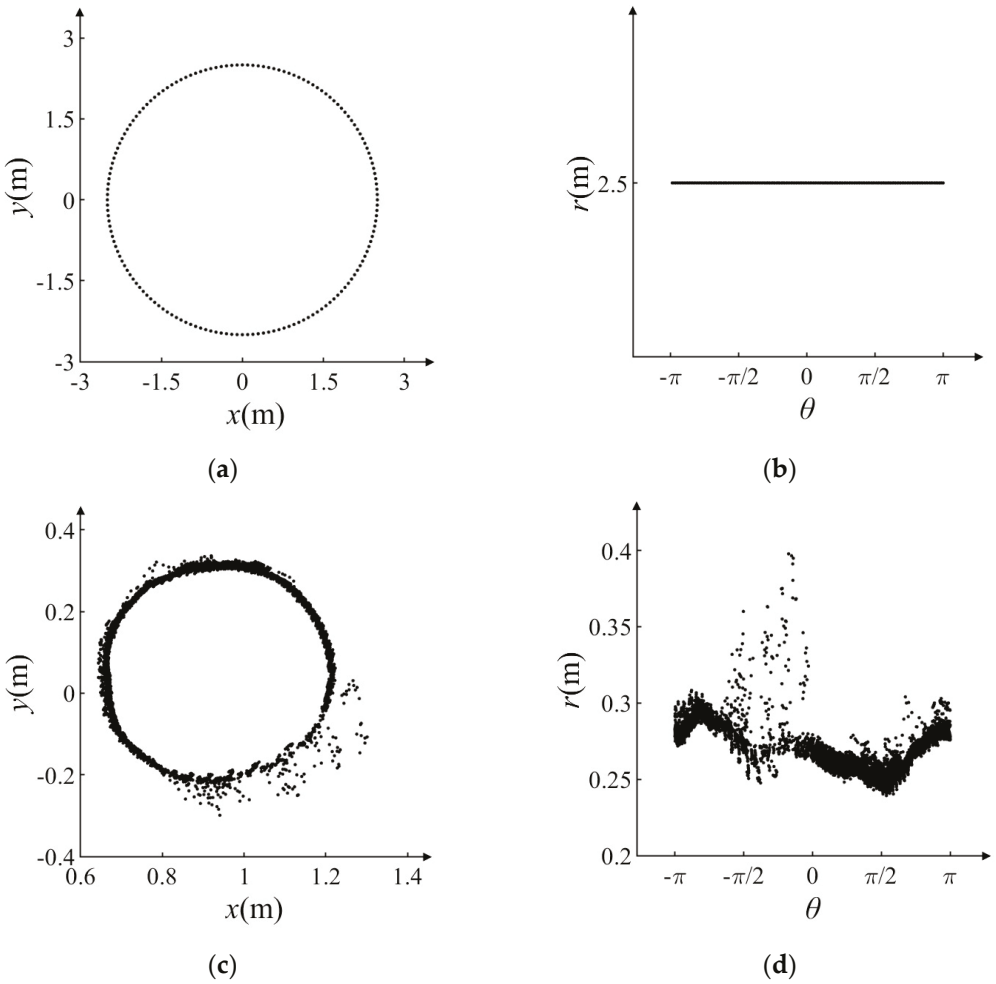



**Figure 3.** Transformation of Cartesian coordinates into polar coordinates. (**a**) Projected horizontal points strictly distributed along the circle in Cartesian coordinates; (**b**) transformed polar coordinates with equal polar radius values; (**c**) Cartesian coordinates with outliers; (**d**) transformed polar coordinates in a wave-like shape.

To transform Cartesian coordinates into polar coordinates, it is crucial to first locate the circle center as the pole. In general, the center should lie within the sliced points. To exclude incorrectly detected centers outside the sliced points, this paper restricts the location of the center within the minimum convex polygon (red line in Figure 4a) formed by applying the convex hull algorithm. Additionally, this area is further divided into rasterized grids for quick center identification. Typically, the center should be located at the farthest point from all projected horizontal points within the convex hull. Therefore, by traversing each grid point individually, the grid point with the farthest distance from the projected points is selected as the initial center (shown as a black asterisk in Figure 4a). It is important to note that if no restrictions are imposed on locating the center point within the convex hull, then the farthest grid point may fall outside of the sliced points. As illustrated in Figure 4b, this farthest grid point may be located in areas I, II, III, or IV. This is why building a convex hull is necessary to confine the center point within the minimum convex polygon.

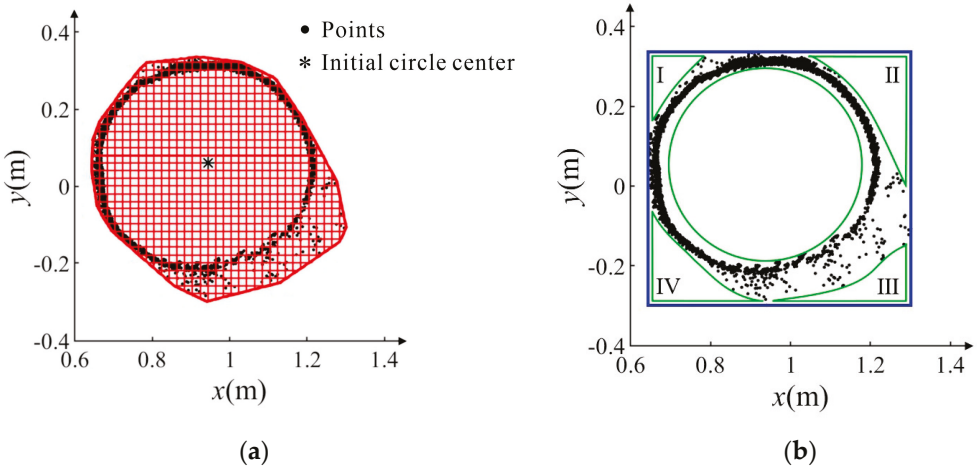



**Figure 4.** Initial circle center detection. (**a**) Convex hull building and rasterization; (**b**) the circle center that may appear at the outside areas (I, II, III, or IV) without convex hull building.

#### 2.2.2. Polar Coordinate Transformation and High/Low Outlier Identification

Once the initial center is obtained, the horizontal points can be centralized. Using Equation (1), the Cartesian coordinates can then be transformed into polar coordinates.

$$\begin{cases} r = \sqrt{x^2 + y^2} \\ \theta = \arctan(y/x) \end{cases} \quad (1)$$

where $r$ is the polar radius, while $\theta$ is the polar angular. Ideally, all circle points should have equal $r$ values. However, due to the influence of the instrument itself or adjacent shrubs, as shown in Figure 5a, the transformed polar coordinates generally take on a wave-like shape with several outliers, as depicted in Figure 5b.

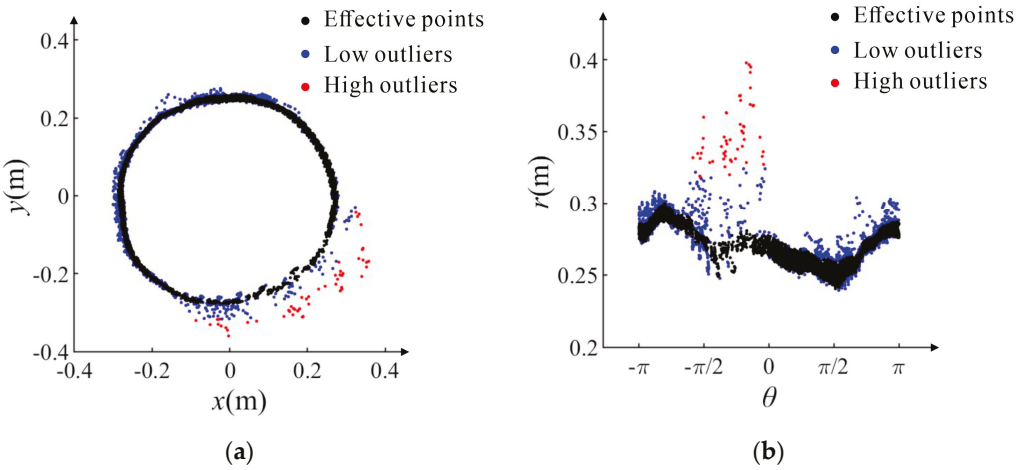



**Figure 5.** Polar coordinate transformation. (**a**) Projected horizontal points with outliers; (**b**) low and high outlier identification.

This paper categorizes these outliers into two groups: high and low outliers. High outliers are those points that deviate significantly from the wave-like curve, represented by the red points in Figure 5b. Conversely, low outliers are points that are mixed in with the wave-like curve, illustrated by the blue points in Figure 5b. The classification rule is defined by Equation (2).

$$\begin{cases} if\ \Delta r_i \geq \delta,\ p_i \in high\ outlier \\ if\ \Delta r_i < \delta\ \&\ abs\left(r_{p_i} - \hat{r}_{p_i}\right) \geq \eta,\ p_i \in low\ outlier \end{cases} \quad (2)$$

where $\Delta r_i$ is the difference between the point and minimum polar radius in each section, as shown in Figure 6. Here, each section is divided by 0.05 radians. Due to the existence of

outliers, the values of $\Delta r_i$ will be fluctuant. $\delta$ is a self-adaptive threshold, which will help remove the point with a higher $\Delta r_i$. $r_{p_i}$ is the polar radius of $p_i$, while $\hat{r}_{p_i}$ is the polar radius correspondingly fitted by the moving least squares algorithm. $\eta$ is a self-adaptive threshold that is used for identifying the low outliers. Both $\delta$ and $\eta$ can be calculated self-adaptively in this paper.

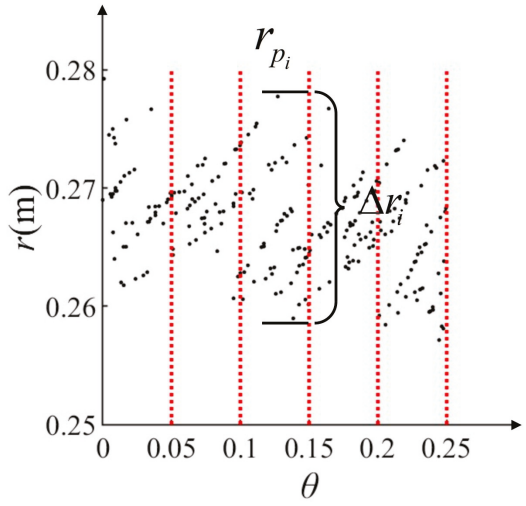



**Figure 6.** Polar radius difference calculation in each section.

As mentioned above, $\Delta r_i$ will be varied due to the existence of outliers. Some points with a larger $\Delta r_i$ will be detected as high outliers. Thus, it is important to calculate the threshold $\delta$. This paper first calculates the maximum polar radius difference, $\Delta r_i^{\max}$, in each section, that is, $S = \{\Delta r_1^{\max} \cdots, \Delta r_i^{\max} \cdots, \Delta r_n^{\max}\}$, which is distributed as shown in Figure 7.

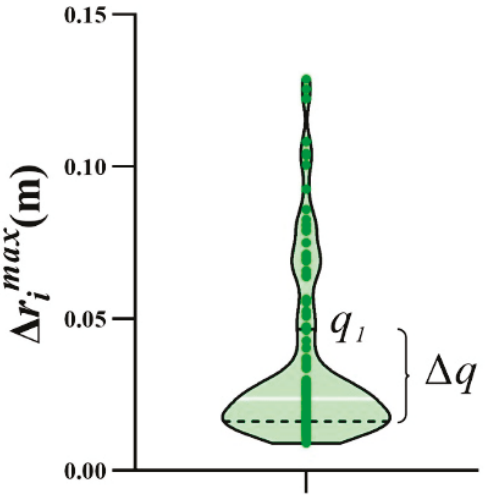



**Figure 7.** Distribution of the maximum polar radius difference. In this figure, the two black dotted lines represent the upper quartile ($q_1$) and lower quartile. $\Delta q$ is the interquartile range. The white line represents the median value.

From Figure 7, it can be found that there are several extremely larger $\Delta r_i^{\max}$ values. Here, this paper excludes these larger $\Delta r_i^{\max}$ values based on the threshold calculated as $q_1 + 1.5\Delta q$. $q_1$ represents the upper quartile, while $\Delta q$ is the interquartile range. After excluding the larger $\Delta r_i^{\max}$, a new set can be achieved, that is, $S' = \{\Delta r_1^{\max} \cdots, \Delta r_i^{\max} \cdots, \Delta r_m^{\max}\}$, $m \leq n$. Then, the data set, $S'$, is sorted from smallest to largest. Subsequently, the difference between the two successive data (eg., $\Delta r_i^{\max}$ and $\Delta r_{i+1}^{\max}$) is calculated as $\sigma r_i$. Hereafter, a new set, $S'' = \{\sigma r_1 \cdots, \sigma r_i \cdots, \sigma r_{m-1}\}$, can be achieved. This paper further calculates the

mean value ($\overline{\sigma r}$) and standard deviation ($\widetilde{\sigma r}$) of $S''$ and detects the first $\sigma r_i$ that is larger than $\overline{\sigma r} + 3\widetilde{\sigma r}$. Finally, the corresponding $\Delta r_i^{\max}$ is set as the high outlier detection threshold, that is, $\delta = \Delta r_i^{\max}$. For instance, if $\sigma r_5$ is the first value larger than $\overline{\sigma r} + 3\widetilde{\sigma r}$, $\delta$ is equal to $\Delta r_5^{\max}$.

As shown in Figure 5b, there are several low outliers mixed with the wave-like curve. This paper adopted the moving least squares algorithm to delete these low outliers. As shown in Figure 8a, the moving window in this paper is set to four neighboring sections. The fitted polar radius for $p_i$ can be calculated as $\hat{r}_{p_i}$ using the points fallen within the moving window. In terms of $\hat{r}_{p_i}$, all the differences between the polar radiuses and $\hat{r}_{p_i}$ can be calculated as shown in Figure 8b. From Figure 8b, the upper quartile, $Q_1$, and interquartile range, $\Delta Q$, can be calculated. In this paper, $\eta$ is set to $Q_1 + 1.5\Delta Q$, that is, if the difference between the polar radius and $\hat{r}_{p_i}$ is larger than $\eta$, the corresponding point will be identified as a low outlier.

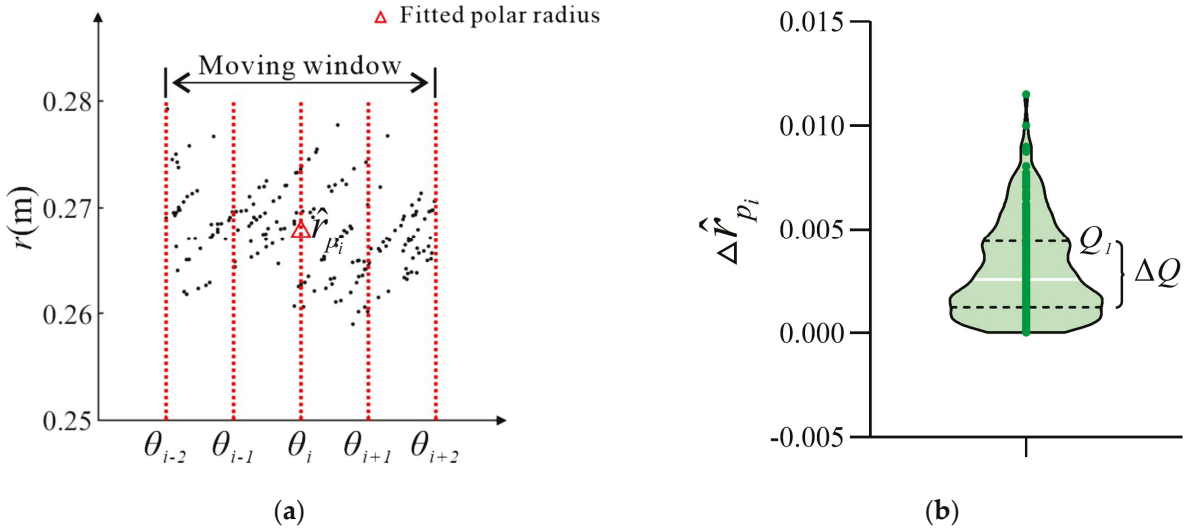



**Figure 8.** Process of low outlier identification. (**a**) Fitted polar radius calculation using moving least squares algorithm; (**b**) threshold calculation for low outlier detection. In (**b**), the two black dotted lines represent the upper quartile ($Q_1$) and lower quartile. $\Delta Q$ is the interquartile range. The white line represents the median value.

#### 2.2.3. DBH Calculation Based on the Definite Integral of Arc Length

After removing all high and low outliers, "clean" polar coordinates can be obtained, as shown in Figure 9a. These polar coordinates can also be converted back into Cartesian coordinates, as demonstrated in Figure 9b. It can be observed from Figure 9b that the point distribution is closer to a circle. By employing the least squares fitting algorithm, a rough center can be estimated and represented by the red fork in Figure 9b. It should be noted that this paper refers to this estimated center as a "rough center" because it is based on an initial center (the black asterisk point shown in Figure 4a). This initial center is generally not accurate. To achieve precise DBH estimation results, the newly calculated center depicted in Figure 9b is used to transform the Cartesian coordinates presented in Figure 9b into polar coordinates, as illustrated in Figure 9c. Consequently, DBH can be estimated according to Equation (3).

$$\begin{cases} DBH = 2\frac{L}{\alpha} \\ L = \sum_{i=1}^{n} L_i, \alpha = \sum_{i=1}^{n} \alpha_i \end{cases} \tag{3}$$

$$L_i = \int_{\theta_i}^{\theta_{i+1}} \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2} d\theta \tag{4}$$

where $L$ is the arc length, while $\alpha$ is the corresponding radian. In complex forest environments, it is generally hard to obtain complete stem points due to the scanning angle

especially when adopting single-scanning mode. As a result, the wave-like curve is generally fragmentized as shown in Figure 9d. Then, the arc length, $L$, can be calculated by summing each fragmentized arc length, $L_i$, while $L_i$ can be calculated based on the definite integral of arc length from $\theta_i$ to $\theta_{i+1}$ according to Equation (4).

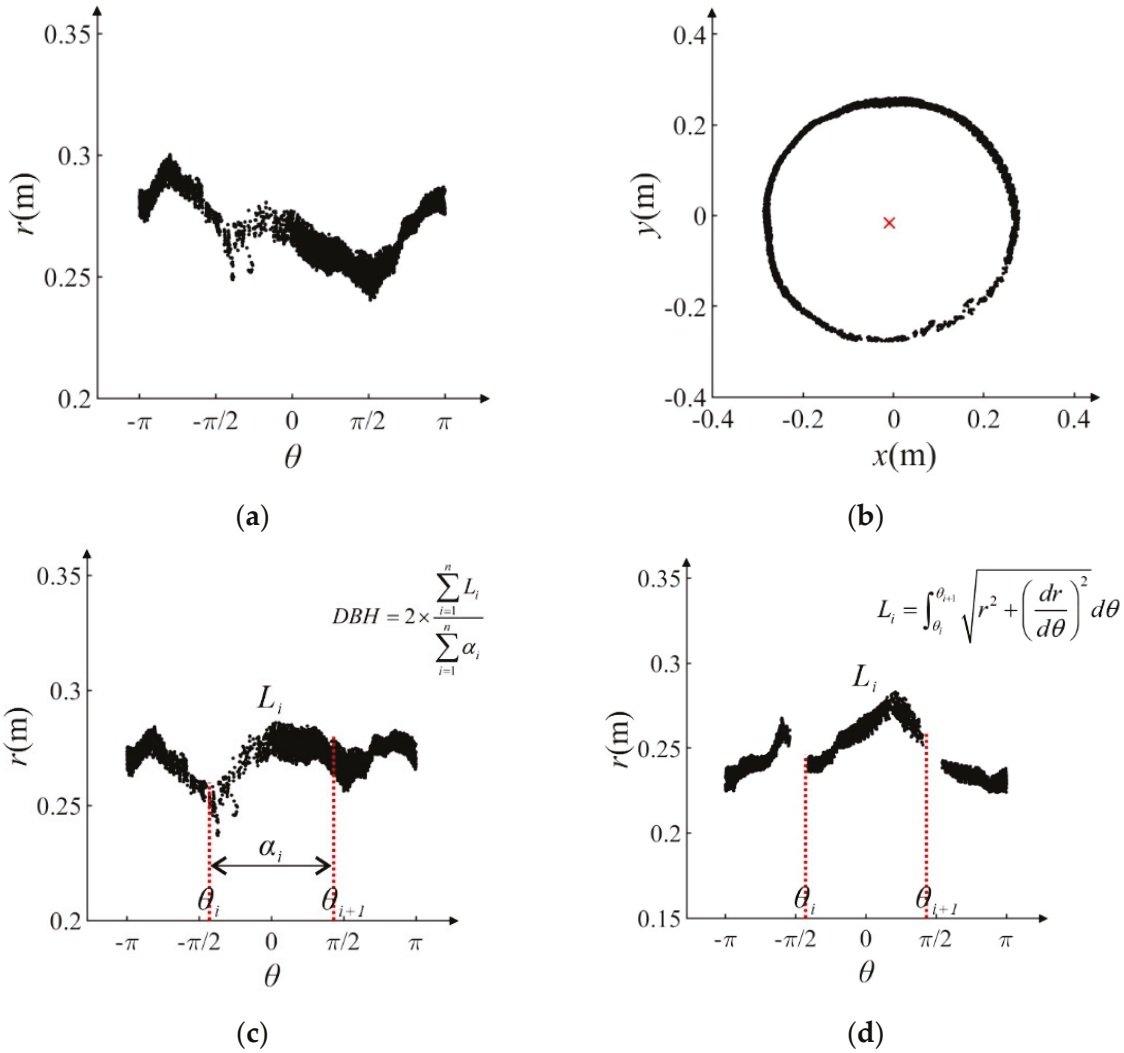



**Figure 9.** Process of DBH calculation. (**a**) Polar coordinates after outlier removal; (**b**) Cartesian coordinates after outlier removal; (**c**) DBH calculation based on arc length; (**d**) arc length calculation when encountering incomplete stem points.

## 3. Experimental Results

To visually demonstrate the effectiveness of this method, six out of the twenty DBH estimation results were tabulated in Figure 10. These six cross-sectional points correspond to the six trees depicted in Figure 1. From Figure 10, it can be found that in terms of complete cross-sectional points, the proposed method can fit a circle-like shape well (red circles in Figure 10a,b). In terms of noisy cross-sectional points, although there are many outliers in Figure 10c,d, the proposed method can estimate the DBH well. This is because the proposed method can exclude outliers effectively in the polar coordinate system. In terms of incomplete cross-sectional points, as depicted in Figure 10e,f, although only small part of cross-sectional points remained, the proposed method achieved satisfactory DBH estimation results.

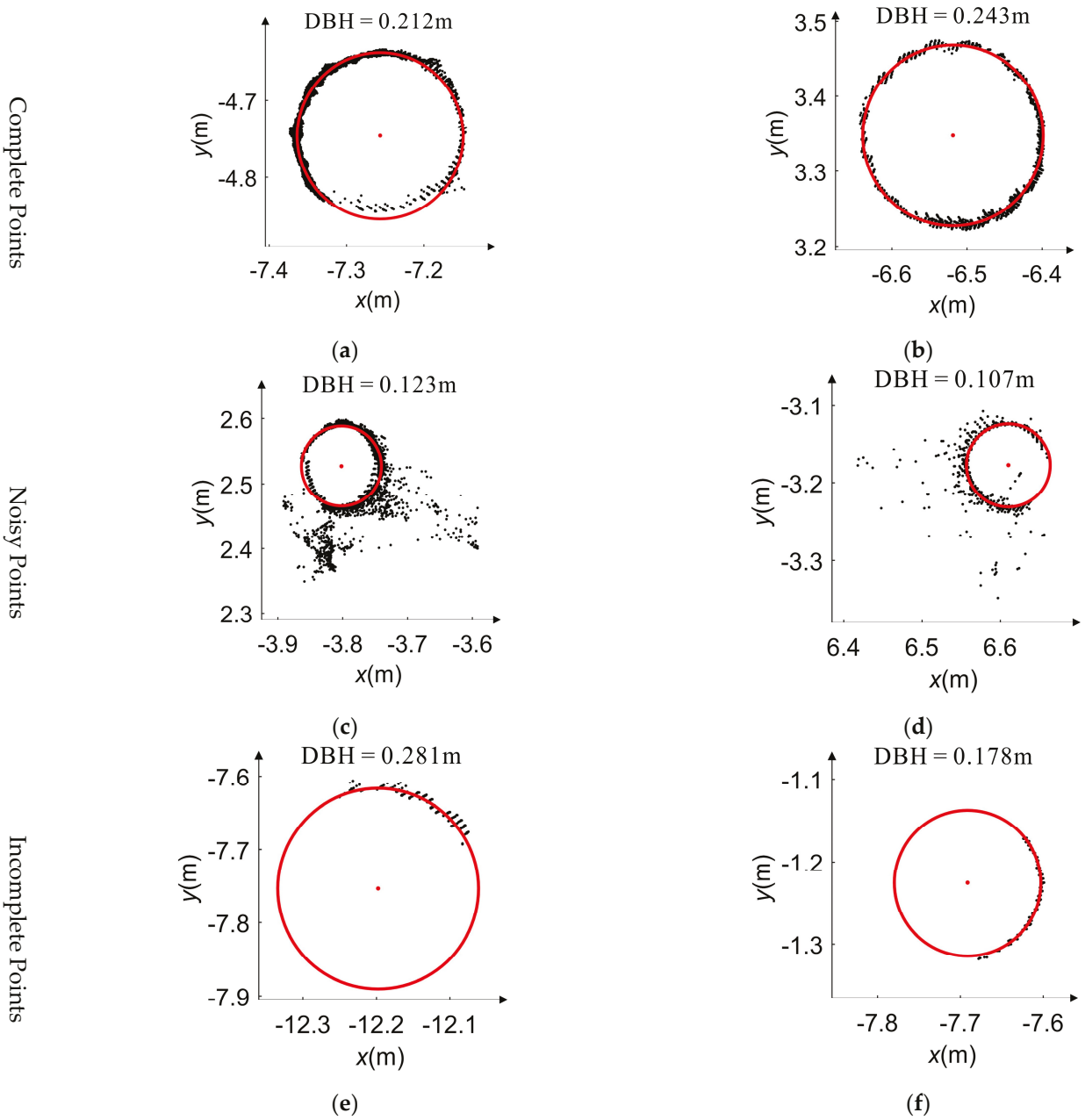



**Figure 10.** DBH estimation results in different cases. (**a**,**b**) DBH estimation results of complete cross-sectional points; (**c**,**d**) DBH estimated results of noisy cross-sectional points at breast height; (**e**,**f**) DBH estimation results of incomplete cross-sectional points at breast height. The red point and red circle represent the circle center and the fitted circle, respectively.

This paper estimated all the DBHs for these 20 trees. The estimated results were shown in Figure 11. Meanwhile, this paper also calculated the DBH estimation results using four other methods proposed by Olofsson et al. [31], Liu et al. [33], Liu et al. [30], and Mokroš et al. [46]. Olofsson et al. [31] introduced a modified RANSAC DBH estimation method. They made several adjustments to the traditional RANSAC approach. Firstly, they deemed a circle invalid if it had more than 1% of the points inside. Secondly, they restricted the range of circle radii to be between 2 cm and 30 cm, considering circles with smaller or larger radii invalid. Liu et al. [33] proposed a random Hough transform DBH estimation method. In their approach, they randomly selected three non-collinear points with a minimum spacing of 0.02 m between them. Each set of three points was then used to fit a circle, and only circles that satisfied the radius constraints were considered

in the Hough parameter space. Liu et al. [30] calculated DBH values at different heights. They applied the RANSAC algorithm for DBH estimation on each set of sliced points. By detecting incorrectly calculated DBH values based on the PauTa criterion, they obtained the final DBH estimation. Mokroš et al. [46] obtained the DBH estimation results using a cylinder fitting technique. In their method, the cylinder element was obtained at 1.3 m above the ground. From Figure 11, it can be found that compared with other methods, there are more points of the proposed method that closely surround the 1:1 line. This means that the estimated DBH values are close to the referenced DBH values. This paper calculated the coefficients of determination ($R^2$) between estimated DBH values and reference values of different methods. It is easy to find that the method in this paper performed much better than the other four methods. Thus, it can be concluded that this paper's method can achieve better performance in DBH estimation.

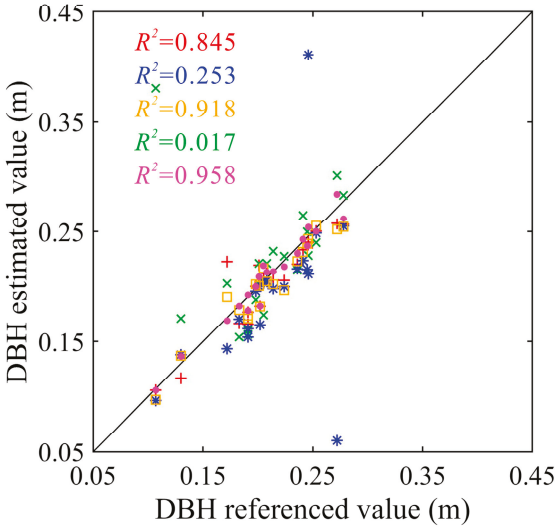



**Figure 11.** The coefficient of determination of different methods between estimated DBH values and referenced DBH values. The red crosses represent the estimation results obtained by Olofsson et al. [31]. The blue stars represent the estimation results obtained by Liu et al. [33]. The orange boxes represent the estimation results obtained by Liu et al. [30]. The green forks represent the estimation results obtained by Mokroš et al. [46]. The pink points represent the proposed method's estimation results.

To quantitatively evaluate the performance of this method, five accuracy indicators were used for assessing the performance of the proposed method, including bias, relative bias, root mean squared error, relative root mean squared error, and the concordance correlation coefficient. They are defined in Equations (5)–(9):

$$Bias = \frac{\sum_{i=1}^{N} (DBH_{est}^{i} - DBH_{ref}^{i})}{N} \tag{5}$$

$$rBias = \frac{Bias}{\overline{DBH_{ref}}} \times 100\% \tag{6}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (DBH_{est}^{i} - DBH_{ref}^{i})^2}{N}} \tag{7}$$

$$rRMSE = \frac{RMSE}{\overline{DBH_{ref}}} \times 100\% \tag{8}$$

$$\begin{cases} CCC = \dfrac{\frac{2}{N}\sum_{i=1}^{N}(DBH_{est}^{i} - \overline{DBH_{est}})(DBH_{ref}^{i} - \overline{DBH_{ref}})}{\sigma_{est}^{2} + \sigma_{ref}^{2} + (\overline{DBH_{est}} - \overline{DBH_{ref}})^{2}} \\ \sigma_{est}^{2} = \dfrac{1}{N}\sum_{i=1}^{N}(DBH_{est}^{i} - \overline{DBH_{est}})^{2} \\ \sigma_{ref}^{2} = \dfrac{1}{N}\sum_{i=1}^{N}(DBH_{ref}^{i} - \overline{DBH_{ref}})^{2} \end{cases} \quad (9)$$

where *Bias*, *rBias*, *RMSE*, *rRMSE*, and *CCC* represent bias, relative bias, root mean squared error, relative root mean squared error, and the concordance correlation coefficient, respectively. $DBH_{est}^{i}$ and $DBH_{ref}^{i}$ represent the *i*-th estimated and reference DBH values. $\overline{DBH_{est}}$ and $\overline{DBH_{ref}}$ represent the mean estimated and reference DBH values. $N$ is the number of trees for DBH estimation.

The mean accuracy metric calculation results are tabulated in Table 1. It is easy to find that the proposed method performs much better than the other methods in terms of *RMSE*, *rRMSE*, and *CCC*. The *Bias* of the proposed method is −0.75 mm, which is only a little larger than the results by Mokroš et al. [46]. However, the *RMSE* and *rRMSE* values of Mokroš et al. [46] are more than three times those of the proposed method. This means that the method by Mokroš et al. [46] cannot achieve satisfactory performance in different DBH estimation cases. Comparatively, the proposed method achieved the smallest *RMSE* and *rRMSE* values. This indicates that the proposed method has stronger robustness. When it comes to *CCC*, the *CCC* of the proposed method is 0.98, which is much closer to 1 than that of the other four methods. This means that the DBH values estimated in this paper are much closer to the referenced DBH values.

**Table 1.** Accuracy metrics of different methods.

| DBH Estimation Methods | *Bias* (mm) | *rBias* (%) | *RMSE* (mm) | *rRMSE* (%) | *CCC* |
|---|---|---|---|---|---|
| Olofsson et al. [31] | −5.67 | −2.70 | 17.54 | 8.36 | 0.91 |
| Liu et al. [33] | −7.21 | −3.44 | 14.00 | 6.67 | 0.94 |
| Liu et al. [30] | 14.10 | 6.72 | 65.19 | 31.06 | 0.12 |
| Mokroš et al. [46] | −0.41 | −0.20 | 30.91 | 14.73 | 0.76 |
| The proposed method | −0.75 | −0.36 | 8.68 | 4.13 | 0.98 |

## 4. Discussion

To visually compare the performance of the proposed method in some extreme cases, such as when there are cross-sectional points with a huge amount of outliers and cross-sectional points with larger data gaps, several trees scanned in practice were used for testing. Different scenarios of cross-sectional points were chosen, including noisy, incomplete, and irregular ones, as displayed in Figure 12. The DBH estimation was carried out using five methods (that of Olofsson et al. [31], Liu et al. [33], Liu et al. [30], and Mokroš et al. [46], and the proposed method), with DBH values obtained from tape or caliper measurements serving as references. The results of the five methods and the referenced values are presented in Figure 12. When dealing with noisy cross-sectional points, it was observed that Liu et al. [33] incorrectly identified the circle, leading to an overestimated DBH value (Figure 12b). Olofsson et al. [31] and Mokroš et al. [46] provided underestimations (Figure 12a,d), while the proposed method yielded the closest estimate to the true value (Figure 12e). For incomplete points (Figure 12g–l), all methods except the proposed one tended to underestimate the DBH. Regarding irregular cross-sectional points, where points are not strictly distributed in circular shapes, the proposed method demonstrated superior accuracy compared with the other four methods (Figure 12q). Overall, based on the extreme cases illustrated in Figure 12, it can be concluded that the proposed method delivers satisfactory DBH estimation performance with strong robustness.

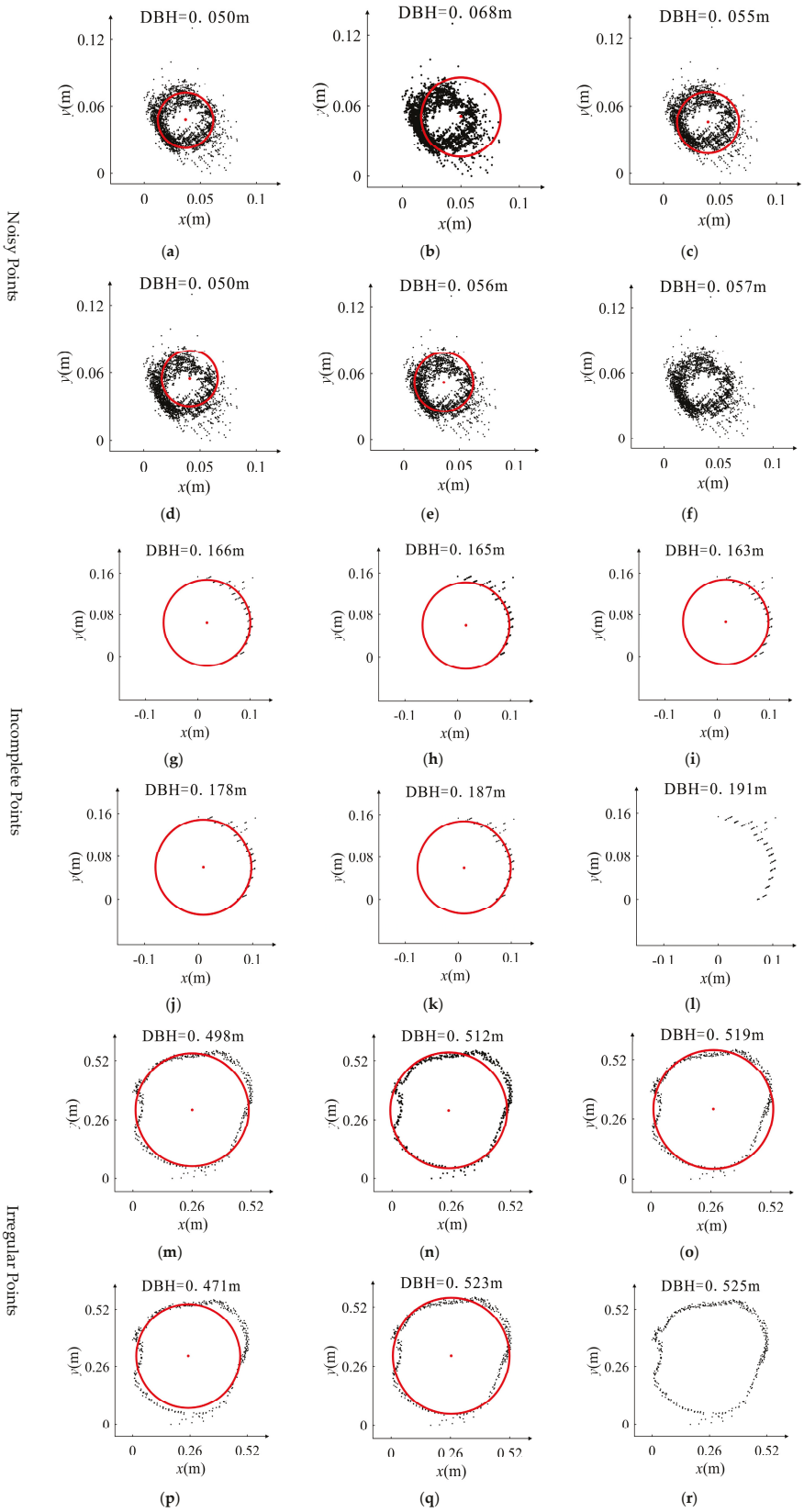
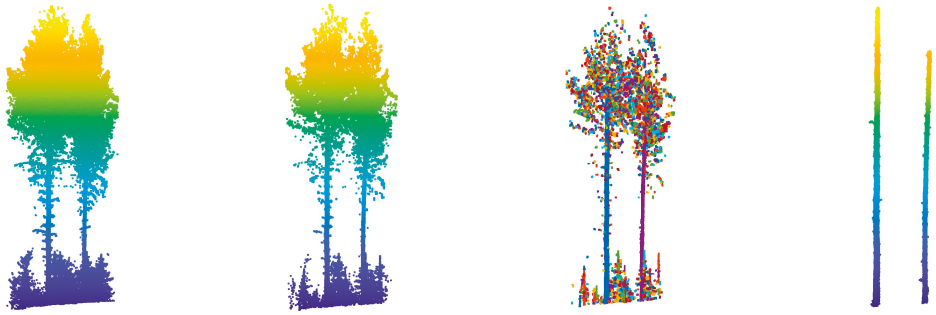



**Figure 12.** (**a**–**f**) The noisy cross-sectional points; (**g**–**l**) the incomplete cross-sectional points; (**m**–**r**) the irregular cross-sectional points. The results of each case are in accordance with those of Olofsson et al. [31], Liu et al. [33], Liu et al. [30], Mokroš et al. [46], and the proposed method, and those measured manually. The black points in the figure represent the projected horizontal points, while the red circle and the red point represent the fitted circle and estimated circle center, respectively.

As can be seen from the discussion above, the proposed approach can produce satisfactory DBH estimation results for a variety of forest stand situations, including locations that are noisy, incomplete, or irregular. It must be acknowledged, nonetheless, that each of the aforementioned forest stands is made up of manually divided individual trees. Correct individual tree segmentation results are necessary for accurate DBH estimation.

Figure 13a depicts individual trees surrounded by dense low bushes or small trees. In this instance, obtaining correct individual tree segmentation results is tough, making reliable DBH estimation results much more difficult. To attain a satisfactory DBH estimation result, some 'boosting' processes should be added. As illustrated in Figure 13b, the normal vector constraint is first used to remove some of the leaves. Next, the DBSCAN clustering technique is used to retrieve various clustering object primitives, as shown in Figure 13c. The final stems were then extracted by analyzing the linear and height features of each object primitive, as shown in Figure 13d. After extracting the stems, DBH can be calculated using our proposed approach.



(**a**) (**b**) (**c**) (**d**)

**Figure 13.** DBH estimation in dense forest environment with dense low bushes or small trees. (**a**) Original tree points; (**b**) result after normal vector constraint; (**c**) DBSCAN clustering result; (**d**) stem extraction result.

## 5. Conclusions

DBH plays a vital role in forest inventory, enabling accurate tree height estimation, AGB calculation, and more. However, estimating DBH remains challenging due to issues like low accuracy, particularly in the presence of noisy cross-sectional points and incomplete data. To enhance the effectiveness and robustness of DBH estimation, this paper introduces a reliable method using terrestrial LiDAR points through polar coordinate transformation and progressive outlier removal. In the proposed method, the initial center is first estimated by rasterizing the convex hull, and then the Cartesian coordinates are transformed into polar coordinates. Subsequently, outliers are identified as low or high outliers and gradually removed. Following outlier removal, the DBH is calculated based on the definite integral of arc length. In total, 20 publicly available individual trees were used for testing. The experimental results demonstrate that our proposed method consistently outperforms other methods across various accuracy indicators. The *CCC* value of the proposed method is 0.98, indicating that estimated DBH values closely align with the referenced DBH values. To further test the robustness of the proposed method, several extreme cases are adopted for testing. The proposed method successfully achieves accurate DBH estimations by progressively removing outliers and effectively resisting their impact. Compared with the other four methods, the proposed method achieved better performance in different DBH estimation cases. These results confirm that the proposed method demonstrates exceptional effectiveness and robustness in DBH estimation.

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

**Funding:** The authors thank the National Natural Science Foundation of China (NSF) (42161060, 41801325, and 42174055), Outstanding Young Talents Funding of Jiangxi Province (20232ACB213017), Double Thousand Plan of Jiangxi Province (DHSQT42023002), the China Post-Doctoral Science Foundation (2019M661858), and the Natural Science Foundation of Jiangxi Province (20192BAB217010) for their financial support.

**Data Availability Statement:** Data are contained within the article.

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

## References

1. Weiser, H.; Schaefer, J.; Winiwarter, L.; Krasovec, N.; Fassnacht, F.E.; Hoefle, B. Individual tree point clouds and tree measurements from multi-platform laser scanning in German forests. *Earth Syst. Sci. Data* **2022**, *14*, 2989–3012. [CrossRef]
2. Li, J.; Wu, H.; Xiao, Z.; Lu, H. 3D modeling of laser-scanned trees based on skeleton refined extraction. *Int. J. Appl. Earth Obs. Geoinform.* **2022**, *112*, 102943. [CrossRef]
3. Hui, Z.; Cheng, P.; Yang, B.; Zhou, G. Multi-level self-adaptive individual tree detection for coniferous forest using airborne LiDAR. *Int. J. Appl. Earth Obs. Geoinform.* **2022**, *114*, 103028. [CrossRef]
4. Ravaglia, J.; Fournier, R.A.; Bac, A.; Vega, C.; Cote, J.; Piboule, A.; Remillard, U. Comparison of three algorithms to estimate tree stem diameter from terrestrial laser scanner data. *Forests* **2019**, *10*, 5997. [CrossRef]
5. Zhu, X.; Skidmore, A.K.; Darvishzadeh, R.; Niemann, K.O.; Liu, J.; Shi, Y.; Wang, T. Foliar and woody materials discriminated using terrestrial LiDAR in a mixed natural forest. *Int. J. Appl. Earth Obs. Geoinform.* **2018**, *64*, 43–50. [CrossRef]
6. Wang, Y.; Lehtomaki, M.; Liang, X.; Pyorala, J.; Kukko, A.; Jaakkola, A.; Liu, J.; Feng, Z.; Chen, R.; Hyyppa, J. Is field-measured tree height as reliable as believed a comparison study of tree height estimates from field measurement airborne laser scanning and terrestrial laser scanning in a boreal forest. *ISPRS J. Photogramm.* **2019**, *147*, 132–145. [CrossRef]
7. Bruggisser, M.; Hollaus, M.; Otepka, J.; Pfeifer, N. Influence of ULS acquisition characteristics on tree stem parameter estimation. *ISPRS J. Photogramm.* **2020**, *168*, 28–40. [CrossRef]
8. Arumae, T.; Lang, M. Estimation of canopy cover in dense mixed-species forests using airborne lidar data. *Eur. J. Remote Sens.* **2018**, *51*, 132–141. [CrossRef]
9. Disney, M.I.; Vicari, M.B.; Burt, A.; Calders, K.; Lewis, S.L.; Raumonen, P.; Wilkes, P. Weighing trees with lasers: Advances challenges and opportunities. *Interface Focus* **2018**, *8*, 201700482. [CrossRef] [PubMed]
10. Brede, B.; Calders, K.; Lau, A.; Raumonen, P.; Bartholomeus, H.M.; Herold, M.; Kooistra, L. Non-destructive tree volume estimation through quantitative structure modelling: Comparing UAV laser scanning with terrestrial LiDAR. *Remote Sens. Environ.* **2019**, *233*, 111355. [CrossRef]
11. Brede, B.; Terryn, L.; Barbier, N.; Bartholomeus, H.M.; Bartolo, R.; Calders, K.; Derroire, G.; Moorthy, S.M.K.; Lau, A.; Levick, S.R.; et al. Non-destructive estimation of individual tree biomass: Allometric models terrestrial and UAV laser scanning. *Remote Sens. Environ.* **2022**, *280*, 113180. [CrossRef]
12. Kukenbrink, D.; Gardi, O.; Morsdorf, F.; Thurig, E.; Schellenberger, A.; Mathys, L. Above-ground biomass references for urban trees from terrestrial laser scanning data. *Ann. Bot.* **2021**, *128*, 709–724. [CrossRef] [PubMed]
13. Cysneiros, V.C.; Pelissari, A.L.; Gaui, T.D.; Fiorentin, L.D.; Carvalho, D.C.D.; Silveira Filho, T.B.; Machado, S.A. Modeling of tree height-diameter relationships in the Atlantic Forest: Effect of forest type on tree allometry. *Can. J. For. Res.* **2020**, *50*, 1289–1298. [CrossRef]
14. Kafuti, C.; Van den Bulcke, J.; Beeckman, H.; Van Acker, J.; Hubau, W.; De Mil, T.; Hatakiwe, H.; Djiofack, B.; Fayolle, A.; Panzou, G.J.L.; et al. Height-diameter allometric Eqss of an emergent tree species from the Congo Basin. *For. Ecol. Manag.* **2022**, *504*, 119822. [CrossRef]
15. Calders, K.; Verbeeck, H.; Burt, A.; Origo, N.; Nightingale, J.; Malhi, Y.; Wilkes, P.; Raumonen, P.; Bunce, R.G.H.; Disney, M. Laser scanning reveals potential underestimation of biomass carbon in temperate forest. *Ecol. Solut. Evid.* **2022**, *3*, e12197. [CrossRef]
16. Tziaferidis, S.R.; Spyroglou, G.; Fotelli, M.N.; Radoglou, K. Allometric models for the estimation of foliage area and biomass from stem metrics in black locust. *Iforest* **2022**, *15*, 281–288. [CrossRef]
17. Demol, M.; Calders, K.; Verbeeck, H.; Gielen, B. Forest above-ground volume assessments with terrestrial laser scanning: A ground-truth validation experiment in temperate, managed forests. *Ann. Bot.* **2021**, *128*, 805–819. [CrossRef] [PubMed]
18. Luo, Y.; Wang, X.; Ouyang, Z.; Lu, F.; Feng, L.; Tao, J. A review of biomass Eqss for China's tree species. *Earth Syst. Sci. Data* **2020**, *12*, 21–40. [CrossRef]
19. Wang, D.; Kankare, V.; Puttonen, E.; Hollaus, M.; Pfeifer, N. Reconstructing stem cross section shapes from terrestrial laser scanning. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 272–276. [CrossRef]
20. Brede, B.; Lau, A.; Bartholomeus, H.M.; Kooistra, L. Comparing RIEGL RiCOPTER UAV LiDAR derived canopy height and DBH with terrestrial LiDAR. *Sensors* **2017**, *1*, 2371. [CrossRef] [PubMed]
21. Brunner, A.; Gizachew, B. Rapid detection of stand density tree positions and tree diameter with a 2D terrestrial laser scanner. *Eur. J. For. Res.* **2014**, *133*, 819–831. [CrossRef]
22. Henning, J.G.; Radtke, P.J. Detailed stem measurements of standing trees from ground-based scanning lidar. *For. Sci.* **2006**, *52*, 67–80. [CrossRef]

23. Hu, C.; Pan, S.; Zhang, H.; Li, P. Trunk model establishment and parameter estimation for a single tree using multistation terrestrial laser scanning. *IEEE Access* **2020**, *8*, 102263–102277. [CrossRef]
24. Koren, M.; Huncaga, M.; Chuda, J.; Mokrog, M.; Surovy, P. The influence of cross-section thickness on diameter at breast height estimation from point cloud. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 4959. [CrossRef]
25. Kalwar, O.P.P.; Hussin, Y.A.; Weir, M.J.C.; de Bie, C.A.J.M.; Karna, Y. Deriving forest plot inventory parameters using terrestrial laser scanning in the tropical rainforest of Malaysia. *Int. J. Remote Sens.* **2021**, *42*, 884–901. [CrossRef]
26. Liang, X.; Litkey, P.; Hyyppa, J.; Kaartinen, H.; Vastaranta, M.; Holopainen, M. Automatic stem mapping using single-scan terrestrial laser scanning. *IEEE Trans. Geosci. Remote Sens.* **2011**, *50*, 661–670. [CrossRef]
27. Liang, X.; Kankare, V.; Yu, X.; Hyyppa, J.; Holopainen, M. Automated stem curve measurement using terrestrial laser scanning. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 1739–1748. [CrossRef]
28. Srinivasan, S.; Popescu, S.C.; Eriksson, M.; Sheridan, R.D.; Ku, N. Terrestrial laser scanning as an effective tool to retrieve tree level height crown width and stem diameter. *Remote Sens.* **2015**, *7*, 1877–1896. [CrossRef]
29. Kuzelka, K.; Marusak, R.; Surovy, P. Inventory of close-to-nature forest stands using terrestrial mobile laser scanning. *Int. J. Appl. Earth Obs. Geoinform.* **2022**, *115*, 103104. [CrossRef]
30. Liu, L.; Zhang, A.; Xiao, S.; Hu, S.; He, N.; Pang, H.; Zhang, X.; Yang, S. Single tree segmentation and diameter at breast height estimation with mobile LiDAR. *IEEE Access* **2021**, *9*, 24314–24325. [CrossRef]
31. Olofsson, K.; Holmgren, J.; Olsson, H. Tree stem and height measurements using terrestrial laser scanning and the RANSAC algorithm. *Remote Sens.* **2014**, *6*, 4323–4344. [CrossRef]
32. Reddy, R.S.; Jha, C.S.; Rajan, K.S. Automatic tree identification and diameter estimation using single scan terrestrial laser scanner data in central Indian forests. *Indian Soc. Remote* **2018**, *46*, 937–943. [CrossRef]
33. Liu, G.; Wang, J.; Dong, P.; Chen, Y.; Liu, Z. Estimating individual tree height and diameter at breast height (DBH) from terrestrial laser scanning (TLS) data at plot level. *Forests* **2018**, *9*, 398. [CrossRef]
34. Panagiotidis, D.; Abdollahnejad, A.; Slavík, M. Assessment of stem volume on plots using terrestrial laser scanner: A precision forestry application. *Sensors* **2021**, *21*, 301. [CrossRef] [PubMed]
35. Koren, M.; Mokros, M.; Bucha, T. Accuracy of tree diameter estimation from terrestrial laser scanning by circle-fitting methods. *Int. J. Appl. Earth Obs. Geoinform.* **2017**, *63*, 122–128.
36. Monika, M.L.; Guang, Z. Retrieving forest inventory variables with terrestrial laser scanning (TLS) in urban heterogeneous forest. *Remote Sens.* **2012**, *4*, 1–20.
37. Panagiotidis, D.; Abdollahnejad, A. Accuracy assessment of total stem volume using close-range sensing: Advances in precision forestry. *Forests* **2021**, *12*, 7176. [CrossRef]
38. You, L.; Wei, J.; Liang, X.; Lou, M.; Pang, Y.; Song, X. Comparison of numerical calculation methods for stem diameter retrieval using terrestrial laser data. *Remote Sens.* **2021**, *13*, 17809. [CrossRef]
39. Stovall, A.E.L.; Vorster, A.G.; Anderson, R.S.; Evangelista, P.H.; Shugart, H.H. Non-destructive aboveground biomass estimation of coniferous trees using terrestrial LiDAR. *Remote Sens. Environ.* **2017**, *200*, 31–42. [CrossRef]
40. You, L.; Tang, S.; Song, X.; Lei, Y.; Zang, H.; Lou, M.; Zhuang, C. Precise measurement of stem diameter by simulating the path of diameter tape from terrestrial laser scanning data. *Remote Sens.* **2016**, *8*, 717. [CrossRef]
41. Hackenberg, J.; Spiecker, H.; Calders, K.; Disney, M.; Raumonen, P. SimpleTree-an efficient open source tool to build tree models from TLS clouds. *Forests* **2015**, *6*, 4245–4294. [CrossRef]
42. Raumonen, P.; Kaasalainen, M.; Åkerblom, M.; Kaasalainen, S.; Kaartinen, H.; Vastaranta, M.; Holopainen, M.; Disney, M.; Lewis, P. Fast automatic precision tree models from terrestrial laser scanner data. *Remote Sens.* **2013**, *5*, 491–520. [CrossRef]
43. Du, S.; Lindenbergh, R.; Ledoux, H.; Stoter, J.; Nan, L. AdTree: Accurate detailed and automatic modelling of laser-scanned trees. *Remote Sens.* **2019**, *11*, 2074. [CrossRef]
44. Ye, N.; van Leeuwen, L.; Nyktas, P. Analysing the potential of UAV point cloud as input in quantitative structure modelling for assessment of woody biomass of single trees. *Int. J. Appl. Earth Obs. Geoinform.* **2019**, *81*, 47–57. [CrossRef]
45. Liang, X.; Hyyppä, J.; Kaartinen, H.; Lehtomäki, M.; Pyörälä, J.; Pfeifer, N.; Holopainen, M.; Brolly, G.; Francesco, P.; Hackenberg, J.; et al. International benchmarking of terrestrial laser scanning approaches for forest inventories. *ISPRS J. Photogramm.* **2018**, *144*, 137–179. [CrossRef]
46. Mokroš, M.; Mikita, T.; Singh, A.; Tomaštík, J.; Chudá, J.; Wężyk, P.; Kuželka, K.; Surový, P.; Klimánek, M.; Zięba-Kulawik, K.; et al. Novel low-cost mobile mapping systems for forest inventories as terrestrial laser scanning alternatives. *Int. J. Appl. Earth Obs. Geoinform.* **2021**, *104*, 102512. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

forests

MDPI

*Article*

# Extracting the DBH of Moso Bamboo Forests Using LiDAR: Parameter Optimization and Accuracy Evaluation

**Longwei Li [1,2,3,†], Linjia Wei [1,†], Nan Li [2,3,*], Shijun Zhang [1], Zhicheng Wu [2], Miaofei Dong [2] and Yuyun Chen [4]**

[1] School of Resources and Environmental Engineering, Anhui University, Hefei 230601, China; lilw@chzu.edu.cn (L.L.); x22301048@stu.ahu.edu.cn (L.W.); x23301161@stu.ahu.edu.cn (S.Z.)
[2] School of Geographic Information and Tourism, Chuzhou University, Chuzhou 239000, China; 2021210074@mail.chzu.edu.cn (Z.W.); 2021210113@mail.chzu.edu.cn (M.D.)
[3] Anhui Province Key Laboratory of Physical Geographic Environment, Chuzhou 239000, China
[4] Shanghai Ubiquitous Navigation Technology Co., Ltd., Shanghai 201799, China; chenyuyun@ubinavi.com.cn
* Correspondence: linan@chzu.edu.cn
† These authors contributed equally to this work.

**Abstract:** The accurate determination of the Diameter at Breast Height (DBH) of Moso bamboo is crucial for estimating biomass and carbon storage in Moso bamboo forests. In this research, we utilized handheld LiDAR point cloud data to extract the DBH of Moso bamboo and enhanced the accuracy of diameter fitting by optimizing denoising parameters. Specifically, we fine-tuned two denoising parameters, neighborhood point number and standard deviation multiplier, across five gradient levels for denoising. Subsequently, DBH fitting was conducted on data processed with varying denoising parameters, followed by a precision evaluation to investigate the key factors influencing the accuracy of Moso bamboo DBH fitting. The research results indicate that a handheld laser was used to scan six plots, from which 132 single Moso bamboo trees were selected. Out of these, 122 single trees were successfully segmented and identified, achieving an accuracy rate of 92.4% in identifying single Moso bamboo trees, with an average accuracy of 95.64% in extracting DBH for individual plants; the mean error was ±1.8 cm. Notably, setting the minimum neighborhood point to 10 resulted in the highest fitting accuracy for DBH. Moreover, the optimal standard deviation multiplier threshold was found to be 1 in high-density forest plots and 2 in low-density forest plots. Forest condition and slope were identified as the primary factors impacting the accuracy of Moso bamboo DBH fitting.

**Keywords:** Moso bamboo; handheld LiDAR; individual tree detection; DBH; error analysis

## 1. Introduction

Moso bamboo (*Phyllostachys edulis*) is an evergreen plant that is commonly found in tropical and subtropical regions [1]. It holds significant economic importance in China due to its widespread distribution [2–4]. The edible bamboo shoots and versatile material properties of Moso bamboo make it highly valuable [5]. The extensive root system of Moso bamboo plays a crucial ecological role in windbreaks, sand stabilization, water conservation, and soil preservation [6]. Moso bamboo exhibits rapid growth, achieving maturity within a short span of 50 to 60 days from the emergence of shoots [7]. Its distinctive mechanism of biomass accumulation highlights its substantial potential for carbon sequestration, making a significant contribution to the global carbon sink [8,9].

Accurately estimating the individual parameters of Moso bamboo is crucial for determining its carbon storage capacity. Forestry researchers evaluate tree health by monitoring the growth of individual trees in a forest, often using quantitative indicators like DBH to ensure consistency in forest inventory results [10]. Laser scanning technology employs a laser emitter to emit pulsed laser beams and scan the surrounding environment, collecting precise three-dimensional point cloud data by measuring distances to target surfaces. This

process effectively captures detailed spatial information, including the shape, size, and depth of objects in a scanned area. [11]. This technology is particularly adept at extracting individual tree parameters for forestry applications [12]. Polewski et al. (2019) introduced an algorithm for matching unmarked airborne laser scanning (ALS) and backpack LiDAR data, enhancing the measurement of forest resources through unmarked data matching [13]. Cai et al. (2021) enhanced DBH fitting accuracy by filtering out noise points with varying intensities in point cloud data [14]. Brolly et al. (2021) utilized LiDAR technology to estimate individual tree structural parameters by creating digital terrain models (DTMs), tree canopy surface models, and trunk models to derive stand parameters [15]. Cai et al. (2018) developed a method to rapidly determine bamboo age by analyzing the relationship between point cloud data echo intensity and bamboo age, achieving an accuracy of 92.5% [16]. Li et al. (2022) conducted a study comparing crown heights measured from individual trees in sample sites with those extracted from LiDAR point cloud data, resulting in a Root Mean Square Error (*RMSE*) of 1.33 m and a coefficient of determination ($R^2$) of 0.96 [17].

Several studies have utilized Terrestrial Laser Scanning (TLS) data to analyze the extraction of DBH in Moso bamboo. Huang et al. (2021) employed TLS technology to capture point cloud data from Moso bamboo forest locations [18]. By conducting a thorough examination of bamboo stem shapes and forest conditions, they introduced a novel stem recognition method that enhanced the identification rate of individual Moso bamboo plants in point cloud data and the accuracy of DBH estimation. Similarly, Jiang et al. (2022) conducted TLS multi-station scanning to gather point cloud data from Moso bamboo forests [19]. By fitting a circular cross-section of bamboo stems and establishing a longitudinal axis of bamboo stems, they derived the DBH and pole length of Moso bamboo, minimizing parameter errors caused by bamboo pole bending. These studies primarily focus on enhancing the accuracy of DBH estimation by investigating the influence of Moso bamboo's growth structure. Additionally, the quality of point cloud data significantly impacts the accuracy of DBH estimation in Moso bamboo. Noise points within the data are identified as a key factor affecting the quality of point cloud data [20].

This study investigates the factors influencing the denoising of point cloud data, focusing on two main questions: (1) How does the fitting accuracy of Moso bamboo DBH change under various denoising parameter conditions? (2) What are the optimal denoising parameters? The analysis in this paper aims to identify the sources of errors and offer insights for parameter extraction for individual Moso bamboo using handheld LiDAR technology.

## 2. Materials and Methods

### *2.1. Study Area*

The study area is situated in the Taipinghu and Wushi Towns, Huangshan City, in the southern region of Anhui Province, China (Figure 1a–c). This area experiences a subtropical humid monsoon climate, with an average annual rainfall of 1274.7 mm, 1711.0 h of sunshine per year, and an average annual temperature of 16.6 °C in 2022 (Huangshan Statistical Bureau, 2023, Huangshan, China). The study area is rich in forest resources and benefits from favorable climate conditions that promote vegetation growth. Notable tree species present include *Phyllostachys edulis*, *Cunninghamia lanceolata*, *Pinus massoniana*, and *Liquidambar formosana*, among others (Huangshan Forestry Bureau, 2021, Huangshan, China).

### *2.2. Data Collection*

#### 2.2.1. Field Data

Six sample sites of Moso bamboo forests were selected in the study area (Figure 1c), each measuring 20 m × 20 m (Figure 2a). Basic information on the sample sites can be found in Table 1. Trees with DBH greater than 5 cm were individually tagged and measured using a diameter ruler, with species category information recorded. Additionally, the coordinates of the center point and four corners of each sample were obtained using the Real-Time

Kinematic (RTK) technique of the Global Navigation Satellite System (GNSS) receiver for later matching with the results of individual tree segmentation.

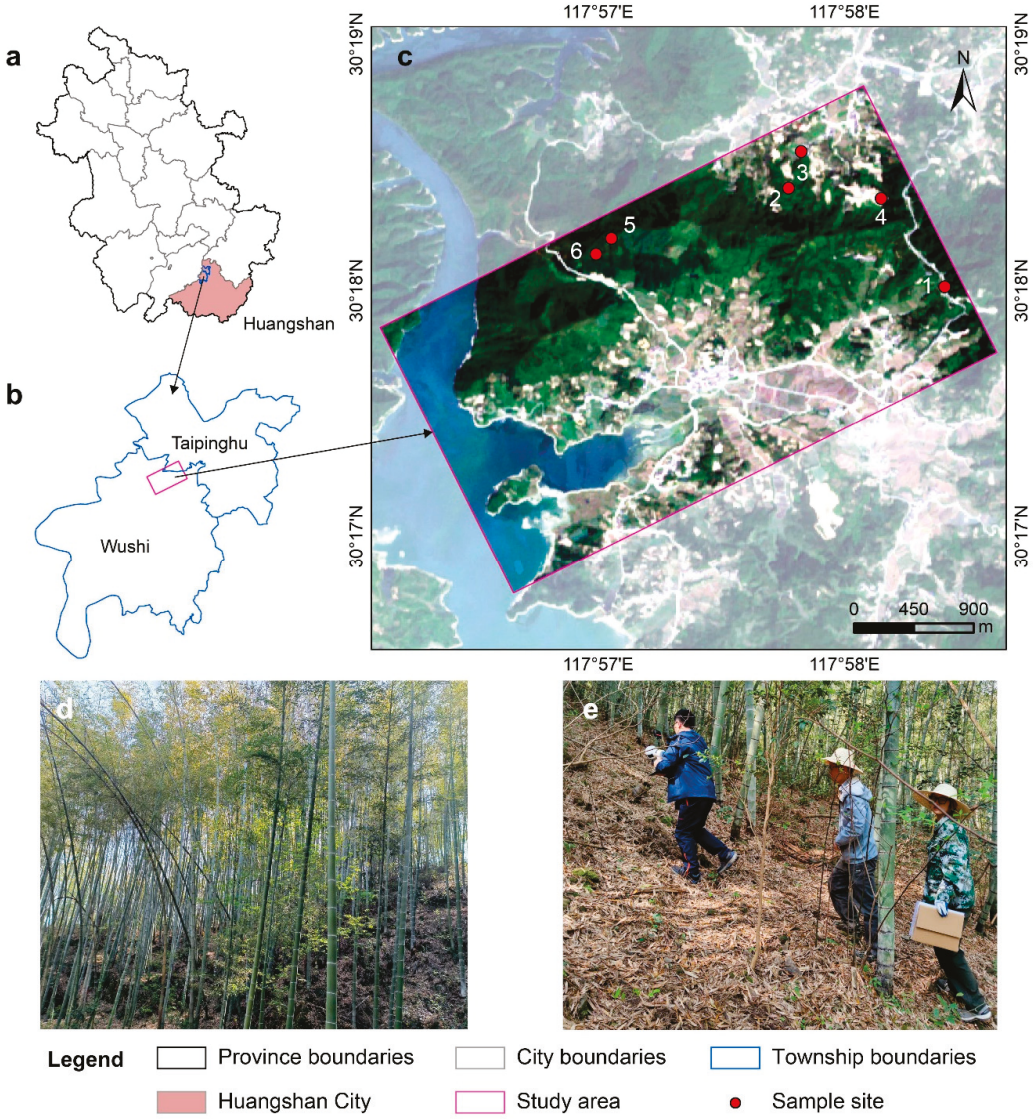



**Figure 1.** Study area. (**a**) Anhui Province, China. (**b**) Taipinghu and Wushi townships in Huangshan City. (**c**) Sample sites. (**d**) Forest condition map of site 3. (**e**) Collecting point cloud data using LiGrip-V100.

**Table 1.** Summary of sample site information for Moso bamboo forest.

| Serial Number | Sample Size | Number of Plants | Under-Forest Condition | Average Slope |
|---|---|---|---|---|
| Site 1 | 20 m × 20 m | 27 | Few shrubs | 5° |
| Site 2 | 20 m × 20 m | 18 | Shrubby | 9° |
| Site 3 | 20 m × 20 m | 27 | Shrubby | 13° |
| Site 4 | 20 m × 20 m | 20 | Shrubby | 25° |
| Site 5 | 20 m × 20 m | 20 | Few shrubs | 15° |
| Site 6 | 20 m × 20 m | 20 | Few shrubs | 5° |

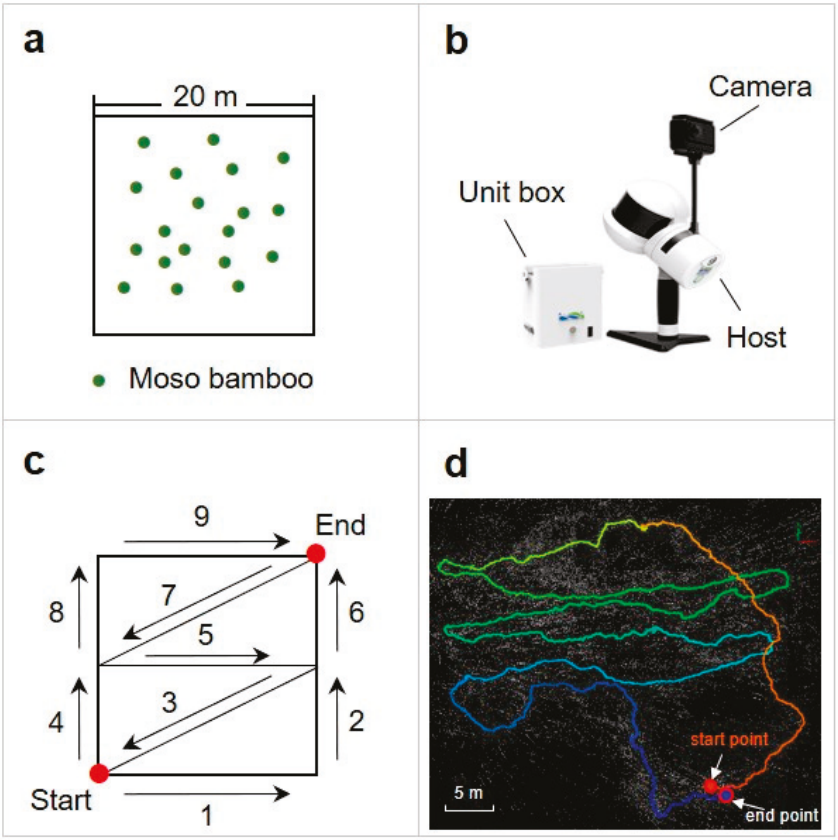



**Figure 2.** Sample site information. (**a**) Schematic diagram of sample site. (**b**) LiGrip-V100 handheld rotary laser scanner. (**c**) *Z*–shaped route. (**d**) Scanning trajectory.

#### 2.2.2. Handheld LiDAR Data

The handheld LiDAR equipment was Green Valley International's model, the LiGrip-V100 handheld rotary laser scanner (Green Valley, Bejing, China), with specifications shown in Table 2. Given the dense Moso bamboo forest terrain with a certain slope, a LiGrip-V100 handheld rotary laser scanner was employed with a double *Z*–shaped route to optimize point cloud data acquisition (Figure 2b,c). To maintain data quality, a constant scanning speed was maintained while ensuring equipment stability to prevent rotational movements. The walking speed was maintained between 0.8 and 1.2 m/s depending on the terrain. The cloud profiles of the sample site are illustrated in Figure 3.

**Table 2.** Handheld LiDAR parameters.

| Performance Indicators | Parameters |
|---|---|
| Laser Sensor | VLP–16 |
| LiDAR Accuracy | ±3 cm |
| Relative Accuracy | ≤3 cm |
| Absolute Accuracy | 5 cm |
| Size | L270 mm × W210 mm × H120 mm |
| Laser Wavelength | 903 nm |
| Scan Rate | 300000 pts/s |
| View Angle Range | 280°~360° (Horizontal); –90°~90° (Vertical) |
| Scan Range | 100 m |

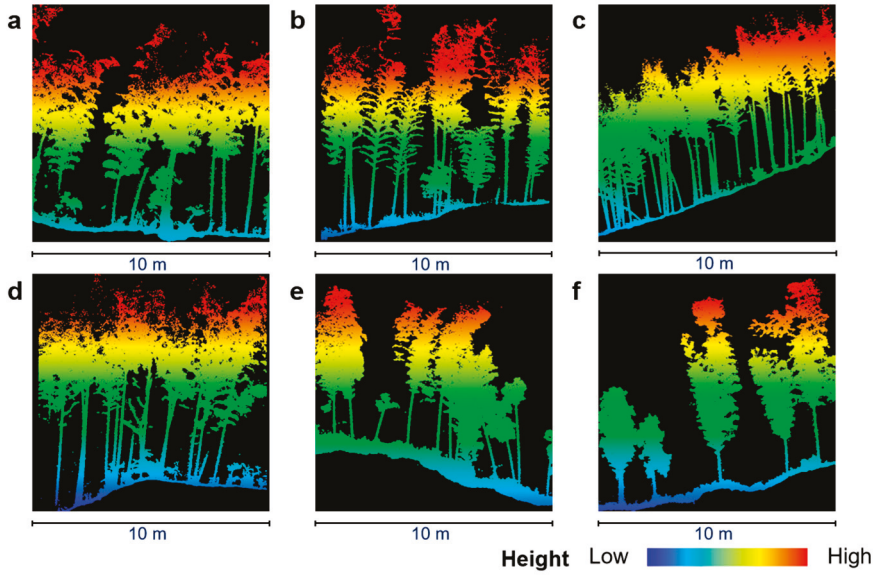



**Figure 3.** Sample site cloud profiles. (**a**–**f**) refer to sample sites 1–6, respectively.

### *2.3. Moso Bamboo DBH Extraction Methods*

The process of extracting the DBH of Moso bamboo included denoising, ground point classification, point cloud normalization, individual tree segmentation, extraction of single tree DBH, accuracy verification, and denoising parameter optimization (Figure 4).

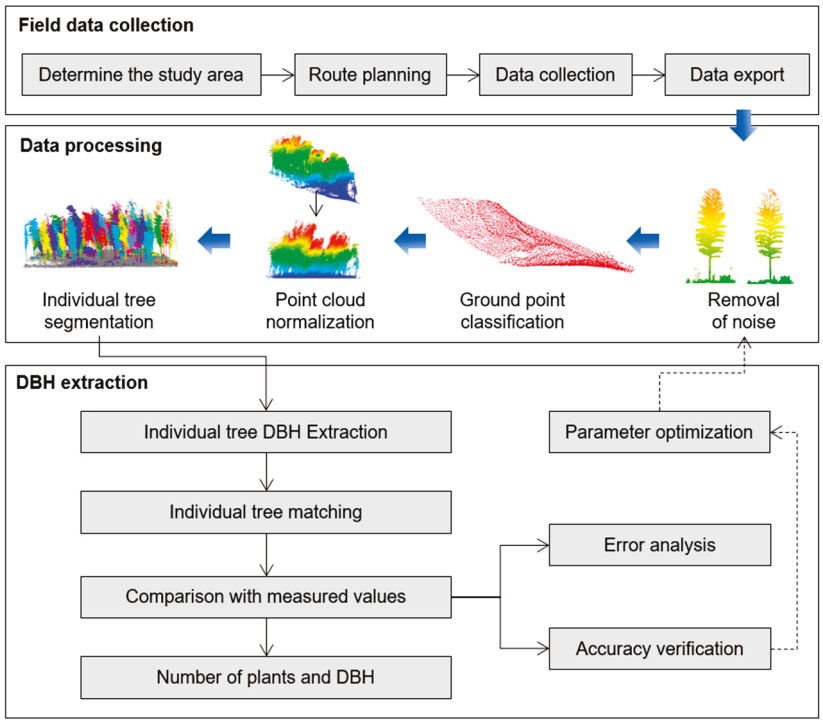



**Figure 4.** Framework for Moso bamboo DBH extraction.

#### 2.3.1. Removal of Noise

The handheld LiDAR data was processed using LiDAR360 V7.0 software (Green Valley, Beijing, China). Initially, the point cloud data underwent cropping based on the sample extent and filtering to remove redundancy and noise (Figure 5). Statistical filtering was used for denoising in this study. The principle was to search for a specified number of neighboring points for each point, calculating the mean of the distances from each

point to its neighbors. The median and standard deviation of these distance means were then calculated. Points whose mean distance was greater than the maximum distance (maximum distance = median + number of standard deviations × standard deviation) were identified as noise and excluded [21]. The parameters impacting the denoising outcome included the number of neighboring points and the number of standard deviations.

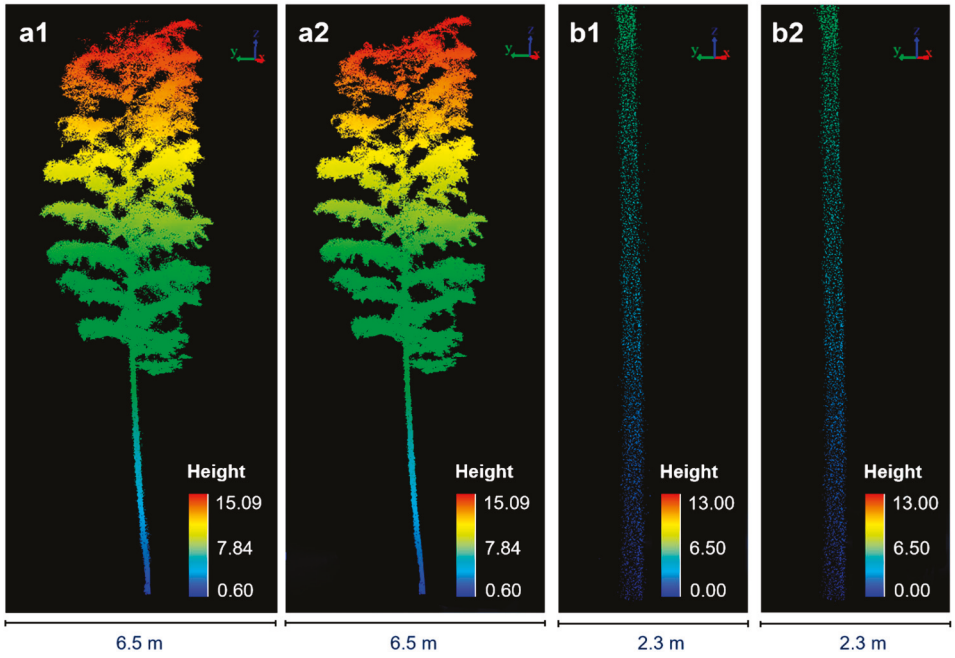



**Figure 5.** Removal of noise. (**a1**,**a2**) are whole plant of Moso bamboo; (**b1**,**b2**) are trunk of Moso Bamboo; (**a1**,**b1**) are point clouds before denoising; (**a2**,**b2**) are point clouds after denoising.

### 2.3.2. Ground Point Classification and Normalization

The improved progressive triangulated irregular network densification algorithm was applied for ground point classification to mitigate the impact of terrain factors on individual tree segmentation and DBH extraction. The main process consisted of point cloud gridding, selecting seed points, constructing a triangulated irregular network (TIN) using the seed points, and iteratively densifying the TIN. A digital elevation model (DEM) with a resolution of 0.5 m was generated through irregular triangular mesh interpolation [22]. Normalized point cloud data were obtained by subtracting the absolute elevation Z values of the data from the ground points' elevation to eliminate topography influence on tree height estimation [23] (Figure 6).

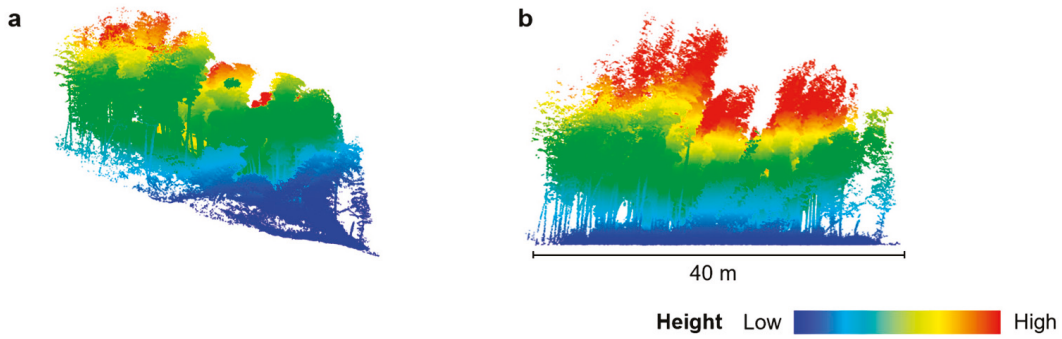



**Figure 6.** Point cloud data normalization. (**a**) Before normalization. (**b**) After normalization.

2.3.3. Individual Tree Segmentation

The point cloud segmentation algorithm was utilized for individual tree segmentation [24]. This methodology was based on the spatial separation characteristics between tree canopies. Initially, local maxima detection was performed on the discrete point cloud, assuming that these local highest points represented the treetop. Using this point as a seed point, the point cloud was segmented using a region-growing algorithm. This process was iterated until all significant trees were segmented (Figure 7).




**Figure 7.** Individual tree segmentation result, adjacent trees are distinguished by different colors.

2.3.4. DBH Extraction

The DBH of bamboo was extracted using the method of least squares circle fitting. This method achieves the optimal circle center and radius by minimizing the squared error of the sum of distances from all data points to the circumference [25]. In this study, the point cloud of the trunk at 1.3 m above ground was selected for circle fitting (Figure 8).

$$f(x_a, y_a, R) = \sum d_i^2 \tag{1}$$

$$r_i = \sqrt{(x_i - x_a)^2 + (y_i - y_a)^2} \tag{2}$$

where $d_i$ is the distance of each point from the center of the fitted circle at 1.3 m ($d_i = r_i - R$), $x_a$ and $y_a$ are the coordinates of the center of the determined fitted circle, $x_i$ and $y_i$ are the coordinates of the center of the iteratively fitted circle, $r_i$ is the radius of the circle fitted at different points, and $R$ is the radius of the determined fitted circle.

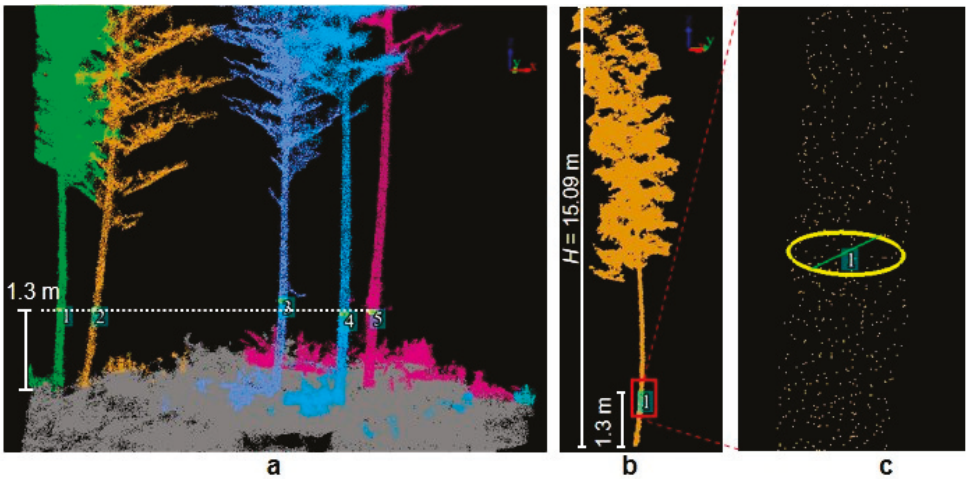



**Figure 8.** Extraction of DBH. (**a**,**b**) DBH fitting height. (**c**) is the partial enlargement of the red square and the yellow circle is the fitted circle.

2.3.5. Denoising Parameter Optimization

In order to identify the best denoising parameters for the point cloud data of a Moso bamboo forest, this study initially set the standard deviation multiplier at 5. Subsequently, varying numbers of minimum neighborhood points (10, 20, 30, 40, 50) were tested for denoising. The DBH was then fitted, and metrics such as precision (*P*), coefficient of determination ($R^2$), *RMSE*, and relative Root Mean Square Error (*rRMSE*) were calculated to compare the fitted values with the measured values. This analysis aimed to determine the optimal parameters for the minimum neighborhood point. Following this, using the identified optimal parameters, the standard deviation multiplier was adjusted to 1, 2, 3, 4, and 5 for further denoising processes. The DBH was fitted once again, and a comparison was made with the measured values to ascertain the most suitable denoising parameters.

2.3.6. Accuracy Assessment

According to the measured RTK coordinates of the Moso bamboo and the point cloud coordinates, a correspondence was established between the measured DBH and the fitted DBH for accuracy verification. This study utilized *P*, $R^2$, *RMSE*, and *rRMSE* to assess the accuracy. A higher $R^2$ indicated a stronger correlation between the measured and fitted values, while a lower *RMSE* signified greater predictive value. The formula is as follows:

$$P = 1 - \frac{1}{n}\sum_{i=1}^{n}\frac{|W_i - w_i|}{W_i} \tag{3}$$

$$R^2 = \frac{\sum_{i=1}^{n}(w_i - \overline{w_i})(W_i - \overline{W_i})}{\sqrt{\sum_{i=1}^{n}(w_i - \overline{w_i})^2 \sum_{i=1}^{n}(W_i - \overline{W_i})^2}} \tag{4}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(W_i - \overline{W_i})^2} \tag{5}$$

$$rRMSE = \frac{RMSE}{\frac{1}{n}\sum_{i=1}^{n} w_i} \tag{6}$$

where *P* represents precision, $R^2$ represents the coefficient of determination, *RMSE* represents the Root Mean Square Error, *rRMSE* represents the relative Root Mean Square Error, *n* is the number of correctly segregated individual Moso bamboo plants, $W_i$ is the fitted DBH, $w_i$ is the measured DBH, $\overline{W_i}$ is the mean value of $W_i$, and $\overline{w_i}$ is the mean value of $w_i$.

## 3. Results

### *3.1. Identification Results and Analysis of Moso Bamboo*

Table 3 displays the number of plants measured, recognized by the handheld LiDAR, and undetected and the detection rate across the six sample sites. The identification accuracy of Moso bamboo varied among the plots, with a 100% identification rate in site 6, 88.89% in sites 1 and 3, and an average of 93% overall. The accuracy of Moso bamboo identification was influenced by factors such as RTK accuracy and the quality of the point cloud data. Site 6 had a sparse distribution of Moso bamboo and fewer understory shrubs, exhibiting high accuracy in single Moso bamboo identification. Conversely, sites 1 and 3 had dense bamboo and poor RTK signals in the understory, resulting in the low identification of single trees.

**Table 3.** Identification results of Moso Bamboo.

| Serial Number | Number of Plants Measured | Number of Plants Identified | Number of Undetected Strains | Recognition Rate (%) |
|---|---|---|---|---|
| Site 1 | 27 | 24 | 3 | 88.89 |
| Site 2 | 18 | 17 | 1 | 94.44 |
| Site 3 | 27 | 24 | 3 | 88.89 |
| Site 4 | 20 | 19 | 1 | 95 |
| Site 5 | 20 | 18 | 2 | 90 |
| Site 6 | 20 | 20 | 0 | 100 |

### 3.2. Parameter Optimization Results and Analysis

The quality of the point cloud data for Moso bamboo forests was enhanced by optimizing denoising parameters prior to fitting the chest diameter. This study investigated the optimal minimum neighborhood point parameter for statistical filtering [26] method to denoise the point cloud data of Moso bamboo forests. Five parameter values were tested for denoising, with a fixed standard deviation multiplier of 5. The extracted DBHs were then compared with the measured values, as shown in Table 4. The results revealed that a minimum neighboring point of 10 yielded the highest accuracy in fitting the cloud DBH of Moso bamboo points in sites 1, 2, 3, 4, and 6, with accuracy rates of 96.37%, 90.44%, 86.80%, 92.39%, and 94.44%, respectively. The $R^2$ values ranged from 0.344 to 0.959, with site 1 achieving the highest fit. The *RMSE* value for site 1 was 0.416, with an *rRMSE* of 4.38%, indicating the best overall fit. Site 5 showed a high accuracy of 90.03% for DBH fit at a minimum neighborhood point of 30, only slightly higher than the accuracy at a minimum neighborhood point of 10. Overall, the most accurate fitting of Moso bamboo diameter was achieved with a minimum neighborhood point of 10, suggesting that noise points could be more precisely removed under this parameter while retaining effective data points that revealed the structure of Moso bamboo.

**Table 4.** DBH fitting results after denoising with different minimum neighborhood point parameters.

| Serial Number | Minimum Number of Neighborhood Points | Accuracy (%) | $R^2$ | *RMSE* | *rRMSE* (%) |
|---|---|---|---|---|---|
| Site 1 | 10 | 96.37 | 0.959 | 0.416 | 4.38 |
| | 20 | 94.39 | 0.872 | 0.632 | 6.24 |
| | 30 | 93.87 | 0.906 | 0.700 | 7.36 |
| | 40 | 93.64 | 0.839 | 0.713 | 7.5 |
| | 50 | 94.48 | 0.898 | 0.647 | 6.80 |
| Site 2 | 10 | 90.44 | 0.344 | 1.431 | 16.03 |
| | 20 | 89.39 | 0.444 | 0.916 | 15.69 |
| | 30 | 88.08 | 0.214 | 1.567 | 17.56 |
| | 40 | 88.40 | 0.205 | 1.613 | 18.07 |
| | 50 | 87.02 | 0.227 | 1.593 | 17.85 |
| Site 3 | 10 | 86.80 | 0.519 | 1.572 | 13.38 |
| | 20 | 85.95 | 0.435 | 1.689 | 14.38 |
| | 30 | 86.68 | 0.511 | 1.606 | 13.67 |
| | 40 | 83.54 | 0.586 | 1.802 | 15.35 |
| | 50 | 82.92 | 0.330 | 1.989 | 16.94 |
| Site 4 | 10 | 92.39 | 0.637 | 1.073 | 9.32 |
| | 20 | 89.66 | 0.206 | 1.492 | 12.97 |
| | 30 | 91.83 | 0.621 | 1.007 | 8.75 |
| | 40 | 91.37 | 0.623 | 1.164 | 10.12 |
| | 50 | 90.67 | 0.539 | 1.220 | 10.61 |

**Table 4.** *Cont.*

| Serial Number | Minimum Number of Neighborhood Points | Accuracy (%) | $R^2$ | *RMSE* | *rRMSE* (%) |
|---|---|---|---|---|---|
| Site 5 | 10 | 89.68 | 0.691 | 1.513 | 14.02 |
| | 20 | 88.55 | 0.576 | 1.220 | 17.89 |
| | 30 | 90.03 | 0.728 | 1.398 | 12.96 |
| | 40 | 89.30 | 0.681 | 1.545 | 14.32 |
| | 50 | 88.65 | 0.634 | 0.586 | 14.70 |
| Site 6 | 10 | 94.44 | 0.865 | 0.702 | 7.26 |
| | 20 | 93.09 | 0.819 | 0.759 | 7.77 |
| | 30 | 93.87 | 0.832 | 0.725 | 7.50 |
| | 40 | 92.47 | 0.770 | 0.883 | 9.14 |
| | 50 | 92.78 | 0.808 | 0.784 | 8.11 |

Based on a minimum neighborhood point set to 10, denoising was performed with five standard deviation multiples, as shown in Table 5. Sites 3, 4, 5, and 6 exhibited the highest accuracy of fit for DBH at a standard deviation multiplier of 2, with accuracies of 91.05%, 92.63%, 92.12%, and 94.72%, respectively, corresponding to $R^2$ values of 0.645, 0.637, 0.740, and 0.872. The *RMSE* and *rRMSE* for the fitted and measured values of DBH for site 6 were the smallest, at 0.668 and 6.91%, respectively. On the other hand, sites 1 and 2 showed the highest accuracy of fitting DBH at a standard deviation multiplier of 1, with accuracies of 96.37% and 90.44%, respectively. The denser growth of Moso bamboo in sites 1 and 2 led to more compact and noisy point cloud data, where a standard deviation multiplier of 1 effectively removed noise and yielded higher-quality point cloud data. In contrast, sites 3, 4, 5, and 6 displayed sparse growth of Moso bamboo, with relatively fewer noise points in the point cloud data. Adjusting the standard deviation multiplier to 1 impacted the structure of the Moso bamboo point cloud itself, influencing the accuracy of DBH fit, while a higher standard deviation multiplier preserved more noise points, also affecting the fit of DBH.

**Table 5.** DBH fitting results after denoising with different standard deviation multiples.

| Serial Number | Standard Deviation Multiplier | Accuracy (%) | $R^2$ | *RMSE* | *rRMSE* (%) |
|---|---|---|---|---|---|
| Site 1 | 1 | 96.37 | 0.959 | 0.416 | 4.38 |
| | 2 | 94.06 | 0.896 | 0.728 | 7.66 |
| | 3 | 93.60 | 0.778 | 0.870 | 9.15 |
| | 4 | 94.20 | 0.215 | 0.761 | 8.00 |
| | 5 | 94.91 | 0.902 | 0.633 | 6.66 |
| Site 2 | 1 | 90.44 | 0.344 | 1.431 | 16.03 |
| | 2 | 89.58 | 0.411 | 1.357 | 15.20 |
| | 3 | 88.30 | 0.241 | 1.539 | 17.24 |
| | 4 | 88.06 | 0.282 | 1.492 | 16.71 |
| | 5 | 89.09 | 0.424 | 1.365 | 15.29 |
| Site 3 | 1 | 86.80 | 0.519 | 1.57 | 13.38 |
| | 2 | 91.05 | 0.645 | 1.157 | 9.85 |
| | 3 | 88.39 | 0.508 | 1.408 | 11.99 |
| | 4 | 87.89 | 0.567 | 1.480 | 12.60 |
| | 5 | 88.97 | 0.486 | 1.390 | 11.84 |
| Site 4 | 1 | 92.39 | 0.515 | 1.073 | 9.32 |
| | 2 | 92.63 | 0.637 | 1.071 | 9.31 |
| | 3 | 92.51 | 0.367 | 1.154 | 10.03 |
| | 4 | 92.49 | 0.449 | 1.159 | 10.08 |
| | 5 | 92.52 | 0.499 | 1.118 | 9.72 |

**Table 5.** *Cont.*

| Serial Number | Standard Deviation Multiplier | Accuracy (%) | $R^2$ | *RMSE* | *rRMSE* (%) |
|---|---|---|---|---|---|
| Site 5 | 1 | 89.68 | 0.691 | 1.513 | 14.02 |
| | 2 | 92.12 | 0.740 | 1.316 | 12.20 |
| | 3 | 89.70 | 0.578 | 1.710 | 15.85 |
| | 4 | 89.98 | 0.544 | 1.771 | 16.42 |
| | 5 | 89.21 | 0.583 | 1.737 | 16.11 |
| Site 6 | 1 | 94.44 | 0.865 | 0.702 | 7.26 |
| | 2 | 94.72 | 0.872 | 0.668 | 6.91 |
| | 3 | 94.36 | 0.857 | 0.731 | 7.56 |
| | 4 | 93.23 | 0.816 | 0.822 | 8.50 |
| | 5 | 94.10 | 0.868 | 0.680 | 7.03 |

### *3.3. DBH Extraction Results and Analysis*

The results presented in Figure 9 are based on the optimal parameter fitted to the diameter of Moso bamboo chests. The correlations between the fitted and measured values of Moso bamboo diameter at each sample site were 0.959, 0.344, 0.654, 0.515, 0.740, and 0.872, respectively. Notably, site 1 exhibited the highest correlation between the fitted and measured values of diameter, with the lowest *RMSE* and *rRMSE* values of 0.416 and 4.38%, respectively. Following closely, site 6 showed an *RMSE* of 0.668 and *rRMSE* of 6.91%. Conversely, site 2 displayed the lowest correlation, resulting in the highest *RMSE* and *rRMSE* values of 1.431 and 16.03%.

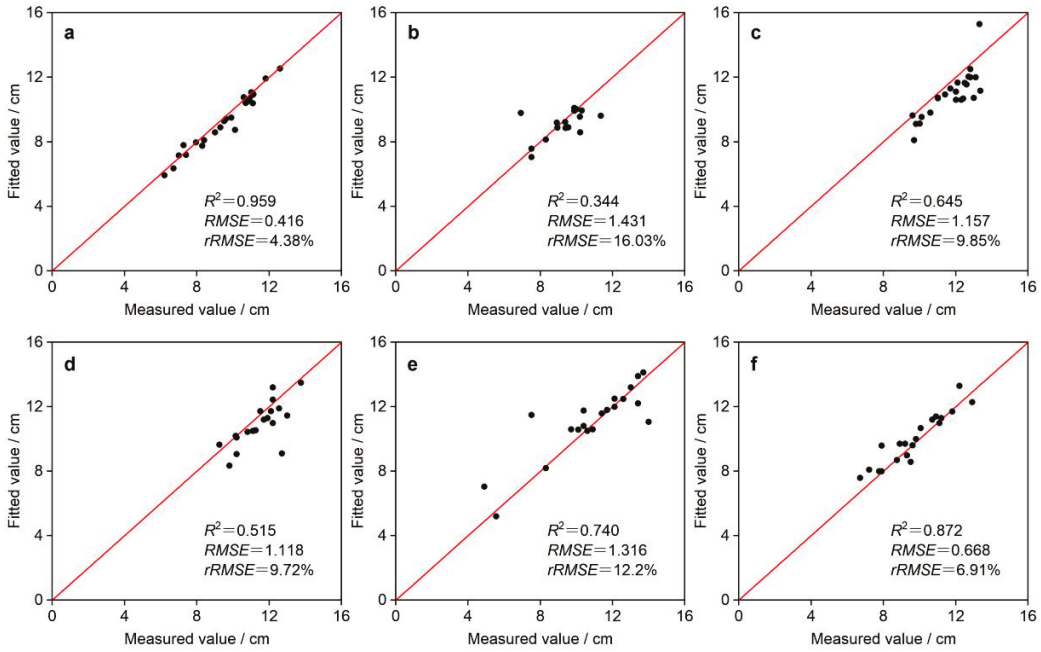



**Figure 9.** Regression analysis of fitted and measured values of bamboo DBH in the sample site. (**a**–**f**) represent site 1-6, respectively. Note: straight line represents $y = x$.

Site 1 exhibited a smaller slope with fewer overgrown shrubs shading the bamboo pole area, resulting in a higher-quality point cloud collection. In contrast, the discrepancy between the fitted and measured values of DBH for site 2 can be attributed to the presence of a greater variety of mixed shrubs in the understory. Additionally, the presence of intertwined and fallen Moso bamboo plants in site 2 may have led to inaccuracies in contour representation during single wood splitting, resulting in errors when fitting the DBH.

### 3.4. Analysis of Errors

In order to investigate the sources of error, eight Moso bamboo plants with significant discrepancies in the fitted values of DBH were chosen for analysis (Table 6).

**Table 6.** Error analysis table.

| Serial ID | Measured DBH (cm) | Fitted DBH (cm) |
|---|---|---|
| MZ02dm01 | 6.9 | 8.8 |
| MZ03dm04 | 13.4 | 11.2 |
| MZ03dm25 | 13.0 | 10.7 |
| MZ04dm10 | 12.7 | 9.1 |
| MZ04dm11 | 9.8 | 8.3 |
| MZ05dm03 | 7.5 | 11.5 |
| MZ05dm05 | 5.0 | 10.5 |
| MZ06dm12 | 7.9 | 9.7 |

The measured DBH of Moso bamboo MZ2dm01 was 6.9 cm, while the fitted DBH was 8.8 cm, indicating a difference of 1.9 cm between the measured and fitted values. This error was attributed to noisy point cloud data at 1.3 m, resulting in unclear contours (Figure 10a). Similarly, the measured DBH of Moso bamboo MD3dm04 was 13.4 cm, whereas the fitted DBH was 11.2 cm, showing a discrepancy of 2.2 cm. This discrepancy was due to excessive stray irradiation in the area, causing the laser light to be absorbed by branches and trunks, leading to inaccurate echo signals and an inability to accurately capture the profile of the single Moso bamboo plant (Figure 10b). Additionally, the measured DBH of Moso bamboo MD3dm25 was 13.0 cm, while the fitted DBH was 10.7 cm. The discrepancy in this case was caused by the close proximity of two Moso bamboo plants, resulting in misidentification. Subsequently, another Moso bamboo plant was manually refitted with a DBH of 11.5 cm, achieving an accuracy of 88.46% (Figure 10c). The measured DBH of Moso bamboo MD4dm10 was 12.7 cm, while the fitted DBH was 9.1 cm. Unfortunately, the point cloud data for this specimen were significantly missing (Figure 10d). In the case of MD4dm11, the measured DBH was 9.8 cm, with a fitted DBH of 8.3 cm. Similar to MD4dm10, the point cloud data were also missing (Figure 10e). Moving on to Moso bamboo MD5dm03, the measured diameter was 7.5 cm, but the fitted diameter was 11.5 cm, resulting in a large fitted value. The outline of the Moso bamboo in the point cloud appeared fuzzy and overlapping, possibly due to rapid scanner movements causing multiple scans of the same area, leading to an abundance of repetitive data points that impacted accuracy (Figure 10f). Continuing with MD5dm05, the measured DBH was 5.0 cm, while the fitted DBH was 10.6 cm, showing a large fitted value caused by fuzzy overlap in the point cloud data at 1.3 m (Figure 10g). Lastly, for Moso bamboo MD6dm12, the measured DBH was 7.9 cm, with a fitted DBH of 9.7 cm. The data indicated a significant break at 1.3 m, likely resulting from a splicing error during the normalization process due to the steep slope of the sample site (Figure 10h).

The challenges in accurately fitting the Moso bamboo DBH using handheld LiDAR data can be attributed to two main factors. Firstly, the dense forest conditions in the scanned sample plots, characterized by numerous low shrubs in the understory, lead to laser absorption and inadequate return, resulting in a point cloud with lower density and reduced quality. This hindered the precise representation of Moso bamboo outlines post-data processing. Secondly, the presence of tilted and overlapping Moso bamboo in some areas caused ambiguity in the point cloud data, further complicating accurate outline depiction. Additionally, in sample plots with steep slopes, maintaining consistent scanning speed and smooth travel paths proved challenging, impacting point cloud alignment and leading to stitching errors. Consequently, the accurate representation of Moso bamboo outlines became compromised.

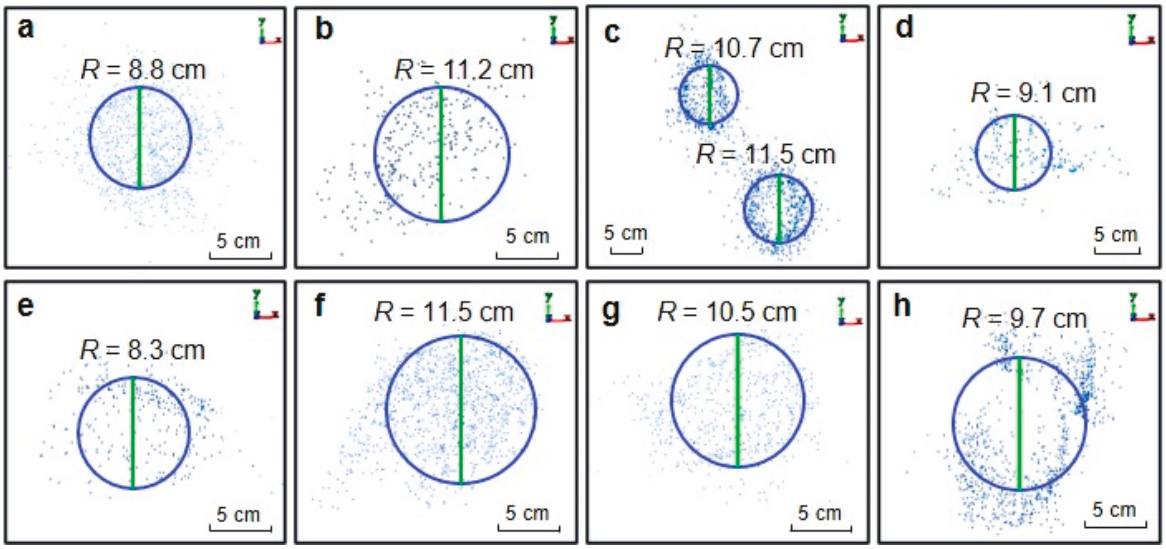



**Figure 10.** Error analysis for DBH fitting. (**a**–**h**) represent eight individual Moso bamboo with the largest DBH fitting error.

After excluding points with large errors, the results of the fitted and measured values of DBH are presented in Figure 11 and Table 7. The extraction accuracy of DBH for sample sites 2–6 generally improved, with an increase in $R^2$ and a decrease in both *RMSE* and *rRMSE*. Overall, after removing points with large errors, all evaluation parameters were significantly optimized.

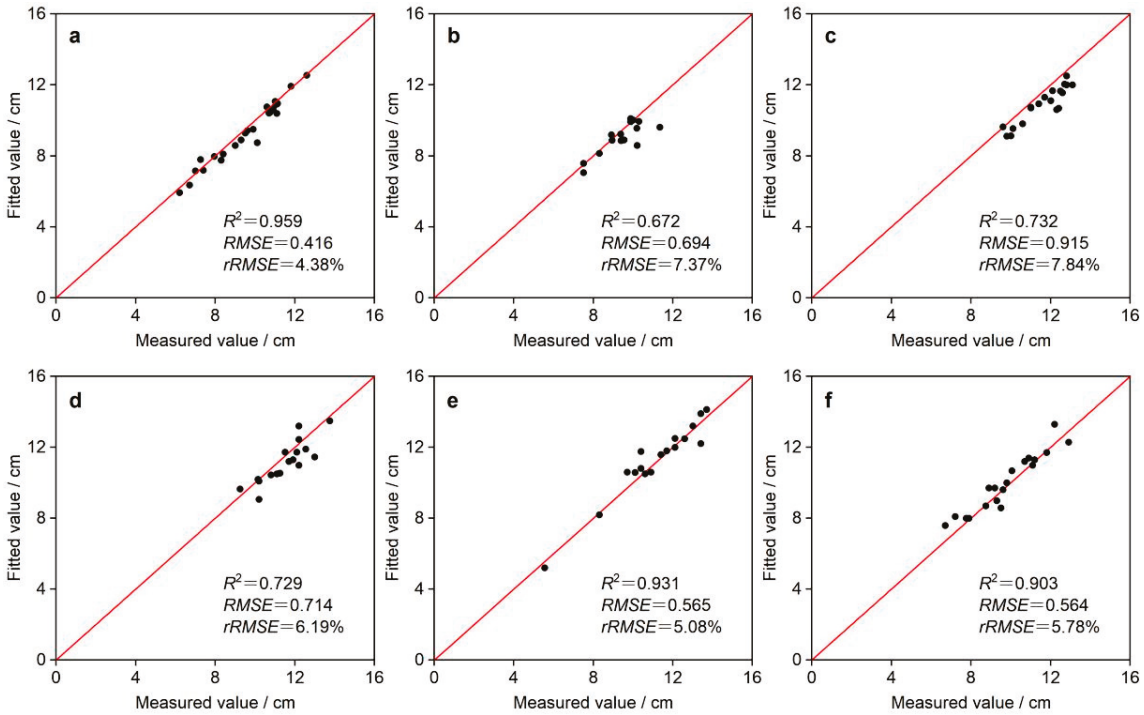



**Figure 11.** Regression analysis chart between the fitted and measured values of corrected DBH. (**a**–**f**) represent site 1-6, respectively. Note: straight line represents $y = x$.

**Table 7.** Precision evaluation results before and after removing points with significant errors.

| Sample Site | Pre-Deletion | | | | Post-Deletion | | | |
|---|---|---|---|---|---|---|---|---|
| | Accuracy (%) | $R^2$ | *RMSE* | *rRMSE* (%) | Accuracy (%) | $R^2$ | *RMSE* | *rRMSE* (%) |
| Site 1 | 96.37 | 0.959 | 0.416 | 4.38 | 96.37 | 0.959 | 0.416 | 4.38 |
| Site 2 | 90.44 | 0.344 | 1.431 | 16.03 | 95.42 | 0.672 | 0.694 | 7.37 |
| Site 3 | 91.05 | 0.645 | 1.157 | 9.85 | 94.24 | 0.732 | 0.915 | 7.84 |
| Site 4 | 92.63 | 0.637 | 1.071 | 9.31 | 95.48 | 0.729 | 0.714 | 6.19 |
| Site 5 | 92.12 | 0.740 | 1.316 | 12.20 | 96.71 | 0.931 | 0.565 | 5.08 |
| Site 6 | 94.72 | 0.872 | 0.668 | 6.91 | 95.60 | 0.903 | 0.564 | 5.78 |

## 4. Discussions

This study utilized a handheld LiDAR device to gather point cloud data from samples of Moso bamboo. The data underwent denoising, ground point classification, normalization, and single-wood segmentation operations. Subsequently, the single Moso bamboo DBH was extracted using the least-squares circle fitting method and compared with the measured DBH for accuracy evaluation.

Compared to terrestrial laser scanning and backpack laser scanning, handheld LiDAR is capable of scanning more complex forest environments. Terrestrial laser scanning, when scanning dense vegetation, often requires setting up multiple sites in different locations and spending a significant amount of time on data processing. On the other hand, backpack laser scans from a higher height than the scanning personnel, making it susceptible to damage from branches and leaves at lower branch heights or complex forest conditions. This can lead to sensor damage and a decrease in point cloud quality due to obstruction from branches and leaves. In contrast, the handheld LiDAR scanner is compact, lightweight, easy to carry, and simple to operate. It enables the quick collection of point cloud data in complex secondary forests, resulting in higher measurement efficiency.

The utilization of handheld LiDAR scanners for arborimetric measurements is a significant advancement in modern forestry and ecological research. This study has demonstrated the effectiveness of using handheld LiDAR scanners to measure the DBH of Moso bamboo. By utilizing point cloud data from these portable devices, additional parameters such as tree height, LAI, and canopy spread can be extracted. These variables play a crucial role in modeling tree growth patterns, understanding vertical canopy structure, and improving estimates of forest carbon sequestration—a critical aspect of global change ecology. Moreover, detailed characterizations of individual trees aid in creating accurate forest biomass maps and enhance our ability to monitor ecosystem services like carbon storage and habitat suitability. Apart from ecological applications, the high-resolution data obtained from handheld LiDAR systems can also support precision forestry initiatives, capturing the intricacies and diversity of forest environments. Therefore, future research should focus on exploring the potential of handheld LiDAR scanners in various dendrometric applications, including evaluating their suitability for measuring unconventional parameters, validating their accuracy against traditional methods, and incorporating them into innovative ecological models that simulate forest responses to environmental stressors.

In this study, only five parameter intervals were selected through the experimental screening of suitable denoising parameters. Subsequent processing can continue to further refine these parameters in order to improve the accuracy of DBH extraction. Precision forestry demands high accuracy in parameters such as single Moso bamboo identification and diameter extraction. The next step could involve attempting to fuse airborne and handheld LiDAR point cloud data to obtain higher-quality and more comprehensive point cloud data. This could pave the way for further research on parameters such as Moso bamboo tree height, bamboo age, biomass, and other related factors.

## 5. Conclusions

This research utilized handheld LiDAR technology to capture point cloud data of Moso bamboo forests in Taipinghu and Wushi Township, Huangshan City. This study conducted DBH fitting post-denoising using various parameters, assessed the accuracy, and examined the errors of the fitted values. The findings indicated a 93% accuracy in identifying Moso bamboo with a handheld LiDAR scanner and achieved an average DBH fitting accuracy of 95.64% after eliminating points with significant errors. The mean error was $\pm$1.8 cm. Notably, setting the number of neighboring points to 10 resulted in the highest fitting accuracy for Moso bamboo diameter. This study observed varying outcomes in the extraction of samples based on different stand densities, suggesting an optimal threshold of 1 for the standard deviation multiplier in dense Moso bamboo sample sites and 2 for sparse Moso bamboo sample sites. Additionally, forest condition and slope were identified as key factors influencing the error in fitting DBH to handheld LiDAR data. Issues such as shrubs obstructing the laser, intertwined bamboo growth affecting point cloud quality, and errors in data from samples with steep slopes were noted.

**Author Contributions:** Conceptualization, L.L., L.W. and N.L.; Methodology, L.L. and L.W.; Software, L.W. and S.Z.; Validation, L.W., S.Z. and Z.W.; Data curation, L.W., Z.W. and M.D.; Formal analysis, L.W., S.Z. and Z.W.; Funding acquisition, L.L. and M.D.; Investigation, L.L., L.W., N.L., S.Z., Z.W., M.D. and Y.C.; Visualization, L.L., L.W. and N.L.; Writing—original draft preparation, L.L., L.W. and N.L.; Writing—review and editing, L.L., L.W., N.L., S.Z., Z.W. and M.D.; Supervision, N.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was financially supported by the National Natural Science Foundation of China (grant no. 42101387), Natural Science Research Project for Anhui Universities (grant no. 2023AH030094), Chuzhou University Research and Development Fund for the Talent Startup Project (grant no. 2022XJZD08), Anhui Province Key Laboratory of Physical Geographic Environment (grant no. 2022PGE004), and the National College Student Innovation and Entrepreneurship Training Project (grant no. 202310377002).

**Data Availability Statement:** The raw data supporting the conclusions of this article will be made available by the authors on request.

**Acknowledgments:** The authors would like to thank Guiying Li, Tianzhen Wu, Wenlin Yang, Shitao Li, Hongfeng Xu, Kai Jian, and Xiaoyu Sun for their support in the field investigation.

**Conflicts of Interest:** Author Yuyun Chen is employed by the company Shanghai Ubiquitous Navigation Technology Co., Ltd. The remaining authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The Company Shanghai Ubiquitous Navigation Technology Co., Ltd. had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Liu, C.; Zhou, Y.; Qin, H.; Liang, C.; Shao, S.; Fuhrmann, J.J.; Chen, J.; Xu, Q. Moso bamboo invasion has contrasting effects on soil bacterial and fungal abundances, co-occurrence networks and their associations with enzyme activities in three broadleaved forests across subtropical China. *For. Ecol. Manag.* **2021**, *498*, 119549. [CrossRef]
2. Li, L.; Zhu, H.; Wu, T.; Wei, L.; Li, N. New landscape-perspective exploration of Moso bamboo forests under on/off-year phenomena and human activities. *Front. For. Glob. Change* **2023**, *6*, 1204329. [CrossRef]
3. Zheng, Y.; Fan, S.; Guan, F.; Xia, W.; Wang, S.; Xiao, X. Strip Clearcutting Drives Vegetation Diversity and Composition in the Moso Bamboo Forests. *For. Sci.* **2022**, *68*, 27–36. [CrossRef]
4. Yuen, J.Q.; Fung, T.; Ziegler, A.D. Carbon stocks in bamboo ecosystems worldwide: Estimates and uncertainties. *For. Ecol. Manag.* **2017**, *393*, 113–138. [CrossRef]
5. Scurlock, J.M.O.; Dayton, D.C.; Hames, B. Bamboo: An overlooked biomass resource? *Biomass Bioenergy* **2000**, *19*, 229–244. [CrossRef]
6. Li, L. Phenology Examination, Classification and Aboveground Biomass Estimation of Moso Bamboo Forests Using Time Series Remote Sensing Data. Ph.D. Thesis, Zhejiang A&F University, Hangzhou, China, 2020.
7. Zhou, G.; Meng, C.; Jiang, P.; Xu, Q. Review of Carbon Fixation in Bamboo Forests in China. *Bot. Rev.* **2011**, *77*, 262–270. [CrossRef]

8. Shendryk, I.; Broich, M.; Tulbure, M.G.; Alexandrov, S.V. Bottom-up delineation of individual trees from full-waveform airborne laser scans in a structurally complex eucalypt forest. *Remote Sens. Environ.* **2016**, *173*, 69–83. [CrossRef]
9. Li, L.; Li, N.; Lu, D.; Chen, Y. Mapping Moso bamboo forest and its on-year and off-year distribution in a subtropical region using time-series Sentinel-2 and Landsat 8 data. *Remote Sens. Environ.* **2019**, *231*, 111265. [CrossRef]
10. Proudman, A.; Ramezani, M.; Digumarti, S.T.; Chebrolu, N.; Fallon, M. Towards real-time forest inventory using handheld LiDAR. *Robot. Auton. Syst.* **2022**, *157*, 104240. [CrossRef]
11. Zhao, L.; Zhang, X.; Sun, H. Application of LiDAR data to forest parameters estimation. *World For. Res.* **2010**, *23*, 61–64.
12. Pham, D.D.; Suh, Y.S. Remote length measurement system using a single point laser distance sensor and an inertial measurement unit. *Comput. Stand. Interfaces* **2017**, *50*, 153–159. [CrossRef]
13. Polewski, P.; Yao, W.; Cao, L.; Gao, S. Marker-free coregistration of UAV and backpack LiDAR point clouds in forested areas. *ISPRS J. Photogramm. Remote Sens.* **2019**, *147*, 307–318. [CrossRef]
14. Cai, S.; Xin, Y.; Duan-Mu, J. Extraction of DBH from filtering out low intensity point cloud by backpack laser scanning. *For. Eng.* **2021**, *37*, 12–19.
15. Brolly, G.; Kiraly, G.; Lehtomaki, M.; Liang, X. Voxel-based automatic tree detection and parameter retrieval from terrestrial laser scans for plot-wise forest inventory. *Remote Sens.* **2021**, *13*, 542. [CrossRef]
16. Cai, Y.; Xu, W.; Liang, D.; Deng, S.; Li, C. Distinguishing *Phyllostachys edulis* age based on laser scanning intensity. *Chin. J. Lasers* **2018**, *45*, 272–280.
17. Li, L.; Sun, Y.; Wan, F.; Xu, X.; Huang, X. Research on canopy height inversion of Moso bamboo forests on Ta-pieh Mountains based on LiDAR. *J. Bamboo Res.* **2022**, *41*, 10–16.
18. Huang, L.; Guan, F. Retrieving parameters of individual Moso bamboo using terrestrial laser scanning data. *J. Northeast For. Univ.* **2021**, *49*, 67–70, 114.
19. Jiang, R.; Lin, J.; Li, T. Refined Aboveground Biomass Estimation of Moso Bamboo Forest Using Culm Lengths Extracted from TLS Point Cloud. *Remote Sens.* **2022**, *14*, 5537. [CrossRef]
20. Yang, B.; Liang, F.; Huang, R. Progress, challenges and perspectives of 3D LiDAR point cloud processing. *Acta Geod. Cartogr. Sin.* **2017**, *46*, 1509–1516.
21. Lu, D.; Zou, G. Comparative research on denoising algorithms of 3D laser point cloud. *Bull. Surv. Mapp.* **2019**, *S2*, 102–105.
22. Zhao, X.; Guo, Q.; Su, Y.; Xue, B. Improved progressive TIN densification filtering algorithm for airborne LiDAR data in forested areas. *ISPRS J. Photogramm. Remote Sens.* **2016**, *117*, 79–91. [CrossRef]
23. Lee, H.; Slatton, K.C.; Roth, B.E.; Cropper, W.P. Adaptive clustering of airborne LiDAR data to segment individual tree crowns in managed pine forests. *Int. J. Remote Sens.* **2010**, *31*, 117–139. [CrossRef]
24. Li, W.; Guo, Q.; Jakubowski, M.K.; Kelly, M. A New Method for Segmenting Individual Trees from the Lidar Point Cloud. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 75–84. [CrossRef]
25. Liu, C.; Xing, Y.; Duanmu, J.; Tian, X. Evaluating different methods for estimating diameter at breast height from terrestrial laser scanning. *Remote Sens.* **2018**, *10*, 513. [CrossRef]
26. Haris, B.; Jasmin, V.; Walter, B.; De Cubber, S.; Bruno, S. Fast Statistical Outlier Removal Based Method for Large 3D Point Clouds of Outdoor Environments. *IFAC-PapersOnLine* **2018**, *51*, 348–353.


**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.

forests

MDPI

*Article*

# Evaluation of Accuracy in Estimating Diameter at Breast Height Based on the Scanning Conditions of Terrestrial Laser Scanning and Circular Fitting Algorithm

**Yongkyu Lee and Jungsoo Lee ***

Department of Forest Management, Kangwon National University, Chuncheon 24341, Republic of Korea; kong445566@kangwon.ac.kr

* Correspondence: jslee72@kangwon.ac.kr

**Abstract:** A growing societal interest exists in the application of lidar technology to monitor forest resource information and forestry management activities. This study examined the possibility of estimating the diameter at breast height (DBH) of two tree species, *Pinus koraiensis* (*PK*) and *Larix kaempferi* (*LK*), by varying the number of terrestrial laser scanning (TLS) scans (1, 3, 5, 7, and 9) and DBH estimation methods (circle fitting [CF], ellipse fitting [EF], circle fitting with RANSAC [RCF], and ellipse fitting with RANSAC [REF]). This study evaluates the combination that yields the highest estimation accuracy. The results showed that for *PK*, the lowest RMSE of 0.97 was achieved when REF was applied to the data from nine scans after noise removal. For *LK*, the lowest RMSE of 1.03 was observed when applying CF to the data from seven scans after noise removal. Furthermore, ANOVA revealed no significant difference in the estimated DBH from nine scans when more than three scans were used for CF and RCF and more than five for EF and REF. These results are expected to be useful in establishing efficient and accurate DBH estimation plans using TLS for forest resource monitoring.

**Keywords:** LiDAR; TLS; DBH; RANSAC; circular fitting; circle fitting; ellipse fitting

## 1. Introduction

Traditional monitoring of forest resources for the construction of forest structural parameters typically involves in situ measurements using diameter tapes, calipers, and hypsometers [1]. In particular, diameter at breast height (DBH) is a fundamental parameter for calculating forest biomass and is used in various forest resource-related models, including stem taper, growth prediction, and height estimation [2–4].

Research related to DBH measurement has been conducted on comparisons of measurement results based on measurement height and tools, such as calipers and diameter tapes, as well as analyses of differences due to field crews [5–7]. Since the 2000s, with the advancement of remote sensing techniques, data collection using various platforms has become feasible, leading to diverse studies on DBH estimation modeling using satellite imagery, drone imagery, and 'light detection and ranging (LiDAR)' [8,9]. In particular, LiDAR technology, which is capable of acquiring high-resolution three-dimensional data, has been extensively researched for extracting information on forest structural parameters [10].

Research on DBH estimation using LiDAR utilizes data collected from various platforms and sensors including terrestrial laser scanning (TLS), mobile laser scanning (MLS), and airborne laser scanning (ALS). ALS has been commonly used for decades, but recently, there has been concurrent research on the utilization of TLS and MLS for ground-based data collection. Because it involves data acquisition from aircraft, ALS allows for efficient data collection; however, its low point density limits its analysis of forest structures. However, TLS and MLS collect data from the lower parts of a forest, making them more suitable

for analyzing forest structures. MLS, utilizing localization algorithms (e.g., simultaneous localization and mapping), is more effective in covering large areas and addressing occlusions owing to its ability to move while scanning. However, its accuracy is lower than that of TLS owing to cumulative errors in scanner positioning and movement. TLS, being stationary, can gather high-density and accurate point data on clouds compared to other platforms. However, occlusions can occur owing to terrain and other obstructions, necessitating adjustments in the number and positions of scans to effectively collect data on point clouds [11,12].

In particular, TLS enables high-accuracy DBH estimation based on dense and precise point cloud data, compared to LiDAR systems on other platforms [13]. Consequently, numerous studies have applied various algorithms using TLS to estimate DBH, such as the Hough transform, circle fitting algorithm, ellipse fitting algorithm, linear least square circle fitting, nonlinear least square circle fitting, minimum bounding box method, centroid method, and maximum distance method. These studies have also proposed new algorithms in this field [14–19]. Additionally, the random sample consensus (RANSAC) algorithm was presented as an effective method for eliminating the influence of noise [20]. The accuracy of DBH estimation using TLS can vary under different conditions. According to previous research, higher slopes may result in lower accuracy of DBH estimation [21]. Moreover, because TLS operates from a fixed point and rotates to acquire data, high stand density and uneven vertical distribution can lead to significant occlusions [22,23]. Consequently, the number of TLS scans significantly impacts data quality. Scanning methods are primarily categorized into single-scanning and multi-scanning, depending on the number of scans. Single-scanning was typically conducted at the center of the plot. Consequently, although data acquisition and preprocessing are rapid, this method is vulnerable to occlusions caused by the terrain and objects [24]. Multi-scanning involves scanning more than twice and aligning the collected data for acquisition. In multi-scanning, there is a trade-off relationship in which an increased scanning frequency leads to higher costs but also improves accuracy [21].

Previous studies have primarily focused on applying various algorithms to single-point clouds acquired in forest environments. However, these approaches often fail to thoroughly examine the relationship between the number of TLS scans, point cloud quality, and DBH estimation accuracy. Therefore, this study aimed to analyze the patterns of accuracy variation relative to the number of TLS scans in the same study area. Additionally, the impact of the presence and treatment of noise and various DBH estimation methods on accuracy was investigated, intending to contribute to the development of forest point cloud strategies for DBH estimation.

## 2. Materials and Methods

In this study, point clouds were collected using the TLS. The datasets were created from data scanned one, three, five, seven, and nine times following the alignment and ground control points (GCP) correction processes. The point cloud constructed for each scanning frequency was processed using a Cloth Simulation Filter (CSF) algorithm to extract ground points and construct the mesh data of the terrain. Subsequently, the distance between the mesh and points was calculated to extract the DBH point cloud at breast height (ranging from 1.15 to 1.25 m). The location of each tree was determined using the position coordinates collected using the Real-Time Kinematic (RTK) system during the field surveys. For each tree, DBH was estimated and evaluated using circle fitting (CF), ellipse fitting (EF), RANSAC circle fitting (RCF), and RANSAC ellipse fitting (REF) on the respective DBH point cloud. Finally, for each tree, the Completeness of the Stem Point Cloud (CPC), Field DBH, Point Density, and Slope were calculated, and their relationship with bias was analyzed (Figure 1).

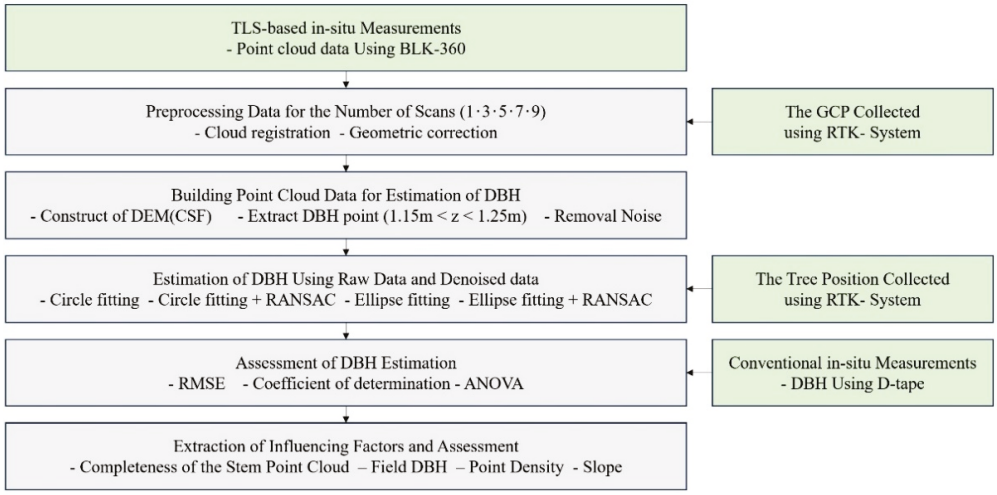



**Figure 1.** Workflow Diagram of the Research Methodology for DBH Estimation Using Terrestrial Laser Scanning (TLS).

### *2.1. Study Area*

This research was conducted in Kangwon-do within the academic forest of Kangwon National University. Five plots each for *Pinus koraiensis* (*PK*) and five for *Larix kaempferi* (*LK*) were installed. The plots were configured as circles with a radius of 11.3 m. The *PK* and *LK* plots were artificial forests of the fifth age class, with 185 *PK* trees and 78 *LK* trees distributed across the five plots. The average slope of the five plots was approximately 21° for *PK* and 40° for *LK*, indicating variation in stand density and slope between species. Additionally, *LK* had a lower stand density than *PK*, and owing to the characteristics of the species, a higher amount of light was passing through the canopy, leading to well-developed understory vegetation [25] (Figure 2).

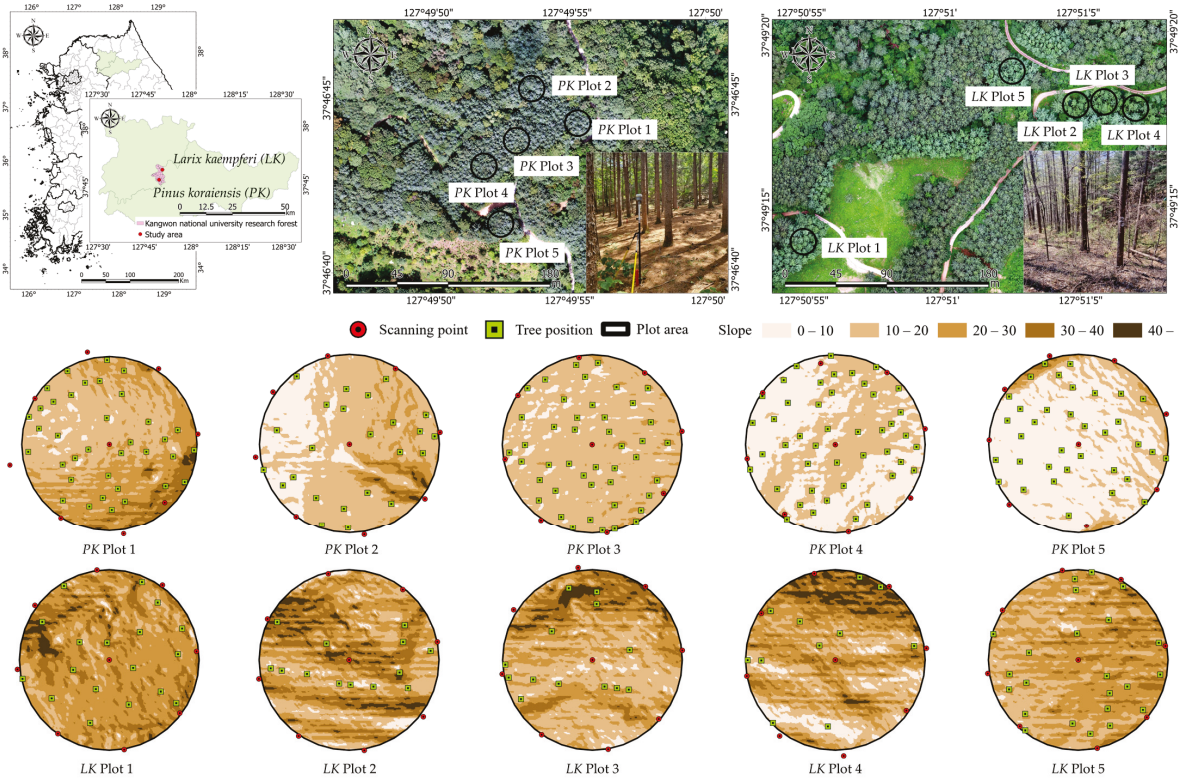



**Figure 2.** Location map and drone orthophoto of the study area, along with schematics showing slope, individual tree positions, and scanning locations for each research plot.

### 2.2. Data Acquisition and Preprocessing

#### 2.2.1. Field Survey for DBH and Individual Tree Locations

During the field survey, data on the locations and DBH of individual trees were collected. The DBH was measured using a diameter tape at a height of 1.2 m from the ground in accordance with the standards of the National Forest Resources Survey of South Korea [26]. The locations of individual trees were determined using an RTK system composed of two R12i units (Trimble Inc., Westminster, CO, USA). One of the two units was installed in a cleared area on a forested road and served as the base station. The other unit, which was used as a rover, was employed within the forest to acquire points in close proximity to each tree. The location of individual trees acquired using the RTK system was used to match the estimated Diameter at Breast Height (DBH) from the LiDAR data with the Field DBH measured using a diameter tape. The average DBH was 27.4 cm at *PK* and 32.3 cm at *LK* (Table 1).

**Table 1.** The statistics of each plot using conventional in situ measurements.

| Species | Plot | Number of Trees (*n*) | Stem Density (Stems/ha) | DBH (cm) | | | | Basal Area ($m^2$/ha) |
|---|---|---|---|---|---|---|---|---|
| | | | | Min | Max | Mean | Std | |
| *PK* | 1 | 37 | 925 | 12.9 | 43.0 | 27.6 | 8.2 | 1258 |
| | 2 | 25 | 625 | 12.5 | 50.3 | 30.0 | 8.3 | 925 |
| | 3 | 44 | 1100 | 15.0 | 47.0 | 26.5 | 7.0 | 1438 |
| | 4 | 44 | 1100 | 12.0 | 42.0 | 25.9 | 6.0 | 1408 |
| | 5 | 35 | 875 | 16.0 | 42.8 | 28.5 | 6.0 | 1230 |
| | Total | 185 | 925 | 12.0 | 50.3 | 27.4 | 7.1 | 1252 |
| *LK* | 1 | 18 | 450 | 22.3 | 38.2 | 31.2 | 4.8 | 692 |
| | 2 | 13 | 325 | 20.1 | 40.5 | 31.3 | 6.5 | 501 |
| | 3 | 13 | 325 | 16.9 | 46.5 | 33.2 | 8.6 | 532 |
| | 4 | 11 | 275 | 23.9 | 47.5 | 37.3 | 8.7 | 506 |
| | 5 | 23 | 575 | 16.2 | 42.8 | 31.0 | 7.5 | 879 |
| | Total | 78 | 390 | 16.2 | 47.5 | 32.3 | 7.3 | 622 |

#### 2.2.2. Data Collection and Preprocessing

Data collection using TLS includes scanning for point cloud acquisition and establishing GCPs for georeferencing. The collection of point cloud data was conducted using the BLK 360 (Leica Geosystems AG., Heerbrugg, Switzerland). The BLK 360 is mounted on a tripod for use. It can complete scanning within 5 min from a single position and accurately collect data within a radius of 10–15 m. The specifications of the BLK 360 include a shooting range of 0.6–60 m, a point measurement rate of over 360,000 points per second, and a ranging accuracy of 4 mm at 10 m [27,28]. The scanning positions were set up once at the center of the standard plot and at eight directional points located 11.3 m from the center (Figure 3). Each scanning location was determined using a tape measure and a compass from the center of the plot, and stakes were installed for easy identification at each location. In addition, the positions of the installed stakes were recorded using an RTK system to acquire the GCPs.

Preprocessing of the TLS data was performed in the order of alignment, followed by GCP correction. Registration was performed on data captured 1, 3, 5, 7, and 9 times to compare the results of DBH estimation according to the number of scans. Data from 2, 4, 6, and 8 scans were excluded from the analysis owing to the asymmetric positioning of the scanning locations when including the central position. The root mean square error (RMSE) for all registration processes was maintained below 5 mm. The scanning positions used for each number of scans correspond to those shown in Figure 3. The GCP correction was performed individually for the point cloud constructed for each number of scans, matching the stakes identifiable in each point cloud to the GCPs. All the GCP correction processes were completed with an RMSE of less than 5 cm. Finally, the noise points located below the ground and points situated beyond 13 m from the central coordinates (x and y) and outside

the plot were removed. Cyclone (2023.0.2) software was used for alignment, whereas cloud computing was employed for geometric correction and point removal.

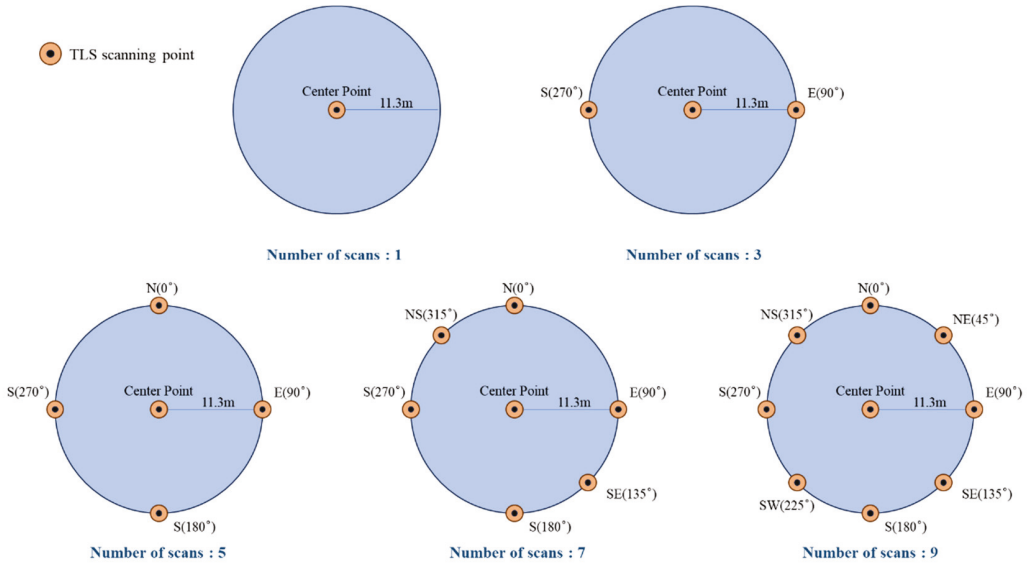



**Figure 3.** Scanning positions of the TLS according to the number of scans.

#### 2.2.3. Noise Point Removal for DBH Estimation

The constructed point clouds contain points distributed across the understory vegetation and canopy, which are unnecessary for DBH estimation and necessitate the preprocessing of these points. Firstly, the Cloth Simulation Filter (CSF) algorithm was applied to extract ground points, and the extracted ground points were used to construct terrain mesh information [29,30]. The parameters for the CSF filter were set as follows: cloth resolution at 0.1, max iterations at 900, and classification threshold at 0.1. The ground mesh constructed using CSF was set above the actual ground at locations where individual trees existed, but ground points did not. To address this, the smooth mesh tool was utilized to create a uniform ground mesh. The parameters for the smooth mesh were set at 20 iterations and a smoothing factor of 0.2. Furthermore, the distance of each point from the ground was calculated, and considering the measurement height for DBH, points located at a distance of 1.15 m–1.25 m from the ground were extracted. The extracted points were defined as 'raw data'. 'Raw data' includes branches and understory vegetation that could interfere with the estimation of DBH [31]. Subsequently, noise points that were unnecessary for the DBH estimation were removed through visual inspection. The data from which these points were removed was defined as 'denoised data' and used to construct the DBH point cloud (Figure 4). This process was performed using cloud compare (v2.13 beta) software.

### *2.3. Estimation and Evaluation of DBH Using Circular Fitting*

Algorithm-based circular fitting, as opposed to manual measurement of DBH, eliminates subjectivity and offers high reproducibility when the same data are used [32]. Therefore, the least squares method for CF and EF algorithms for DBH estimation was applied in this study [33]. The least squares method seeks the optimal solution where the sum of the squared distances between the given points and the drawn circle is minimized. Additionally, the RANSAC algorithm, known for its high performance in noisy data environments, was applied to each fitting algorithm, resulting in four DBH estimation outcomes, including RCF and REF. The optimal solutions for CF and EF were estimated using the x, y coordinates of points within 30 cm from the center of individual trees and the normal equation.

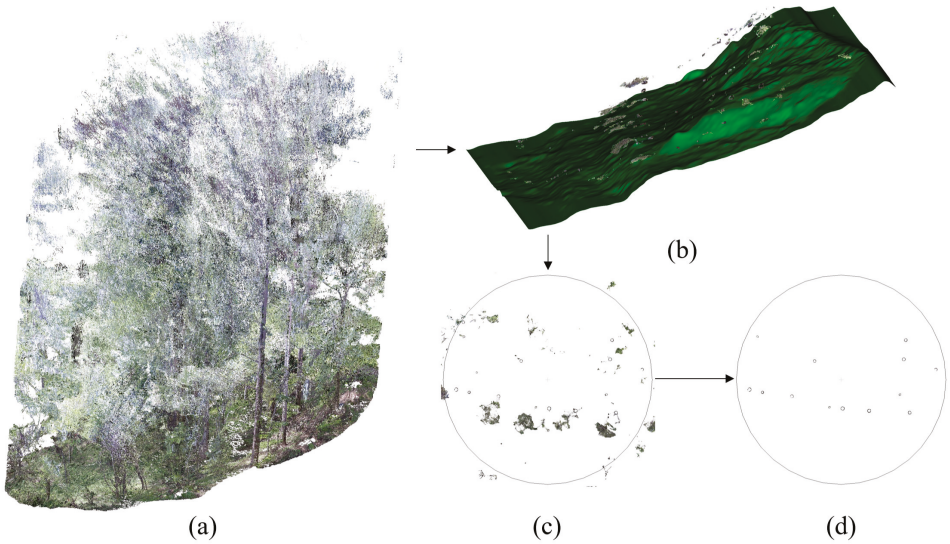



**Figure 4.** Constructing point clouds for DBH: (**a**) a scanned plot using TLS; (**b**) construction of the ground mesh using CSF and extracted points at 1.15–1.25 m height range from the ground; (**c**) top view of constructed DBH point cloud (raw data); (**d**) top view of denoised DBH point cloud (denoised data).

The RANSAC algorithm is a method for predicting noise removal from the given data [34]. The RANSAC algorithm procedure is primarily divided into the hypothesis phase and the verification phase, and it operates by iterating through these two stages. The hypothesis phase involves extracting n data points from the entire dataset and applying a prediction model, whereas the verification phase evaluates the applied results and stores the model with the highest performance. In this study, the hypothesis phase involved randomly selecting 10 points to apply the CF and EF. The verification phase was evaluated based on the number of points in the entire dataset whose distances from the estimated circle were below the threshold (0.6 cm). These two phases were repeated 100 times, and the circle with the highest number of points within the threshold was selected [20]. The threshold was set based on the error range of the BLK360 when used at a shooting distance of 10 m.

The evaluation of the DBH estimation results involved checking the possibility of estimating DBH according to the circular fitting method and the number of scanning sessions. Cases were defined as unfeasible for estimation when the circular fitting method yielded negative radius values or when the DBH was estimated to be outside the 6–60 cm range. Accuracy was evaluated by comparing the DBH estimated using the circular fitting method with the Field DBH. This comparison considered factors such as noise removal, tree species, and number of scans. Subsequently, the RMSE and R^2 were calculated based on these comparisons. Additionally, to determine the optimal number of scans for each circular fitting method, analysis of variance (ANOVA) was conducted. This analysis tested for significant differences at the 0.05 level in the mean absolute values of the bias between the number of scans for each circular fitting method. Furthermore, post hoc tests were performed to examine whether there were significant differences in the mean absolute values of bias when comparing one, three, five, and seven scans against the highest number of scans, which was nine.

### *2.4. Construction and Evaluation of Influencing Factors on DBH Estimation*

The influencing factors were CPC, Point Density, Field DBH, and Slope. Because the TLS is scanned from a single location, it is difficult to collect complete data for a plot owing to occlusion by terrain and objects. Therefore, in the case of a single scan or an insufficient number of scans, the point cloud for DBH is formed as an incomplete circle. The Circular fitting method can estimate DBH with an incomplete circle. However, the

incompleteness of a circle can affect DBH estimation [35]. Consequently, the CPC was selected as the influencing factor. The CPC involves constructing the central point of each tree and calculating the angle with all points of the DBH point cloud, rounding off to the nearest degree. Subsequently, the CPC was calculated in percentage terms by removing duplicate values of the calculated angles and dividing the count of unique values by 360.

The point density increased as the distance between the LiDAR and subject decreased, and additional scanning was conducted at various locations. Point density was added as an influencing factor, considering the relationship between the scanning locations and positions of individual trees [21]. Point density was calculated by dividing the number of points in the DBH point cloud by the length of the circular arc, resulting in the number of points per centimeter. The length of the circular arc was calculated using the computed CPC.

Field DBH and Slope were selected as influencing factors based on prior research that indicated that DBH and slope affect estimation accuracy [35,36]. The slope was calculated by constructing a 1-m buffer zone around the central point of the DBH point cloud and computing the average of the pixels located within the buffer zone. The slope for each plot was determined using the digital elevation model (DEM) constructed for that plot. The DEM was based on point cloud data collected from nine scans. Utilizing the constructed ground mesh, a DEM with a resolution of 10 cm was created (Figure 5). The CPC and Point Density were calculated using Python (3.7.2), while Slope was constructed using the Cloud Compare and ArcGIS Pro (3.0.2) programs.

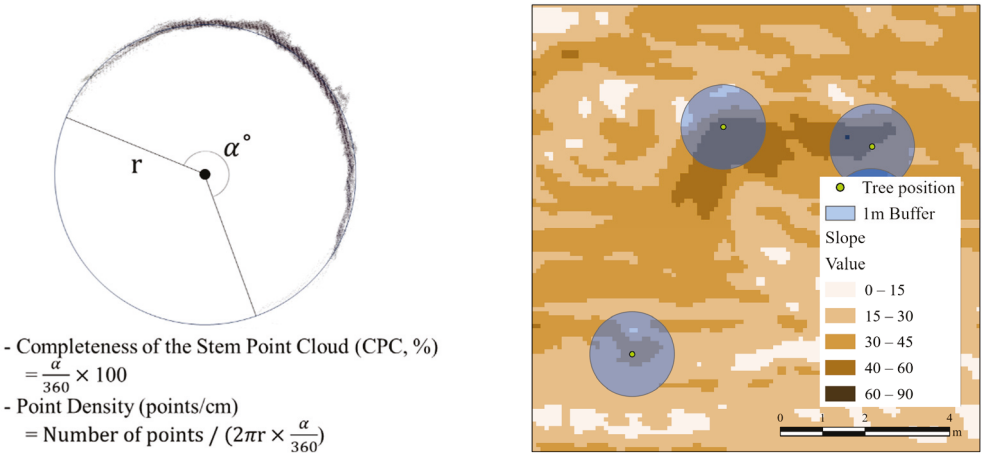



**Figure 5.** Extraction of influencing factors: (**a**) examples of calculating the 'completeness of the stem point cloud (CPC, %)' and 'point density'; (**b**) calculating average slope values within a 1-m radius around individual tree.

## 3. Results and Discussion

### *3.1. Evaluation of the Possibility of DBH Estimation Based on the Circular Fitting Method and Number of Scans*

In *PK*, DBH estimation was possible for all individual trees using five scans. The number of trees that could be estimated was the same for both denoised and raw data. In contrast, for *LK*, DBH estimation was possible for all individual trees using the three scans, except when EF was applied to the raw data (Table 2). When a single scan was conducted, the number of trees for which DBH estimation was possible was the same for both CF and RCF. In addition, the distribution was higher for CF and RCF than for REF and EF. DBH estimation using CF and RCF was possible even when a small number of points were collected. However, with EF and REF, when a few points were collected, the DBH was estimated to be less than 6 cm, rendering the estimation infeasible (Figure 6a). In contrast, in *LK*, with abundant understory vegetation, using EF on raw data from nine scans led to an overestimation in some cases. Specifically, two individual trees were estimated to have DBHs exceeding 60 cm owing to noise points, rendering the DBH estimation unfeasible

(Figure 6b). Consequently, to estimate the DBH for all individual trees, it is necessary to consider the environmental conditions of the study area and the circular fitting algorithm. To estimate the DBH of all individual trees in each plot, noise points were removed and at least five scanning sessions were performed. The CF was more suitable than the EF for data from fewer than five scans. Additionally, because EF is sensitive to noise points, it is essential to include a process for removing these points or to consider the application of the RANSAC algorithm.

**Table 2.** Number of trees for which DBH estimation was possible based on noise removal status, number of scans, and circular fitting algorithm.

| Species | | *PK* (*n* = 185) | | | | | *LK* (*n* = 78) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Scans | | 1 | 3 | 5 | 7 | 9 | 1 | 3 | 5 | 7 | 9 |
| Raw data | CF | 174 | 184 | 185 | 185 | 185 | 74 | 78 | 78 | 78 | 78 |
| | RCF | 174 | 184 | 185 | 185 | 185 | 74 | 78 | 78 | 78 | 78 |
| | EF | 164 | 184 | 185 | 185 | 185 | 65 | 74 | 76 | 76 | 76 |
| | REF | 168 | 184 | 185 | 185 | 185 | 72 | 78 | 77 | 76 | 78 |
| Denoised data | CF | 174 | 184 | 185 | 185 | 185 | 74 | 78 | 78 | 78 | 78 |
| | RCF | 174 | 184 | 185 | 185 | 185 | 74 | 78 | 78 | 78 | 78 |
| | EF | 164 | 184 | 185 | 185 | 185 | 71 | 78 | 78 | 78 | 78 |
| | REF | 168 | 184 | 185 | 185 | 185 | 72 | 78 | 78 | 78 | 78 |

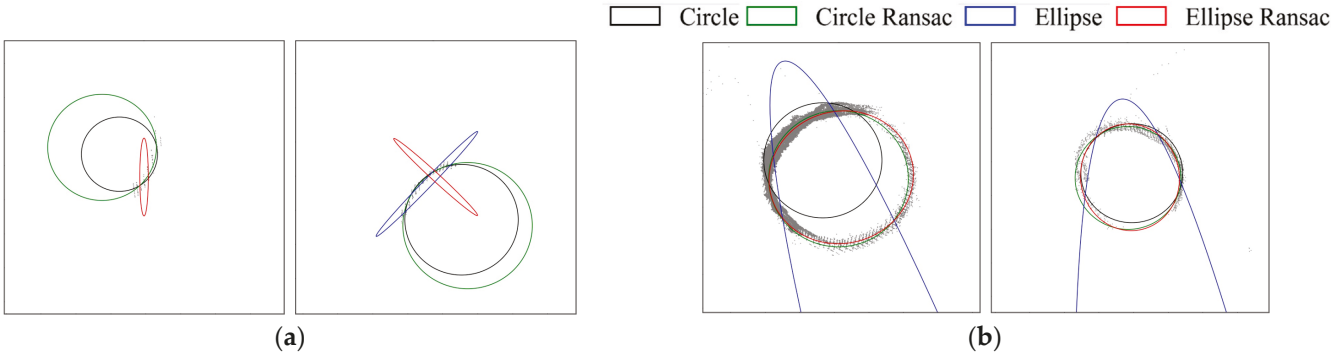



**Figure 6.** Cases of DBH estimation using the circular fitting method being unfeasible: (**a**) a case of underestimation below 6 cm using EF and REF in a single scan data; (**b**) a case of overestimation above 60 cm using EF in the data of nine scans.

### *3.2. Accuracy Evaluation of DBH Estimation*

For both tree species, the highest accuracy was achieved when the REF was applied to the data scanned seven times, resulting in an RMSE of 1.01 *PK* and 1.12 *LK*. In the denoised data, the highest accuracy for *PK* was observed when REF was applied to data scanned nine times, with an RMSE of 0.97. For *LK*, the highest accuracy was achieved using CF on data scanned seven times, resulting in an RMSE of 1.03. The aforementioned RMSE values were distributed similarly to the range of 0.5 to 1.6 cm for RMSE obtained in previous studies on DBH estimation [13,37–39] (Table 3).

As a result of estimating DBH using a single scan, a difference in accuracy was observed between CF and EF. In single-scanning data, the RMSE of EF was approximately 1.3 to 3.7 times higher than that of CF. Therefore, CF is deemed more suitable for single-scanning data. In contrast, in the multi-scanning data, the RMSE decreased for both the raw and denoised data as the number of scans increased. In particular, the accuracies of EF and REF improved significantly as the number of scans increased. In four cases based on noise removal status and tree species, REF achieved the lowest RMSE in three of the cases. Accordingly, EF is more significantly influenced by the number of scans than CF. EF

requires optimization of more variables, such as the center point, major axis, minor axis, and rotation angle, compared with CF during the fitting process. Therefore, sufficient data acquisition is necessary for accurate DBH estimation using EF.

**Table 3.** Evaluation results of accuracy by tree species and number of scans according to noise removal status and circular fitting method.

| Species | | | *PK* | | | | | *LK* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of Scans | | | 1 | 3 | 5 | 7 | 9 | 1 | 3 | 5 | 7 | 9 |
| Raw data | CF | RMSE | 3.27 | 1.53 | 1.55 | 1.55 | 1.46 | 3.21 | 2.92 | 2.23 | 2.59 | 1.64 |
| | | $r^2$ | 0.81 | 0.96 | 0.96 | 0.96 | 0.96 | 0.83 | 0.86 | 0.91 | 0.89 | 0.96 |
| | RCF | RMSE | 2.95 | 1.59 | 1.85 | 1.80 | 1.79 | 2.54 | 2.02 | 1.57 | 1.33 | 1.28 |
| | | $r^2$ | 0.84 | 0.95 | 0.94 | 0.94 | 0.94 | 0.89 | 0.94 | 0.96 | 0.97 | 0.98 |
| | EF | RMSE | 4.71 | 2.24 | 1.54 | 1.45 | 1.32 | 7.12 | 3.53 | 1.92 | 1.90 | 1.72 |
| | | $r^2$ | 0.70 | 0.91 | 0.96 | 0.96 | 0.97 | 0.53 | 0.78 | 0.93 | 0.93 | 0.95 |
| | REF | RMSE | 4.58 | 1.75 | 1.17 | 1.01 | 1.03 | 6.06 | 2.72 | 1.38 | 1.12 | 1.33 |
| | | $r^2$ | 0.67 | 0.94 | 0.98 | 0.99 | 0.99 | 0.59 | 0.89 | 0.98 | 0.99 | 0.98 |
| Denoised data | CF | RMSE | 3.08 | 1.12 | 1.10 | 1.06 | 1.03 | 1.68 | 1.25 | 1.14 | 1.03 | 1.09 |
| | | $r^2$ | 0.84 | 0.98 | 0.98 | 0.99 | 0.99 | 0.95 | 0.99 | 0.99 | 0.99 | 0.99 |
| | RCF | RMSE | 2.69 | 1.29 | 1.31 | 1.25 | 1.18 | 2.00 | 1.57 | 1.35 | 1.30 | 1.13 |
| | | $r^2$ | 0.87 | 0.97 | 0.98 | 0.98 | 0.98 | 0.93 | 0.97 | 0.98 | 0.98 | 0.99 |
| | EF | RMSE | 4.13 | 1.56 | 1.27 | 1.11 | 1.02 | 6.19 | 2.65 | 1.11 | 1.09 | 1.10 |
| | | $r^2$ | 0.78 | 0.96 | 0.98 | 0.99 | 0.99 | 0.63 | 0.90 | 0.99 | 0.99 | 0.99 |
| | REF | RMSE | 4.05 | 1.73 | 1.16 | 1.20 | 0.97 | 6.52 | 2.46 | 1.22 | 1.84 | 1.26 |
| | | $r^2$ | 0.74 | 0.94 | 0.98 | 0.98 | 0.99 | 0.53 | 0.91 | 0.98 | 0.95 | 0.98 |

However, in some cases, the RMSE increases with the number of scans. Such cases were predominantly observed when the raw data were used. For example, when using raw data from single-scanning data, there were no noise points, allowing all four circular fitting methods to correctly estimate the DBH. However, as the number of scans increased, noise points were progressively added, leading to gradually larger errors in DBH estimation by CF and EF. Notably, EF was more significantly affected by noise points than CF (Figure 7).

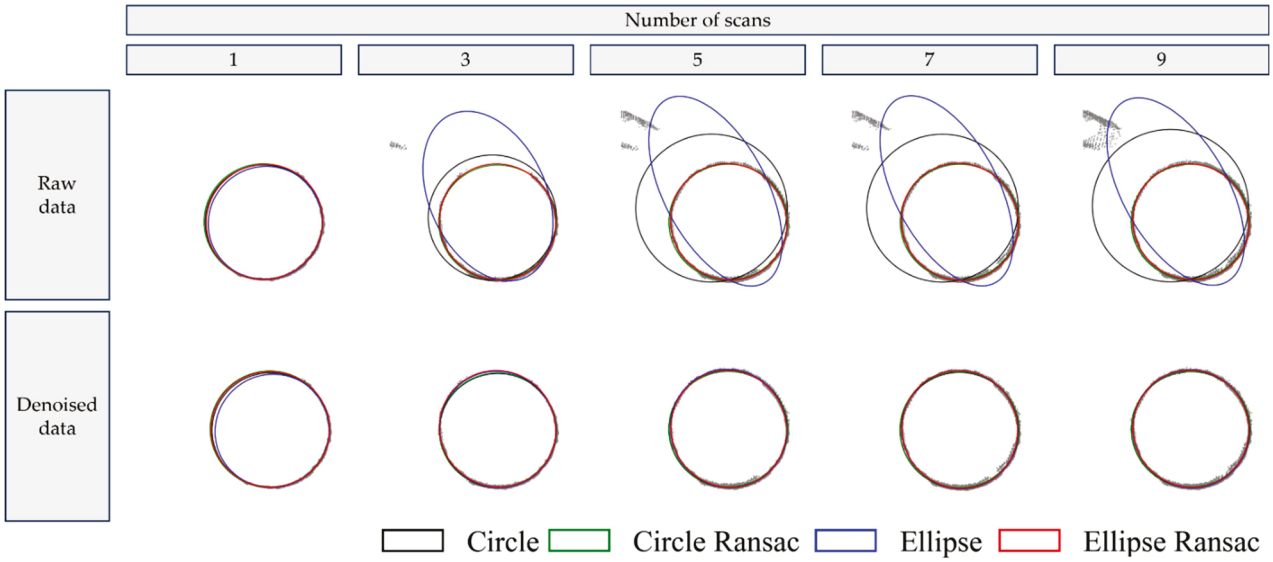



**Figure 7.** Results of applying circular fitting algorithms with increasing number of scans in raw and denoised data (*LK*, Plot 3, No. 13).

Using fitting methods where the RANSAC algorithm was applied, it was possible to accurately estimate the DBH, even in the presence of some noise points. Consequently, in the cases mentioned above, as the number of scans increased, the point cloud from understory vegetation and branches around the individual trees affected the accuracy of DBH estimation. It is considered appropriate to improve this by applying the RANSAC algorithm or using denoised data. However, in data without noise, the application of the RANSAC algorithm sometimes led to reduced accuracy. This is because the RANSAC algorithm randomly samples some data for fitting, which can result in a biased algorithm focused on a subset of data. The number of samples is a crucial parameter in the RANSAC algorithm; it can lead to bias towards a subset of the data if it is too small and an increased likelihood of including surrounding noise data if it is too high. Therefore, depending on the acquired data and fitting algorithm used, there is a need to evaluate whether to apply the RANSAC algorithm and review its parameters.

### *3.3. Optimal Number of Scans for Estimating DBH Using TLS*

ANOVA was conducted for each of the four circular fitting methods to evaluate the impact of the number of scans on accuracy. The analysis results showed that the absolute values of bias differed significantly according to the number of scans in all the Circular Fitting methods. The F-statistic for the CF method was approximately 12.9, whereas that for the EF method was much higher at 81.9. This indicates that the EF method is more sensitive than the CF method to changes in the number of scans with respect to the absolute value of the CF method. Additionally, when the RANSAC algorithm was applied to both the CF and EF methods, the F-statistic decreased, indicating that the variation in bias due to the number of scans was reduced by the RANSAC algorithm. These results suggest that the RANSAC algorithm can mitigate the impact of the number of scans on the bias (Table 4).

Tamhane's T2 post hoc test was used to verify the changes in the absolute values of bias according to the number of scans. The absolute values of the bias obtained from nine scans were compared with those obtained from one, three, five, and seven scans. Significant differences were observed between the absolute values of bias for one and nine scans in CF and RCF, and between one, three, and nine scans for EF and REF. Consequently, three scans for CF and RCF and five scans for EF and REF were sufficient to produce bias estimates that were as accurate as those from nine scans, without any significant difference. This implies that reliable DBH estimates can be obtained even with fewer scans. Therefore, from an efficiency standpoint, it is advisable to conduct three scans for CF and RCF, and five scans for EF and REF (Table 5).

**Table 4.** ANOVA analysis results of the number of scans for each circular fitting using denoised data.

| Dependent Variable | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| Absolute Deviation of CF | Between Groups | 30.43 | 4 | 7.61 | 12.94 | 0.00 |
| | Within Groups | 752.27 | 1280 | 0.59 | | |
| | Total | 782.70 | 1284 | | | |
| Absolute Deviation of RCF | Between Groups | 24.07 | 4 | 6.02 | 7.21 | 0.00 |
| | Within Groups | 1067.70 | 1280 | 0.83 | | |
| | Total | 1091.78 | 1284 | | | |
| Absolute Deviation of EF | Between Groups | 1021.90 | 4 | 255.48 | 81.95 | 0.00 |
| | Within Groups | 3990.38 | 1280 | 3.12 | | |
| | Total | 5012.28 | 1284 | | | |
| Absolute Deviation of REF | Between Groups | 604.61 | 4 | 151.15 | 53.44 | 0.00 |
| | Within Groups | 3620.47 | 1280 | 2.83 | | |
| | Total | 4225.08 | 1284 | | | |

**Table 5.** Determining the optimal number of scans using Tamhane's T2 post hoc test.

| Dependent Variable | Number of Scanning (I) | Number of Scanning (J) | MeanDifference (I–J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower Bound | Upper Bound |
| Absolute Deviation of CF | 9 | 1 | −0.42 * | 0.08 | 0.00 | −0.64 | −0.21 |
| | | 3 | −0.11 | 0.06 | 0.52 | −0.29 | 0.06 |
| | | 5 | −0.04 | 0.06 | 1.00 | −0.22 | 0.14 |
| | | 7 | −0.00 | 0.06 | 1.00 | −0.17 | 0.16 |
| Absolute Deviation of RCF | 9 | 1 | −0.39 * | 0.09 | 0.00 | −0.63 | −0.15 |
| | | 3 | −0.17 | 0.07 | 0.20 | −0.37 | 0.04 |
| | | 5 | −0.06 | 0.07 | 0.99 | −0.27 | 0.15 |
| | | 7 | −0.05 | 0.07 | 1.00 | −0.25 | 0.15 |
| Absolute Deviation of EF | 9 | 1 | −2.41 * | 0.23 | - | −3.07 | −1.74 |
| | | 3 | −0.36 * | 0.11 | 0.01 | −0.66 | −0.06 |
| | | 5 | −0.09 | 0.07 | 0.84 | −0.28 | 0.09 |
| | | 7 | −0.03 | 0.06 | 1.00 | −0.20 | 0.14 |
| Absolute Deviation of ECF | 9 | 1 | −1.86 * | 0.21 | 0.00 | −2.47 | −1.26 |
| | | 3 | −0.42 * | 0.10 | 0.00 | −0.71 | −0.13 |
| | | 5 | −0.04 | 0.06 | 1.00 | −0.22 | 0.14 |
| | | 7 | −0.07 | 0.08 | 0.99 | −0.29 | 0.15 |

* The mean difference is significant at the 0.05 level.

### *3.4. Evaluation of Factors Affecting the Accuracy of DBH Estimation*

The average CPC by species showed an increasing pattern in both species as the number of scans increased, with *Pk* increasing to 45%, 71%, 88%, 94%, and 99%, and *LK* increasing to 46%, 75%, 91%, 95%, and 98%. The stand densities for *PK* and *LK* were 925 and 390 trees/ha, respectively, with *PK* having approximately 2.3 times more trees. However, the average CPC of the two species varied by less than 5% depending on the number of scans, indicating a minor difference. As the number of scans increased, the standard deviation of the CPC in *PK* showed a pattern of increase and then decrease, with values of 8.14, 15.31, 12.93, 8.29, and 3.91, indicating an increase from one to three scans, followed by a decrease. Similarly, *LK* exhibited the same pattern with values of 6.59, 13.27, 9.99, 7.75, 5.28. However, compared to *PK*, *LK* had a lower standard deviation in one scan but a higher standard deviation in nine scans. Consequently, it was inferred that in *LK*, it was more challenging to construct a complete DBH point cloud than in *PK*. In areas with high stand density but low understory vegetation, increasing the number of scans enabled the construction of most of the information about stands in the PCD. However, in *LK*, where the understory vegetation obscured visibility, there were limitations to constructing a complete DBH point cloud, even with an increased number of scans. In summary, for efficient scanning, it is appropriate to aim for approximately five scans, where the CPC reaches approximately 89%, beyond which the rate of increase decreases. Nine or more scans were necessary to obtain a complete point cloud (Figure 8).

Boxplots of the constructed factors and biases are shown in Figure 9. When the CPC was below 60%, the variance in the bias was significantly distributed, with ellipse fitting showing greater variance than circle fitting. When the CPC was above 60%, the bias between the different circular fitting algorithms was similarly distributed, and the bias decreased. Furthermore, as the CPC approached 100%, the bias and variance decreased, allowing for a more accurate estimation of diameter at breast height (DBH). Therefore, for consistent DBH estimation, the data collection must comprise at least 60% CPC, and it can be observed that the performance of DBH estimation is proportional to the CPC.

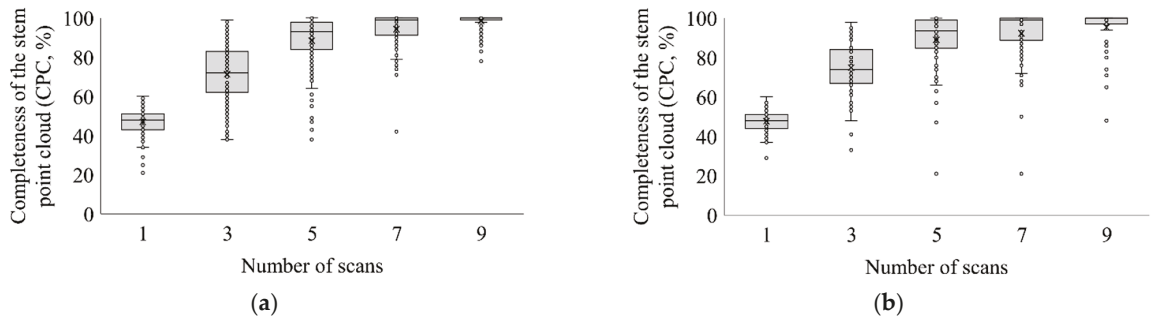



**Figure 8.** Result of calculated completeness of the stem point cloud depending on number of scans: (**a**) *PK*; (**b**) *LK*.

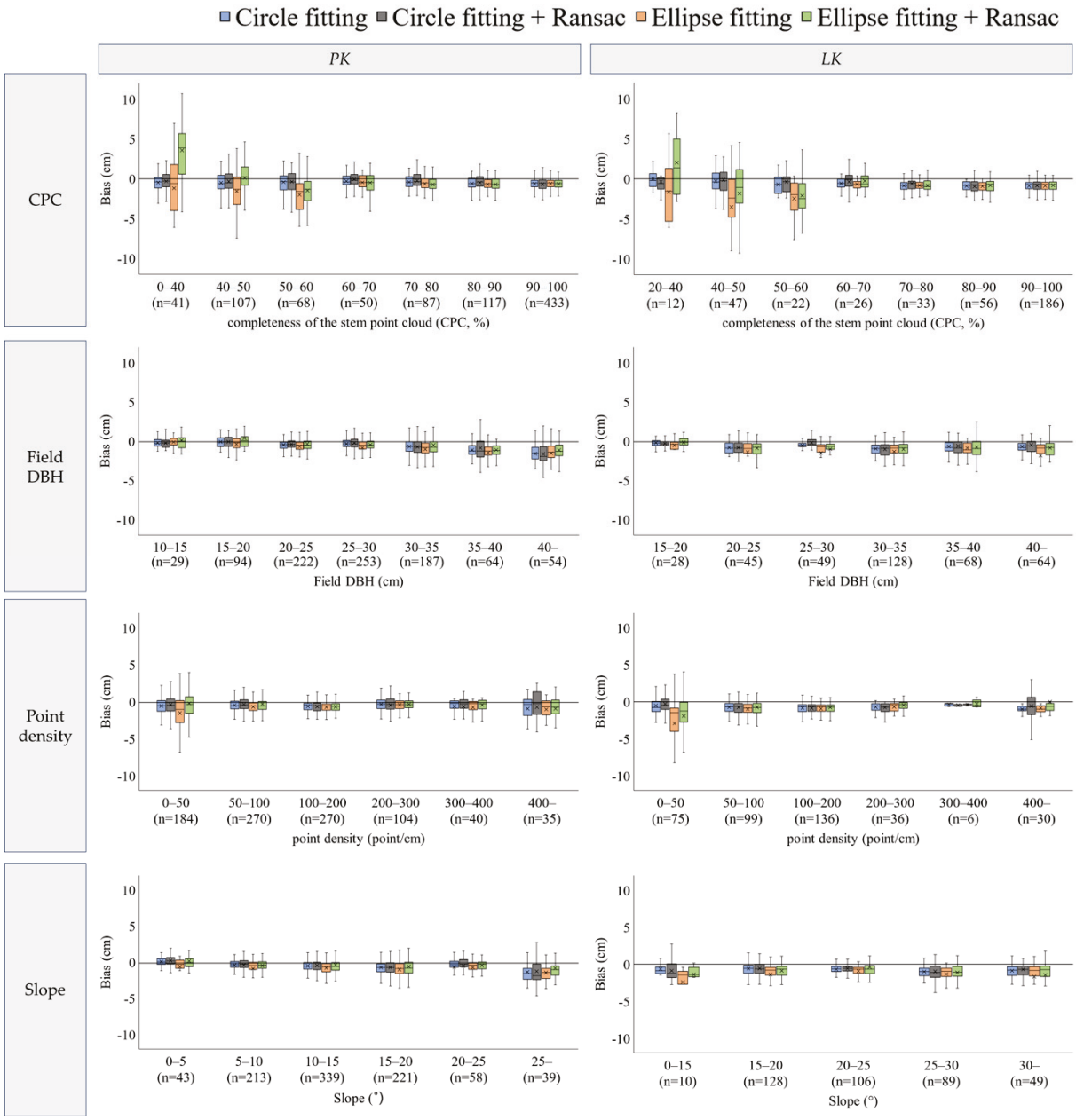



**Figure 9.** Boxplots of bias (Estimation DBH—Field DBH) depending on influencing factors, species, and fitting method using denoised data.

The relationship between Field DBH and bias showed that as Field DBH increased, the mean value of the bias gradually decreased, whereas the variance did not change significantly. In this study, Field DBH was measured using a diameter tape, and the results are consistent with those of previous studies that estimated and evaluated DBH using LiDAR systems and a diameter tape [35]. Diameter tape, a measuring tool designed based on the assumption of a circular shape, can result in overestimation when the cross-section of breast height is in the form of an ellipse or an irregular shape. Moreover, this error increases with DBH, even in ellipses with the same eccentricity [6]. Therefore, the underestimation of DBH using point cloud data in cases of larger-field DBH is considered to be a limitation inherent in the characteristics of the diameter tape, making it challenging to attribute this to errors in the LiDAR system.

The accuracy of estimating the DBH based on point density improved as the point density increased. However, at the highest point density levels exceeding 400 points, the accuracy decreased. Previous research related to point density and DBH estimation has indicated that lower point densities can lead to increased errors [40]. Additionally, when using MLS, it was observed that the error pattern initially decreased and then increased again with an increase in point density [35]. The point density varies owing to multiple factors, including the distance and angle from the scanner to the target, resolution of the equipment, slope, size of the target, and scanning settings. Additionally, in multi-scan scenarios, the point density can vary depending on the number and position of the scans [41]. Therefore, errors arising from the specifications of the equipment and alignment process, particularly in multi-scan scenarios with high point density, need to be considered.

The slope influenced the accuracy of the measurements, with the bias variance increasing and the accuracy decreasing as the slope exceeded 15°. The slope affects the extraction of ground points and the construction of the DEM. According to previous research, higher slopes are associated with significant errors in the DEM [42]. Ground points and DEM are crucial for estimating the forest resource parameters using LiDAR systems. In this study, ground points were used to extract points at the breast height locations (1.15–1.25 m). As a result, the higher error in DBH estimation in larch forests than in pine forests can be partly attributed to steeper slopes. In the future, for accurate forest resource estimation using LiDAR systems, it will be necessary to review methods for extracting accurate ground points under various environmental conditions.

## 4. Conclusions

This study aimed to compare the accuracy of DBH estimation using terrestrial laser scanning (TLS) based on the number of scans and circular fitting algorithms. The number of scans was set to 1, 3, 5, 7, and 9 to construct the point cloud at the plot level, and ground points were extracted to establish points at breast height. DBH was measured using CF, RCF, EF, and REF applied to the constructed breast height points, and the optimal number of scans was determined using ANOVA. Additionally, each tree's CPC, Field DBH, point density, and slope information were constructed as influencing factors for DBH estimation, and their relationships with bias were analyzed.

The number of scans in the TLS was proportional to the accuracy; however, the increase in accuracy gradually diminished. However, excessive scanning can lead to reduced accuracy due to scanner limitations, cumulative errors in the alignment process, and scanning noise from understory vegetation and branches. Particularly, in a single scan, more than 90% of individual trees could be estimated for DBH, and in five scans, the identification of all individual trees was possible. Statistical analysis showed no significant difference in bias between the data from the nine scans and three or more scans. Thus, considering the increasing cost of data collection and preprocessing with more scans, a careful selection of the number of scans is necessary based on these results.

The optimal circular fitting algorithm varied depending on the number of scans and presence of data noise. CF allowed for a more accurate DBH estimation with fewer scans and a lower CPC than EF. The EF performed better with higher scan counts and complete

data collection. The effectiveness of the RANSAC algorithm for data with noisy points, such as understory vegetation, was also clearly observed. Therefore, selecting a circular fitting algorithm based on site and data collection status can enable effective DBH estimation.

Finally, when evaluating the main factors affecting DBH estimation, it was found that the variance in bias decreased when the CPC was above 60%, and as Field DBH increased, the DBH estimated using LiDAR was underestimated. However, this underestimation is considered to be a result of the overmeasurement of Field DBH in large trees owing to the characteristics of the diameter tape. Therefore, when evaluating data collected using LiDAR, it is necessary to consider the characteristics of the conventional measuring equipment. Finally, as the point density increased, the accuracy increased until it reached 400 points/cm, after which it decreased. Hence, for accurate and precise DBH estimation, data collection and preprocessing must be conducted to maintain an appropriate level of point density.

This study examined the variations in accuracy depending on the data construction methods centered around DBH. Recently, research using LiDAR systems to estimate various forest-related parameters such as tree height, biomass, forest structure, and 3D modeling was conducted [43–47]. Therefore, future research should not only explore DBH but also how data construction methods influence the estimation of various forest parameters.

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

**Funding:** This research was funded by the Korea Forestry Service (Korea Forestry Promotion Institute), 2019151D10-2323-0301 and 2021359A00-2223-BD01.

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

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

## References

1. Yim, J.S.; Kim, D.H.; Ko, C.U.; Kim, D.G.; Cho, H.J. Design and Implementation of System for Estimating Diameter at Breast Height and Tree Height using LiDAR point cloud data. *JKSCI* **2023**, *28*, 99–110.
2. McTague, J.P.; Weiskittel, A. Evolution, history, and use of stem taper equations: A review of their development, application, and implementation. *Can. J. For. Res.* **2021**, *51*, 210–235. [CrossRef]
3. Chen, J.; Yang, H.; Man, R.; Wang, W.; Sharma, M.; Peng, C.; Parton, J.; Zhu, H.; Deng, Z. Using machine learning to synthesize spatiotemporal data for modelling DBH-height and DBH-height-age relationships in boreal forests. *For. Ecol. Manag.* **2020**, *466*, 118104. [CrossRef]
4. Semenzato, P.; Cattaneo, D.; Dainese, M. Growth prediction for five tree species in an Italian urban forest. *Urban For. Urban Green.* **2011**, *10*, 169–176. [CrossRef]
5. Brokaw, N.; Thompson, J. The H for DBH. *For. Ecol. Manag.* **2000**, *129*, 89–91. [CrossRef]
6. Binot, J.M.; Pothier, D.; Lebel, J. Comparison of relative accuracy and time requirement between the caliper, the diameter tape and an electronic tree measuring fork. *For. Chron.* **1995**, *71*, 197–200. [CrossRef]
7. McRoberts, R.E.; Hahn, J.T.; Hefty, G.J.; Cleve, J.R.V. Variation in forest inventory field measurements. *Can. J. For. Res.* **1994**, *24*, 1766–1770. [CrossRef]
8. Zhang, W.; Ke, Y.; Quackenbush, L.J.; Zhang, L. Using error-in-variable regression to predict tree diameter and crown width from remotely sensed imagery. *Can. J. For. Res.* **2010**, *40*, 1095–1108. [CrossRef]
9. Kattenborn, T.; Hernández, J.; Lopatin, J.; Kattenborn, G.; Fassnacht, F.E. Pilot study on the retrieval of DBH and diameter distribution of deciduous forest stands using cast shadows in uav-based orthomosaics. *ISPRS Ann.* **2018**, *4*, 93–99.
10. Pitkänen, T.P.; Raumonen, P.; Kangas, A. Measuring stem diameters with TLS in boreal forests by complementary fitting procedure. *ISPRS J. Photogramm. Remote Sens.* **2019**, *147*, 294–306. [CrossRef]
11. Donager, J.J.; Sánchez Meador, A.J.; Blackburn, R.C. Adjudicating perspectives on forest structure: How do airborne, terrestrial, and mobile Lidar-derived estimates compare? *Remote Sens.* **2021**, *13*, 2297. [CrossRef]
12. Holopainen, M.; Kankare, V.; Vastaranta, M.; Liang, X.; Lin, Y.; Vaaja, M.; Yu, X.; Hyyppä, J.; Hyyppä, H.; Kaartinen, H. Tree mapping using airborne, terrestrial and mobile laser scanning—A case study in a heterogeneous urban forest. *Urban For. Urban Green.* **2013**, *12*, 546–553. [CrossRef]

13. Liu, G.; Wang, J.; Dong, P.; Chen, Y.; Liu, Z. Estimating Individual Tree Height and Diameter at Breast Height (DBH) from Terrestrial Laser Scanning (TLS) Data at Plot Level. *Forests.* **2018**, *9*, 398. [CrossRef]
14. Singh, A.; Kushwaha, S.K.P.; Nandy, S.; Padalia, H. An approach for tree volume estimation using RANSAC and RHT algorithms from TLS dataset. *Appl. Geomat.* **2022**, *14*, 785–794. [CrossRef]
15. Ritter, T.; Schwarz, M.; Tockner, A.; Leisch, F.; Nothdurft, A. Automatic Mapping of Forest Stands Based on Three-Dimensional Point Clouds Derived from Terrestrial Laser-Scanning. *Forests* **2017**, *8*, 265. [CrossRef]
16. Liu, C.; Xing, Y.; Duanmu, J.; Tian, X. Evaluating different methods for estimating diameter at breast height from terrestrial laser scanning. *Remote Sens.* **2018**, *10*, 513. [CrossRef]
17. Koreň, M.; Mokroš, M.; Bucha, T. Accuracy of tree diameter estimation from terrestrial laser scanning by circle-fitting methods. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *63*, 122–128. [CrossRef]
18. Kong, J.; Ding, X.; Liu, J.; Yan, L.; Wang, J. New hybrid algorithms for estimating tree stem diameters at breast height using a two dimensional terrestrial laser scanner. *Sensors* **2015**, *15*, 15661–15683. [CrossRef]
19. Zhou, J.; Zhou, G.; Wei, H.; Zhang, X.; Wang, X. Evaluation of three methods for estimating diameter at breast height from terrestrial laser scanning data. In Proceedings of the IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 6674–6677.
20. Olofsson, K.; Holmgren, J.; Olsson, H. Tree Stem and Height Measurements using Terrestrial Laser Scanning and the RANSAC Algorithm. *Remote Sens.* **2014**, *6*, 4323–4344. [CrossRef]
21. Torralba, J.; Carbonell-Rivera, J.P.; Ruiz, L.Á.; Crespo-Peremarch, P. Analyzing TLS scan distribution and point density for the estimation of forest stand structural parameters. *Forests* **2022**, *13*, 2115. [CrossRef]
22. Wang, Y.; Lehtomäki, M.; Liang, X.; Pyörälä, J.; Kukko, A.; Jaakkola, A.; Liu, J.; Feng, Z.; Chen, R.; Hyyppä, J. Is field-measured tree height as reliable as believed—A comparison study of tree height estimates from field measurement, airborne laser scanning and terrestrial laser scanning in a boreal forest. *ISPRS J. Photogramm. Remote Sens.* **2019**, *147*, 132–145. [CrossRef]
23. Cabo, C.; Ordóñez, C.; López-Sánchez, C.A.; Armesto, J. Automatic dendrometry: Tree detection, tree height, and diameter estimation using terrestrial laser scanning. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *69*, 164–174. [CrossRef]
24. Liang, X.; Litkey, P.; Hyyppa, J.; Kaartinen, H.; Vastaranta, M.; Holopainen, M. Automatic Stem Mapping Using Single-Scan Terrestrial Laser Scanning. *IEEE Trans. Geosci. Remote Sens.* **2011**, *50*, 661–670. [CrossRef]
25. Lee, Y.K.; Son, Y.W.; Oh, J.S. Effects of Nitrogen and Phosphorus Fertilization on Aboveground Biomass and Distribution of Nutrient Content of Pinus rigida and Larix kaempferi Plantations in Yangpyeong area, Gyeonggi Province. *J. Korea For. Energy* **2004**, *23*, 1–18.
26. Korea Forest Service. *The 8th National Forest Inventory and Forest Health Monitoring—Field Manual*; Korea Forest Service: Daejeon, Republic of Korea, 2021; p. 37.
27. Pokswinski, S.; Gallagher, M.R.; Skowronski, N.S.; Loudermilk, E.L.; Hawley, C.; Wallace, D.; Everland, A.; Wallace, J.; Hiers, J.K. A simplified and affordable approach to forest monitoring using single terrestrial laser scans and transect sampling. *MethodsX* **2021**, *8*, 101484. [CrossRef] [PubMed]
28. Calders, K.; Adams, J.; Armston, J.; Bartholomeus, H.; Bauwens, S.; Bentley, L.P.; Chave, J.; Danson, F.M.; Demol, M.; Disney, M. Terrestrial laser scanning in forest ecology: Expanding the horizon. *Remote Sens. Environ.* **2020**, *251*, 112102. [CrossRef]
29. Zhang, W.; Qi, J.; Wan, P.; Wang, H.; Xie, D.; Wang, X.; Yan, G. An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. *Remote Sens.* **2016**, *8*, 501. [CrossRef]
30. Fareed, N.; Flores, J.P.; Das, A.K. Analysis of UAS-LiDAR Ground Points Classification in Agricultural Fields Using Traditional Algorithms and PointCNN. *Forests* **2023**, *15*, 483. [CrossRef]
31. Heinzel, J.; Huber, M.O. Detecting tree stems from volumetric TLS data in forest environments with rich understory. *Remote Sens.* **2016**, *9*, 9. [CrossRef]
32. Wang, F.; Heenkenda, M.K.; Freeburn, J.T. Estimating tree diameter at breast height (DBH) using an iPad Pro LiDAR Sensor. *Remote Sens. Lett.* **2022**, *13*, 568–578. [CrossRef]
33. Gander, W.; Golub, G.H.; Strebel, R. Least-squares fitting of circles and ellipses. *BIT Numer. Math.* **1994**, *34*, 558–578. [CrossRef]
34. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **1981**, *24*, 381–395. [CrossRef]
35. Xie, Y.; Yang, T.; Wang, X.; Chen, X.; Pang, S.; Hu, J.; Wang, A.; Chen, L.; Shen, Z. Applying a Portable Backpack Lidar to Measure and Locate Trees in a Nature Forest Plot Accuracy and Error Analyses. *Remote Sens.* **2022**, *14*, 1806. [CrossRef]
36. Xie, Y.; Zhang, J.; Chen, X.; Pang, S.; Zeng, H.; Shen, Z. Accuracy assessment and error analysis for diameter at breast height measurement of trees obtained using a novel backpack LiDAR system. *For. Ecosyst.* **2020**, *7*, 33. [CrossRef]
37. Mokros, M.; Mikita, T.; Singh, A.; Tomaštík, J.; Chudá, J.; Wężyk, P.; Kuželka, K.; Surový, P.; Klimánek, M.; Zięba-Kulawik, K.; et al. Novel low-cost mobile mapping systems for forest inventories as terrestrial laser scanning alternatives. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *104*, 102512.
38. Panagiotidis, D.; Abdollahnejad, A.; Slavik, M. 3D point cloud fusion from UAV and TLS to assess temperate managed forest structures. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *112*, 102917. [CrossRef]
39. Ruhan, A.; Du, W.; Ying, H.; Wei, B.; Shan, Y.; Dai, H. Estimation of Aboveground Biomass of Individual Trees by Backpack LiDAR Based on Parameter-Optimized Quantitative Structural Models (AdQSM). *Forests* **2023**, *14*, 475. [CrossRef]

40. Srinivasan, S.; Popescu, S.C.; Eriksson, M.; Sheridan, R.D.; Ku, N.W. Terrestrial Laser Scanning as an Effective Tool to Retrieve Tree Level Height, Crown Width, and Stem Diameter. *Remote Sens.* **2015**, *7*, 1877–1896. [CrossRef]
41. Liu, S.; Tan, K.; Tao, P.; Yang, J.; Zhang, W.; Wang, Y. Rigorous Density Correction Model for Single-Scan TLS Point Clouds. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5701318. [CrossRef]
42. Bauwens, S.; Bartholomeus, H.; Calders, K.; Lejeune, P. Forest Inventory with Terrestrial LiDAR: A Comparison of Static and Hand-Held Mobile Laser Scanning. *Forests* **2016**, *7*, 127. [CrossRef]
43. Du, S.; Lindenbergh, R.; Ledoux, H.; Stoter, J.; Nan, L. AdTree: Accurate, detailed, and Automatic Modelling of Laser-Scanned Trees. *Remote Sens.* **2019**, *11*, 2074. [CrossRef]
44. Raumonen, P.; Kaasalainen, M.; Åkerblom, M.; Kaasalainen, S.; Kaartinen, H.; Vastaranta, M.; Holopainen, M.; Disney, M.; Lewis, P. Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. *Remote Sens.* **2013**, *5*, 491–520. [CrossRef]
45. Delagrange, S.; Jauvin, C.; Rochon, P. PypeTree: A Tool for Reconstructing Tree Perennial Tissues from Point Clouds. *Sensors* **2014**, *14*, 4271–4289. [CrossRef] [PubMed]
46. Wang, F.; Sun, Y.; Jia, W.; Li, D.; Zhang, X.; Tang, Y.; Guo, H. A Novel Approach to Characterizing Crown Vertical Profile Shapes using Terrestrial Laser Scanning (TLS). *Remote Sens.* **2023**, *15*, 3272. [CrossRef]
47. Wang, F.; Sun, Y.; Jia, W.; Zhu, W.; Li, D.; Zhang, X.; Tang, Y.; Guo, H. Development of estimation models for individual tree aboveground biomass based on TLS-derived parameters. *Forests* **2023**, *14*, 351. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

forests

MDPI

*Article*

# Rethinking Productivity Evaluation in Precision Forestry through Dominant Height and Site Index Measurements Using Aerial Laser Scanning LiDAR Data

**Iván Raigosa-García [1], Leah C. Rathbun [2], Rachel L. Cook [1,*], Justin S. Baker [1], Mark V. Corrao [3,4] and Matthew J. Sumnall [5]**

[1] Department of Forestry and Environmental Resources, North Carolina State University, 2820 Faucette Dr., Raleigh, NC 27695, USA; oiraigos@ncsu.edu (I.R.-G.); jsbaker4@ncsu.edu (J.S.B.)
[2] USDA Forest Service Rocky Mountain Region, 1617 Cole Blvd, Lakewood, CO 80401, USA; leah.rathbun@usda.gov
[3] Department of Forest, Rangeland, and Fire Sciences, College of Natural Resources, University of Idaho, Moscow, ID 83844, USA; mcorrao@uidaho.edu or mcorrao@nmi2.com
[4] Northwest Management Inc., Moscow, ID 83843, USA
[5] Department of Forest Resources and Environmental Conservation, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA; msumnall@vt.edu
* Correspondence: rlcook@ncsu.edu; Tel.: +1-919-515-5979

**Abstract:** Optimizing forest plantation management has become imperative due to increasing forest product demand, higher fertilization and management costs, declining land availability, increased competition for land use, and the growing demands for carbon sequestration. Precision forestry refers to the ability to use data acquired with technology to support the forest management decision-making process. LiDAR can be used to assess forest metrics such as tree height, topographical position, soil surface attributes, and their combined effects on individual tree growth. LiDAR opens the door to precision silviculture applied at the tree level and can inform precise treatments such as fertilization, thinning, and herbicide application for individual trees. This study uses ALS LiDAR and other ancillary data to assess the effect of scale (i.e., stand, soil type, and microtopography) on dominant height and site index measures within loblolly pine plantations across the southeastern United States. This study shows differences in dominant height and site index across soil types, with even greater differences observed when the interactions of microtopography were considered. These results highlight how precision forestry may provide a unique opportunity for assessing soil and microtopographic information to optimize resource allocation and forest management at an individual tree scale in a scarce higher-priced fertilizer scenario.

**Keywords:** LiDAR; site index; precision forestry

## 1. Introduction

### *1.1. Loblolly's Pine Importance*

Loblolly pine (*Pinus taeda* L.) is the most prevalent planted forest species in the southern US and covers 14.4 million planted hectares across 13 states [1]. In the context of plantation management, site index is a measure of site quality that reflects the potential maximum height productivity of a fully stocked, even-aged stand, and is generally considered to be independent of stand density [2].

### *1.2. Dominant Height as a Measure of Productivity*

For any given tree, total height is directly related to the tree's above-ground biomass and total volume, while merchantable height is based on local specifications for given wood products [2–4]. Dominant height is defined as the average total height of the dominant and codominant trees for a given area and is often used to calculate site index [5].

Due to the difficulty in measuring the total height for all trees within a stand, different definitions for dominant height have been created, largely founded on field-plot sampling methodologies. One of the earliest definitions of dominant height was the average height of the tallest one hundred trees per hectare [6,7]. Other sources define dominant height as the average height for the largest one hundred trees per hectare, determined by the greatest diameter at breast height (DBH), measured at 4.5 feet from the ground) [8] In general, definitions for dominant height for a given area can be placed into three categories [9]:

1. Mean height for a given percentage of the tallest trees;
2. Mean height for a given number of the tallest trees as defined by size/density;
3. The tallest single tree height.

Regardless of the definition used, sampling methods are necessary to obtain an estimate of dominant height. These sampling methods can introduce bias when the tallest trees are unevenly distributed across the stand [10]. A suggested approach to correcting this bias involves the random placement of large field plots (>1 ha) to measure at least 100 trees per plot, across the stand [8]. With regard to field sampling, bias can be addressed if the plot size can appropriately capture the spatial distribution of tree heights within the given area [8,11].

### *1.3. Stand Characterization with Remote Sensing Data*

Remotely sensed data such as satellite imagery or LiDAR data provide a unique opportunity to assess dominant height and site index at finer resolution scales than a stand level, further classifying a stand into multiple strata to inform the estimations of the site index. Ultimately, allowing forest managers to obtain a more detailed understanding of how tree growth varies across different micro-topographical positions within a stand will create opportunities for increased management action effectiveness that may lead to cost savings. The acquisition of large-scale, high-resolution ALS data additionally provides an opportunity to advance the role of precision forestry in assessing site quality and forest inventory; however, the standardization of a definition for "dominant height" across any region is needed given the new era of big data in forestry.

Remotely sensed data can mitigate some challenges related to sampling for dominant height within a stand. For example, ALS point clouds can provide wall-to-wall geospatial measurements and the locations of tree canopies within a stand. It is also possible to use ALS point clouds to estimate additional metrics, such as crown length and height to live crown, for individual trees, as well as total stand biomass [2,12–16]. Commonly applied methods for determining total tree height from ALS data include the identification of the highest LiDAR returns within a pre-defined area assumed to represent an individual tree crown [16]. However, total tree height estimation using individual tree crown (ITC) methodologies from ALS data can provide results that are comparable to field measurements [15,17–19]. For example, Sumnall et al. [15] found an overall R2 value of 0.87 and a Normalized Root Mean Square Error (NRMSE) of 6.52% between field measurements and LiDAR-derived measurements for total tree height. With the ability of ALS to provide wall-to-wall coverage and more accurate estimates of individual total tree height than field measurements within a stand, the definition of dominant height can be reexamined across different scales. In most situations, estimating DBH from ALS data is still a challenge as data density plays a significant role in model accuracy and tends to decrease with closer proximity to the ground [12]. To solve this point density issue, terrestrial laser scanning (TLS) has been used to estimate DBH and individual tree volume [13,20–22]; however, scaling these methods from a few hectares to full stands or the landscape scale has been challenging [20]. Nevertheless, the use of the ultra-high-density drone LiDAR can produce clouds with thousands of points per square meter, with low-altitude flights and wide scan angles that resolve individual stem and branch structures [20].

### *1.4. Use of Remote Sensing in Precision Forestry*

Precision forestry uses information collected from various sensors to optimize spatial and temporal inputs aimed at maximizing forestland productivity. The goal of precision forestry is to improve stand productivity and operational efficiencies through high-resolution data [23]. Several data sources can be used in precision forestry, including satellite imagery, photogrammetric detection and ranging (PhoDAR), and light detection and ranging (LiDAR) data. Airborne laser scanning (ALS), which is one method of collecting LiDAR data, penetrates the canopy from above and provides feedback about the bare earth, forest structure, and understory cover. LiDAR data can be used to measure individual tree height, which can then be integrated with existing biometric equations [24] to estimate stand- or tree-level metrics, ecosystem services such as $CO_2$ sequestration, and other attributes that are important for economic valuation and management planning. Moreover, when combined with precise digital terrain models (DTMs) and soil maps, these high-resolution data products provide a unique opportunity to investigate the drivers of tree height at multiple scales and could replace costly ground-based forest inventory measurements. Furthermore, ALS produces high-quality metrics that can be used for creating and validating allometric models that inform the decision process of forest management [17,18].

The goal of this study is to enhance the characterization of forest stand productivity by assessing forest inventory methodologies that can benefit from the use of LiDAR-derived individual tree heights and better within-stand variability measurements to inform stand-level dominant height and site index estimates for loblolly pine plantations located across the southern USA. Our objectives were as follows:

1. Assess how scale impacts these estimates for loblolly pine plantations;
2. Compare different definitions of dominant height;
3. Evaluate methods for stand stratification (e.g., soils, topography, and soils plus topography), relative to simulations of traditional field inventory sample plot data.

This study makes several contributions to the literature. First, we provide an application that uses ALS LiDAR and topographic data to characterize within-stand productivity for loblolly pine plantations—a globally important supply source of timber. Second, our analysis demonstrates that while dominant height definitions may be consistent, the percentile definition does a better job representing dominant height in contrasting tree densities. The study also reveals that soil and topography play an important role in dominant height, even at the sub-stand level, and should not be overlooked. We also develop and demonstrate a method to stratify and weight site productivity, and provide tools to incorporate soils and microtopography in management decisions. Finally, we discuss the implications of our findings for forest managers interested in precision silviculture.

## 2. Materials and Methods

### *2.1. Site Locations*

Three study sites were located in the southeastern United States, inclusive of Texas, Alabama, and North Carolina (Figure 1). Eight separate forest stands were captured within these study sites. The sites in Texas and Alabama each contain two stands, with four additional stands located in North Carolina. Latitude ranges from 30°32″ N to 36°15″ N and longitude from 94°43″ W to 77°58″ W. Stands at each site are close in proximity, under 10 km apart.

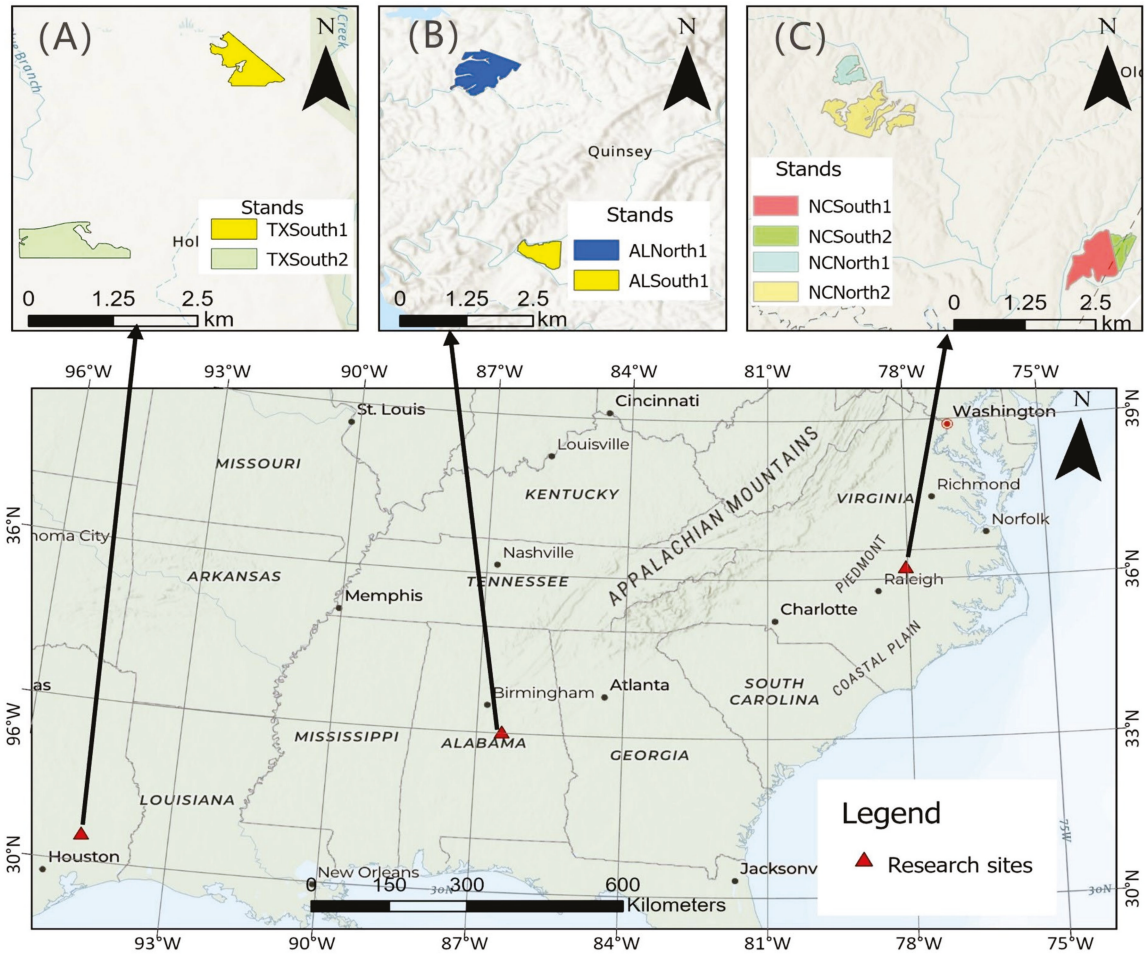



**Figure 1.** Location of the three study sites. Study site locations are indicated with a red triangle and stand locations are identified in the insets ((**A**)—Texas; (**B**)—Alabama; (**C**)—North Carolina).

Across this region, mean annual precipitation ranges from 1298 to 1154 mm and the mean annual temperature from 20.1 °C to 15.6 °C (Table 1). The mean annual maximum temperature varies from 35.6 to 33.1 °C, and the mean annual minimum temperature from 3.4 °C to −2.1 °C [25].

**Table 1.** Climate data for each study site covering the years 1981 to 2022 [25].

| Location | Mean Annual Precipitation (mm) | Mean Annual Temperature (°C) | Mean Annual Maximum Temperature (°C) | Mean Annual Minimum Temperature (°C) |
|---|---|---|---|---|
| Alabama | 1298 | 17.3 | 33.8 | 0.2 |
| North Carolina | 1154 | 15.6 | 33.1 | −2.1 |
| Texas | 1241 | 20.1 | 35.6 | 3.4 |

All stands were planted with loblolly pine between the years 1994 and 2005 and varied in size from 13 to 56.5 hectares (Table 2). The majority of stands received a thinning at 16 to 19 years after establishment; two stands, one in Texas and one in North Carolina, received a thinning at 22 and 27 years after establishment, respectively. Across all stands, thinning treatments included removing every fifth row. Residual density ranged between 298 and 733 trees per hectare. The stand-level difference in elevation varies between 7.2 and 41.4 m, with the stands in Alabama located at higher elevations than those in Texas or North Carolina.

**Table 2.** Stand overview information including management actions and elevation.

| State | Stand | Area (ha) | Year Planted | Age at Treatment (Years) | Tree Density after Thinning (Trees/ha) | Elevation Range above Sea Level (m) |
|---|---|---|---|---|---|---|
| AL | ALNorth1 | 49.1 | 2001 | 19 | 733 | 127.1–168.5 |
| | ALSouth1 | 20.4 | 2002 | 18 | 941 | 118.8–146.0 |
| NC | NCNorth1 | 16.7 | 2003 | 18 | 443 | 58.8–88.9 |
| | NCNorth2 | 56.5 | 2005 | 16 | 403 | 55.7–88.9 |
| | NCSouth1 | 51.5 | 1994 | 27 | 298 | 52.0–78.9 |
| | NCSouth2 | 13.0 | 2004 | 17 | 401 | 60.4–79.5 |
| TX | TXSouth1 | 52.3 | 2004 | 17 | 577 | 49.3–57.4 |
| | TXSouth2 | 35.8 | 1999 | 22 | 411 | 42.0–49.2 |

### 2.2. Soils and Geology

Soil maps were provided by the Forest Productivity Cooperative (FPC) Site Productivity Optimization for Trees (SPOT) system's soil code [26]. These soil maps provide a comprehensive assessment of factors known to affect tree response to forest management practices because the SPOT codes characterize soil based on a combination of factors such as dominant profile texture, depth to an increase in clay, drainage class, three soil modifiers, geology, and physiographic province. Data sources include the SSURGO database, which is an effort deriving from more than 100 years of work [27]. The sites located in Alabama are in the southern Piedmont and contain igneous, sedimentary, and metamorphic rocks that contain calcium (Ca), iron (Fe), potassium (K), and magnesium (Mg). They have a fine or coarse loamy texture with varying degrees of drainage ranging from somehow poorly to well drainage. The North Carolina sites are situated in the transitional area between the Piedmont and the Atlantic coastal plain. Soil texture varies from silty to fine loamy and is well-drained. The sites located in Texas are in the Gulf coastal plain physiographic province and the Lissie formation; these soil units were found to be severely deficient in phosphorus and have a coarse or fine loamy texture that varies from poorly drained to moderately well drained (Table 3).

**Table 3.** Soil occurrences described by the FPC SPOT codes at each stand location.

| Site | Stand | Soil Type (FPC SPOT Codes) |
|---|---|---|
| AL | ALNorth1 | B4SyxfSrPD and C2WextScPD |
| | ALSouth1 | B1WeaoBgPD, B4SyxfBgPD, C2WextBgPD, and C2WextScPD |
| NC | NCNorth1 | B2WekoAuPD and E2WykoAuPD |
| | NCNorth2 | B2WekoAuPD |
| | NCSouth1 | B2WexvAuPD, B3WekoAuPD, E2WykoAmPD, and E2WyxvAuPD |
| | NCSouth2 | B2WexvAuPD, B3WekoAuPD, and E2WyxvAuPD |
| TX | TXSouth1 | B3WeioLbGF, C3MeioLbGF, C3PyaoLbGF, and C3WeioLbGF |
| | TXSouth2 | C2MeioLbGF, C3MeioLbGF, C3PyaoLbGF, C4WoioLbGF, C6MyioLbGF, and C6PyanLbGF |

### 2.3. Data Collection and Processing

Data were collected using multiple aerial LiDAR sensors across the three study sites over the years 2020 and 2021. Table 4 provides the details for the LiDAR acquisitions. The choice of sensors was dictated by data availability rather than a deliberate intent to compare various sensors simultaneously. All the flights were conducted in a parallel pattern at 400 m above the ground in Alabama, 180 m above the ground in Texas, and 60 m above the ground in North Carolina. All airborne acquisitions used onboard aircraft navigation and real-time kinematic geographic positioning system with a precision of

≤10 cm and positional accuracy of ≤5 cm using RiACQUIRE [28]. Re-processing tasks, such as calibration scan data and strip adjustment, were performed in RiProcess [29].

**Table 4.** LiDAR acquisition details including dates of collection, and sensor and flight information. For more details and outcomes from the flight in North Carolina, please see Sumnall et al. [15].

| Location | Dates of Collection | Sensor | Pulse Density | Flight Line Overlap |
|---|---|---|---|---|
| AL | January 2020 | Riegl 1560ii | 26 pulses/m$^2$ | 60% |
| NC | July 2021 | Riegl MiniVux1 | 20 pulses/m$^2$ | 80% |
| TX | January–March 2021 | Riegl VQ-480 II | 289 to 493 pulses/m$^2$ | 55% |

All analyses were completed within R v.4.1.2 [30]. The LiDAR point cloud was processed to ensure all outliers were removed using the LidR package [31]. We applied the statistical outlier removal (SOR) tool to identify outliers in the point cloud (e.g., extreme ground points and extreme canopy points). Filter parameters for this process included the use of one normal distribution and a k-nearest neighbor value of eight points [31]. The LAScatalog processing engine was used to segment the LiDAR point cloud into smaller data files [32] and default parameters for segmentation size were used. After outliers were removed, we processed the LiDAR point cloud to create two products used in the analysis: (1) the digital terrain model (DTM) and (2) the individual tree crown (ITC) segmentation layer.

We performed ground classification of the LiDAR point cloud using the classify ground function in R, with a progressive morphological filter with window sizes of 3, 12, and 3 m and elevation difference threshold sizes of 0.1 and 1.5 m [32,33]. The resulting DTM represents a pixel size of 0.25 m. We generated the DTMs using the grid_terrain function and the TIN algorithm within the LidR package. Using the DTM, we calculated the difference between elevation at a central point and the average elevation around it for a given radius or topographical position index (TPI) [34] for each 0.25 pixel. To calculate TPI, we used the TPI function within the spatialEco package [35] with an adaptive window size. The size of the window was different for each site and based on the rate of elevation change for a given unit area. Hence, we used a smaller window size for sites with a smaller elevation change and a larger window size for sites with a greater elevation change. We performed a sensitivity analysis to determine the correct window size by conducting comparisons between the TPI raster file and known topographical structures within the stands. The sensitivity analysis indicated that an ideal window size was correlated with ten times the range of elevation within the stand.

From the DTM, a topographic position index (TPI) layer was created to describe three microtopography classes: valley, flat, or ridge. For each pixel in the DTM, microtopography classes were calculated [34] as follows:

1. Valleys—value of less than one negative standard deviation of the TPI;
2. Flat—value greater than or equal to one negative standard deviation but less than one half standard deviation of the TPI;
3. Ridge—value greater than one half standard deviation of the TP.

We delineated individual tree crowns (ITC) within the LiDAR point cloud using the ITC segmentation method created by Sumnall et al. [15] for loblolly pine. The detection of stem locations in this method relies on exploiting the high pulse density of the LiDAR data. ITCS delineation is based on segmenting the canopy height model (CHM) [12] using previously detected stems as initial markers. The author reported an average proportion of missed alive trees of around 8%; however, for low tree densities (618 trees/ha), the method only missed 1–3%. For each tree crown, we used the tallest returned pulse to determine the total height and the center of the tree. A microtopography class was assigned to each tree center. Microtopographic classifications were determined as the mode value for all pixels

contained within a 2.76 m radius circle around the tree center using the majority option in the exact_extract function from the exactextractr package [36]. In addition, for each tree center, we assigned a corresponding FPC SPOT soil code. We removed all trees contained within a buffer zone of 15 m from the edge of each stand to address any edge effects on tree height growth for this analysis.

### *2.4. Data Analysis*

#### 2.4.1. Dominant Height Definitions

Three definitions of dominant height were calculated for each stand using all trees identified from the ITC methodology. In addition, plots were simulated for each stand and the dominant height values were determined for trees found within each plot. The simulation included the generation of a grid of 150 m in size randomly placed within each stand, with a 12.6 m radius circular plot (1/20 hectare) at the "node or centroid" of each grid cell (total plots within each stand are provided in Table 5). The size and shape used in the simulation are common in forest inventories for pine plantations [37]. Bar graphs were created to assess the differences across the dominant height definitions for each stand and the plot simulation. Based on these differences, one definition was selected for further analysis of how scale impacts dominant height and site index. The definitions for dominant height were as follows:

1. Top HT 85—mean height of the 85th percentile of tallest trees;
2. Top HT 100—mean height of the tallest 100 trees per hectare within the stand;
3. Top HT 40—mean height of the tallest 40 trees per hectare within the stand.

**Table 5.** Summary of LiDAR-derived population tree detection value results by stand, including the size of the stand, total and mean trees detected, and mean tree height with the corresponding standard deviation.

| Location | Stand | Area (ha) | Number of Trees | Mean Trees per Hectare | Mean Tree Height (m) | Number of Plots Simulated |
|---|---|---|---|---|---|---|
| AL | ALNorth1 | 49.1 | 35,537 | 724 | 17.3 (1.9) | 22 |
| | ALSouth1 | 20.4 | 19,154 | 938 | 17.2 (1.6) | 9 |
| NC | NCSouth1 | 51.5 | 15,136 | 294 | 21.2 (1.7) | 26 |
| | NCSouth2 | 13.0 | 5260 | 404 | 17.0 (1.4) | 9 |
| | NCNorth1 | 16.7 | 7296 | 437 | 17.3 (1.5) | 9 |
| | NCNorth2 | 56.5 | 22,723 | 400 | 17.2 (1.7) | 31 |
| TX | TXSouth1 | 52.3 | 19,061 | 365 | 17.7 (1.2) | 24 |
| | TXSouth2 | 35.8 | 18,585 | 519 | 22.0 (2.0) | 16 |

#### 2.4.2. Scale

To assess the effects of scale on dominant height, each stand was stratified into four scales:

1. Full stand: dominant trees are selected from the tree population for each stand without stratification.
2. Soil classification: dominant trees are selected from the tree population for each soil type.
3. Microtopography classification: dominant trees are selected from the tree population for each microtopography type.
4. Combined soil and microtopography classification: dominant trees are selected from the tree population for each combination of soil and microtopography.

Each tree center found from the ITC methodology was classified for each scale and the area within each stand was determined for each classification (Figure 2).

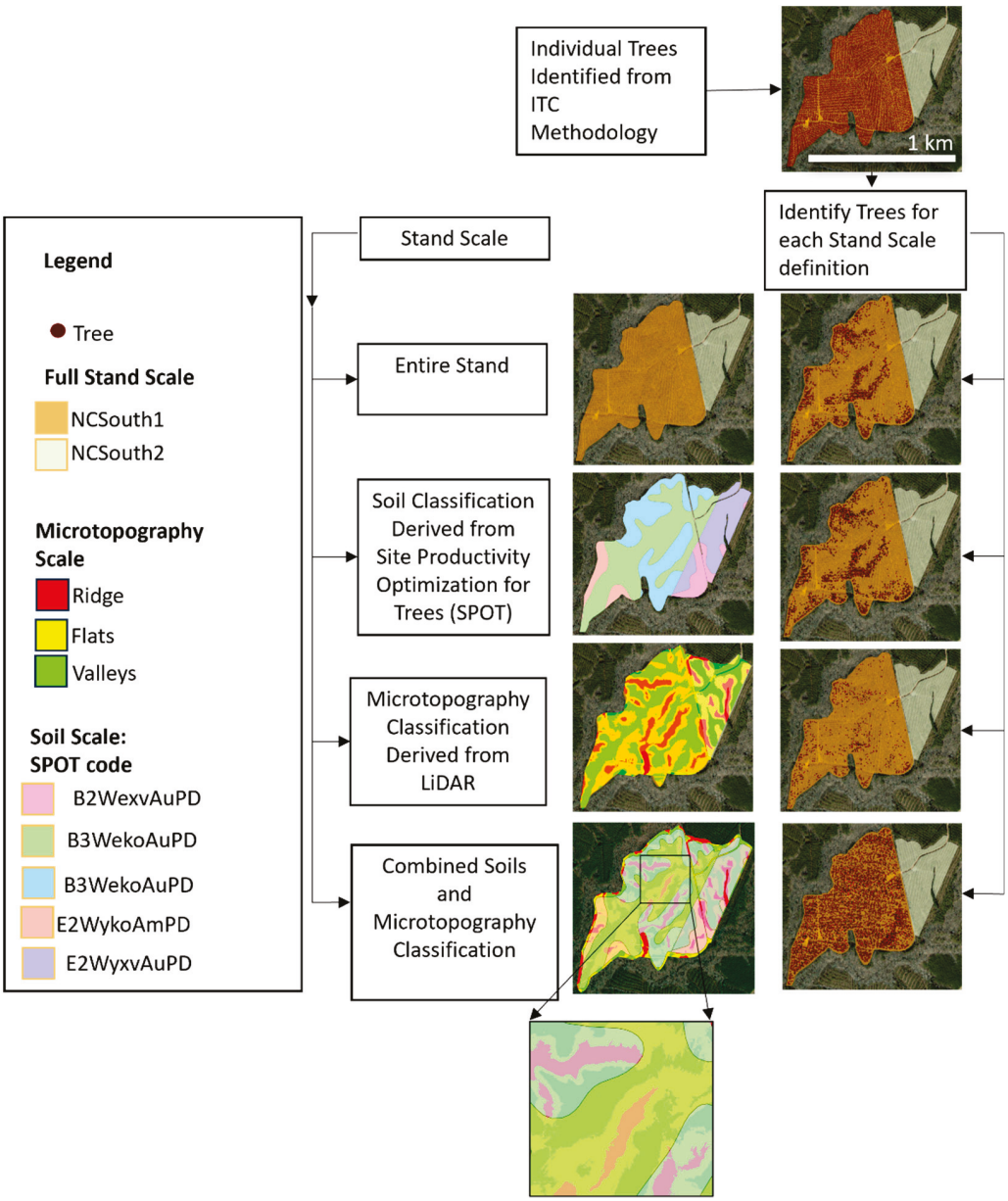



**Figure 2.** Stand scale and tree selection process for stand NCSouth1.

To explore the impacts of scale on dominant height, dominant height was calculated for each class within each scale (e.g., soils, microtopography, and the interaction of soils and microtopography). Boxplots were created to compare the range of tree heights used to compute the dominant height value for each soil type for each stand. These boxplots were compared to the trees used at the full stand scale. In addition, Dunnett's tests [38] were conducted to determine if there was a difference between the means for each soil class compared to the full stand values for each stand. The tests were conducted at the stand level to ensure the effect of age was removed. Dunnett's tests were also conducted to compare each microtopography scale to the full stand values for each stand. For each test, an alpha level of 0.05 was used to determine statistical significance.

To investigate the interaction of soils and microtopography, we developed heat maps for each study site. The heat maps visualize the percent differences between the mean dominant height values at the combined soil and microtopography scales compared to the full stand scale. Percent difference was used to normalize the information across the [39] different aged stands.

2.4.3. Site Index

To assess the impacts of incorporating scale into stand-level site index estimates, a weighted site index was calculated for each scale and compared to the full stand estimate for the site index. The site index was calculated using the equation for loblolly pine by Dieguez-Aranda et al. [39]:

$$SI = \frac{26.14 + X_0}{1 + \frac{1455}{X_0 t^{-1.107}}} \quad (1)$$

where $t$ is a base age value of 25 and $X_0$ is as follows:

$$X_0 = \frac{Y_0 - 26.14 + \sqrt{(Y_0 - 26.14)^2 + 4 * 1455 Y_0 t_0{}^{-1.107}}}{2} \quad (2)$$

where $t_0$ is the stand age and $Y_0$ is the dominant height. Site index values were computed for each class within a scale and a weighted average was determined for each stand. Weights were based on the area for each class within a scale:

$$Weighted\ SI = \sum_{i=1}^{n} \frac{SI_i A_i}{\sum_{i=1}^{n} A_i} \quad (3)$$

The Wilcoxon test [40] was conducted to determine if there was a difference among weighted site index estimations across all stands for each scale. For each test, an alpha level of 0.05 was used to determine statistical significance.

## 3. Results

The number of population trees detected from the LiDAR point cloud varied from 294 to 938 per hectare, with North Carolina having the lowest-density stands and Alabama having the highest (Table 5). Mean tree height was consistent at around 17 m for the majority of the stands, except for two, NCSouth1 and TXSouth2, which were 21.2 and 22.0, respectively. These two stands are also the oldest in the study. The standard deviation of mean tree height varied from 1.2 m to 2.0 m across all stands, while the total number of trees varied from 5260 to 35,527 per stand.

### *3.1. Dominant Height Definition*

The dominant height definitions impact the dominant height value and its corresponding variability (Figure 3). As expected, when dominant height is defined by a given number of trees per unit area, the mean value increases when the given number of trees per unit area decreases. This trend is seen across all stands when comparing the forty trees per hectare to the hundred trees per hectare definition. Also as expected, the spread in the data increases with an increasing number of trees. Across all stands, the forty trees per hectare definition provided the largest mean dominant height value. When the definition is based upon the 85th percentile, the mean dominant height value fluctuates to either higher or lower than the 100 trees per hectare value. Across the study sites, tree density varies from 365 to 938 trees per hectare, allowing for the percentile method to be adjusted based on density changes, unlike the other two definitions. Because a percentile definition better reflects the relative differences across stands with large fluctuations in tree density, the 85th percentile definition will be used to investigate the impacts of soil and microtopography on dominant height and site index.

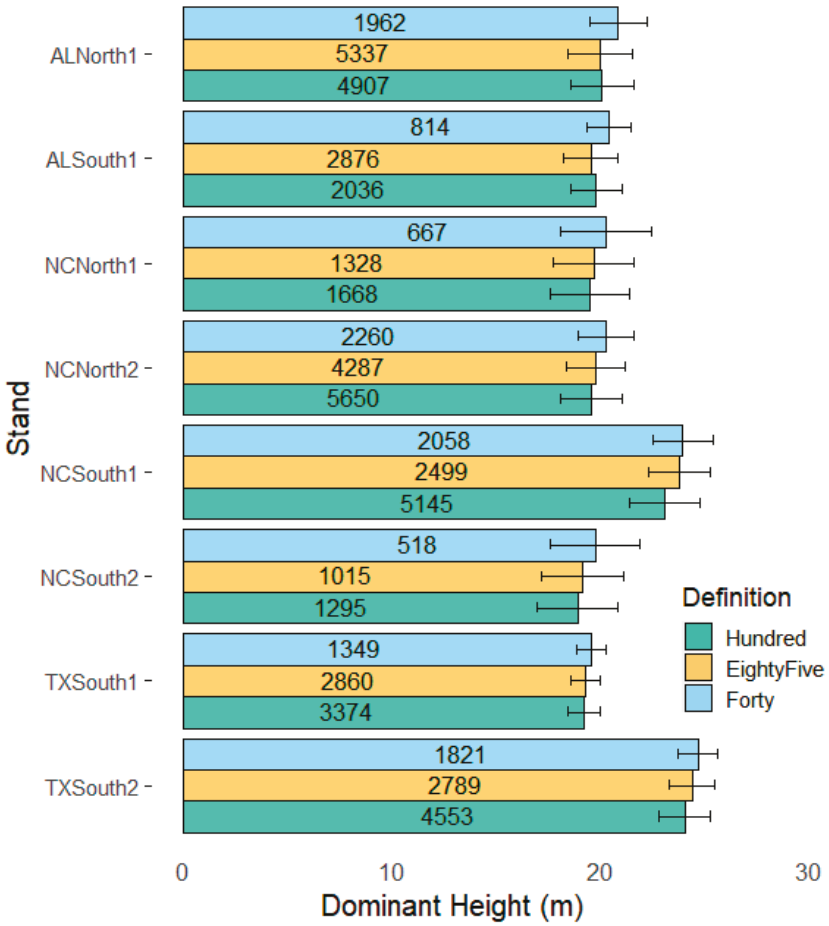



**Figure 3.** Dominant height values were computed at the stand scale using the following three definitions: (1) one hundred trees per hectare, (2) forty trees per hectare, and (3) the eighty-fifth percentile. The error bars represent the 95% confidence interval for the mean value from the simulated plot inventory data. The numbers listed inside the bars represent the number of trees used to calculate the dominant height per definition within a stand.

### *3.2. Scale*

#### 3.2.1. Soils

The number of soil types and their impact on dominant height varies by stand (Figure 4). The range of height for trees within the 85th percentile varied by state, with North Carolina having the widest range and Alabama having the narrowest range. In stand ALSouth1, two soil types were found to be statistically different from the full stand estimate for mean dominant height and two soil types were not: C2WextScPD and B1WeaoBgPD ($p$-values of 0.94 and 0.64). Soil type C2WextScPD, found in stand ALNorth1, was not significantly different from the full stand ($p$-value 0.94) but B4SyxfSrPD was ($p$-value < 0.01). Across the four stands within North Carolina, there was one soil type that was not statistically significant for one stand but was significant for another, B2WexvAuPD ($p$-values < 0.01 and 0.93). Stand NCNorth2 had only one soil type and was not included in the analysis. The boxplots show that two of the soil classes in North Carolina have greater variability when compared to the others, E2WykoAuPD and B2WexvAuPD. The greatest differences in soil types can be seen in the state of Texas. Three soil types were found to not be statistically significant when compared to the full stand in Texas: C3WeioLbGF ($p$-value of 0.08) in stand TXSouth1, and C6MyioLbGF and C3PyaoLbGF ($p$-values of 0.61 and 0.17) in stand TXSouth2. Soil type C3PyaoLbGF was statistically significant for TXSouth1, but not for another stand ($p$-value < 0.01).

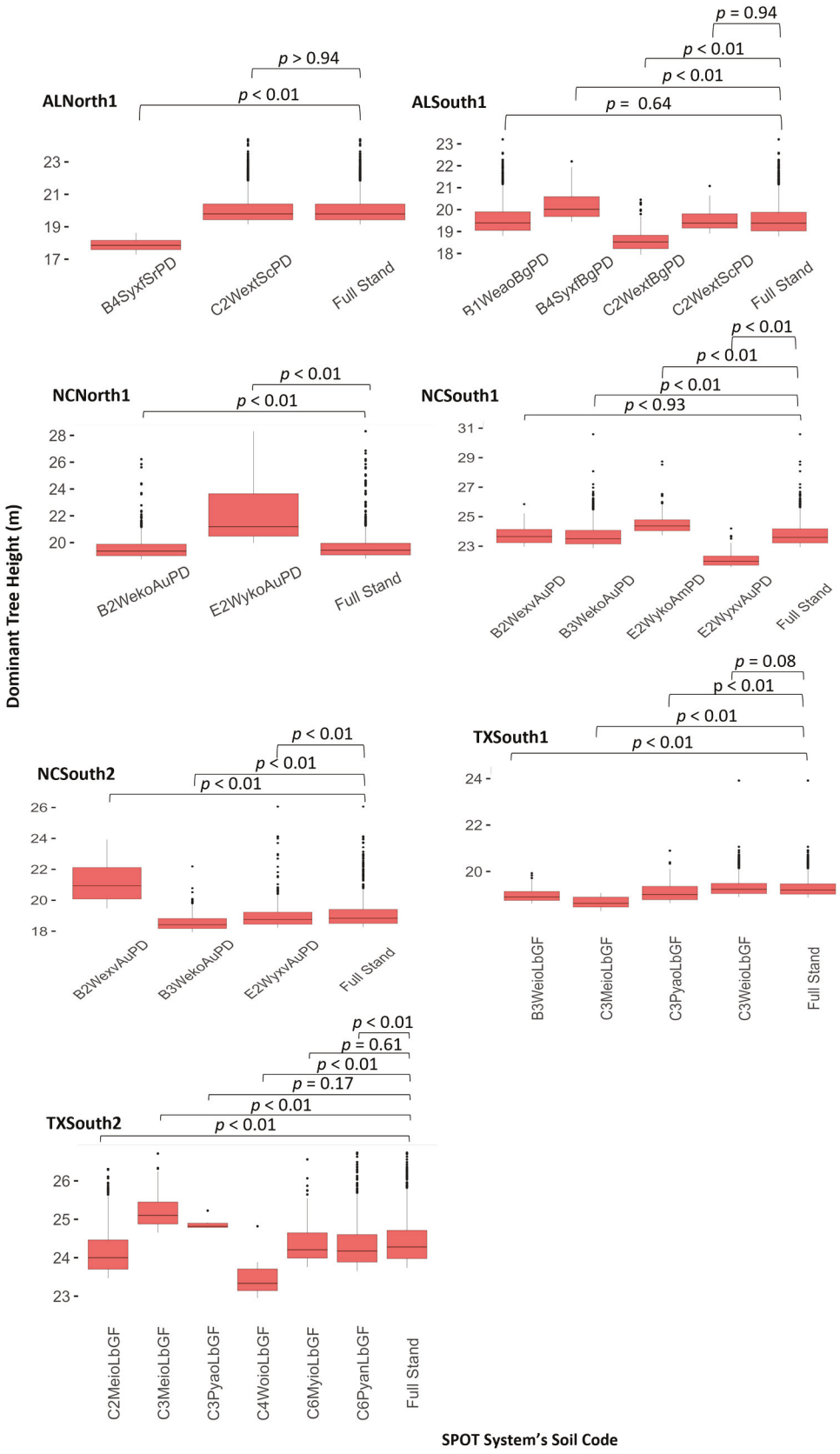



**Figure 4.** Average dominant height values for each soil type within a stand were compared. Any soil types with less than five (5) values were excluded from the graphic. Dunnett's test results were calculated using the means for each soil class compared to the full stand scale values for each stand.

### 3.2.2. Microtopography

The results for microtopography indicate that Alabama has the smallest range of height values for the 85th percentile trees within the microtopography scale (Figure 5). The *p*-values indicate that all microtopography class estimates for mean dominant height are statistically different from the full stand estimates in North Carolina. In stand ALSouth1, the mean dominant height for those trees found on a flat is not statistically different from the full stand estimate (*p*-value of 0.37), while those on mounds or in valleys show statistical differences in height. All microtopography class estimates are statistically different from the full stand estimates in stand ALNorth1. A different trend is seen in Texas, where the mean dominant height for those trees found in valleys and on a flat is not different from the full stand estimate in the stand TXSouth1 (*p*-values of 0.22 and 0.07). For stand TXSouth2, the mean dominant height for trees found on a flat is not statistically different from the full stand estimate (*p*-value of 0.39), while those on mounds or in valleys show statistical differences in height. In all states, the mean dominant height for trees in a valley is greater than the full stand estimate, with a very large difference found for Texas.

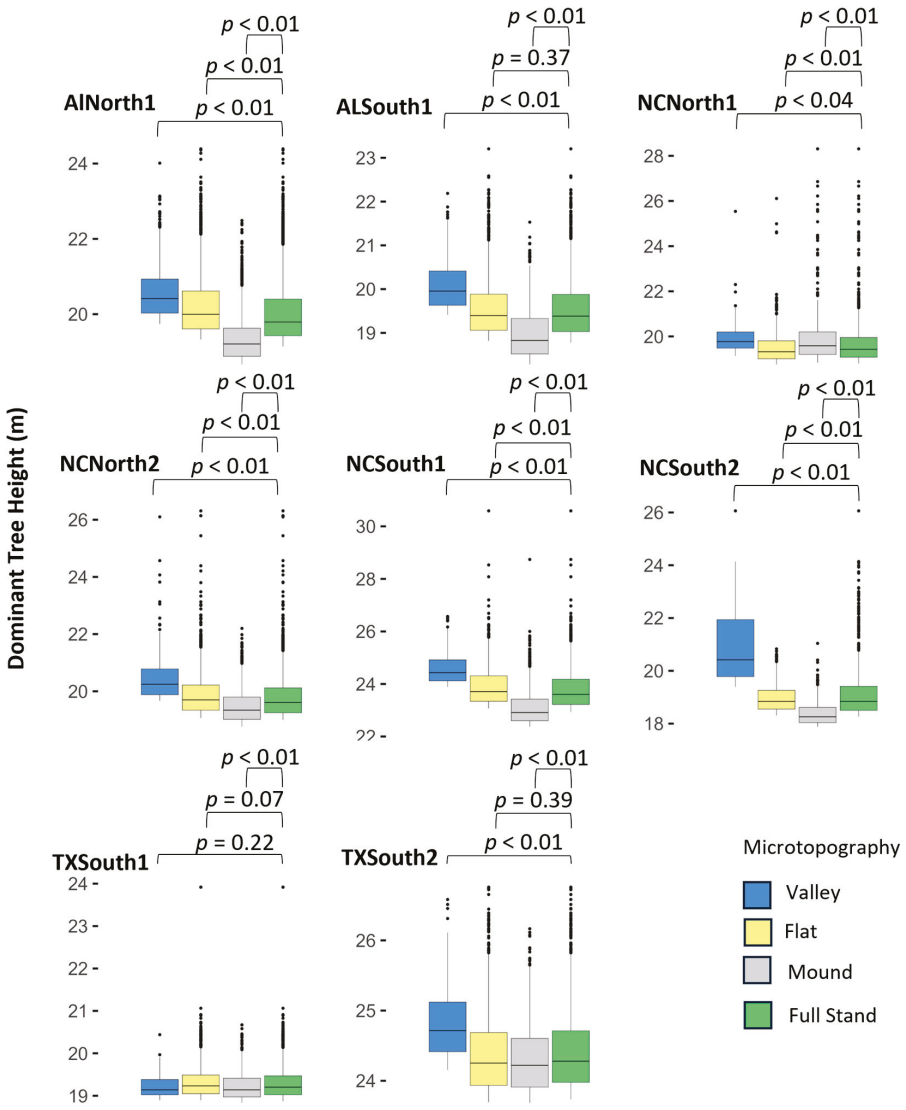



**Figure 5.** Average dominant height value for each microtopography class within a stand. The *p*-values provided at the top of the graph are for the null hypothesis that the mean differences are equal to zero. Dunnett's test results were conducted between the means for each soil class compared to the full stand scale values for each stand.

3.2.3. Soils and Microtopography

The heat maps below provide a visual representation of how soils, microtopography, and the interaction of soils and microtopography impact dominant height within the same region. For stand ALNorth1 in Alabama (Table 6), the percent change in average dominant height differs by 11 percent, while stand ALSouth1's percent change is 8 percent. Soil type C2WextScPD appears in both stands and shows no difference from the full stand estimate in either stand. Soil type B4SyxfSrPD shows the largest departure from the full stand estimate at −11%. Microtopography differences are consistent across both stands for estimates obtained for trees on a mound or in a valley, but the percent change for those trees in a flat area is different between the two stands. The mean dominant height for those trees in a valley is 3 percent greater than the full stand estimate, compared to the mean height of those in a mound, which is 3 percent less than the full stand estimate. The mean dominant height for those trees found on soil type B4SyxfBgPD in a valley is less than the full stand estimate, which is different from trees found in a valley on any other soil type. Dominant trees on a mound in the B4SyxfBgPD soil type are, on average, the shortest, and those in flat areas in the B4SyxfBgPD soil type are the tallest on average.

**Table 6.** Heat map for the stands in Alabama with each cell value representing the percent difference in average dominant height. The average dominant height for each scale was compared to the overall average for the full stand (labeled in grey). Green indicates a percent difference less than 3, red indicates a percent difference between 3 and 5 percent, and blue indicates a percent difference greater than 5 percent. Black indicates no trees in that combination of soil and microtopography.

| State | Stand | Soil | Flat | Mound | Valley | Sub Total |
|---|---|---|---|---|---|---|
| AL | ALNorth1 | B4SyxfSrPD | | −9% | | −11% |
| | | C2WextScPD | 1% | −3% | 3% | 0% |
| | | Sub Total | 3% | −3% | 3% | 20.2 m |
| | ALSouth1 | B1WeaoBgPD | 0% | −3% | 3% | 0% |
| | | B4SyxfBgPD | 4% | −12% | −2% | 3% |
| | | C2WextBgPD | −7% | −7% | | −5% |
| | | C2WextScPD | −1% | −8% | 2% | 0% |
| | | Sub Total | 0% | −3% | 3% | 19.6 m |

A greater range in percent difference is seen across the soil types from North Carolina as compared to Alabama. Two soil types show large positive percent differences: E2WykoAuPD (13 percent) and B2WexvAuPD in stand NCSouth2 (11 percent); however, B2WexvAuPD shows a slightly smaller percent difference in NCSouth1 (3 percent). Soil type E2WyxvAuPD shows a large negative percent difference at −7 percent. Trees located in a valley show a percent increase in average dominant height when compared to the full stand estimate; those in a mound have a negative percent difference; and those on a flat show little percent difference. Dominant trees on a mound in the E2WykoAuPD soil type show a large increase in the percent difference while across all other soil types, dominant trees located on a mound are shorter on average when compared to the full stand (Table 7).

The stands in Texas show the least amount of variability (Table 8). Percent differences are more pronounced across different soil types than microtopography. The stands in Texas have the smallest elevation range. Dominant trees located on flat surfaces within the C3MeioLbGF soil type are the shortest on average when compared to the full stand and those within the C3MeioLbGF soil type are the tallest.

*3.3. Site Index*

The Wilcoxon test indicates that when compared to the full stand estimate for site index, the weighted estimates using the soils and the interaction of soils and microtopography are statistically different for an alpha level of 0.05. In addition, the interaction of soils and microtopography is statistically different from all other scales but not when compared to the plot simulation values (Figure 6 and Table 9).

**Table 7.** Heat map for the stands in North Carolina with each cell value representing the percent difference in average dominant height. The average dominant height for each scale was compared to the overall average for the full stand (labeled in grey). Green indicates a percent difference less than 3, red indicates a percent difference between 3 and 5 percent, and blue indicates a percent difference greater than 5 percent. Black indicates no trees in that combination of soil and microtopography.

| State | Stand | Soil | Flat | Mound | Valley | Sub Total |
|---|---|---|---|---|---|---|
| NC | NCNorth1 | B2WekoAuPD | −1% | 0% | 1% | −1% |
| | | E2WykoAuPD | 8% | 16% | | 13% |
| | | Sub Total | −1% | −1% | 2% | 19.7 m |
| | NCNorth2 | B2WekoAuPD | 0% | −2% | 3% | 0% |
| | | Sub Total | 0% | −2% | 3% | 19.8 m |
| | NCSouth1 | B2WexvAuPD | 1% | −5% | −3% | 0% |
| | | B3WekoAuPD | 0% | −3% | 3% | 0% |
| | | E2WykoAmPD | 3% | 2% | 5% | 3% |
| | | E2WyxvAuPD | −7% | −7% | | −7% |
| | | Sub Total | 0% | −3% | 3% | 23.8 m |
| | NCSouth2 | B2WexvAuPD | 1% | −2% | 18% | 11% |
| | | B3WekoAuPD | −2% | −5% | 0% | −3% |
| | | E2WyxvAuPD | −1% | −4% | 6% | −1% |
| | | Sub Total | −1% | −4% | 10% | 19.1 m |

**Table 8.** Heat map for the stands in Texas with each cell value representing the percent difference in average dominant height. The average dominant height for each scale was compared to the overall average for the full stand (labeled in grey). Green indicates a percent difference less than 3, red indicates a percent difference between 3 and 5 percent, and blue indicates a percent difference greater than 5 percent. Black indicates no trees in that combination of soil and microtopography.

| State | Stand | Soil | Flat | Mound | Valley | Sub Total |
|---|---|---|---|---|---|---|
| TX | TXSouth1 | B3WeioLbGF | −1% | 1% | −3% | −2% |
| | | C3MeioLbGF | −6% | −2% | 0% | −3% |
| | | C3PyaoLbGF | −1% | 0% | 1% | −1% |
| | | C3WeioLbGF | 0% | 0% | −1% | 0% |
| | | Sub Total | 0% | 0% | 0% | 19.3 m |
| | TXSouth2 | C2MeioLbGF | −2% | −2% | 2% | −1% |
| | | C3MeioLbGF | 4% | 3% | 2% | 3% |
| | | C3PyaoLbGF | #NAME? | 0% | 0% | 2% |
| | | C4WoioLbGF | −4% | −5% | 0% | −4% |
| | | C6MyioLbGF | −1% | 0% | 2% | 0% |
| | | C6PyanLbGF | 0% | −1% | 0% | 0% |
| | | Sub Total | 0% | 0% | 2% | 24.4 m |

**Table 9.** Comparison of weighted site index estimations across all stands for each scale with the Wilcoxon test. *p*-values listed in bold indicating statistical significance.

| Stand Scale | Soils | Microtopography | Soils and Microtopography | Plot Simulation |
|---|---|---|---|---|
| Full Stand | **0.0225** | 0.0547 | **0.0078** | 0.2500 |
| Soils | | 0.6406 | **0.0078** | 0.3828 |
| Microtopography | | | **0.0225** | 0.3125 |
| Soils and Microtopography | | | | 0.6406 |

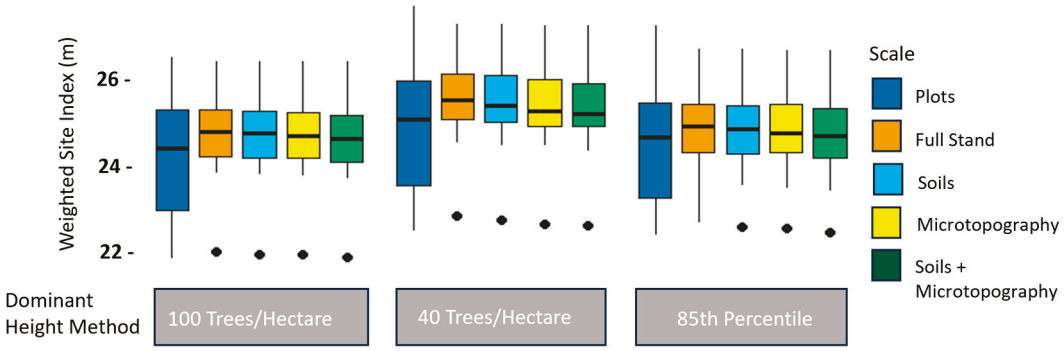



**Figure 6.** Weighted dominant height values computed using stand-scale level (i.e., full stand, soil classification, microtopography classification, and combined soil and microtopography classification) within three definitions (i.e., forty trees per hectare, the eighty-fifth percentile, and a hundred trees per hectare).

## 4. Discussion

### *4.1. Productivity Definition*

Measures of site quality provide land managers with information that is critical to forestland decision-making processes. Improvements in the accuracy and resolution of site quality offer the potential to make better decisions regarding tree density choices, potential localized growth rates, and nutrient availability or response potential from additives such as fertilizer or herbicide treatments. An increased understanding of these conditions within a forest stand provides opportunities to minimize costs and increase profitability, as well as influence additional landowner objectives such as forest health and aesthetics, wildlife preservation, and/or hydrologic function. Profitability is often assessed as the amount of current or future merchantable volume available within a stand at a given point in time. Traditionally, site index has been used as a measure of site quality and has been defined and spatially confined to within stand boundaries within the southeastern U.S. Burkhart and Tomé [5] define a stand as a contiguous group of trees, sufficiently uniform in age–class distribution, species composition, and structure, growing on a site of sufficiently uniform quality to be a distinguishable unit. Stand boundaries in the southern U.S. typically reflect localized geographic features such as roads, land ownership, fire breaks, water bodies, and other physical land features, rather than uniform productivity. The use of remote sensing data such as satellite imagery, PhoDAR point clouds, and LiDAR point clouds creates a unique opportunity to assess information at a scale finer than the stand level. The use of LiDAR point clouds to create tree maps and assess microtopography moves the metrics available to forestland management decisions closer to those available in precision agriculture, thus providing the ability to assess conditions and apply management activities at finer resolutions. This creates the need to understand how traditional metrics such as dominant height and site index perform at scales other than the stand scale.

### *4.2. Dominant Height*

Dominant height performance is linked directly to the definition used for the calculation. Sharma et al. [41] investigated seven definitions for dominant height across thinned and unthinned loblolly pine stands and determined that, in most cases, all seven definitions produced different results. This study indicates that for stands with sharp differences in density (e.g., trees per hectare), a more consistent result is achieved using a definition for dominant height that is based on a percentile of the trees than when using a dominance selection based on trees per hectare counts [42].

Traditionally, dominant height was determined by measuring dominant trees within a stand. In some cases, dominant-height trees were selected by placing field plots within the stand. Determining the correct plot size can be challenging, and bias can be introduced if the plot size does not appropriately capture the distribution of dominant height within

the stand [10,41]. This study found that the dominant height was shorter for the simulated plots when compared to using the full stand tree list developed from the LiDAR point cloud. This difference in dominant height estimate is likely due to selection bias as the grid of plots was spaced evenly throughout the stand, capturing a measure of within-stand height heterogeneity. García and Batho [43] define this selection effect as the phenomenon of when larger trees within the stand are not located on the plot.

### 4.3. Influence of Site Variability on Dominant Height

Tree height in the southern U.S. is often influenced by microtopography and soil variations, which are foundational to defining microsite conditions. Microtopography is defined as influential topographic variability for individual trees, and it has been described as small-scale (<2 m) surface variation within an elevation range [44]. Additionally, variations in soil type (texture, structure, etc.) are often present within traditionally defined stand boundaries. Between soil types, different soil properties affect vegetation. For example, soil depth has been found to be an indicator of effective rooting volume [45]. Similarly, Koirala et al. [46] found strong associations between tree height and water-holding capacity in surface soil conditions related to soil permeability and aeration in northeastern Texas. Studies from other geographic regions, such as those by Zhao, Y. et al. [47], have observed that relative bulk density can limit total tree height for Douglas-fir and lodge pole pine in British Columbia. The influences of soil type and condition suggest that a combination of shallow soil depth, low rainfall, and less fertile horizons may slow tree growth [48]. Thus, understanding the soil attributes across a landscape would lead to improved site index predictions [49]. Even in relatively homogenous sites, when other resources are limited, small variations in soil nutrients or water availability can positively influence microsite productivity [49].

In addition to soils, microtopographic conditions impact dominant height at a finer scale. Small variations in soil drainage within pine plantations located within the southeastern U.S, Western Gulf Coast, are not always caused by broad changes in slope but rather small topographic changes associated with randomly distributed “pimple mounds”, circular to elliptical soil mounds ranging from 0.5 to 1.5 m in height and 10 to 30 m in diameter [50]. Additionally, the presence of a floodplain can lead to an accumulation of soil organic carbon (SOC) and an increased cation exchange capacity (CEC), providing vegetation with additional nutrient uptake abilities. This contrasts with higher elevation, ridge, and spur positions, where a lower CEC may exist due to rainwater leaching, erosion, and past land-use practices that further impact the availability of nutrients such as potassium (K), which is critical to tree growth [51,52]. Tree growth for some species has been shown to increase in relation to sediment deposition areas and the presence of more neutral pH levels [53], which can positively affect CEC. Additionally, during seasonal or extreme droughts, microtopographic conditions can influence water availability to trees in lowland areas, where alluvial or erosive soil deposition has increased soil depth and texture, thereby influencing the water-holding capacity of some areas and supporting the recovery and resilience of soil health [54]. These microtopographic changes and local topography influences influence tree growth, soil fertility, and water availability at multiple scales, even within individual inventory field plots [55,56]. Site index maps have been used to improve dominant tree height estimates for growth and yield modeling [49,57–60]. Therefore, the inclusion of high-resolution ALS data for tree height, tree density, and microtopography could help to predict soil features, like water availability and texture, since soil nutrition partly depends on the topographical position. The influence of soil and microtopography on tree height has been acknowledged throughout many forested regions, and thus the heterogeneity within a forester-drawn stand is also an aspect of forest management that needs to be considered. If dominant-height trees are not found evenly throughout a stand, there may be LiDAR-derived values that can be used to delineate stands to more biologically representative areas. Similarly, Liu and Burkhart [61] found that increments of total height were spatially autocorrelated.

Using ancillary data to map soil types and define dominant trees at a different scale can lead to better dominant height estimates. We found that soil type influenced dominant height estimates. For example, the C2WextScPD soil type in Alabama produced a higher dominant height than the B4SyxfSrPD soil type. The C2WextScPD soil type has a stronger nutrient availability of C, Fe, K, and Mg, a well-drained soil profile characterized by a coarse loamy texture, and a shallower depth due to the argillic horizon. In contrast, the B4SyxfSrPD soil type has good nutrient availability for K, drains poorly, and possesses a deeper depth due to the increase in clay. Kelting et al. [62] found that for young loblolly pine, site preparation treatments, and soil productivity indicators explain 87% of the variation in tree height; and Kelting et al. [63] found that surface soil depth had a weak positive effect on tree growth. At the study site in Texas, the lowest dominant height values were found in well-drained soil with a coarse loamy texture, while the soils with wetter drainage conditions had higher dominant height values. Pangle and Seiler [64] found that soil moisture limitations in coarse loamy soils can limit microbial activity in loblolly pine. These results were somewhat anticipated given the body of literature defining the influences of soil texture, structure, CEC, and water-holding capacity on vegetation development within both commercial forest and agricultural systems.

Incorporating microtopography metrics resulted in an improvement in the overall location accuracy of trees exhibiting a dominant height. We found that across all three states, dominant-height trees located in a valley or lowland positions showed greater total height-by-age than those found on a mound or in a flat area of similar age. These results are similar to those described by Borders et al. [52], who found that for loblolly pine sites located near each other in the Dixon State Forest near Waycross, Georgia, USA, slight elevation changes create distinct site characteristics. Lorio et al. [65] observed that for loblolly pine located on wet sites, variations in mycorrhizal surface area were dependent upon localized differences in microtopography (pimple mounds up to 2 m). The heat maps developed in this study highlight differences in combinations of soil type and microtopography impact in terms of measured total height. Changes in total height are observed within and between stands when either the soil type or the microtopography are held constant. Additionally, the weighted site index Wilcoxon tests indicate significant differences in weighted site index when the interaction of soil and microtopography are considered together from location to location. Care should be taken when considering estimates of tree dominance, and ultimately site index, at any scale to ensure that the area sampled for each class is large enough to provide a representative sample and considers the heterogeneity of soil type and microtopography throughout [66].

### *4.4. Precision Forestry Using Remote Sensing Data*

Conventionally, forest management has focused on using sampled, mean, and stand-level estimates of forest attributes for planning and management activities within growth and yield models of projected performance. The research here suggests LiDAR data can enable a total stand population to be assessed and identify differences in valuable metrics such as dominant tree height and relative spatial positioning at scales finer than the stand-level. This differentiation enables us to consider tailored silviculture planning and management practices such as targeted fertilizer applications to improve forestland resource investments. Incorporating ancillary data such as soil type maps and microtopography can improve estimates of variables such as dominant height and site index, which are often important drivers within growth and yield models. Future research should consider the potential of using measured tree growth from LiDAR populations to improve soil attribute mapping for aspects such as depth and water-holding capacity. From the results of this research, it is suggested that, within the southeastern U.S., stands of loblolly can potentially be managed at different scales to optimize resource allocation in scenarios of increasing fertilizer prices and uncertainty in timber prices [4].

## 5. Conclusions

ALS LiDAR data collected at or above the densities presented in this work offer a robust opportunity to characterize total forest stand populations and provide information at finer scales than previously available to inform forest management practices. The interaction of soil conditions and microtopography can play a meaningful role in the average height of dominant trees within loblolly stands similar to those in this study. However, the results of this work indicate that slope position and tree density also play key roles in the height and growth within plantations similar to these. Forestland managers should consider soil and microtopography within management decisions because, when included in a weighted site index measure, the resulting values were statistically different from standard practices within these sites. Additionally, it may be beneficial to employ stratification techniques when measuring site quality and productivity to prevent bias in measurements of dominant height when employing them for determining site index, especially in stands with a wide variety of soil types and microtopographies [67]. Publicly available LiDAR data, regardless of their low pulse densities, can provide valuable microtopography information (<2 m DTM), depending on the composition and density of the understory vegetation. However, LiDAR data with an increased pulse energy and greater pulse densities has been shown to capture more accurate tree heights and improved tree counts compared to publicly available datasets. The cost of these mostly private collections may, however, be prohibitive for some project scales and restrict the ability to create high-quality individual tree segmentation maps. Lastly, while access to repeated LiDAR datasets may be limited in many geographic regions, opportunities for using PhoDAR or surface models of tree canopies in conjunction with LiDAR datasets may provide an option for assessing tree height changes over time at a stand or sub-stand resolution.

**Author Contributions:** Conceptualization, I.R.-G., L.C.R. and R.L.C.; methodology, I.R.-G., L.C.R. and R.L.C.; software, I.R.-G., L.C.R. and M.J.S.; validation, I.R.-G. and L.C.R.; formal analysis, I.R.-G. and L.C.R.; investigation, I.R.-G. and L.C.R.; resources, L.C.R., R.L.C. and J.S.B.; data curation, I.R.-G.; writing—original draft preparation, I.R.-G.; writing—review and editing, L.C.R., R.L.C., M.V.C. and J.S.B.; visualization, I.R.-G.; supervision, L.C.R., R.L.C. and J.S.B.; project administration, R.L.C.; funding acquisition, R.L.C. and J.S.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** Funding for this work was provided in part by the McIntire-Stennis Program of the National Institute of Food and Agriculture, U.S. Department of Agriculture Accession No. 1025518, and the National Science Foundation Center for Advanced Forestry Systems Award 1916552. The use of trade names in this paper does not imply endorsement by the associated agencies of the products named nor criticism of similar ones not mentioned.

**Data Availability Statement:** The raw data supporting the conclusions of this article will be made available by the authors upon request.

**Acknowledgments:** We appreciate the support from the Forest Productivity Cooperative. We gratefully acknowledge the support provided by the Department of Forestry and Environmental Resources at North Carolina State University, Virginia Department of Forestry, the Department of Forest Resources and Environmental Conservation at Virginia Polytechnic Institute an State University, the Departamento de Silvicultura, Facultad de Ciencias Forestales, Universidad de Concepcion, and the Federal University of Lavras.

**Conflicts of Interest:** The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Rachel L. Cook reports financial support was provided by National Science Foundation. Ivan Raigosa-Garcia reports financial support was provided by National Institute of Food and Agriculture. The additional authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References


1. Oswalt, S.N.; Smith, W.B.; Miles, P.D.; Pugh, S.A. *Forest Resources of the United States, 2017*; General Technical Report; US Department of Agriculture, Forest Service: Washington, DC, USA, 2019.
2. Kandare, K.; Ørka, H.O.; Dalponte, M.; Næsset, E.; Gobakken, T. Individual Tree Crown Approach for Predicting Site Index in Boreal Forests Using Airborne Laser Scanning and Hyperspectral Data. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *60*, 72–82. [CrossRef]
3. Tasissa, G.; Burkhart, H.E.; Amateis, R.L. Volume and Taper Equations for Thinned and Unthinned Loblolly Pine Trees in Cutover, Site-Prepared Plantations. *South. J. Appl. For.* **1997**, *21*, 146–152. [CrossRef]
4. Restrepo, H.I. Modeling Loblolly Pine (*Pinus taeda* L.) Stand Dynamics and Its Associated Financial Implications for Forestland Owners in the Southeastern U.S. Ph.D. Thesis, University of Georgia, Athens, GA, USA, 2019.
5. Burkhart, H.E.; Tomé, M. *Modeling Forest Trees and Stands*, 1st ed.; Springer: Dordrecht, The Netherlands, 2012; ISBN 978-90-481-3170-9.
6. Hart, H.M.J. *Stamtal en Dunning: Een Oriënteerend Onderzoek Naar de Beste Plantwijdte en Dunningswijze Voor den Djati*; Veenman & Zonen: Wageningen, The Netherlands, 1928.
7. Becking, J.H. Thinning Research in Forestry. *Neth. J. Agric. Sci.* **1953**, *1*, 122–129. [CrossRef]
8. Rennolls, K. Forest Height Growth Modelling. *For. Ecol. Manag.* **1978**, *71*, 217–225. [CrossRef]
9. Loetsch, F.; Zöhrer, F.; Khaller, K.Ė. *Inventory Data Collected by Terrestrial Measurements and Observations, Data Processing in Forest Inventory. The Sample Plot, Plotless Sampling and Regeneration Survey. List Sampling with Unequal Prbabilities and Planning, Performance and Field Checking of Forest Inventries*; BLV Verlagsgesellschaft: Munich, Germany, 1973; ISBN 3-405-10812-8.
10. Orrego, S.; Montes, C.; Restrepo, H.I.; Bullock, B.P.; Zapata, M. Modeling Height Growth for Teak Plantations in Colombia Using the Reducible Stochastic Differential Equation Approach. *J. For. Res.* **2021**, *32*, 1035–1045. [CrossRef]
11. Strand, L. Numerical Constructions of Site-Index Curves. *For. Sci.* **1964**, *10*, 410–414.
12. Silva, C.A.; Hudak, A.T.; Vierling, L.A.; Loudermilk, E.L.; O'Brien, J.J.; Hiers, J.K.; Jack, S.B.; Gonzalez-Benecke, C.; Lee, H.; Falkowski, M.J.; et al. Imputation of Individual Longleaf Pine (*Pinus palustris* Mill.) Tree Attributes from Field and LiDAR Data. *Can. J. Remote Sens.* **2016**, *42*, 554–573. [CrossRef]
13. Rahman, M.A.; Bakar, M.A.; Razak, K.; Rasib, A.; Kanniah, K.; Kadir, W.W.; Omar, H.; Faidi, A.; Kassim, A.; Latif, Z.A. Non-Destructive, Laser-Based Individual Tree Aboveground Biomass Estimation in a Tropical Rainforest. *Forests* **2017**, *8*, 86. [CrossRef]
14. Narin, O.G.; Vatandaslar, C.; Abdikan, S. Estimating Stand Top Height Using Freely Distributed ICESat-2 LiDAR Data: A Case Study from Multi-species Forests in Artvin. *Forestist* **2022**, *72*, 294–298. [CrossRef]
15. Sumnall, M.J.; Albaugh, T.J.; Carter, D.R.; Cook, R.L.; Hession, W.C.; Campoe, O.C.; Rubilar, R.A.; Wynne, R.H.; Thomas, V.A. Estimation of Individual Stem Volume and Diameter from Segmented UAV Laser Scanning Datasets in *Pinus taeda* L. Plantations. *Int. J. Remote Sens.* **2023**, *44*, 217–247. [CrossRef]
16. Rahman, M.Z.A.; Kadir, W.H.W.; Rasib, A.W.; Ariffin, A.; Razak, K.A. Integration of High Density Airborne LiDAR and High Spatial Resolution Image for Landcover Classification. In Proceedings of the 2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS, Melbourne, Australia, 21–26 July 2013; pp. 2927–2930.
17. Kellner, J.R.; Armston, J.; Birrer, M.; Cushman, K.C.; Duncanson, L.; Eck, C.; Falleger, C.; Imbach, B.; Král, K.; Krůček, M.; et al. New Opportunities for Forest Remote Sensing Through Ultra-High-Density Drone Lidar. *Surv. Geophys.* **2019**, *40*, 959–977. [CrossRef]
18. Corrao, M.V.; Sparks, A.M.; Smith, A.M.S. A Conventional Cruise and Felled-Tree Validation of Individual Tree Diameter, Height and Volume Derived from Airborne Laser Scanning Data of a Loblolly Pine (*P. taeda*) Stand in Eastern Texas. *Remote Sens.* **2022**, *14*, 2567. [CrossRef]
19. Gonzalez-Benecke, C.A.; Fernández, M.P.; Gayoso, J.; Pincheira, M.; Wightman, M.G. Using Tree Height, Crown Area and Stand-Level Parameters to Estimate Tree Diameter, Volume, and Biomass of Pinus radiata, Eucalyptus globulus and Eucalyptus nitens. *Forests* **2022**, *13*, 2043. [CrossRef]
20. Krůček, M.; Král, K.; Cushman, K.C.; Missarov, A.; Kellner, J.R. Supervised Segmentation of Ultra-High-Density Drone Lidar for Large-Area Mapping of Individual Trees. *Remote Sens.* **2020**, *12*, 3260. [CrossRef]
21. Brach, M.; Chormański, J. Terrestrial Laser Scanning (TLS) as a Detection Method of the Natural River Valley Microtopography–Case Study of the Upper Biebrza. *Ann. Wars. Univ. Life Sci. Land Reclam.* **2014**, *46*, 276. [CrossRef]
22. Panagiotidis, D.; Abdollahnejad, A.; Slavík, M. Assessment of Stem Volume on Plots Using Terrestrial Laser Scanner: A Precision Forestry Application. *Sensors* **2021**, *21*, 301. [CrossRef]
23. University of Washington, College of Forest Resources; University of Washington, College of Engineering; University of Washington, Institute of Forest Resources. Precision Forestry. In Proceedings of the First International Precision Forestry Cooperative Symposium, Seattle, WA, USA, 17–20 June 2001.
24. Dash, J.; Pont, D.; Brownlie, R.; Dunningham, A.; Watt, M.; Pearse, G. Remote sensing for precision forestry. *N. Z. J. For.* **2016**, *60*, 15–24.
25. NASA Langley Research Center. NASA Prediction of Worldwide Energy Resource (POWER). 2022. Available online: https://power.larc.nasa.gov/ (accessed on 22 January 2024).

26. Cook, R.; Fox, T.R.; Allen, H.L.; Cohrs, C.W.; Ribas-Costa, V.; Trlica, A.; Ricker, M.; Carter, D.R.; Rubilar, R.; Campoe, O.; et al. Forest Soil Classification for Intensive Pine Plantation Management: "Site Productivity Optimization for Trees" System. *For. Ecol. Manag.* **2024**, *556*, 121732. [CrossRef]
27. Soil Survey Staff, United States Department of Agriculture, Natural Resources Conservation Service Web Soil Survey. 2024. Available online: https://websoilsurvey.nrcs.usda.gov/ (accessed on 23 January 2024).
28. RIEGL. *RIEGL RiACQUIRE 2022*, 1.3.7; RIEGL: Horn, Austria, 2022. Available online: http://www.riegl.com/products/software-packages/riacquire(accessed on 15 December 2022).
29. RIEGL. *RIEGL RiProcess 2022*, 1.8; RIEGL: Horn, Austria, 2022. Available online: http://www.riegl.com/products/software-packages/riprocess/(accessed on 15 December 2022).
30. R Core Team. *A Language and Environment for Statistical Computing 2021*, 4.1.2; R Core Team: Vienna, Austria, 2021. Available online: https://www.r-project.org/(accessed on 22 January 2022).
31. Roussel, J.-R.; Auty, D.; Boissieu, F.D. *lidR 2022*, 4.0.0; R Core Team: Vienna, Austria, 2022. Available online: https://cran.r-project.org/web/packages/lidR/index.html(accessed on 12 March 2022).
32. Roussel, J.-R.; Goodbody, T.; Tompalski, P. *The lidR Package 2021*, 2022; R Core Team: Vienna, Austria, 2022. Available online: https://r-lidar.github.io/lidRbook/(accessed on 23 November 2023).
33. Zhang, K.; Chen, S.C.; Whitman, D.; Shyu, M.L.; Yan, J.; Zhang, C. A Progressive Morphological Filter for Removing Nonground Measurements from Airborne LIDAR Data. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 872–882. [CrossRef]
34. De Reu, J.; Bourgeois, J.; Bats, M.; Zwertvaegher, A.; Gelorini, V.; De Smedt, P.; Chu, W.; Antrop, M.; De Maeyer, P.; Finke, P.; et al. Application of the topographic position index to heterogeneous landscapes. *Geomorphology* **2013**, *186*, 39–49. [CrossRef]
35. Evans, J.S.; Murphy, M.A. spatialEco. Spatial Analysis and Modeling with Ecological Data. 2021. Available online: https://cran.r-project.org/web/packages/spatialEco/index.html (accessed on 5 June 2022).
36. Hollister, J.W.; Hijmans, R.J.; Mattiuzzi, M.; Simpson, G.L.; Mosher, S.; Dyck, A.R.; Little, D.; Michonneau, F. exactextractr: Fast Extraction of Grids and Polygons from Raster Datasets. 2022. Available online: https://cran.r-project.org/web/packages/exactextractr/index.html (accessed on 5 June 2022).
37. Lema-Tapias, A. *Elementos Teórico-Prácticos Sobre Inventarios Forestales (Estadística y Planeación)*; Universidad Nacional de Colombia, Sede Medellín: Medellín, Colombia, 2003; ISBN 958-33-5661-1.
38. Dunnett, C.W. A Multiple Comparison Procedure for Comparing Several Treatments with a Control. *J. Am. Stat. Assoc.* **1955**, *50*, 1096–1121. [CrossRef]
39. Dieguez-Aranda, U.; Burkhart, H.E.; Amateis, R.L. Dynamic Site Model for Loblolly Pine (*Pinus taeda* L.) Plantations in the United States. *For. Sci.* **2006**, *52*, 262–272. [CrossRef]
40. Wilcoxon, F. Individual Comparisons by Ranking Methods. *Biom. Bull.* **1945**, *1*, 80–83. [CrossRef]
41. Sharma, M.; Amateis, R.L.; Burkhart, H.E. Top Height Definition and Its Effect on Site Index Determination in Thinned and Unthinned Loblolly Pine Plantations. *For. Ecol. Manag.* **2002**, *168*, 163. [CrossRef]
42. Ritchie, M.; Zhang, J.; Hamilton, T. Effects of Stand Density on Top Height Estimation for Ponderosa Pine. *West. J. Appl. For.* **2012**, *27*, 18–24. [CrossRef]
43. García, O.; Batho, A. Top Height Estimation in Lodgepole Pine Sample Plots. *West. J. Appl. For.* **2005**, *20*, 64–68. [CrossRef]
44. Moser, K.; Ahn, C.; Noe, G. Characterization of microtopography and its influence on vegetation patterns in created wetlands. *Wetlands* **2007**, *27*, 1081–1097. [CrossRef]
45. Rhoton, F.E.; Lindbo, D.L. A Soil Depth Approach To Soil Quality Assessment. *J. Soil Water Conserv.* **1997**, *52*, 66–72.
46. Koirala, A.; Montes, C.R.; Bullock, B.P. Modeling Dominant Height Using Stand and Water Balance Variables for Loblolly Pine in the Western Gulf, US. *For. Ecol. Manag.* **2021**, *479*, 118610. [CrossRef]
47. Zhao, Y.; Krzic, M.; Bulmer, C.E.; Schmidt, M.G.; Simard, S.W. Relative Bulk Density as A Measure of Compaction and Its Influence on Tree Height. *Can. J. For. Res.* **2010**, *40*, 1724–1735. [CrossRef]
48. Socha, J.; Tymińska-Czabańska, L.; Bronisz, K.; Zięba, S.; Hawryło, P. Regional Height Growth Models for Scots Pine in Poland. *Sci. Rep.* **2021**, *11*, 10330. [CrossRef] [PubMed]
49. Bravo-Oviedo, A.; Tomé, M.; Bravo, F.; Montero, G.; del Río, M. Dominant Height Growth Equations Including Site Attributes in the Generalized Algebraic Difference Approach. *Can. J. For. Res.* **2008**, *38*, 2348–2358. [CrossRef]
50. Seifert, C.L.; Cox, R.T.; Forman, S.L.; Foti, T.L.; Wasklewicz, T.A.; McColgan, A.T. Relict Nebkhas (Pimple Mounds) Record Prolonged Late Holocene Drought in The Forested Region of South-Central United States. *Quat. Res.* **2009**, *71*, 329–339. [CrossRef]
51. Borders, B.E.; Will, R.E.; Markewitz, D.; Clark, A.; Hendrick, R.; Teskey, R.O.; Zhang, Y. Effect of Complete Competition Control and Annual Fertilization on Stem Growth and Canopy Relations for A Chronosequence of Loblolly Pine Plantations in the Lower Coastal Plain of Georgia. *For. Ecol. Manag.* **2004**, *192*, 21–37. [CrossRef]
52. Sanchez, F.G.; Scott, D.A.; Ludovici, K.H. Negligible Effects of Severe Organic Matter Removal and Soil Compaction on Loblolly Pine Growth over 10 Years. *For. Ecol. Manag.* **2006**, *227*, 145. [CrossRef]
53. Marks, C.O.; Yellen, B.C.; Wood, S.A.; Martin, E.H.; Nislow, K.H. Variation in Tree Growth along Soil Formation and Microtopographic Gradients in Riparian Forests. *Wetlands* **2020**, *40*, 1909–1922. [CrossRef]
54. Zhao, H.; Wu, J.; Wang, A.; Guan, D.; Liu, Y. Microtopography Mediates the Climate–Growth Relationship and Growth Resilience to Drought of Pinus Tabulaeformis Plantation in the Hilly Site. *Front. Plant Sci.* **2022**, *13*, 1060011. [CrossRef] [PubMed]

55. Haywood, J.D.; Tiarks, A.E.; Shoulders, E. Loblolly and Slash Pine Height and Diameter are Related to Soil Drainage in Winter on Poorly Drained Silt Loams. *New For.* **1990**, *4*, 81–96. [CrossRef]
56. McNab, W.H. Effects of Landform on Site Index for Two Mesophytic Tree Species in the Appalachian Mountains of North Carolina, USA. *Int. J. For. Res.* **1991**, *2010*, 298674. [CrossRef]
57. Woollons, R.C.; Haywood, A.; McNickle, D.C. Modeling Internode Length and Branch Characteristics for Pinus Radiata in New Zealand. *For. Ecol. Manag.* **2002**, *160*, 243–261. [CrossRef]
58. Monserud, R.A.; Huang, S.; Yang, Y. Biomass and Biomass Change in Lodgepole Pine Stands in Alberta. *Tree Physiol.* **2006**, *26*, 819–831. [CrossRef] [PubMed]
59. Crookston, N.L.; Rehfeldt, G.E.; Dixon, G.E.; Weiskittel, A.R. Addressing Climate Change in the Forest Vegetation Simulator to Assess Impacts on Landscape Forest Dynamics. *For. Ecol. Manag.* **2010**, *260*, 1198–1211. [CrossRef]
60. Weiskittel, A.R.; Hann, D.W.; Kershaw, J.A.; Vanclay, J.K. *Forest Growth and Yield Modeling*; Wiley & Sons, Inc.: Hoboken, NJ, USA, 2011; ISBN 1-119-99852-2.
61. Liu, J.; Burkhart, H.E. Spatial Characteristics of Diameter and Total Height in Juvenile Loblolly Pine (*Pinus taeda* L.) Plantations. *For. Sci.* **1994**, *40*, 774–786. [CrossRef]
62. Kelting, D.L.; Burger, J.A.; Patterson, S.C.; Corporation, W. Early Loblolly Pine Growth Response to Changes in Soil Environment. *N. Z. J. For. Sci.* **2000**, *30*, 206–224.
63. Kelting, D.L.; Burger, J.A.; Patterson, S.C.; Aust, W.M.; Miwa, M.; Trettin, C.C. Soil Quality Assessment in Domesticated Forests–a Southern Pine Example. *For. Ecol. Manag.* **1999**, *122*, 167–185. [CrossRef]
64. Pangle, R.E.; Seiler, J. Influence of Seedling Roots, Environmental Factors and Soil Characteristics on Soil $CO_2$ Efflux Rates in a 2-Year-Old Loblolly Pine (*Pinus taeda* L.) Plantation in the Virginia Piedmont. *Environ. Pollut.* **2002**, *116* (Suppl. S1), S85–S96. [CrossRef] [PubMed]
65. Lorio, P.L.; Howe, V.K.; Martin, C.N. Loblolly Pine Rooting Varies with Microrelief on Wet Sites. *Ecology* **1972**, *53*, 1134–1140. [CrossRef]
66. Mieza, M.S.; Cravero, W.R.; Kovac, F.D.; Bargiano, P.G. Delineation of Site-Specific Management Units for Operational Applications Using the Topographic Position Index in La Pampa, Argentina. *Comput. Electron. Agric.* **2016**, *127*, 158–167. [CrossRef]
67. Cohrs, C.W. Optimizing Pine Plantation Management Via Geospatial Data Science. Ph.D. Thesis, North Carolina State University, Raleigh, NC, USA, 2022.


**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.

forests

MDPI

*Article*

# Registration of TLS and ULS Point Cloud Data in Natural Forest Based on Similar Distance Search

**Yuncheng Deng [1,2,3], Jinliang Wang [1,2,3,4,*], Pinliang Dong [5], Qianwei Liu [6], Weifeng Ma [1,2,3,7], Jianpeng Zhang [1,2,3], Guankun Su [8] and Jie Li [1,2,3]**

1 Faculty of Geography, Yunnan Normal University, 768 Juxian Street, Chenggong District, Kunming 650500, China; dengyckk@user.ynnu.edu.cn (Y.D.); maweifeng@ynnu.edu.cn (W.M.); jpzhang@ynnu.edu.cn (J.Z.); jieli@ynnu.edu.cn (J.L.)
2 Key Laboratory of Resources and Environmental Remote Sensing for Universities in Yunnan, Kunming 650500, China
3 Center for Geospatial Information Engineering and Technology of Yunnan Province, Kunming 650500, China
4 Southwest United Graduate School, Kunming 650092, China
5 Department of Geography and the Environment, University of North Texas, 1155 Union Circle #305279, Denton, TX 76203, USA; pinliang.dong@unt.edu
6 Yunnan Institute of Military-Civilian Integration Development, Kunming 650034, China; 13759400628@163.com
7 Power China Kunming Engineering Corporation Limited, Kunming 650051, China
8 New Coordinates Technology Co., Ltd., Kunming 650100, China; dyc_kk@163.com
* Correspondence: jlwang@ynnu.edu.cn

**Abstract:** Multiplatform fusion point clouds can effectively compensate for the disadvantages of individual platform point clouds in forest parameter extraction, maximizing the potential of LiDAR technology. However, existing registration algorithms often suffer from insufficient feature extraction and limited registration accuracy. To address these issues, we propose a ULS (Unmanned Aerial Vehicle Laser Scanning)-TLS (Terrestrial Laser Scanning) point cloud data registration method based on Similar Distance Search (SDS). This method enhances coarse registration by accurately retrieving points with similar features, leading to high overlap in the rough registration stage and further improving fine registration precision. (1) The proposed method was tested on four natural forest plots, including *Pinus densata* Mast., *Pinus yunnanensis* Franch., *Pices asperata* Mast., *Abies fabri* (Mast.) Craib, and demonstrated high registration accuracy. Both coarse and fine registration achieved superior results, significantly outperforming existing algorithms, with notable improvements over the TR algorithm. (2) In addition, the study evaluated the accuracy of individual tree parameter extraction from fusion point clouds versus single-platform point clouds. While ULS point clouds performed slightly better in some metrics, the fused point clouds offered more consistent and reliable results across varying conditions. Overall, the proposed SDS method and the resulting fusion point clouds provide strong technical support for efficient and accurate forest resource management, with significant scientific implications.

**Keywords:** ULS; TLS; registration; natural forest; similar distance search

## 1. Introduction

Forests play a dominant role in the terrestrial ecosystem. They are an important part of the biosphere and have a close material exchange with the surrounding ecosystem. Forests also play a key role in maintaining the ecosystem balance, known as the "lungs of the earth" [1,2]. LiDAR (Light detection and ranging) can gather information on the intensity, elevation, and coordinates of objects on land, making it valuable for detecting forest structures and managing forest resources effectively [3,4].

Airborne Laser Scanning (ALS) rapidly captures extensive forest point clouds, aiding multi-scale forest monitoring. The emergence of Unmanned Aerial Vehicles (UAVs), noted

for their agility and efficiency, has seen them being used extensively for forest structure analysis [5–7]. Due to the "top-down" scanning approach of ALS and the limitations in penetration ability, the point cloud density is relatively low. This is especially true in natural forest areas, where ALS cannot capture complete information on tree trunks, and more of the point cloud data originate from the forest canopy. For example, parameters such as the diameter at breast height (DBH) and the subbranches of tall trees cannot be directly extracted from ALS point cloud data and need to be estimated by building models with other parameters [8,9]. TLS compensates well for this shortcoming. Mounted on tripods or other carriers at ground level, TLS operates with a "bottom-up" scanning approach to acquire point cloud data, allowing it to capture comprehensive information on the lower trunk and understory shrub structures beneath the canopy. Of course, the scanning approach of TLS also results in the loss of some canopy information [10–12]. It can be seen that the data of ALS and TLS platforms have inherent advantages and disadvantages. The fusion of ULS and TLS will help to compensate for the respective disadvantages of the two kinds of data, obtain more perfect forest point cloud data, accurately extract forest structure parameters, and provide support for improving the efficiency and accuracy of forest resource investigations [13].

To solve the problem of point cloud registration of natural forestland, many scholars have proposed registration methods according to the characteristics of natural forest sample land. Guan et al. [14] utilized point cloud individual tree segmentation results to construct Triangulated Irregular Networks (TIN) for matching and selecting homonymous tree points, and combined this with the ICP (Iterative Closet Point) algorithm [15] to achieve point cloud registration. However, the registration results were significantly affected by the individual tree segmentation outcomes. Paris et al. [16] segmented individual trees from ALS point cloud data, extracted the boundaries of individual tree crowns, and overlaid them with TLS data for registration. However, when stand density is high, the uncertainty of canopy segmentation increases, reducing the applicability of the registration method. Dai et al. [17] realized key point extraction of ULS and TLS platform point cloud data using the mean shift algorithm, combined with the ICP algorithm and CPD (Coherent Point Drift) algorithm [18] to complete registration. The average accuracy was 0.05 m–0.08 m. Polewski et al. [19] detected the position of an individual tree using ULS and TLS point clouds, took the vertical and horizontal distance of an individual tree as the constraint condition, and completed the registration through image matching. However, using the DBH center as the position of an individual tree in the TLS point cloud created certain deviations, and it was not applicable to some specific tree species. Polewski et al. [20] later achieved registration between ALS and backpack LiDAR point clouds based on tree distances. However, they did not consider that trees do not grow perfectly vertically, thus the applicability of this method in complex forest scenarios remains to be further verified. Liu et al. [21] used the rasterized high point of the CHM tree, took the tree height value and the distance between an individual tree and the center of the sample as constraint conditions to find the same name point pairs, and combined the nearest-neighbor iterative algorithm to complete the point cloud registration. The registration accuracy was between 0.18 m and 0.69 m. Previous studies provide a reference for forest area point cloud registration research; however, there are still areas for further improvement: (1) the quantity of point cloud data is large, and the overall registration time cost based on point clouds is too high; (2) some existing registration algorithms rely on good initial positions and high initial overlaps, which cannot be guaranteed in practical applications; (3) registration methods based on the same name feature points are difficult to capture in natural forests, and it is easy to fall into local optima; and (4) the positions of individual trees are rigid characteristic points worth using, but due to the different conditions of actual natural forests and the differences in specific tree species, the applicability of various methods is still worth discussing, and the registration accuracy still has room for improvement.

In summary, the natural forest area of Shangri-La City in Yunnan Province of China was the research subject to carry out research on the registration method of ULS and TLS

point clouds without control points in natural forest areas. This paper proposes a point cloud registration method that uses the tree height points of individual trees as proxies for their positions, searches for similar distances between trees, and achieves ULS-TLS cloud fusion in natural forest areas.

## 2. Data and Method

### *2.1. Overview of the Study Area*

Shangri-La City is located in the northeast of the Diqing Tibetan Autonomous Prefecture, Yunnan Province, with geographic coordinates of 26°52′~28°52′ N, 99°22′~100°19′ E, and an average elevation of about 3459 m [22] (Figure 1). The region experiences a small annual temperature variation, distinct dry and wet seasons, and lacks a clear summer, with an average annual temperature of 5.5 °C. Intense solar radiation leads to sharp temperature increases during the day and significant cooling at night, with the daily temperature range in the dry season reaching up to 30 °C [23].

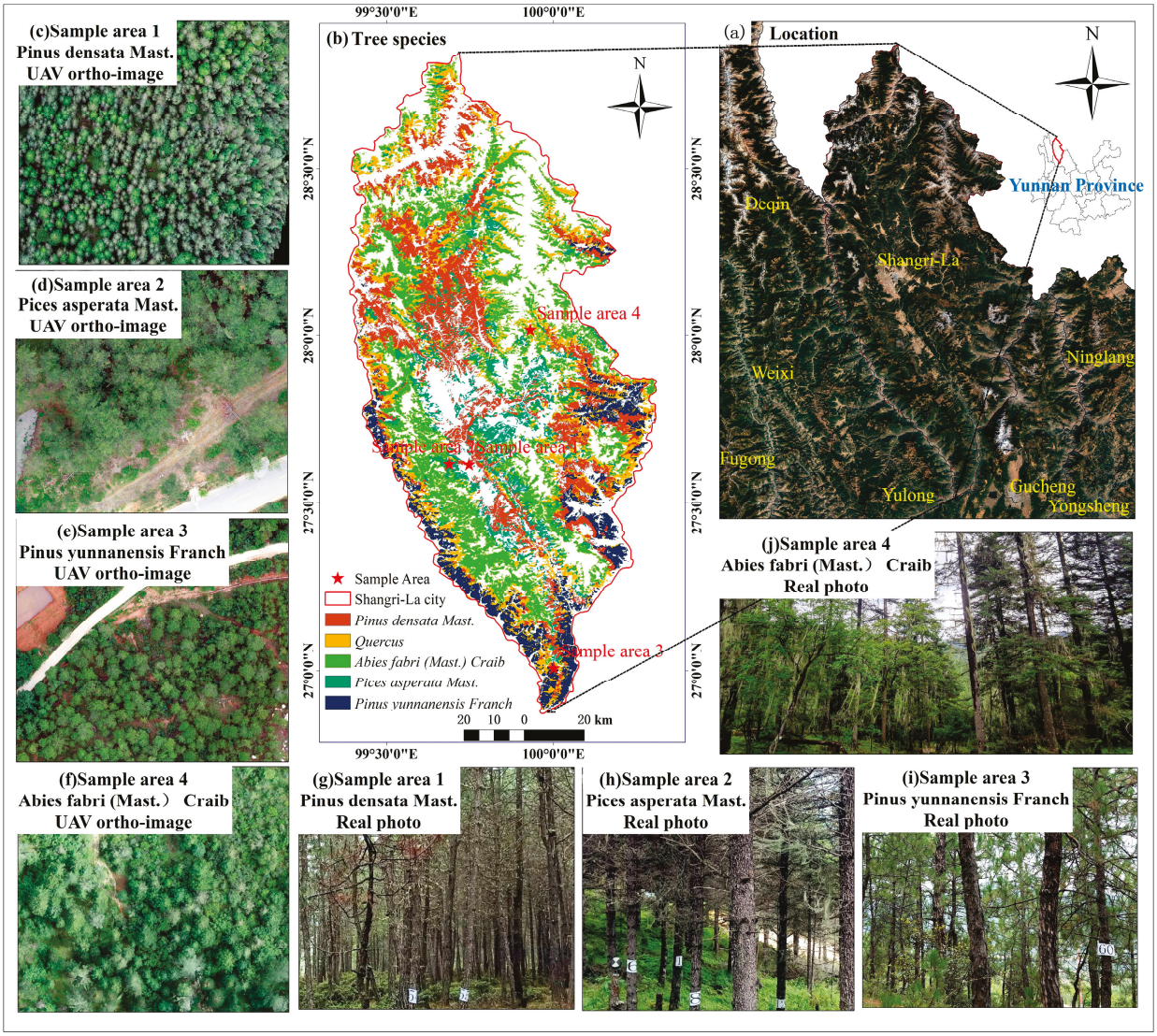



**Figure 1.** Study area location.

Shangri-La is located in the transition zone from the subtropical evergreen broad-leaved forest vegetation area of Yunnan Province to the alpine vegetation area of the Qinghai-Tibet Plateau. It is rich in forest resources, with a forest coverage rate of 78.61%. The city contains *Pinus densata* Mast., *Pinus yunnanensis* Franch., *Pices asperata* Mast.,

*Abies fabri* (Mast.) Craib, and *Quercus aquifolioides* Rehd. & E. H. Wilson [9]. Among them, the first three kinds of trees occupy 74.27% of the land area in the arboreal forest.

This study selected four natural forest sample plots featuring four typical tree species in Shangri-La City to gather TLS and ULS, and measured data as research subjects. It conducted multi-platform point cloud registration research in these natural forest areas. Sample plot information appears in Table 1.

**Table 1.** Sample area information.

| Sample Area | Coordinates | Tree Species | Average Tree Height/m | Average DBH/cm | Elevation/m | Slope/Aspect |
|---|---|---|---|---|---|---|
| 1 | 99°44′ 51.85″ E 27°37′0.90″ N | *Pinus densata* Mast. | 10.88 | 14.87 | 3267.10 | 1°/Southwest 195° |
| 2 | 99°41′16.35″ E 27°37′5.15″ N | *Pices asperata* Mast. | 13.11 | 18.06 | 3612.70 | 2°/Northwest 335° |
| 3 | 100°04′48.43″ E 27°27′54.34″ N | *Pinus yunnanensis* Franch. | 9.31 | 18.40 | 2415.08 | 4°/Northwest 340° |
| 4 | 99°55′50.63″ E 28°00′57.93″ N | *Abies fabri* (Mast.) Craib | 16.36 | 28.63 | 3686.00 | 5°/Southeast 148° |

## 2.2. Data

### 2.2.1. TLS Point Cloud

The study utilized the Leica P40 3D laser scanner as the TLS point cloud data acquisition tool, with a collection system including a tripod, laser scanner, and targets of control points. Complete natural forest sample plot point cloud data were obtained through multi-station stitching. The specific parameters of the instrument are shown in Table 2.

**Table 2.** Leica P40 3D Laser scanner parameter.

| Index | Specification |
|---|---|
| Dimensions | 238 mm × 358 mm × 395 mm |
| Distance accuracy | 1.2 mm + 10 pmm |
| Angle accuracy | Horizontal: 8″; Vertical: 8″ |
| Point accuracy | 3 mm @ 50 m; 6 mm @ 100 m |
| Target acquisition accuracy | 2 mm @ 50 m |
| Laser angle | <0.23 mrad |
| Scan range & Reflectivity | Min distance: 0.4 m; Max range & Reflectivity: 120 m (8%), 180 m (18%), 270 m (34%) |
| Scan rate | 1,000,000 points/s |
| Noise range | 0.4 mm rms @ 10 m; 0.5 mm rms @ 50 m |
| Field angle | Horizontal: 360°; Vertical: 270° |
| Laser line | 1550 nm (Visible); 658 nm (Invisible) |
| Front window laser spot diameter | <3.5 mm |

### 2.2.2. ULS Point Cloud

The acquisition system consists of a global navigation satellite system (GNSS), an inertial measurement unit (IMU), and a laser sensor. The study employed a UAV point cloud data collection system composed of a DJI Matrice M600 Pro drone equipped with a VUX-1UAV laser scanner developed and produced by the Austrian company RIEGL (Horn, Austria), as seen in Figure 2. The specific parameters of the loading platform and sensor are shown in Tables 3 and 4. The flight parameters are shown in Table 5.

**Table 3.** DJI Longitude and Latitude M600 Pro UAV platform parameters.

| Index | Specification [1] |
|---|---|
| Dimensions | 1668 mm × 1518 mm × 727 mm |
| Recommended maximum take-off weight | 15.5 kg |
| Hover accuracy (GPS) | Vertical: ±0.5 m, Horizontal: ±1.5 m |
| Maximum angular velocity of rotation | Pitch axis: 300°/s, Course axis: 150°/s |
| Maximum pitch angle | 25° |
| Maximum ascent speed | 5 m/s |
| Maximum descent speed | 3 m/s |
| Maximum wind speed | 8 m/s |
| Maximum horizontal flight speed | 65 km/h (No wind) |
| Maximum take-off altitude | Airscrew blade 2170/2195: 2500 m/4500 m |
| Hover time | Without load: 38 min; Load 5.5 kg: 18 min |

[1] www.dji.com accessed on 14 June 2023.

**Table 4.** VUX-1UAV LiDAR sensor parameter.

| Index | Specification [1] |
|---|---|
| Field angle | 360° |
| Farthest range | 1050 m |
| Maximum laser emission frequency | 520,000 points/s, 550 kHz |
| Repeated accuracy | 5 mm |
| Maximum echo | 15 |
| Weight | 3.5 kg |

[1] www.ilidar.com accessed on 14 June 2023.

**Table 5.** The flight parameters.

| Sample Area | Flight Time | Number of Flights | Flight Height (m) | Flight Speed (m/s) | Longitudinal Overlap (%) | Lateral Overlap (%) | Scanner Transmitting Frequency (KHz) |
|---|---|---|---|---|---|---|---|
| 1 | 17:43 18 August 2021 | 1 | 60 | 5.3 | 90 | 60 | 400 |
| 2 | 14:51 18 August 2021 | 1 | 80 | 5.2 | 90 | 60 | 400 |
| 3 | 09:56 19 August 2021 | 1 | 80 | 5.2 | 90 | 60 | 400 |
| 4 | 11:04 19 August 2021 | 1 | 60 | 5.4 | 90 | 60 | 400 |

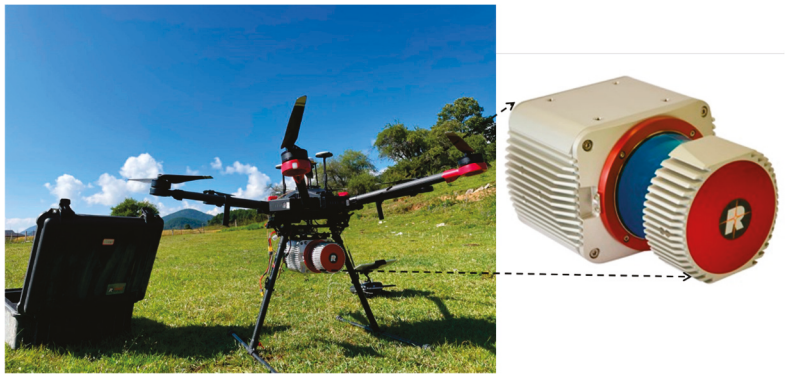


**Figure 2.** ULS point cloud data acquisition system.

2.2.3. The Measured Data of the Plot

In order to further study the influence of different platform data on the extraction of single tree height, the study also artificially measured the tree height of each single tree in the sample field by using the 'Berulay' height meter. Each single tree was measured three times, and the average value was taken as the final single tree height.

*2.3. Method*

To solve the above registration problem, this paper takes advantage of the constant position between individual trees in natural forest areas, uses the point cloud data of different platforms to detect the tree height points of individual trees as approximate individual tree locations, and extracts the same name point pairs of TLS point clouds and ULS point clouds through distance sorting and similar distance searches. The coarse registration of the ULS point cloud and TLS point cloud is completed based on the SVD (Singular Value Decomposition) method of the same name point pair, and the final registration of the TLS and ULS forest area point clouds is completed by combining the nearest point fine registration method. The research technical route shown in Figure 3 includes three parts: data preprocessing, point cloud rough registration, and point cloud fine registration.

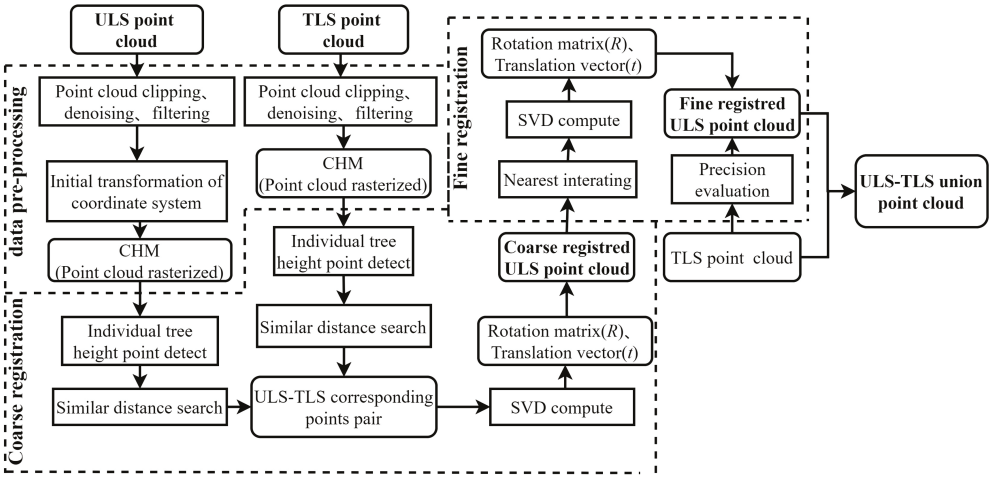



**Figure 3.** ULS-TLS point cloud registration flowchart.

2.3.1. Data Preprocessing

The main preprocessing steps included data cropping and coordinate transformation. To facilitate subsequent data registration and mitigate computational inefficiencies due to computer hardware during the data processing phase, the study trimmed ULS data into a circular plot with a radius of 25 m and TLS data into a circular plot with a radius of 20 m, based on the center coordinates of the sample plots. The center of the sample area was determined using GPS (Global Positioning System). For the ULS data, we calculated the distance from each point in the dataset to the sample area center using the distance formula. Points within a specified radius $R$ from the center were then selected for further analysis. We also recorded the actual center point's projected coordinates using GPS, which facilitated the integration with ULS data. By utilizing GPS, the TLS data could be transformed into projected coordinates, which facilitated a more accurate overlap of the datasets. This approach helped ensure that the datasets aligned correctly, even in the absence of coarse registration. If the datasets do not achieve 100% overlap, we recommend using registration algorithms to resolve the issue. This is a key aspect of our research; the proposed algorithm is specifically designed to address scenarios where the datasets do not fully overlap. Through this algorithm, precise registration of the data is ultimately achieved, ensuring accurate alignment despite initial discrepancies.

Airborne laser scanning data consist of point clouds with projection coordinates. Terrestrial laser scanning data use the position of the scanner at the first measurement

station as the center of the sample plot (0.0), and the scanned point cloud data contain local coordinate information [24]. Therefore, the two kinds of data have different initial coordinate systems. To merge and register the point cloud data of the two platforms, the two kinds of data should be converted to a unified coordinate system. Obviously, more parameters are needed to convert the TLS coordinate system to the ULS point cloud coordinate system [16]. The study first attempted to register the ULS data on the TLS coordinate system using the proposed method, and also explored registering the TLS data on the ULS coordinate system. Both approaches were tested, and the effectiveness of the proposed method was validated.

The Cloth Simulation Filter (CSF) method [25] extracted the ground points. We subtracted the nonground point cloud from the central value of the ground point cloud and updated the converted ULS point cloud coordinates. The specific formula appears in Formula (1).

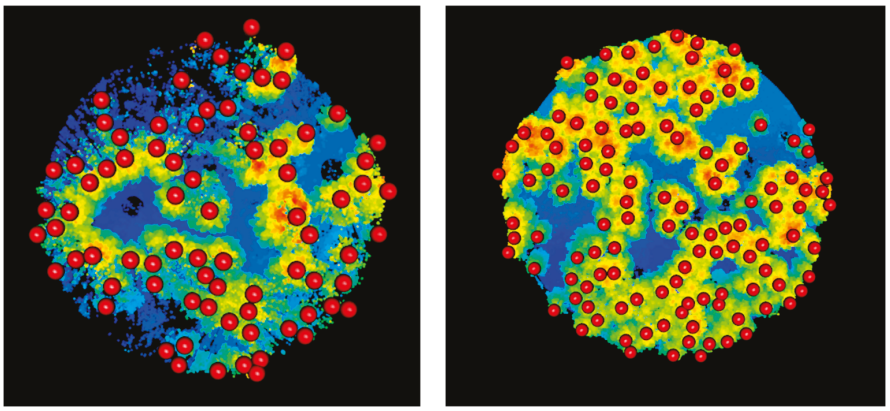
$$ULS(x,y,z)_i = ULS_{nonground}(x,y,z)_i - mean\left(ULS_{ground}(x,y,z)\right),\ i = 1,2,\ldots,n \quad (1)$$

where $ULS(x,y,z)_i$ represents the nonground point of the ULS point cloud after coordinate conversion, $ULS_{nonground}(x,y,z)_i$ represents the ith nonground point of the ULS point cloud before coordinate conversion, and $mean\left(ULS_{ground}(x,y,z)\right)$ represents the ground center point of the ULS point cloud before coordinate conversion. At this point, the ULS point cloud and the TLS point cloud have been converted to the same coordinate system, and the data range of the two platforms has a certain overlap.

#### 2.3.2. Coarse Registration

Point Cloud Rasterization and Tree Height Point Detection

The point cloud is transformed into a two-dimensional raster image, and the individual tree position is further extracted by using the sliding window to detect the local maximum value of the two-dimensional image. The main steps to convert point clouds to raster images are as follows: (1) select the raster image resolution, (2) divide the grid and establish the index relationship between the grid and point cloud, (3) determine the interpolation method and fill the attribute values into the grid, (4) flip the grid, and (5) add coordinates for each grid crossing point [21]. We use the R language (Version: x64 4.0.4) to generate a raster image with a resolution of 0.1 m by means of maximum interpolation. Because the nonground point cloud is selected and the z value of the point cloud is filled into the grid as an attribute value, the maximum value of the raster image is detected through a $3 \times 3$ sliding window, and the detected point is the tree high point $TH_{ULS}$, $TH_{TLS}$, that is, the individual tree position point. Figure 4 shows the detected highest point of the tree.



(**a**) (**b**)

**Figure 4.** Tree height (TH) point detection: (**a**) TLS tree height ($TH_{TLS}$) point detection, (**b**) ULS tree height ($TH_{ULS}$) point detection.

Similar Distance Search

Figure 4 shows that the tree heights detected by TLS and ULS are not the same due to the different scanning methods of other platforms and different field conditions at the sample sites. However, there are still some tree highs with local similarities to the surrounding tree highs. The tree height represents the location of an individual tree, and the same individual tree should be approximately the same distance from the center of the plot in both the ULS and the TLS.

The distance between the tree and its neighboring trees and the distance between the neighboring trees and the center of the plot should be approximately the same both in the ULS and in the TLS (Formula (2)).

$$\begin{cases} D_{TLS_{center}} = \sqrt{\left({x_{TH_{TLS_i}}}^2 + {y_{TH_{TLS_i}}}^2\right)} \approx D_{ULS_{center}} = \sqrt{\left({x_{TH_{ULS_i}}}^2 + {y_{TH_{ULS_i}}}^2\right)} \\ D_{THTLS} = \sqrt{\left(x_{TH_{TLSi}} - y_{TH_{TLSn}}\right)^2 + \left(x_{TH_{TLSi}} - y_{TH_{TLSn}}\right)^2} \\ \approx D_{THULS} = \sqrt{\left(x_{TH_{ULSi}} - y_{TH_{ULSn}}\right)^2 + \left(x_{TH_{ULSi}} - y_{TH_{ULSn}}\right)^2} \end{cases} \quad (2)$$

Based on this idea, a point pair extraction method with the same name based on a similar distance search (SDS) is proposed. The main process is shown in Table 6:

**Table 6.** Coarse registration (SDS) procedure code.

| | |
|---|---|
| | Input: $TH_{TLS}$, $TH_{ULS}$, $D_{ULS_{center}}$ , $D_{TLS_{center}}$ |
| Step 1 | For each tree height $TH_{TLSi}$ in $D_{TLS_{center}}$<br>nearest_ULS = find_nearest($TH_{ULS}$, $D_{TLS_{center}i}$)<br>If nearest_ULS exists, then selected_pair = ($TH_{TLSi}$, *nearest_ULS*)<br>add_to_selected_pairs(*selected_pair*)<br>end if<br>end for |
| Step 2 | For each selected_pair in selected_pairs, enter<br>nearest_neighbor_TLS = find_nearest_neighbor(*selected_pair*, $TH_{TLS}$) nearest_neighbor_ULS = find_nearest_neighbor(*selected_pair*, $TH_{ULS}$)<br>If nearest_neighbor_TLS and nearest_neighbor_ULS are similar, then high_point_distance = calculate_distance_to_center(*nearest_neighbor_TLS, nearest_neighbor_ULS*)<br>If high_point_distance is similar, then<br>match_point_pair = (*nearest_neighbor_TLS, nearest_neighbor_ULS*)<br>add_to_match_point_set(*match_point_pair*)<br>Or else<br>go back to Step 3<br>end if<br>Or else<br>go back to Step 3<br>end if<br>end for |
| Step 3 | Output match_point_set |

Coarse Registration of TLS and ULS

Point cloud registration rotates and translates the ULS to achieve maximum coincidence with the reference cloud (TLS). Based on the set of point pairs with the same name extracted in "SDS", the paper analyzes the rotation matrix $R$ and the translation vector $t$ by SVD, thus achieving the coarse registration of the ULS point cloud and TLS point cloud in the forest area. The principle of SVD is as follows:

Assume $Q = \{q1, q2, \ldots, qn\}$ and $P = \{p1, p2, \ldots p3\}$ are the corresponding point sets in two groups of D-dimensional space, corresponding to the high point set of the same name tree in the research, Q represents the TLS tree high point set, and P represents the ULS tree

high point set. Calculating the rigid transformation information between these two point sets is a problem using the least square method [26], which can be described as Formula (3).

$$(R,t) = \underset{R,t}{argmin} \sum_{i=1}^{n} w_i \left\| (Rp_i + t) - q_i \right\|^2 \tag{3}$$

where $R$ represents the rotation matrix, $t$ represents the translation vector, $w_i$ is the weight of the $i$th point in the point set, and $w_i > 0$.

The steps to solve the optimal rotation matrix and translation vector include the following:

(1) Calculate the weighted centroid of the two sets of points

$$\overline{p} = \frac{\sum_{i=1}^{n} w_i p_i}{\sum_{i=1}^{n} w_i},\ \overline{q} = \frac{\sum_{i=1}^{n} w_i q_i}{\sum_{i=1}^{n} w_i} \tag{4}$$

(2) Calculate the center vector

$$x_i := p_i - \overline{p},\ y_i := q_i - \overline{q},\ i = 1, 2, \ldots, n \tag{5}$$

(3) Calculate the $d \times d$ covariance matrix

$$S = XWY^T \tag{6}$$

where $X$ and $Y$ are $d \times n$ matrices with $x_i$ columns and $y_i$ columns, $W = diag(w_1, w_2, \ldots, w_n)$

(4) Calculate the singular value decomposition $S = U\Sigma V^T$ and obtain the rotation matrix $R$

$$R = V \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & det(VU^T) \end{pmatrix} U^T \tag{7}$$

(5) Calculate the optimal translation vector $t$

$$t = \overline{q} - R\overline{p} \tag{8}$$

The calculated rotation matrix $R$ and translation vector $t$ are substituted into the on-board point set to be configured (ULS point cloud dataset) to calculate the registered on-board point cloud dataset ($ULS_{tr}$).

$$ULS_{tr} = p_i \times R + t \tag{9}$$

The initial registration of the ULS point cloud and TLS point cloud is coarse registration by detecting the highest point of the tree instead of the individual tree position and then searching for similar distances to find the same feature points. Figure 5 (red points represent ULS tree heights and green points represent TLS tree heights) shows that the tree heights before rough registration are disorganized, and the same list of tree points cannot be identified with the naked eye. After rough registration, the ULS point cloud tree heights and TLS point cloud tree heights in the four sample sites selected by the research show similar spatial distribution patterns. The tree heights of the same name for the data of the two platforms can be clearly identified (Figure 5). However, not all ULS tree heights have corresponding TLS tree heights because the top-down scanning mode of ULS and the bottom-up scanning mode of TLS, resulting in tree height detection, cannot completely detect the corresponding tree height. The range of the ULS point cloud in the study is larger than that of the TLS point cloud, so there is no matching TLS point cloud at the edge tree height of the ULS point cloud.

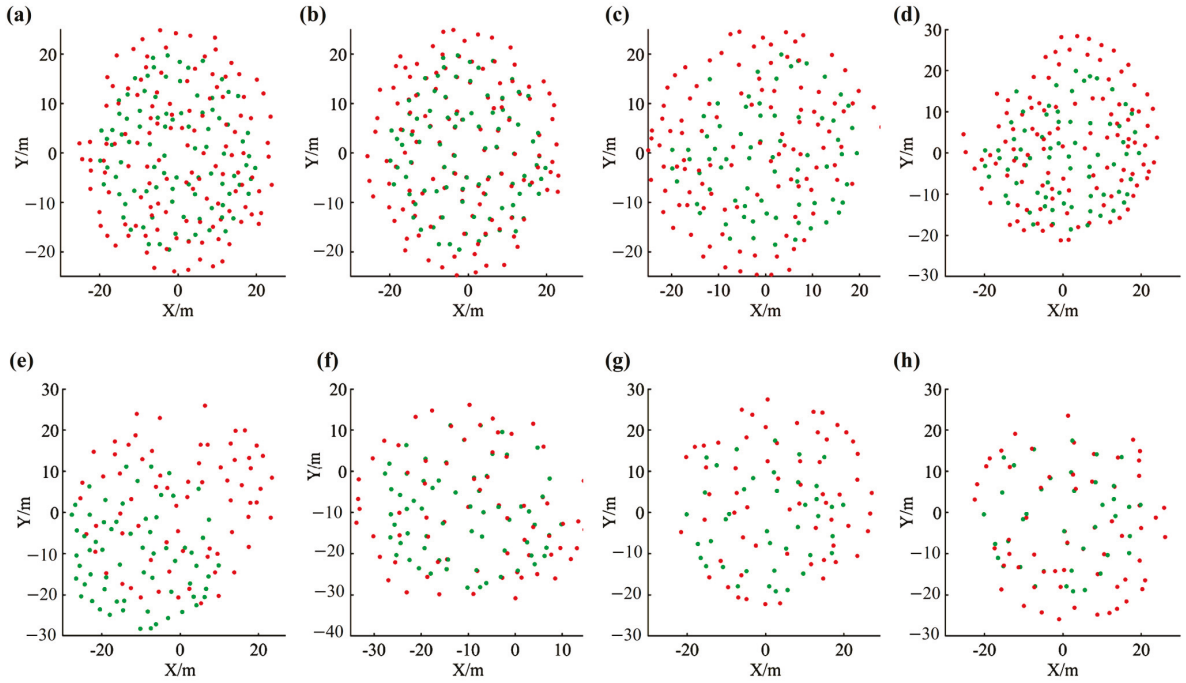



**Figure 5.** Comparison of ULS and TLS tree height point positions before and after coarse registration. ULS (red) and TLS (green). (**a**) Plot 1 before coarse registration, (**b**) plot 1 after coarse registration, (**c**) plot 2 before coarse registration, (**d**) plot 2 after coarse registration, (**e**) plot 3 before coarse registration, (**f**) plot 3 after coarse registration, (**g**) plot 4 before coarse registration, (**h**) plot 4 after coarse registration.

#### 2.3.3. Fine Registration Method Based on Nearest Points

Through the rough registration method of searching for the same feature point pairs, the ULS point cloud and the TLS point cloud after coarse registration already have a high overlap degree, the number of the same name point pairs obtained in the end is different, and the registration results are prone to local optimal solutions. To achieve higher registration accuracy, further fine registration is required [21]. The basic idea of fine registration is to minimize the distance between the same name points by iterative calculation of the nearest points (ICP) [15]. The main steps of fine registration include: (1) searching for the nearest space point to each point on $ULS_{tr}$ after coarse registration in the TLS cluster, (2) performing SVD decomposition on the nearest point pairs to solve for the rotation matrix and translation vector, which are then applied to the ULS point cloud for precise registration, and (3) iteratively calculating errors until finding the rotation matrix and translation vector that minimize the mean square error of the distance between nearest points. Substituting the rotation matrix $R$ and the translation vector $t$ from the precise registration of nearest points into Formula (9) yields the ULS point cloud after final registration. The precisely registered ULS point cloud is merged with the TLS point cloud to create a complete, integrated TLS and ULS fusion point cloud.

## 3. Results

### *3.1. Registration Effect*

The initial registration of ULS and TLS point clouds is realized by SVD decomposition through an approximate distance search to detect point pairs with the same features. To further improve the registration accuracy, the final registration of the ULS and TLS point clouds is completed using the nearest point iteration to achieve fine registration. Figures 6–13 show the registration effect. Figures 6, 8, 10 and 11 show that the ULS point cloud is consistent with the overall range of the TLS point cloud after coordinate conversion ((a), (b), and (c)). No obvious feature points with the same name can be found by artificial

visual inspection either from the canopy or from the ground, and the point cloud data of the two platforms present chaotic arrangements. After coarse registration and fine registration, the ULS point cloud and TLS point cloud achieve a certain degree of overlapping fusion ((d), (e), (f)), and the canopy and plane are fitted together. The slopes of plots 3 (Figure 10) and 4 (Figure 12) are significantly higher than those of plots 1 and 2. Soon, missing ground point clouds will occur in areas with complex terrain conditions. However, the method proposed in this study also shows that the ground points of point clouds in the two plots are well registered together because the features extracted in the study are based on canopy point clouds. Of course, even though the tree heights extracted by the point clouds of the two platforms in the stage of feature extraction of tree heights do not exactly correspond one-to-one, the registration results do not affect the final registration results, proving that the proposed extraction method is also applicable to natural forest sample plots with certain slopes and different stand densities.

To further check the registration effect, the research clipped and extracted individual tree point clouds from ULS point clouds and TLS point clouds in each sample plot to check the local registration effect, as shown in Figures 7, 9, 11 and 13. The local single log registration results show that the fusion of the ULS point cloud and TLS point cloud registration can make up for the respective data disadvantages of the two platforms to a certain extent. Compared with a single TLS point cloud, the fused point cloud can obtain more complete canopy information, and some canopy information missing due to scanning angle or special canopy shape can be supplemented, especially canopy height information. Tree species with high tree heights, such as Yunnan pine (Figure 11) and fir (Figure 13), have serious canopy width losses in TLS point clouds. The canopy information in the fusion point cloud is fully expressed, and it is convenient to extract the information of individual trees more accurately. Compared with a single ULS point cloud, the fused point cloud can compensate for the missing understory information and single trunk information. In the plots with higher stand density and larger canopy width, the ULS cannot completely penetrate the canopy layer, resulting in the loss of single trunk information. The fused point cloud accurately retains the information under the TLS point cloud and stores it in the fusion point cloud.

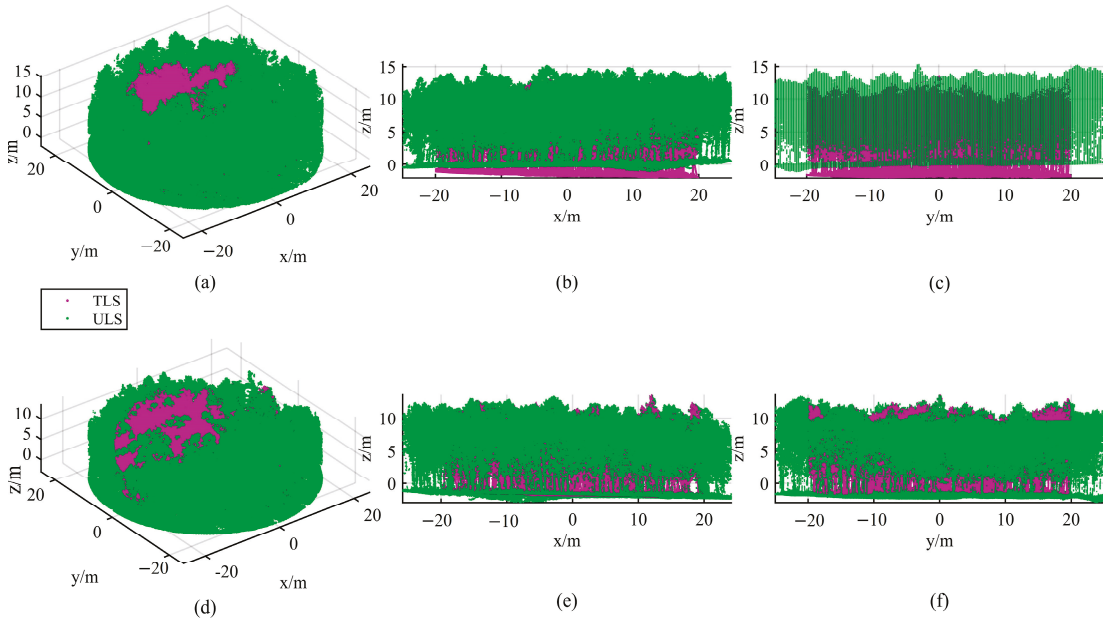



**Figure 6.** Overall registration effect of plot 1 (*Pinus densata* Mast.). Overall effect of precise registration for plot 1 (*Pinus densata* Mast.). (**a**–**c**) represent the ULS and TLS point clouds before registration, and (**d**–**f**) show the point clouds after registration. TLS points are shown in magenta, and ULS points are shown in green.

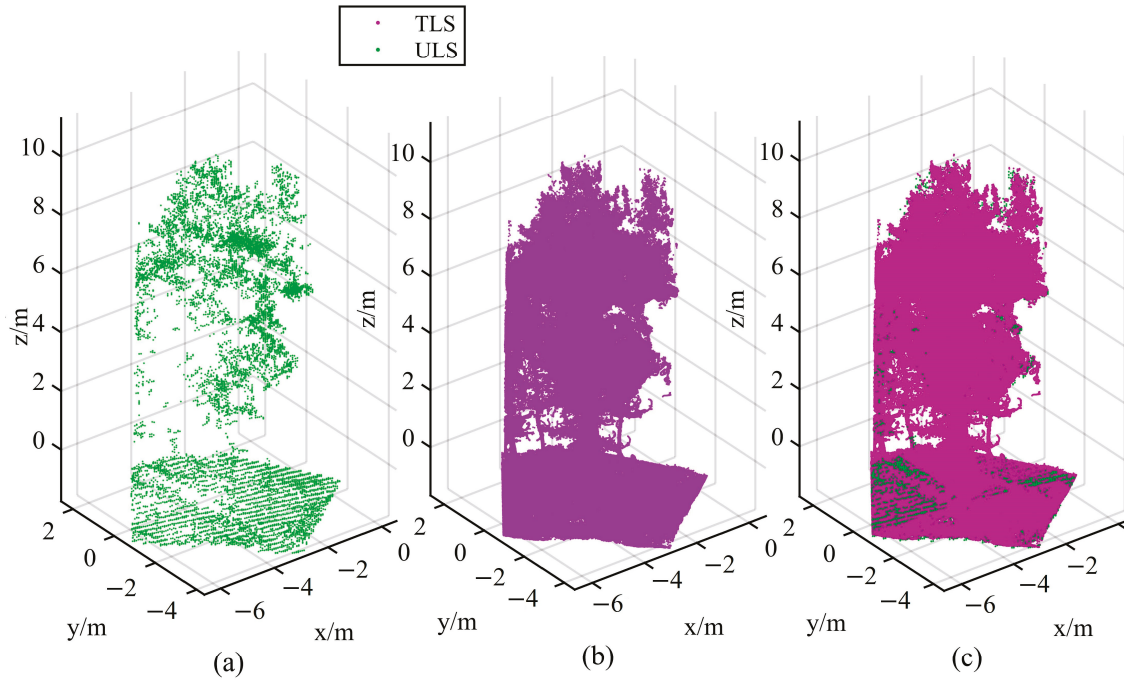



**Figure 7.** Fine registration—local amplification of plot 1 (*Pinus densata* Mast.). Local amplification of fine registration results for plot 1 (*Pinus densata* Mast.). (**a**) ULS point cloud, (**b**) TLS point cloud, (**c**) combined ULS-TLS point cloud after registration.

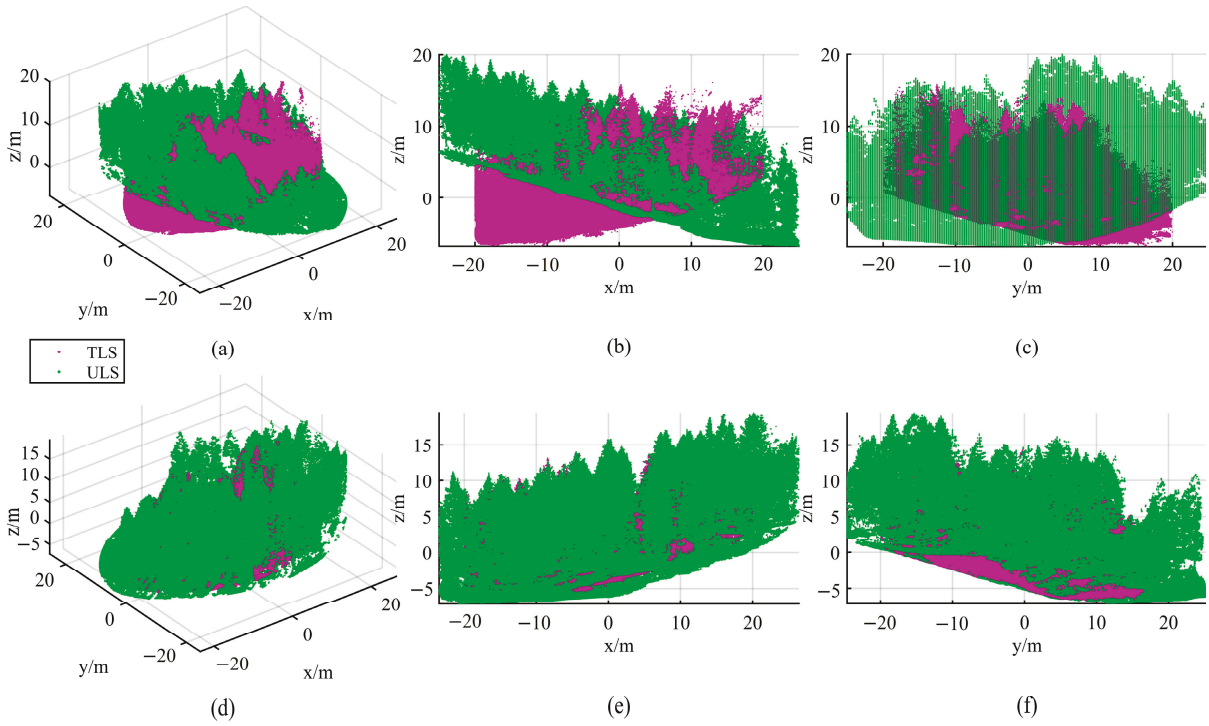



**Figure 8.** Overall registration effect of plot 2 (*Pices asperata* Mast.). Overall effect of precise registration for plot 1 (*Pices asperata* Mast.). (**a**–**c**) represent the ULS and TLS point clouds before registration, and (**d**–**f**) show the point clouds after registration. TLS points are shown in magenta, and ULS points are shown in green.

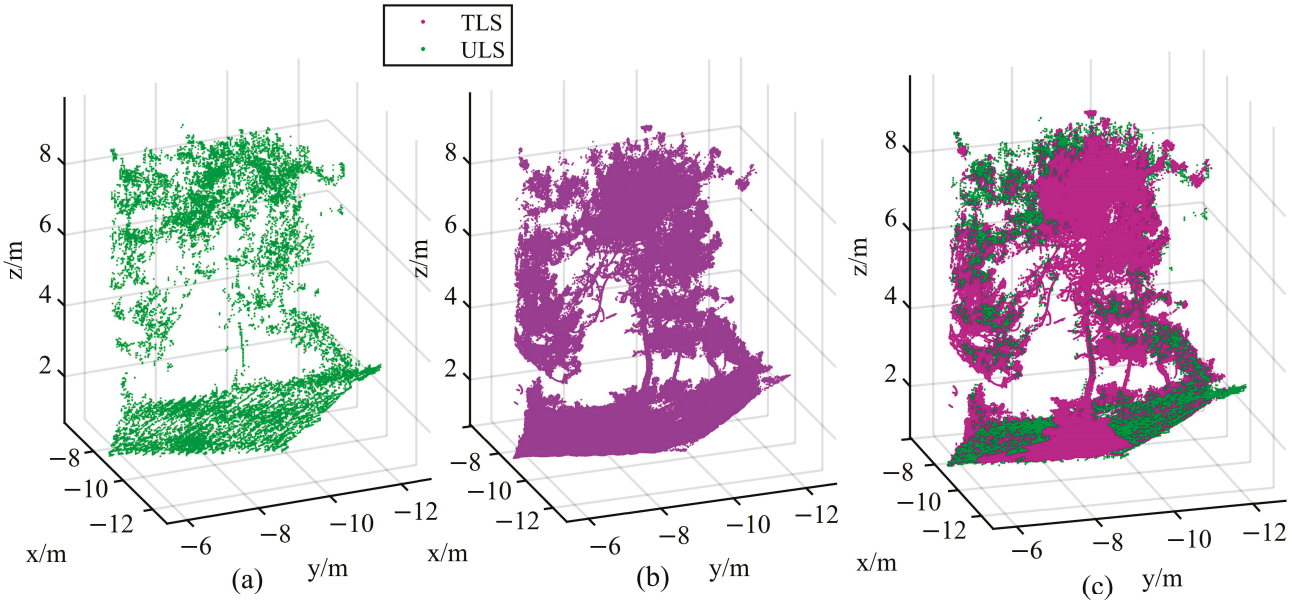



**Figure 9.** Fine registration—Local amplification of plot 2 (*Pices asperata* Mast.). Local amplification of fine registration results for plot 1 (*Pices asperata* Mast.). (**a**) ULS point cloud, (**b**) TLS point cloud, (**c**) combined ULS-TLS point cloud after registration.

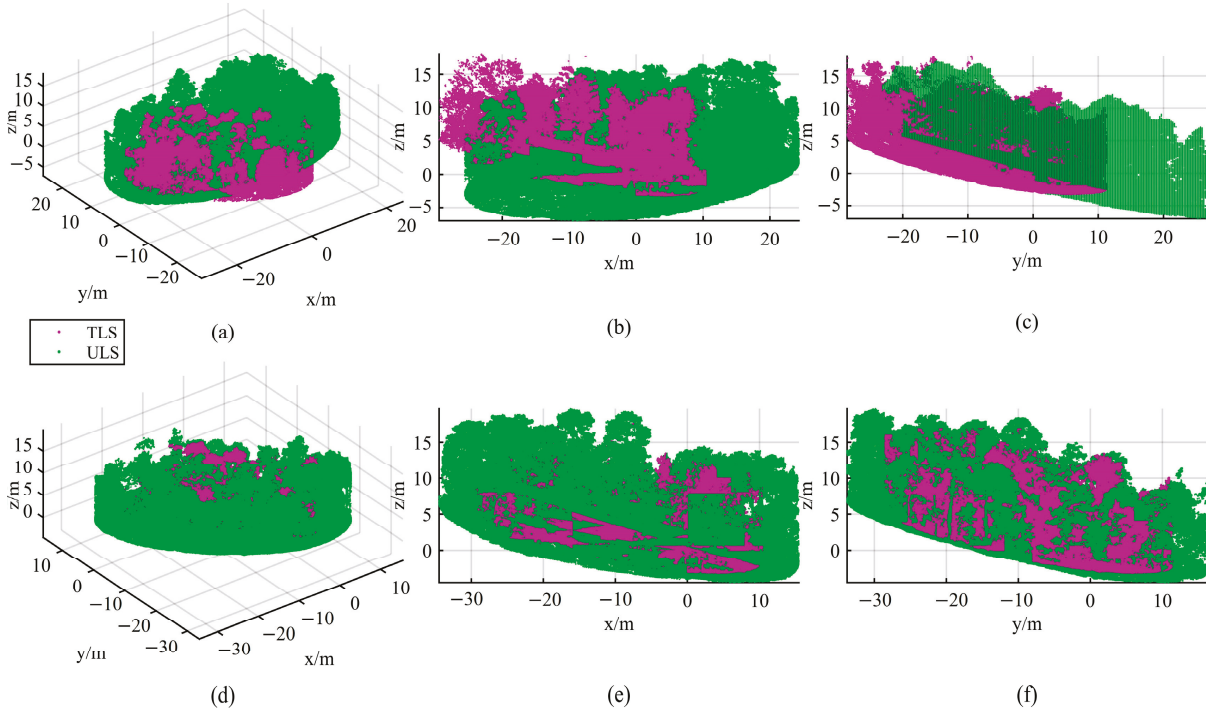



**Figure 10.** Overall registration effect of plot 3 (*Pinus yunnanensis* Franch.). Overall effect of precise registration for plot 1 (*Pinus yunnanensis* Franch.). (**a**–**c**) represent the ULS and TLS point clouds before registration, and (**d**–**f**) show the point clouds after registration. TLS points are shown in magenta, and ULS points are shown in green.

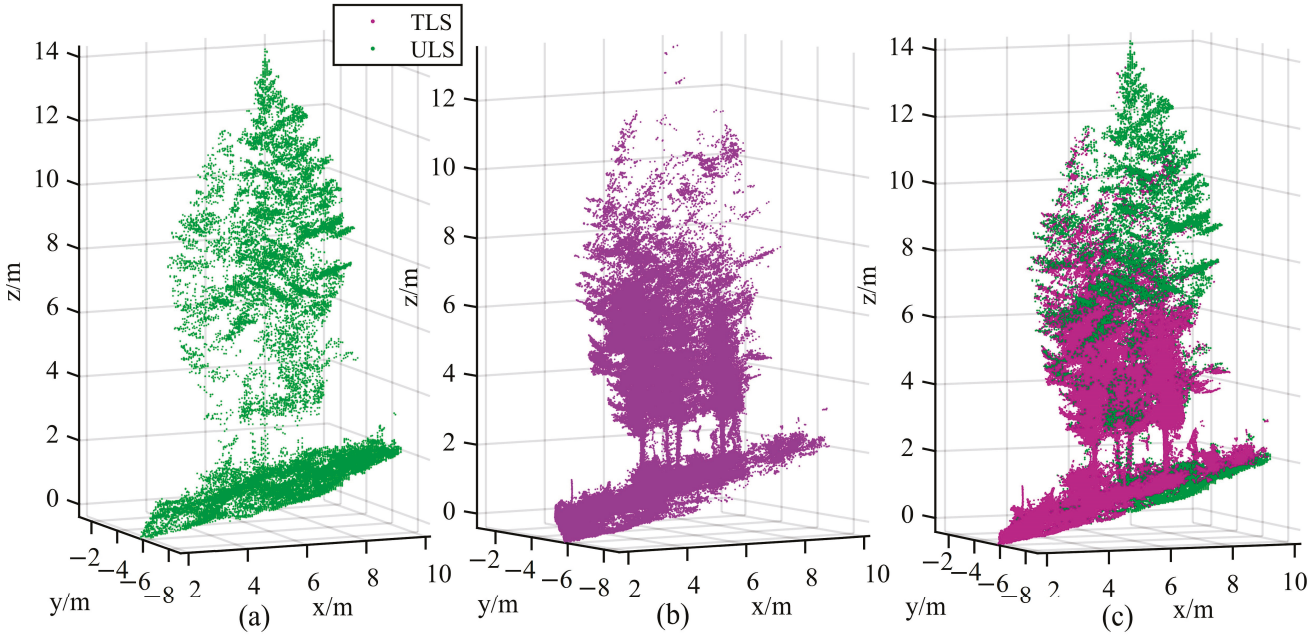



**Figure 11.** Fine registration—local amplification of plot 2 (*Pinus yunnanensis* Franch.). Local amplification of fine registration results for plot 1 (*Pinus yunnanensis* Franch.). (**a**) ULS point cloud, (**b**) TLS point cloud, (**c**) combined ULS-TLS point cloud after registration.

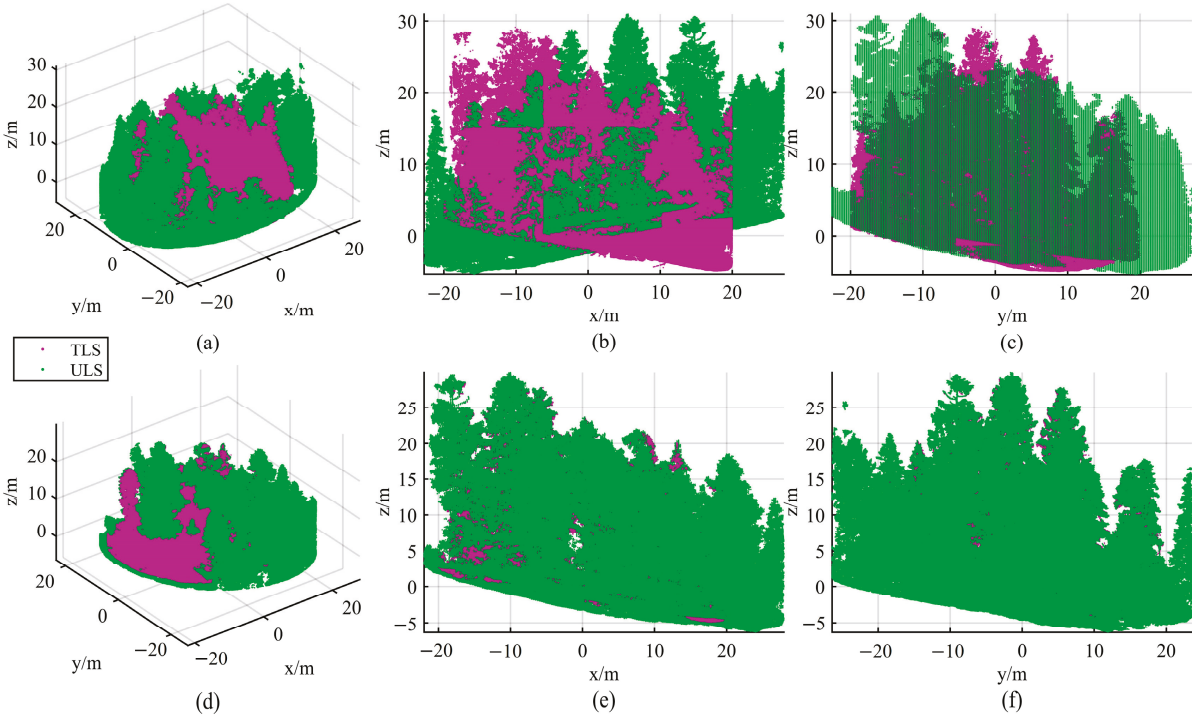



**Figure 12.** Overall registration effect of plot 3 (*Abies fabri* (Mast.) Craib). Overall effect of precise registration for plot 1 (*Abies fabri* (Mast.) Craib). (**a**–**c**) represent the ULS and TLS point clouds before registration, and (**d**–**f**) show the point clouds after registration. TLS points are shown in magenta, and ULS points are shown in green.

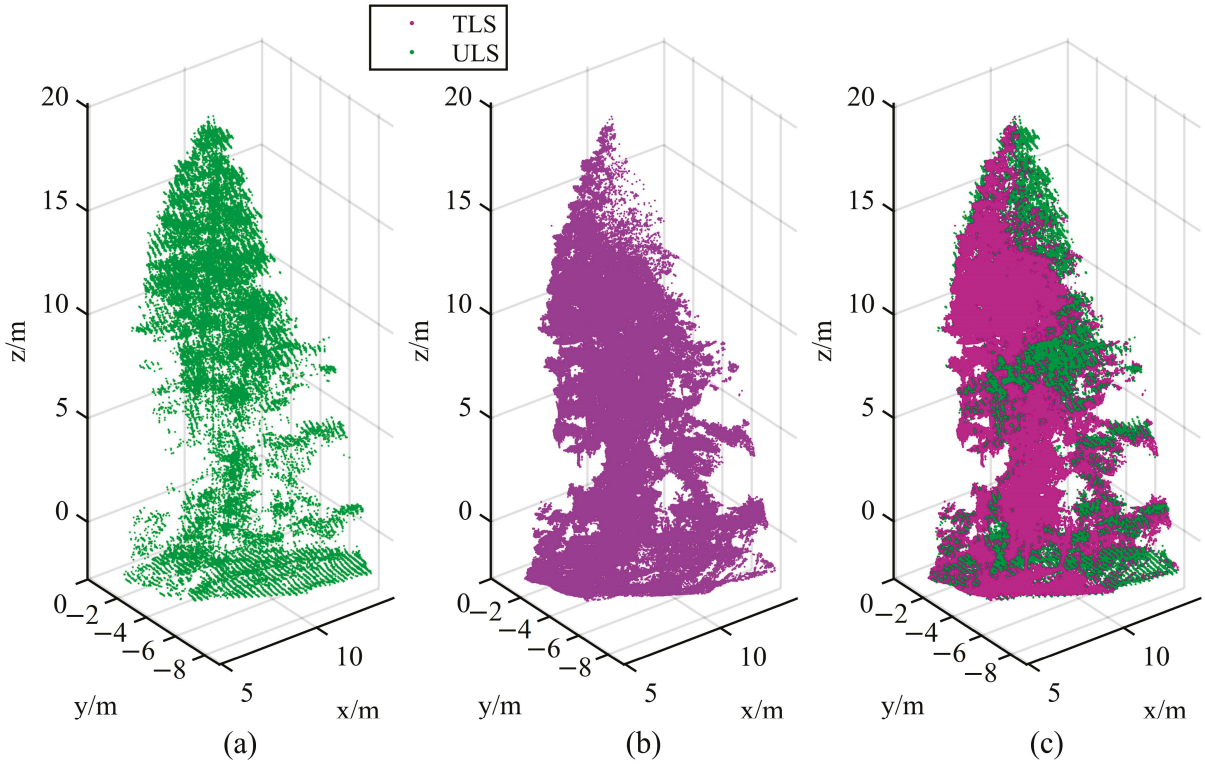



**Figure 13.** Fine registration—local amplification of plot 2 (*Abies fabri* (Mast.) Craib). Local amplification of fine registration results for plot 1 (*Abies fabri* (Mast.) Craib). (**a**) ULS point cloud, (**b**) TLS point cloud, (**c**) combined ULS-TLS point cloud after registration.

### *3.2. Registration Accuracy*

The research selects the nearest point distance as the index for registration accuracy verification, that is, the nearest point in the ULS point cloud and TLS point cloud after searching registration, and calculates the average p of the sum of the distances, which is the final registration accuracy [14,22,27]. Its calculation formula is shown in Formula (9). The distance distribution of the nearest points of the ULS point cloud in the TLS point cloud after coarse registration and fine registration is studied, calculated, and counted, as shown in Figures 14–17, where (a) represents the distance distribution after coarse registration and (b) represents the distance distribution after fine registration. After fine registration, the distribution of the nearest distance tends to be smaller, and the distant points before fine registration are improved.

$$P = \frac{1}{n} \times \sum_{i=1}^{n} d_{knn(i)} \tag{10}$$

where $n$ represents the total number of points and $d_{knn(i)}$ represents the nearest distance of the $i$th point.

The mean ($\mu$) and standard deviation ($\sigma$) of these distances provide a quantitative measure of the registration quality (Table 7). Table 7 shows that coarse registration has already achieved a certain level of accuracy, with relatively low mean nearest point distances and small standard deviations across the plots, indicating that the ULS and TLS point clouds have been effectively merged. However, fine registration further enhances the accuracy, reducing both the mean and standard deviation in each plot, significantly improving the alignment accuracy and consistency, ensuring a tighter match and higher precision fusion of the point clouds.

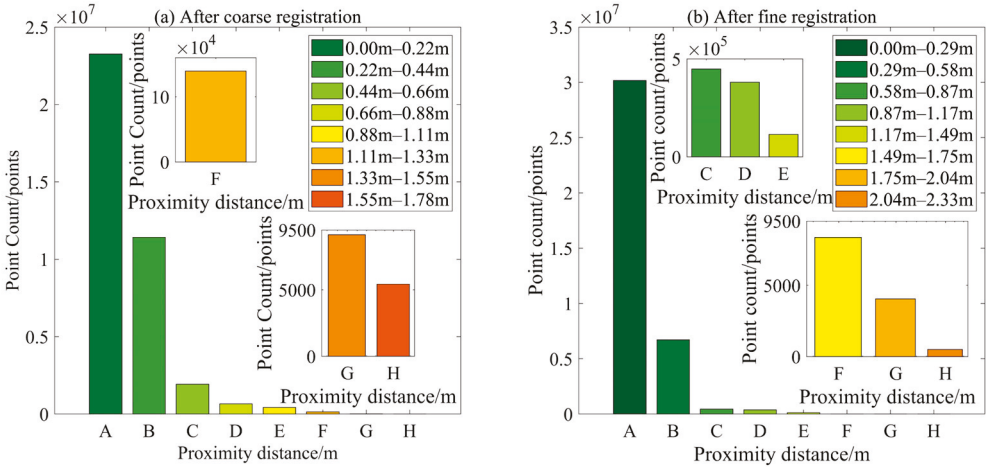



**Figure 14.** Plot 1 ULS-TLS distance statistics of nearest points, (**a**) after coarse registration, (**b**) after fine registration.

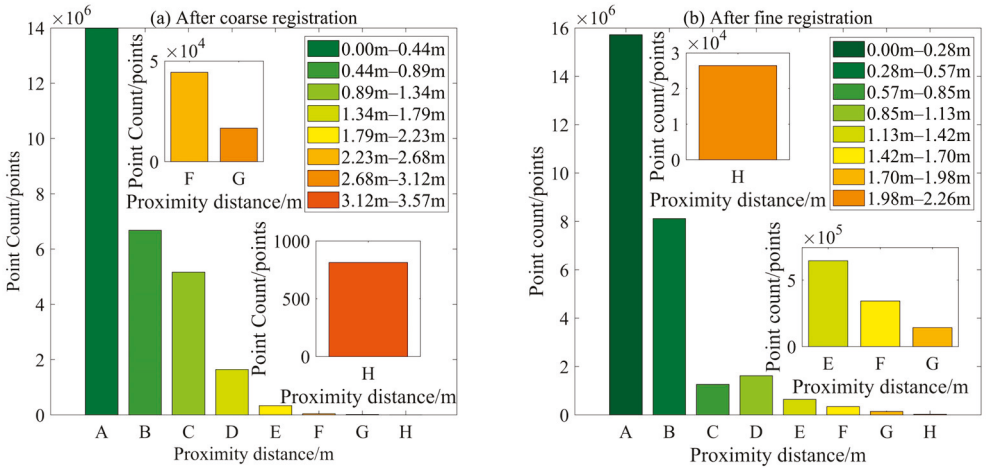



**Figure 15.** Plot 2 ULS-TLS distance statistics of nearest points, (**a**) after coarse registration, (**b**) after fine registration.

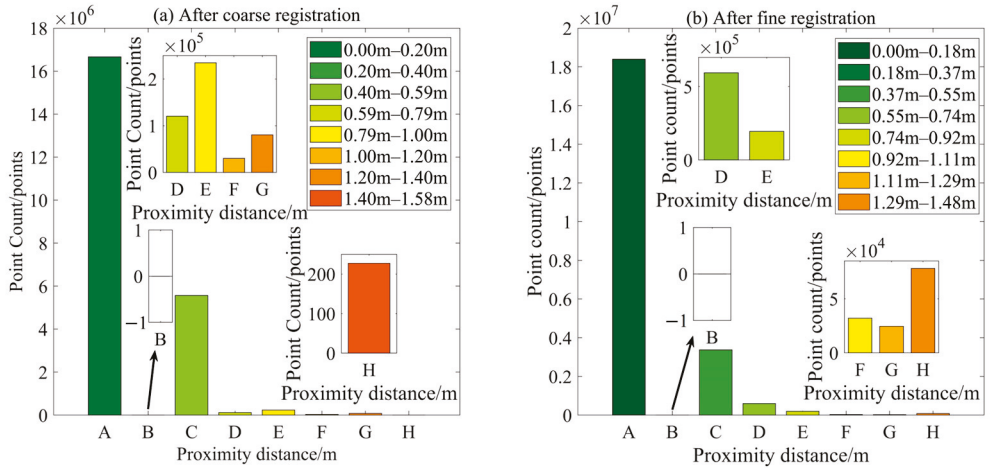



**Figure 16.** Plot 3 ULS-TLS distance statistics of nearest points, (**a**) after coarse registration, (**b**) after fine registration.

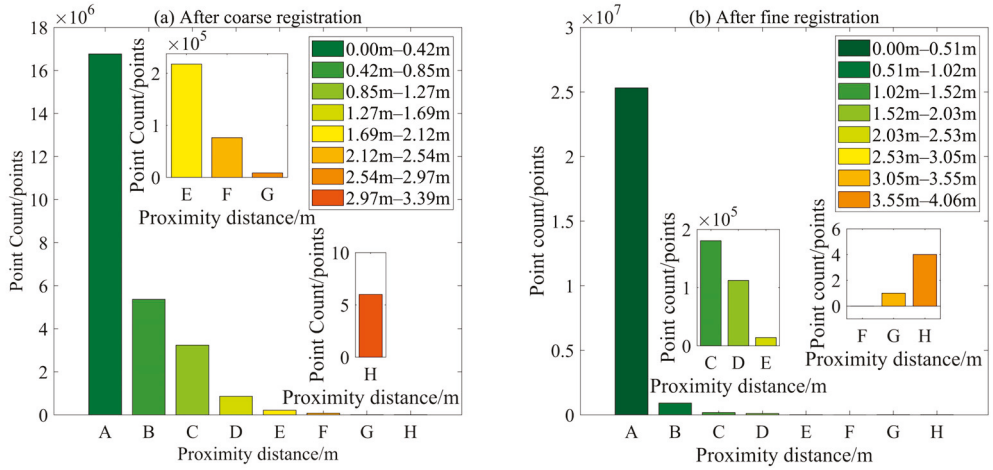



**Figure 17.** Plot 4 ULS-TLS distance statistics of nearest points, (**a**) after coarse registration, (**b**) after fine registration.

**Table 7.** Registration Accuracy Statistics.

| Plot | Coarse Registration | | Fine Registration | |
|---|---|---|---|---|
| | $\mu$/m | $\sigma$/m | $\mu$/m | $\sigma$/m |
| 1 | 0.18 | 0.24 | 0.10 | 0.21 |
| 2 | 0.59 | 0.46 | 0.27 | 0.38 |
| 3 | 0.12 | 0.22 | 0.08 | 0.20 |
| 4 | 0.45 | 0.42 | 0.09 | 0.23 |

## 4. Discussion

### *4.1. The Accuracy Difference between Different Registration Methods*

To prove the validity and reliability of the proposed registration method, two registration methods were selected to compare the registration accuracy. The registration accuracy table is shown in Table 8. Through comparison with other methods, it was found that the coarse registration of the proposed search similarity distance method achieved good results. Compared with the TR (Tree-height Registration) algorithm proposed by Liu et al. [21], which also uses the high point of an individual tree, the precision registration effect of each sample plot improved. Compared with the TR (Tree height Registration) algorithm, the average accuracy improved by 34.88%. The TR algorithm relies on the similarity of the distance between the high point of the tree and the center of the plot and is prone to local optimization. The proposed algorithm not only considers the distance between the individual tree and the center of the plot but also considers the position relationship between the adjacent individual trees. By comparing with the registration results of other methods, the proposed registration method has been proven to have certain validity and reliability, and the registration accuracy improves.

**Table 8.** Comparison of the accuracy of each registration method.

| Plot | Fine Registration Accuracy (m) | | Absolute Accuracy Difference (m) | Relative Accuracy Difference (%) |
|---|---|---|---|---|
| | SDS | TR [1] | SDS vs. TR | SDS vs. TR |
| 1 | 0.10 | 0.12 | 0.02 | 16.67 |
| 2 | 0.27 | 0.48 | 0.21 | 43.75 |
| 3 | 0.08 | 0.10 | 0.02 | 20.00 |
| 4 | 0.09 | 0.22 | 0.13 | 59.10 |

[1] TR (Tree-height Registration) algorithm.

*4.2. Difference in Extracting Single-Wood Parameters of Point Cloud Data before and after Fusion*

Fusion point cloud data were obtained by merging the ULS point cloud data after registration conversion with the original TLS point cloud data. To further explore the difference between individual platform point cloud and fusion point cloud extraction of individual tree parameters, individual tree height extraction was carried out for single TLS point cloud data, single ULS point cloud data, and point clouds after registration fusion.

The study introduced three metrics: *RMSE* (Root Mean Square Error), $R^2$ (Coefficient of Determination), and *MRE* (Mean Relative Error) to evaluate the accuracy of tree height extraction. The evaluation results are shown in Figure 18. From the correlation between the measured tree height and the tree height extracted based on the point clouds of each platform, the tree height extracted from the fusion point cloud and ULS point cloud has a high correlation with the measured value, $R^2$ reaches 0.955 (*RMSE* = 1.897 m, *MRE* = 8.693%), followed by the ULS point cloud, $R^2$ reaches 0.965 (*RMSE* = 1.707 m, *MRE* = 10.082%), and the difference between the two is not large. The lowest is the TLS point cloud data, where $R^2$ is 0.838 (*RMSE* = 3.344 m, *RME* = 10.670%) (Figure 18). The reason for the lower accuracy of tree height extraction from the fused point cloud compared with the ULS point cloud may be due to the selected study area where the tree heights range from 7 m to 30 m, resulting in significant variations in *RMSE* with large differences in tree height. Therefore, when considering the *MRE* metric, it can be observed that the fused point cloud provides higher overall accuracy and greater stability in tree height extraction. Among the three kinds of data, the high precision of the TLS cloud extraction tree is relatively low, which is caused by the difference between the bottom-up scanning mode of TLS and the top-down scanning mode of ULS LiDAR [28,29]. For some tall trees, the scanning range of the TLS is limited, the laser point cannot reach the true high point of the individual tree canopy, and the maximum value in the Z-axis direction is not the true maximum value, that is, the true tree high point [30]. Although ULS can obtain a more complete individual canopy point cloud, part of the noise attached to the upper layer of the tree will become the source of error, which is also the reason why the height estimate of an individual tree extracted based on the ULS point cloud in the study is relatively high. Fusion point clouds make up for the data disadvantages caused by the different scanning methods of a single platform, making the canopy information of TLS point clouds more complete and the attribution of point clouds more clear. Some point clouds attached to the upper canopy are easier to remove by denoising to obtain more accurate tree height values [31].

In fact, actual tree height measurements also contain human errors. Tree heights were measured using a Blume-Leiss altimeter, which requires visual identification of the tree's highest and root points. Such subjective judgments, combined with interference from complex forest environments, increase the margin of human measurement error [31]. Especially for some tall trees, extracting tree heights from LiDAR point cloud data is more accurate than manual measurements. Of course, measuring individual trees by felling them would yield different results, but that approach would cause ecological damage and incur higher measurement costs.

*4.3. Coordinate System Transformation*

In the preprocessing phase, the initial ULS coordinate system is transformed into the same local coordinate system as TLS, with the proposed registration method using TLS as the reference point cloud to register the ULS point cloud within it. This approach has some drawbacks in practical applications. Converting TLS into geographic coordinates offers advantages; however, existing studies indicate that converting TLS to geographic coordinates requires more transformation parameters, making it difficult to operate and presenting significant disadvantages [16,19]. The study did not acquire the necessary transformation parameters during the data acquisition phase, so it ultimately chose to transform ULS into the TLS coordinate system. Additionally, the coordinate transformation principles outlined in Section 2.3.1 were applied. After clipping the point cloud, the centroid of the airborne ground point cloud was computed and added to the ground-based point

cloud, thereby updating the coordinates of the transformed ground-based point cloud. This transformation is expressed by Equation (11):

$$\mathrm{TLS}(\mathrm{x},\mathrm{y},\mathrm{z})_{\mathrm{i}} = TLS_{origin}(x,y,z)_i + \mathrm{mean}\left(ULS_{ground}(x,y,z)\right), i = 1,2,\ldots,n \quad (11)$$

where $TLS(x,y,z)_i$ denotes the transformed TLS coordinates, $TLS_{origin}(x,y,z)_i$ represents the original TLS coordinates, and mean$\left(ULS_{ground}(x,y,z)\right)$ represents the centroid of the airborne ground point cloud prior to coordinate transformation. Following this transformation, both the airborne and ground-based point clouds were aligned to the WGS1984 coordinate system (Figure 19b). The registration method proposed in this study was then employed, achieving point cloud registration with the registered point cloud aligned to the coordinate system of the original ULS (Figure 19d). The results demonstrate that the choice of coordinate system does not affect the applicability of the proposed registration algorithm, as the registration concept is independent of the geographic or local coordinate system.

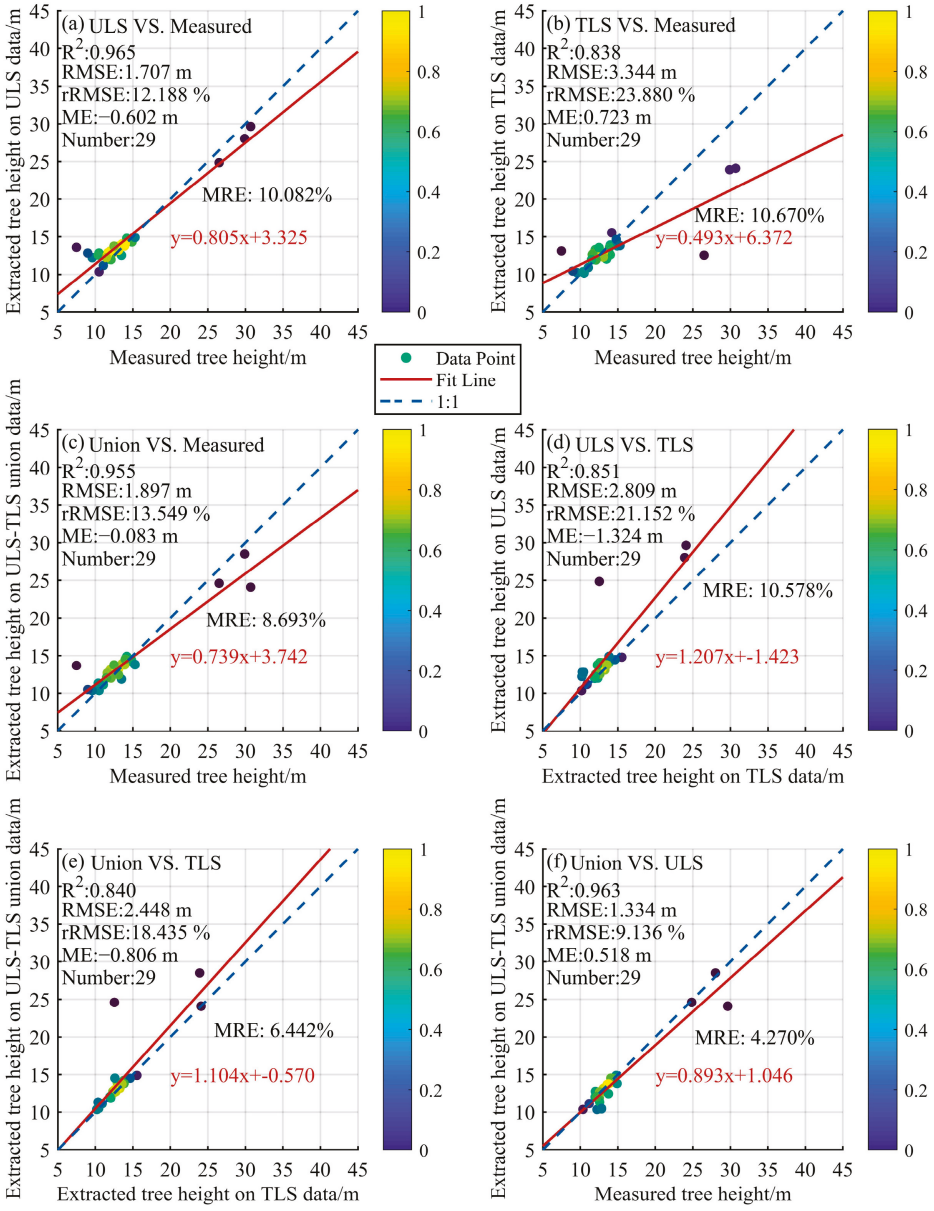



**Figure 18.** Comparison of tree height extraction results before and after point cloud fusion.

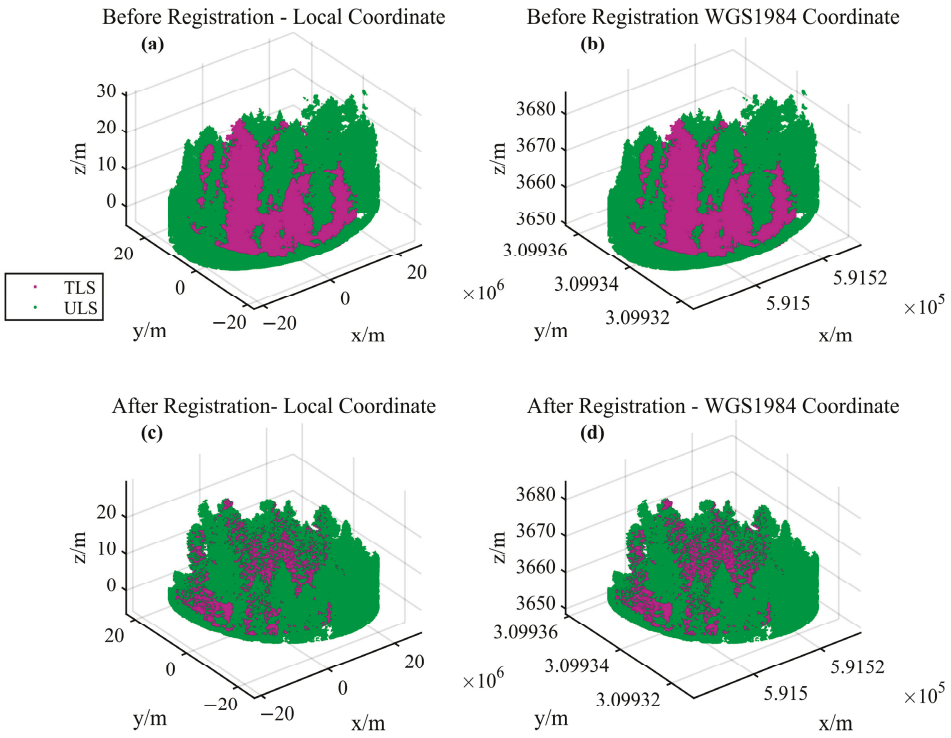



**Figure 19.** The original point cloud and the registered point cloud in two coordinate systems (same views).

### *4.4. Tree Height Detection*

The sliding window technique is utilized to detect local maxima in the CHM, which indicates the positions of individual trees. However, due to different scanning methods used by ULS and TLS, the TLS occasionally fails to capture the true canopy top, resulting in discrepancies in detected local maxima [32,33]. To address this, our proposed method uses tree height as one of the constraints, along with the distance similarity between trees to match detected tree heights, thereby confirming the final point pairs. This approach mitigates errors that could arise from relying solely on parameter extraction and considers the spatial relationships between individual trees, thus avoiding local optimization issues [19]. Quantitative verification demonstrates that this is an effective registration method for point clouds in natural forest areas.

### *4.5. Deficiency and Prospects*

The proposed registration algorithm achieved a good registration effect in natural forest sample plots. Compared with some existing registration algorithms, the accuracy is improved, but there are still some areas that can be further studied and improved. For example, the forest sample plots involved in the research are all natural forest sample plots in the alpine mountains of Shangri-La City; however, they are all natural pure forests with strong homogeneity and individual tree species. To further prove the validity and universality of the proposed registration method or to explore a more suitable registration method, this study will be conducted for mixed forests or more complex stands in future studies.

## 5. Conclusions

LiDAR technology is an effective means of informed, intelligent and efficient management of forest resources. Various platforms' point cloud data have advantages and disadvantages. Rapid and accurate registration and fusion of multiplatform point cloud data will better utilize the advantages of each platform point cloud data and further improve the role of LiDAR in forest surveys. Aiming at the shortcomings of current point cloud registration algorithms in natural forests, a ULS-TLS point cloud registration algorithm, SDS, based on a similar distance search, is proposed. The point cloud was generated

into a DSM to extract the tree's high point instead of the individual tree position. The target tree high point and the similar position of the surrounding tree high point were searched to find the same name feature point pair, achieve the point cloud rough registration by SVD solution, and finally complete the point cloud fine registration by the nearest iteration. The registration method was applied to four natural forest sample plots of four typical tree species in different site environments in Shangri-La City, Yunnan Province, China. The average registration accuracy reached 0.13 m, and compared with the TR algorithm, the accuracy increased by 34.88%, proving that the proposed method has certain validity.

To further illustrate the advantage of fused point cloud data over individual platform point cloud data in extracting individual tree parameters, the study extracted tree height parameters from both pre- and post-fusion point clouds. By comparing these with actual measured parameters, it was found that the tree heights extracted from the fused point clouds had higher correlation with the actual measured tree heights, achieving greater extraction precision. The correlation coefficient $R^2$ increased by 0.004, making the point cloud data of natural forest areas more complete and the expression of individual tree point cloud information more comprehensive.

From a practical standpoint, the successful application of the SDS algorithm in diverse natural forest environments highlights its robustness and adaptability. Forest managers can now leverage this method to obtain more accurate tree height measurements, which are crucial for inventory assessments, biomass estimation, and ecological monitoring. The improved precision in data fusion also means that large-scale forest surveys can be conducted more efficiently, with reduced need for manual intervention and correction.

Moreover, the enhanced accuracy of the fused data opens up new possibilities for advanced forest management practices, such as dynamic monitoring of forest health, detection of subtle changes in tree growth, and more precise modeling of forest structure. These capabilities are essential for sustainable forest management, enabling more informed decision-making processes that can better address conservation goals, resource allocation, and climate change mitigation efforts.

In summary, this study provides a significant methodological advancement with both scientific and practical implications. By effectively combining ULS and TLS point cloud data, the SDS algorithm not only improves the accuracy and reliability of forest resource assessments but also supports more sustainable and efficient forest management practices. The contributions of this research extend beyond technical improvements, offering a valuable tool for a wide range of applications in forestry and environmental management.

**Author Contributions:** Conceptualization, Y.D., P.D. and Q.L.; methodology, Y.D., J.Z. and J.W.; validation, Y.D., J.W. and G.S.; software, Y.D. and W.M.; writing—original draft preparation, Y.D., P.D. and J.Z.; writing—review and editing, Y.D., J.L. and J.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by the Yunnan Provincial Science and Technology Project at Southwest United Graduate School [grant number 202302AO370003]; the Yunnan Province Science and Technology Talents and Platform Plan Project [grant number 202305AO350003]; the National Natural Science Foundation of China [grant number 41961060], and Yunnan Fundamental Research Projects [grant number 202401AT070111].

**Data Availability Statement:** The data that support the fundings of this study are available from the corresponding author upon reasonable request.

**Acknowledgments:** We would like to thank anonymous reviewers and editors for their professional comments and suggestions. We would like to thank Dajiang Shao and Guankun Su for their help in data collecting.

**Conflicts of Interest:** Author Weifeng Ma was employed by the company Power China Kunming Engineering Corporation Limited. Author Guankun Su was employed by the company New Coordinates Technology Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References


1. Watson, J.E.M.; Evans, T.; Venter, O.; Williams, B.; Tulloch, A.; Stewart, C.; Thompson, I.; Ray, J.C.; Murray, K.; Salazar, A. The exceptional value of intact forest ecosystems. *Nat. Ecol. Evol.* **2018**, *2*, 599–610. [CrossRef] [PubMed]
2. Flores, B.M.; Montoya, E.; Sakschewski, B.; Nascimento, N.; Staal, A.; Betts, R.A.; Levis, C.; Lapola, D.M.; Esquivel-Muelbert, A.; Jakovac, C. Critical transitions in the Amazon forest system. *Nature* **2024**, *626*, 555–564. [CrossRef] [PubMed]
3. Lefsky, M.A.; Cohen, W.B.; Acker, S.A.; Parker, G.G.; Spies, T.A.; Harding, D. Lidar Remote Sensing of the Canopy Structure and Biophysical Properties of Douglas-Fir Western Hemlock Forests. *Remote Sens. Environ.* **1999**, *70*, 339–361. [CrossRef]
4. Wang, S.; Liu, C.; Li, W.; Jia, S.; Yue, H. Hybrid model for estimating forest canopy heights using fused multimodal spaceborne LiDAR data and optical imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *122*, 103431. [CrossRef]
5. Blair, J.B.; Rabine, D.L.; Hofton, M.A. The Laser Vegetation Imaging Sensor: A medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 115–122. [CrossRef]
6. Zhu, X.X.; Nie, S.; Wang, C.; Xi, X.H.; Wang, J.S.; Li, D.; Zhou, H.Y. A noise removal algorithm based on OPTICS for photon-counting LiDAR data. *IEEE Geosci. Remote Sens.* **2021**, *18*, 1471–1475. [CrossRef]
7. Du, L.M.; Pang, Y.; Wang, Q.; Huang, C.Q.; Bai, Y.; Chen, D.S.; Lu, W.; Kong, D. A LiDAR biomass index-based approach for tree- and plot- level biomass mapping over forest farms using 3D point clouds. *Remote Sens. Environ.* **2023**, *290*, 113543. [CrossRef]
8. Bucksch, A.; Lindenbergh, R.; Abd Rahman, M.Z.; Menenti, M. Breast Height Diameter Estimation from High-Density Airborne LiDAR Data. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 1056–1060. [CrossRef]
9. Deng, Y.C.; Pan, J.Y.; Wang, J.L.; Liu, Q.W.; Zhang, J.P. Mapping of Forest Biomass in Shangri-La City Based on LiDAR Technology and Other Remote Sensing Data. *Remote Sen.* **2022**, *14*, 5816. [CrossRef]
10. Bauwens, S.; Bartholomeus, H.; Calders, K.; Lejeune, P. Forest inventory with terrestrial lidar: A comparison of static and hand-held mobile laser scanning. *Forest* **2016**, *7*, 127. [CrossRef]
11. Kelbe, D.; van Aardt, J.; Romanczyk, P.; van Leeuwen, M.; Cawse-Nicholson, K. Marker-free registration of forest terrestrial laser scanner data pairs with embedded confidence metrics. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4314–4330. [CrossRef]
12. Zhang, W.; Wan, P.; Wang, T.; Cai, S.; Chen, Y.; Jin, X.; Yan, G. A Novel Approach for the Detection of Standing Tree Stems from Plot Level Terrestrial Laser Scanning Data. *Remote Sens.* **2019**, *11*, 211. [CrossRef]
13. Balestra, M.; Marselis, S.; Sankey, T.T.; Cabo, C.; Liang, X.; Mokroš, M.; Peng, X.; Singh, A.; Stereńczak, K.; Vega, C.; et al. LiDAR Data Fusion to Improve Forest Attribute Estimates: A Review. *Curr. For. Rep.* **2024**, *10*, 281–297. [CrossRef]
14. Guan, H.C.; Su, Y.J.; Sun, X.L.; Xu, G.C.; Li, W.K.; Ma, Q.; Wu, X.Y.; Wu, J.; Liu, L.L.; Guo, Q.H. A marker-free method for registering multiscan terrestrial laser scanning data in forest environments. *ISPRS J. Photogram. Remote Sens.* **2020**, *166*, 82–94. [CrossRef]
15. Besl, P.J.; McKay, N.D. A method for registration of 3-D shapes. *IEEE Transact. Pattern Analys. Mach. Intell.* **1992**, *14*, 239–256. [CrossRef]
16. Paris, C.; Kelbe, D.J.; van Aardt Bruzzone, L. A Novel Automatic Method for the Fusion of ALS and TLS LiDAR Data for Robust Assessment of Tree Crown Structure. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3679–3693. [CrossRef]
17. Dai, W.X.; Yang, B.S.; Liang, X.L.; Dong, Z.; Huang, R.G.; Wang, Y.S.; Li, W.Y. Automated fusion of forest airborne and terrestrial point clouds through canopy density analysis. *ISPRS J. Photogram. Remote Sens.* **2019**, *156*, 94–107. [CrossRef]
18. Myronenko, A.; Song, X.B. Point Set Registration: Coherent Point Drift. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 2262–2275. [CrossRef]
19. Polewski, P.; Erickson, A.; Yao, W.; Coops, N.; Krzystek, P.; Stilla, U. Object-based coregistration of terrestrial photogrammetric and als point clouds in forested areas. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *III–3*, 347–354. [CrossRef]
20. Polewski, P.; Yao, W.; Cao, L.; Cao, S. Marker-free coregistration of UAV and backpack LiDAR point clouds in forested areas. *ISPRS J. Photogram. Remote Sens.* **2019**, *147*, 307–318. [CrossRef]
21. Liu, Q.W.; Wang, J.L.; Ma, W.F.; Zhang, J.P.; Deng, Y.C.; Shao, D.J.; Xu, D.F.; Liu, Y.C. Target-free ULS-TLS point-cloud registration for alpine forest lands. *Comput. Electron. Agric.* **2021**, *190*, 106460. [CrossRef]
22. Liu, G.J.; Wang, J.L.; Dong, P.L.; Chen, Y.; Liu, Z.Y. Estimating Individual Tree Height and Diameter at Breast Height (DBH) from Terrestrial Laser Scanning (TLS) Data at Plot Level. *Forests* **2018**, *9*, 398–417. [CrossRef]
23. Li, L.; Zhou, B.; Liu, Y.; Wu, Y.; Tang, J.; Xu, W.; Wang, L.; Ou, G. Reduction in Uncertainty in Forest Aboveground Biomass Estimation Using Sentinel-2 Images: A Case Study of Pinus densata Forests in Shangri-La City, China. *Remote Sens.* **2023**, *15*, 559. [CrossRef]
24. Wang, C.; Xi, X.H.; Yang, X.B.; Nie, S. *Introduction to LiDAR Remote Sensing*, 1st ed.; Higher Education Press: Beijing, China, 2022; pp. 1–5.
25. Zhang, W.M.; Qi, J.B.; Wan, P.; Wang, H.T.; Xie, D.H.; Wang, X.Y.; Yan, G.J. An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. *Remote Sens.* **2016**, *8*, 501. [CrossRef]
26. Sorkine-Hornung, O.; Rabinovich, M. *Least-Squares Rigid Motion Using SVD*; ETH Zurich: Zurich, Switzerland, 2017.
27. Dai, W.X.; Kan, H.Y.; Tan, R.C.; Yang, B.S.; Guan, Q.F.; Zhu, N.N.; Xiao, W.; Dong, Z. Multisource forest point cloud registration with semantic-guided key points and robust RANSAC mechanisms. *Int. J. Appl. Earth Obs. Geo* **2022**, *115*, 103105.
28. Pang, Y.; Zhao, F.; Li, Z.Y.; Zhou, S.F.; Deng, G.; Liu, Q.W.; Chen, E.X. Forest Height Inversion using Airborne Lidar Technology. *J. Remote Sens.* **2008**, *12*, 152–158.

29. Liu, L.X.; Pang, Y.; Li, Z.Y. Individual Tree DBH and Height Estimation Using Terrestrial Laser Scanning (TLS) in A Subtropical Forest. *Sci. Silvae Sin.* **2016**, *52*, 26–37.
30. Swayze, N.C.; Tinkham, W.T.; Vogeler, J.C.; Hudak, A.T. Influence of flight parameters on UAS-based monitoring of tree height, diameter, and density. *Remote Sens. Environ.* **2021**, *263*, 112540. [CrossRef]
31. Benjamin, B.; Alvaro, L.; Harm, M.B.; Lammert, K. Comparing RIEGL RiCOPTER UAV LiDAR Derived Canopy Height and DBH with Terrestrial LiDAR. *Sensors* **2017**, *17*, 2371. [CrossRef]
32. Zhang, Z.N.; Wang, T.J.; Skidmore, A.K.; Cao, L.; She, G.H.; Cao, L. An improved area-based approach for estimating plot-level tree DBH from airborne LiDAR data. *For. Ecosyst.* **2023**, *10*, 100089. [CrossRef]
33. Liang, X.; Wang, Y.; Pyörälä, J.; Lehtomäki, M.; Yu, X.; Kaartinen, H.; Kukko, A.; Honkavaara, E.; Issaoui, A.; Nevalainen, O.; et al. Forest in situ observations using unmanned aerial vehicle as an alternative of ter-restrial measurements. *For. Ecosyst.* **2019**, *6*, 20. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

forests

MDPI

*Article*

# Multi-Constraint and Shortest Path Optimization Method for Individual Urban Street Tree Segmentation from Point Clouds

**Shengbo Yu [1], Dajun Li [1], Xiaowei Xie [1], Zhenyang Hui [1,2,*], Xiaolong Cheng [3], Faming Huang [4], Hua Liu [1] and Liping Tu [5]**

[1] School of Surveying and Geoinformation Engineering, East China University of Technology, Nanchang 330013, China; 201461003@ecut.edu.cn (S.Y.); djli@ecut.edu.cn (D.L.); xwxie@ecut.edu.cn (X.X.); liuhua@ecut.edu.cn (H.L.)
[2] National Key Laboratory of Uranium Resources Exploration-Mining and Nuclear Remote Sensing, East China University of Technology, Nanchang 330013, China
[3] School of Civil Engineering and Surveying & Mapping Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, China; chengxiaolong@jxust.edu.cn
[4] School of Infrastructure Engineering, Nanchang University, Nanchang 330031, China; faminghuang@ncu.edu.cn
[5] Mineral Exploration Institute of Jiangxi Geological Survey and Exploration, Nanchang 330038, China; m18970899190@163.com
* Correspondence: huizhenyang2008@ecut.edu.cn

**Abstract**

Street trees are vital components of urban ecosystems, contributing to air purification, microclimate regulation, and visual landscape enhancement. Thus, accurate segmentation of individual trees from point clouds is an essential task for effective urban green space management. However, existing methods often struggle with noise, crown overlap, and the complexity of street environments. To address these challenges, this paper introduces a multi-constraint and shortest path optimization method for individual urban street tree segmentation from point clouds. In this paper, object primitives are first generated using multi-constraints based on graph segmentation. Subsequently, trunk points are identified and associated with their corresponding crowns through structural cues. To further improve the robustness of the proposed method under dense and cluttered conditions, the shortest-path optimization and stem-axis distance analysis techniques are proposed to further refine the individual tree extraction results. To evaluate the performance of the proposed method, the WHU-STree benchmark dataset is utilized for testing. Experimental results demonstrate that the proposed method achieves an average F1-score of 0.768 and coverage of 0.803, outperforming superpoint graph structure single-tree classification (SSSC) and nyström spectral clustering (NSC) methods by 17.4% and 43.0%, respectively. The comparison of visual individual tree segmentation results also indicates that the proposed framework offers a reliable solution for street tree detection in complex urban scenes and holds practical value for advancing smart city ecological management.

**Keywords:** individual tree segmentation; point clouds; street tree; multi-constraint segmentation; shortest path optimization

## 1. Introduction

With the rapid development of Light Detection and Ranging (LiDAR) technology, vehicle-mounted systems have enabled efficient acquisition of 3D point-cloud information for street trees [1–3]. As a vital element of the urban landscape, street trees signifi-

cantly contribute to air purification, noise reduction, and the enhancement of ecological quality [4–6]. In the context of urban forest resource surveys and greening management, acquiring precise structural parameters of individual trees holds immense importance for comprehending urban forest dynamics and facilitating refined management and decision making processes [7–11]. Individual tree segmentation constitutes a crucial aspect of forest resource surveys [12–14]. Nevertheless, owing to acquisition configurations and the occlusion among trees, bottom–up LiDAR scanning frequently fails to capture the upper canopy, whereas top–down acquisition encounters difficulties in obtaining lower structural details, thereby rendering high-precision individual segmentation challenging to achieve [15–17]. Consequently, devising an individual tree segmentation method capable of adapting to complex street environments and diverse acquisition strategies bears significant theoretical and practical value.

Currently, existing segmentation methods for vehicle-mounted LiDAR primarily concentrate on trunk-based or crown-based approaches. Trunk-based methods typically identify trunk geometry as seed points and conduct clustering through projection or region growing strategies [18–21]. For example, Mongus et al. [20] pinpointed vertically oriented points within the middle and lower layers of the tree trunk and employed local fitting surfaces to detect concave regions in the vicinity of the trunk. Cabo et al. [21] introduced an automatic trunk identification method, which leverages the independence and vertical continuity of the trunk to identify and extract trunk points, subsequently facilitating the segmentation of individual trees. Zhang et al. [22] utilized hierarchical gridding for data analysis, integrating the features of point cloud elevation distribution and point cloud density, and applied the region growing algorithm to accomplish the segmentation of individual trees. Li et al. [23] proposed a hierarchical clustering method based on trunk-branch constraints to extract individual trees directly from LiDAR point clouds.

Crown-based approaches regard the treetop or crown structure as the pivotal feature for tree identification [24–27]. For example, Zhen et al. [28] initially employed the dynamic window local maximum method to pinpoint the locations of individual trees and subsequently utilized the label-controlled region growth method to demarcate the tree crown boundaries of the detected individual trees. However, this method demonstrates subpar performance in segmenting targets within complex scenes and when dealing with missing street tree crowns. Koch et al. [27] applied filters to locate the tree tops and segment the canopies of both coniferous and broad-leaved trees. They also used the lower limit of the tree height and canopy width regression curve to determine the size of the local maximum search window. Nevertheless, inherent errors and uncertainties are present during the generation of the canopy height model. Hui et al. [28] introduced a multi-level adaptive individual tree segmentation hybrid model that amalgamates the advantages of CHM (canopy height model)-based and 3D points-based methods. This approach first acquires initial tree detection results through the CHM based method and then employs the 3D points to optimize tree clustering. This paper also automatically estimates the canopy width and scale, integrating multi-scale segmentation results to fully leverage the benefits of different scales, thereby completing individual tree segmentation. Hua et al. [29] utilized adjacent roadside trees with overlapping canopies to first isolate the tree canopy region using a binary classifier. Subsequently, they calculated the optimal projection from the 3D point clouds to the two-dimensional image, corresponding to the scenario where the pixel overlap of a individual tree image is minimized.

To sum up, although substantial progress has been achieved in segmenting individual street trees from vehicle-mounted LiDAR point clouds, a notable research gap persists in achieving accurate and generalizable individual tree segmentation across complex urban environments. This gap endures due to the following three unresolved issues:

https://doi.org/10.3390/f17010027

- Data-quality issues stemming from real-world acquisition conditions, such as noise, missing points, and occlusions caused by buildings or vehicles, frequently result in incomplete trunk or crown structures. This incomplete information significantly diminishes segmentation accuracy.
- Algorithmic limitations persist in crown-based methods. Approaches that rely on CHMs, dynamic windows, or region growing are highly sensitive to crown overlap or crown loss, often leading to over-segmentation or under-segmentation.
- Many existing methods perform well only in specific road scenarios or for particular tree species, but their robustness markedly decreases when applied across different cities, vegetation compositions, or LiDAR sampling densities. This lack of universality restricts their applicability in diverse urban environments.

To tackle these challenges, this study aims to develop an urban individual tree segmentation method based on multi-constraint joint shortest path analysis. The primary contributions are as follows:

- Enhancing robustness to noise and incomplete structures. A graph-based object-primitive segmentation approach is first introduced to establish a stable structural foundation. This step mitigates the influence of noise, missing points, and occlusions, ensuring that the subsequent segmentation process can operate reliably even when the point cloud is partially incomplete.
- Improving trunk identification under crown overlap or loss. A multi-constraint trunk extraction strategy is then devised by integrating tree-vertex information, aspect-ratio constraints, and linear-feature characteristics. This design facilitates accurate trunk detection even when crown data are missing or heavily overlapped, thereby reducing both over-segmentation and under-segmentation that commonly plague crown-based methods.
- Strengthening generalization across diverse environments and datasets. Shortest-path analysis is combined with point-to-stem-axis distance to associate crown points with their correct trunks. This mechanism enhances segmentation reliability across different road conditions, tree species, and LiDAR acquisition densities, thereby improving the overall generalization capability of the method.

## 2. Methodology

In general, ground point clouds can compromise the extraction accuracy of roadside trees. To address this, this paper initially employs Cloth Simulation Filter (CSF) [30] to effectively separate ground points from non-ground points, and constructs a digital terrain model (DTM) based on ground points to normalize the point cloud. Subsequently, to enhance computational efficiency, the point clouds below 2.5 m from the ground level are extracted and voxelized, and the influence of low shrub point clouds is eliminated by combining point density and height difference thresholds. Finally, cluster analysis is performed on the remaining point clouds to separate tree trunk point cloud clusters and remove small-scale noise clusters, thereby reducing the complexity of subsequent point cloud segmentation and improving overall the implementation efficiency.

Following the aforementioned point cloud preprocessing, this paper proposes a graph segmentation method to construct a network structure towards the non-ground points as shown in Figure 1. This network structure is then integrated with multiple proposed constraints to segment the point cloud and obtain object primitives. Next, tree trunk point clouds and root nodes are extracted based on the spatial features of the object primitives. By calculating the spatial features (including linear features and aspect ratio features) of each object primitive and combining them with tree vertex constraints, point cloud clusters that conform to the trunk features are selected, and root nodes are further located and extracted.

https://doi.org/10.3390/f17010027

Subsequently, connectivity analysis is performed on the preliminary tree trunk point cloud to obtain a complete tree point cloud. Then, combining shortest path analysis and point-to-stem axis distance constraints, the tree point cloud is optimized and segmented, ultimately yielding a structurally complete and independent street tree point cloud. This approach preserves the integrity of the overall tree structure while effectively suppressing noise interference. Overall, the methodology presented in this paper mainly includes the following three steps: (1) Object primitive acquisition based on graph segmentation. (2) Tree trunk point cloud extraction and root node identification based on the spatial features of object primitives. (3) Individual tree point cloud segmentation based on shortest path and point-to-stem axis distance.

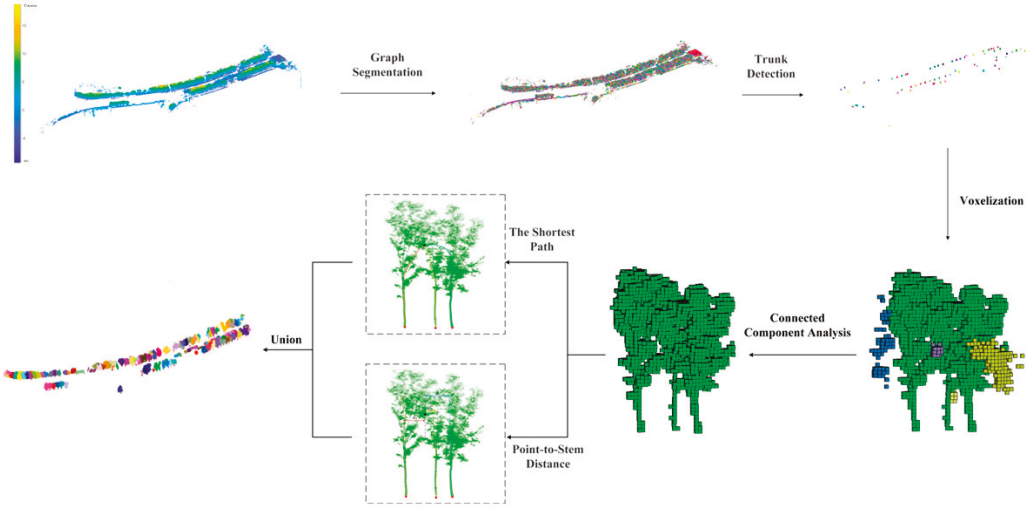



**Figure 1.** The flowchart of the proposed method. The final results of individual tree segmentation are rendered in different colors.

### *2.1. Object Primitive Acquisition Based on Graph Segmentation*

To effectively extract tree trunk point clouds, this study initially employs a graph segmentation approach. The algorithm is based on the geometric primitive segmentation principle [31]. Specifically, a neighborhood graph is constructed as expressed in Equation (1), where each point is treated as a node, and edges are established according to spatial proximity within an adaptive neighborhood radius $R$. Three geometric constraints as shown in Figure 2, namely the neighborhood radius, the deviation of the normal vector angle, and the rate of change in the vertical angle, control the connection weights between nodes. An edge is retained only when all constraint thresholds are satisfied, ensuring that points with similar local geometric properties belong to the same primitive. Subsequently, connected component analysis is performed on the graph $G$ to derive object primitives representing homogeneous geometric elements such as trunks, branches, or crown surfaces. Through this process, non-ground points are segmented into object primitives with similar spatial characteristics.

$$G = (V, E) \tag{1}$$

where $G$ represents the network structure, $V$ is the set of vertices ($v_i$) that constitute the network, $E$ is a collection of edges ($e_{i,j}$). $v_i$ is composed of all the points ($p_i$, $i = 1, 2, ..., N$) in the point cloud, $e_{i,j}$ is the edge formed by any two points ($p_i$, $p_j$).

https://doi.org/10.3390/f17010027

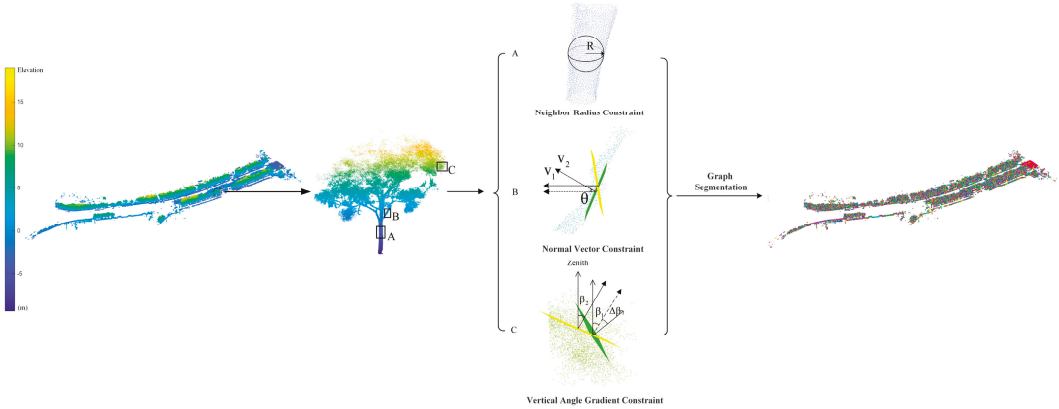



**Figure 2.** Constraints for graph segmentation. In this figure, 'A' represents the neighborhood radius constraint. $R$ is the neighborhood radius. 'B' represents the normal vector angle constraint, where $\theta$ is the angle between the normal vectors $v_1$ and $v_2$ of two neighboring points. And 'C' is the vertical angle change rate constraint, where $\Delta\beta$ is the change in the vertical angles $\beta_1$ and $\beta_2$ between two neighboring points. The right segmentation results are rendered in different colors.

To further clarify the definition of edge weights in the graph segmentation, this study adopts a constraint-based elimination mechanism. During the construction of the neighborhood graph, each potential edge between two nodes is first evaluated under three geometric constraints: the neighborhood radius, the normal vector angular deviation, and the vertical angle change rate. An edge is retained only when all constraints simultaneously satisfy their respective thresholds; otherwise, the edge is directly removed from the graph. This strategy simplifies the weighting process and prevents the introduction of biased cost terms arising from heterogeneous geometric measures. As a result, the retained edges represent stable spatial relationships between points with consistent geometric features, ensuring that subsequent connected component analysis can effectively form object primitives corresponding to homogeneous structures.

Regarding the neighborhood radius constraint, to prevent interference from long edges on the network graph segmentation, this paper uses the neighborhood radius $R$ to restrict the connection relationship between point pairs. The neighborhood radius is obtained by randomly selecting $n$ points in the point cloud and calculating the average maximum distance of their respective $k$ nearest neighbors, where the number of neighboring points $k$ is typically set to 10 [32]. The calculation of the neighborhood radius is illustrated in Equation (2).

$$R = \frac{\sum_{i=1}^{n} \max(d_{i,j})}{n} \tag{2}$$

where $R$ represents the neighborhood radius. $\max(d_{i,j})$ represents the maximum distance between the $i$-th random point and its neighbors. The neighborhood radius can be adaptively adjusted to suit the density and spatial distribution of different point cloud data, thereby improving the automation level of the algorithm.

With regard to the constraint concerning the normal vector angle, to ensure that the points within an object primitive exhibit similar spatial geometric features, this paper calculates the normal vector angle between each point and its adjacent point, and establishes an angle threshold. The calculation of the normal vector angle is presented in Equation (3):

$$cos(\theta) = \frac{v_1 \cdot v_2}{\|v_1\| \|v_2\|} \tag{3}$$

where $\theta$ represents the angle of the normal vector. $v_1$, $v_2$ are the normal vectors of two adjacent nodes, respectively. Typically, the angle between the normal vectors of features such as tree trunks and buildings is relatively small, while the angle between the normal vectors of tree canopies and low vegetation is larger due to their scattered distribution. Therefore, the consistency of normal vectors can be used to segment these features into different object primitives.

Regarding the constraint on the vertical angle change rate, given that tree trunks typically exhibit a columnar shape, relying solely on the normal vector angle threshold may result in under-segmentation. To further enhance the accuracy of tree trunk point cloud extraction, this paper introduces the vertical angle change rate, which characterizes the rate of change in the angle between the normal vector of a point and those of its neighboring points relative to the $z$-axis. The calculation formula is presented in Equation (4):

$$\beta_{rate} = \sqrt{\frac{\sum_{i=1}^{num} \Delta\beta^2}{num}} \tag{4}$$

where $\Delta\beta$ is the difference in vertical angle between the point and those of its neighboring points, and $num$ is the number of neighboring points connected to the point. Neighboring points with a small rate of change in vertical angle, such as tree trunks and building points, can be segmented into the same object primitive, while points with a larger rate of change, such as tree crowns, can be segmented into multiple smaller object primitives.

Under the combined influence of the aforementioned three constraints, the point cloud is segmented into several object primitives, with the tree trunk primitives among them offering a dependable basis for subsequent tree trunk point cloud extraction. To present a clearer overview of the segmentation workflow, the detailed procedure for constructing the neighborhood graph and extracting object primitives is summarized in the algorithm, and the corresponding pseudocode for acquiring object primitives is provided in Algorithm 1.

**Algorithm 1:** The pseudocode of object primitive acquisition based on graph segmentation

**Input**: Point Cloud $P$, Neighbor count $k$, Thresholds $\{T_\theta, T_\beta\}$
Initialize Graph $G = (V, E)$;
Calculate Adaptive Neighborhood Radius $R$;
Select $n$ random points from $P$;
**for** point $p_i \in V$ **do**
Find $k$ nearest neighbors $N_i$ of $p_i$
Calculate Vertical Angle Change Rate $\beta_{rate}$
**for** point $p_j \in N_i$ **do**
Calculate Euclidean distance $d_{ij} = \|p_i - p_j\|$
Calculate Normal Vector Angle $\theta$:
if $d_{ij} \leq R$ and $\theta \leq T_\theta$ and $\beta_{rate} \leq T_\beta$
Add edge $(p_i, p_j)$ to $E$
**end**
**end**
**end**
**Graph Segmentation:**
$O \leftarrow$ Connected Components $(G)$
**Output:** Set of Object Primitives O

https://doi.org/10.3390/f17010027

### 2.2. Tree Trunk Point Cloud Extraction and Root Node Identification Based on Object Primitive Spatial Features

After conducting threshold segmentation on the point cloud and acquiring the object primitive, this paper proceeds to identify object primitives based on the geometric features of the tree trunk to further extract the tree trunk point cloud. The tree trunk manifests as an independent cylinder, with its spatial features primarily marked by pronounced linearity and a substantial aspect ratio. Consequently, this paper analyzes the tree trunk from two perspectives, namely linearity and aspect ratio.

#### 2.2.1. Linear Feature Extraction

The linear characteristics of the object primitive are obtained by calculating the eigenvalues of its covariance matrix, as illustrated in Equation (5):

$$linearity_{obj} = (\lambda_1 - \lambda_2)/\lambda_1 \tag{5}$$

where $\lambda_i$ is the eigenvalue of the object primitive covariance matrix, $\lambda_1 > \lambda_2 > \lambda_3 > 0$. The magnitude of the eigenvalue reflects the degree of dispersion of the point cloud in the corresponding direction. Taking a single plant as an example, compared to the canopy point cloud, the eigenvalue $\lambda_1$ of the trunk and branches is notably larger than $\lambda_2$. Consequently, the trunk and branches usually have higher linear values, whereas the canopy displays lower linear values.

#### 2.2.2. Aspect Ratio Feature Extraction

To further differentiate tree trunks from other non-ground objects, this paper incorporates the aspect ratio feature of object primitives. Tree trunks are generally tall and slender cylinders, with their height markedly exceeding their base width. Consequently, their aspect ratio is generally greater than that of other objects, such as building facades, low vegetation and vehicles. The aspect ratio is defined as the ratio of the height of the smallest bounding box of an object primitive to its base width, calculated using the following Equation:

$$ratio_{obj} = \frac{h_{obj}}{\sqrt{(X_{\max} - X_{\min})^2 + (Y_{\max} - Y_{\min})^2}} \tag{6}$$

where $h_{obj}$ is the height of the object primitive, calculated from the elevation difference between the highest and lowest points of the object primitive. $X_{\max}$, $X_{\min}$, $Y_{\max}$, $Y_{\min}$ represent the value range of the object primitive on the $X$ and $Y$ axes, respectively. The height of the tree trunk object primitive is usually greater than its base width, while the aspect ratio of other objects such as buildings, low vegetation and vehicles is usually flat or slender cuboid, much smaller than that of the tree trunk. Therefore, by setting an appropriate aspect ratio threshold, the tree trunk point cloud can be accurately extracted.

To precisely identify tree trunks, threshold-based criteria are applied to the extracted linearity and aspect ratio features. An object primitive is designated as a tree trunk candidate if its linearity surpasses a predetermined threshold and its aspect ratio exceeds a specified value. In this study, the linearity threshold is set at 0.65 to denote strong linear structures, while the aspect ratio threshold is established at 0.85 to represent tall and slender vertical objects. Object primitives that fail to meet either criterion are classified as non-trunk features and excluded from subsequent processing. These thresholds are empirically determined based on the geometric characteristics of tree trunks and validated through extensive experimentation. They effectively minimize interference from canopy points, buildings, low-lying vegetation, and other objects.

https://doi.org/10.3390/f17010027

2.2.3. Tree Vertex Constraints

After extracting tree trunk point clouds based on linear and aspect ratio features, this paper further incorporates tree vertex constraints to facilitate effective association between tree trunks and canopies.

First, points that exceed a preset height threshold are selected to form a candidate set of tree vertices. These vertices are subsequently subjected to voxelization. Following voxelization, connectivity analysis is conducted to retain only major connected regions with more points than the threshold, thereby eliminating isolated point cloud clusters and mitigating interference from false treetops. Subsequently, within the candidate set, a K-D tree neighborhood search is employed to extract local height maxima. A non-maximum suppression strategy is then employed to ensure that each canopy region retains only one unique treetop candidate point. The core formulas are presented in Equations (7) and (8).

$$P_{top} = \{ p_i \mid z_i > H_{th}, z_i = \underset{p_j \in N_k(p_i)}{max} (z_j) \} \tag{7}$$

$$P_{\text{final}} = \left\{ q_i \in P_{top} \mid \left\| q_i - q_j \right\|_2 > r_{NMS}, \forall q_j \in P_{\text{final}} \right\} \tag{8}$$

where $P_{top}$ is the candidate set of tree vertices that satisfy the height threshold and local maximum conditions, $p_i$ is the $i$-th point in the point cloud, $z_i$ is the $z$ coordinate corresponding to $p_i$, max( $\bullet$) represent the selection of local peaks, $N_k(p_i)$ is the set of $K$ nearest neighbors centered at $p_i$. $P_{\text{final}}$ is the final set of unique tree vertices, $q_i$ is the set of tree vertices after filtering, $\| \ \|_2$ is the Euclidean distance calculation, $r_{\text{NMS}}$ is the non-maximum suppression radius, which controls that only one vertex is retained in a tree canopy region, and $q_j$ is the currently selected final tree vertex.

Finally, the treetop candidate points are matched with candidate tree trunks. Based on geometric constraints such as horizontal distance, tree height structure, and spatial relative position, the system calculates a matching score. In cases where multiple root points are present, the root point with the highest matching score is retained. This ensures the correct correspondence between treetops and trunks, thereby enhancing the accuracy and robustness of individual tree identification.

As shown in Figure 3, the tree trunk extraction and tree vertex constraint based on spatial features can effectively associate the tree trunk and the tree crown, providing reliable and quantifiable initial conditions for subsequent fine segmentation of tree point clouds based on the shortest path and the distance from the point to the stem axis.

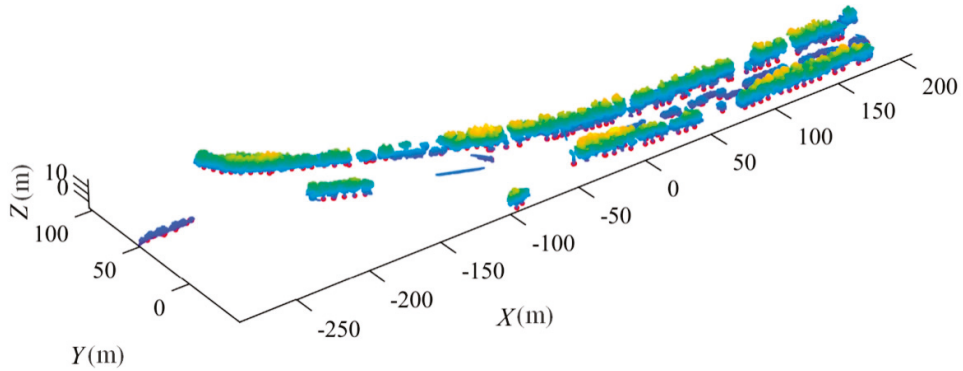



**Figure 3.** Illustration of the treetop constraint process. The point cloud shows the spatial distribution of points after applying height thresholding, voxelization, and connectivity filtering. Local height maxima extracted through K-D tree neighborhood search are retained as treetop candidates after non-maximum suppression. The red points represent the detected trunk nodes used to associate each treetop candidate with its corresponding tree trunk.

https://doi.org/10.3390/f17010027

### *2.3. Individual Tree Point Cloud Segmentation Based on Shortest Path and Point-to-Stem Axis Distance*

After completing the extraction of the trunk point cloud and obtaining the complete point cloud of an individual tree, this paper further introduces a connectivity analysis method. Specifically, the non-ground point cloud is initially voxelized. Subsequently, it is determined whether each voxel contains point cloud data, with voxels containing points assigned a value of 1 and empty voxels assigned a value of 0. Next, using the voxel containing the tree trunk point as the seed, connected voxel blocks are identified, and the point clouds within them are extracted as the initial tree point cloud.

After obtaining the initial results, this paper employs two optimization strategies to segment the tree point cloud: one is tree point cloud segmentation optimization based on the shortest path, and the other is tree point cloud segmentation optimization based on the distance from the point to the stem axis. Finally, the intersection of the segmentation results obtained by the two optimization methods is taken as the final tree point cloud segmentation result of this study.

By integrating voxel connectivity with multiple optimization strategies, this method effectively avoids over-segmentation and under-segmentation issues caused by interference from nearby vegetation or noise points, while enhancing the accuracy and robustness of the segmentation results and ensuring the integrity of the overall tree structure. The overall processing flow is depicted in Figure 4.

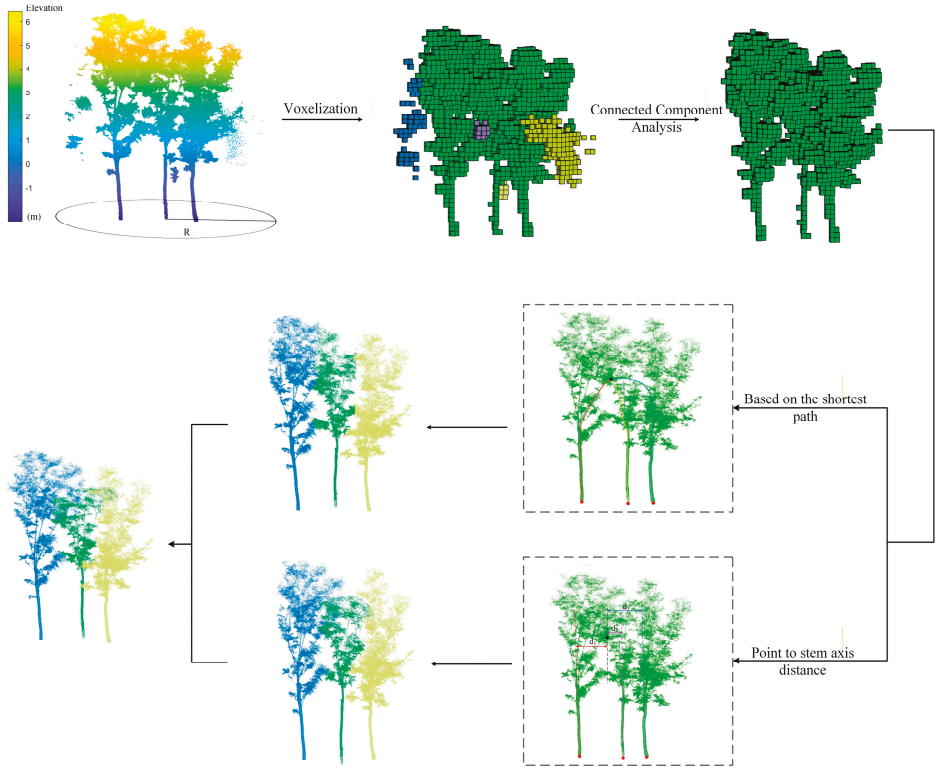



**Figure 4.** Flowchart of the individual tree point clouds segmentation.

#### 2.3.1. Tree Point Cloud Segmentation Based on Shortest Path

In real-world scenarios, tree canopies frequently intersect and overlap (as illustrated in Figure 5). This can result in under-segmentation of the tree point cloud obtained through connectivity analysis, thereby compromising the accuracy of subsequent vegetation parameter extraction and modeling. To address this issue, this paper introduces a point cloud segmentation method based on shortest path analysis to refine the initial tree point cloud.

https://doi.org/10.3390/f17010027



(**a**) (**b**)

**Figure 5.** Optimized segmentation of tree point clouds based on the shortest path. (**a**) The individual tree point clouds before optimized segmentation. (**b**) The result after optimized segmentation. Red dots represent identified root nodes, and curves of different colors represent the shortest distances from that point to the root nodes of adjacent trees.

Specifically, this method initially checks whether a tree point cloud set contains two or more lowest points of the trunk primitives. If it does, a connected graph is constructed using the k-nearest neighbor method to depict the spatial topological relationships between the point clouds. Subsequently, the shortest path length from each node to each lowest point of the trunk is calculated. Theoretically, the shortest path from a point in the canopy to the lowest point of its own trunk should be less than the shortest path from that point to the lowest points of other trunks. Consequently, nodes are allocated to the trunk category corresponding to their shortest paths, thereby achieving optimized segmentation of the tree point cloud.

As depicted in Figure 5b, this method can effectively distinguish between roadside trees with intersecting and overlapping canopies, enabling precise segmentation of multiple trees. Compared to methods that rely solely on connectivity, the shortest path-based optimized segmentation not only improves the accuracy and robustness of segmentation but also significantly enhances the separation of intersecting and overlapping trees, providing a reliable data foundation for subsequent extraction of tree morphological features.

#### 2.3.2. Tree Point Cloud Segmentation Based on Point-to-Stem Axis Distance

In Section 2.2, this paper extracts the tree trunk point cloud based on the spatial features of object primitives. Building upon this, Principal Component Analysis (PCA) is employed to identify the principal orientations of the tree trunk point cloud, enabling the calculation of the stem axis. After identifying and isolating the base of the tree trunk, the axis of each trunk can be established through the orientation of the first principal component derived from PCA, thereby serving as the spatial representation of the stem axis (as illustrated in Figure 6).

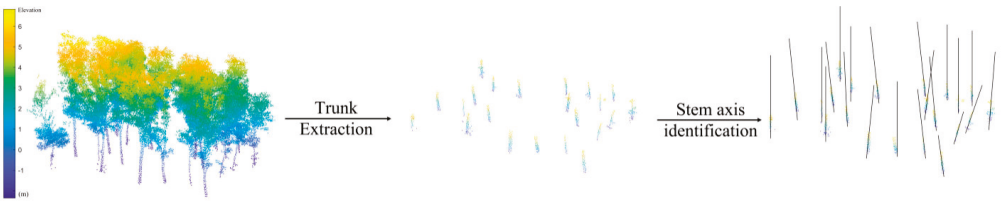



**Figure 6.** Identification of tree trunk axis. The black straight lines depict the stem axes extracted via PCA, serving as the orientation reference for individual trees.

Subsequently, the Euclidean distance from each point to each trunk axis is calculated. Theoretically, the distance from a point within a tree canopy to the trunk axis of

its corresponding tree should be shorter than the distance from that point to the trunk axes of other trees. Based on this principle, this paper assigns labels to points across the entire point cloud according to the distance of each point to these axes, thereby assigning each point to a corresponding tree and achieving optimized segmentation of the point cloud of the entire tree. The segmentation results are shown in Figure 7. This method can effectively distinguish street trees with three intersecting and overlapping canopies, enabling precise separation of multiple trees. Specifically, street trees with three intersecting and overlapping canopies were successfully segmented. Compared to methods that rely solely on connectivity, the segmentation strategy based on the distance from a point to the stem axis better reflects the geometric structural characteristics of the trees. It also provides higher stability and segmentation accuracy in scenarios involving intersecting and overlapping canopies.

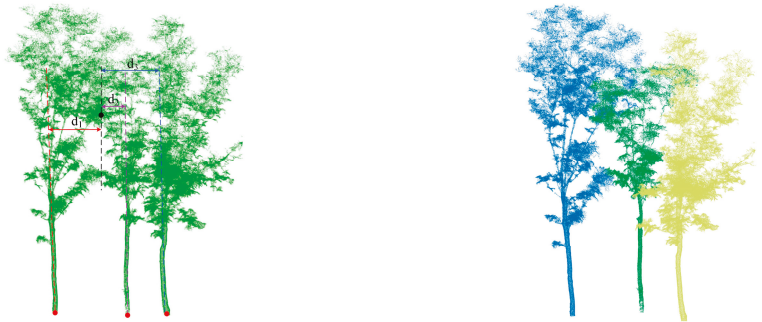



**Figure 7.** Optimized segmentation of tree point clouds based on point-to-stem axis distance. (**a**) The individual tree point clouds before optimized segmentation. (**b**) The result after optimized segmentation. $d_i$ is the distance between the two trees.

Finally, to further enhance the reliability of the segmentation results, this paper integrates the optimization results based on the shortest path and those based on stem-axis distance, taking their intersection as the final segmentation result (Figure 8). The intersection strategy can effectively prevents the retention of redundant point clouds caused by misclassification by a single method, and only retain the regions jointly identified by the two methods, thereby improving the robustness and reliability of the segmentation results while ensuring the integrity of the overall tree structure.

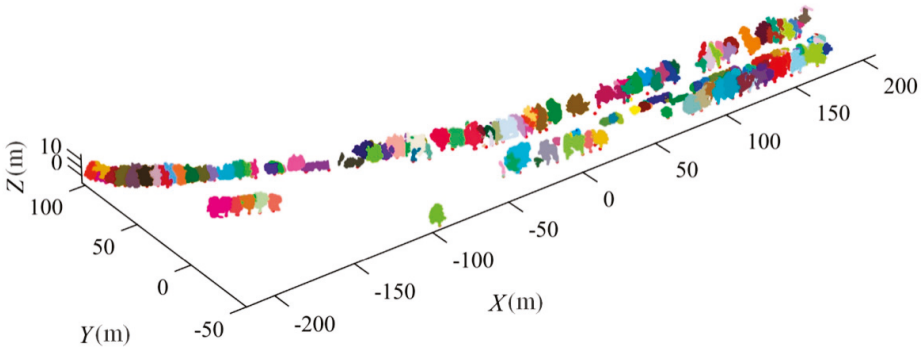



**Figure 8.** Optimized segmentation results. Each individual segmented trees are in different colors, illustrating the effectiveness of the proposed method in separating densely distributed trees across the study area.

https://doi.org/10.3390/f17010027

### 2.4. Accuracy Assessment

To quantitatively evaluate the performance of the proposed method, this paper selects precision (*Prec*), Recall (*Rec*), F1-score (F1 − score), coverage (*Cov*) and Weighted Coverage (*WCov*), which are commonly used in point cloud segmentation instance, as evaluation indicators with IoU > 0.75 as the threshold. The specific formulas are presented in Equations (9)–(13).

$$Prec = \frac{TP}{TP + FP} \tag{9}$$

$$Rec = \frac{TP}{TP + FN} \tag{10}$$

$$\text{F1} - \text{score} = \frac{2Prec \cdot Rec}{Rec + Prec} \tag{11}$$

$$Cov(S_i) = \frac{\left|\hat{S}_i \cap T_i\right|}{|T_i|} \tag{12}$$

$$WCov = \frac{\sum_i w_i \times Cov(S_i)}{\sum_i w_i} \tag{13}$$

where $TP$ represents the number of trees successfully segmented under the threshold of IoU > 0.75. $FP$ represents the number of individual trees that have been incorrectly segmented or missed. $FN$ represents the number of individual trees that were wrongly segmented. $S_{\text{i}}$ represents the $i$-th segmented tree. $\hat{S}_i$ represents the number of point clouds in $S_{\text{i}}$. $T_i$ represents the number of point clouds contained in the ground truth instance of an individual tree. $\hat{S}_i \cap T$ represents the number of intersections between an individual tree and its true value, i.e., the number of correctly segmented point clouds. $w_i$ represents the weight of the $i$-th individual tree, which is the weight of the instance based on the number of ground truth points in the tree instance's point cloud.

It is important to note that the experimental unit in this study is the "individual tree", which constitutes the primary focus of our individual trees segmentation. Specifically, when referring to true positive (*TP*), it is indicating the count of individual trees that have been accurately segmented. Object primitives, on the other hand, serve as intermediate geometric constructs employed during the segmentation process to streamline the extraction of individual trees from point cloud datasets. These primitives are not the ultimate experimental unit; rather, they function as instrumental means to attain precise tree segmentation. Additionally, pixels are not applicable in this study since this paper is working with 3D point cloud data rather than 2D images.

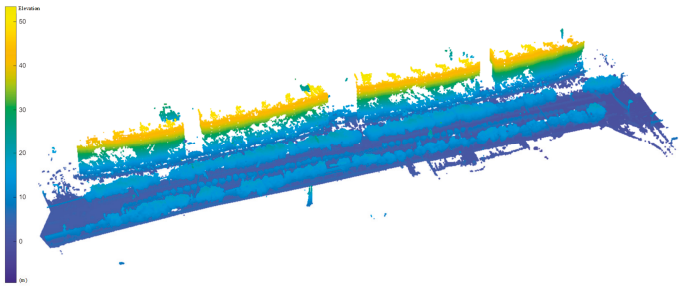
## 3. Experimental Results and Analysis

### 3.1. Experimental Dataset

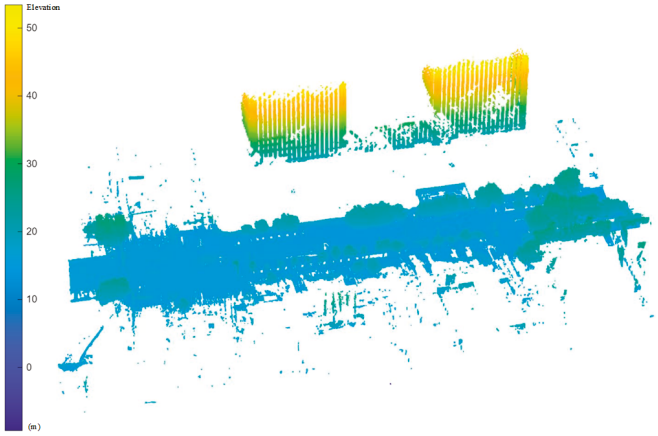
To evaluate the effectiveness of the proposed method, this paper use the WHU-Stree dataset [33] for testing. This publicly available dataset integrates vehicle-mounted LiDAR point cloud data acquired from two cities in China, Shenyang and Nanjing, which are located in different climate zones and exhibit distinct urban structures and vegetation compositions. The dataset covers approximately 66 km of urban roads and contains a total of 21,007 individual tree instances, encompassing more than 50 tree species and including two morphological parameters: tree height and diameter at breast height (DBH). The dataset can be accessed at https://github.com/WHU-USI3DV/WHU-STree (accessed on 1 July 2025).

In this study, three representative road sections were selected as experimental areas for method validation and comparative testing. These areas were deliberately chosen to

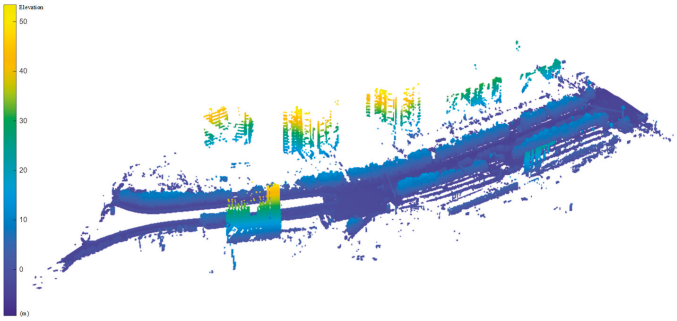
encompass varying levels of tree density, canopy overlap, and environmental heterogeneity, thereby ensuring that the evaluation results accurately reflect diverse street conditions. The selected sections also exhibit differences in point density and occlusion intensity caused by surrounding urban infrastructure, offering a more thorough assessment of the method's robustness. The three experimental areas employed in this paper are illustrated in Figure 9, which displays small yet representative segments of the overall dataset.



(**a**) Site 1




(**b**) Site 2



(**c**) Site 3

**Figure 9.** Three experimental sites. (**a**) Site 1. (**b**) Site 2. (**c**) Site 3. Specifically, Site 1 contains 178 individual trees with a total of 22,064,482 points; Site 2 contains 64 trees with 12,291,874 points; and Site 3 contains 137 trees with 20,914,084 points.

### *3.2. Experimental Results*

Figure 10 presents the street tree extraction results across the three experimental areas. As can be seen from the figure, the proposed method achieves favorable results in street tree segmentation for all three experimental sites (Site 1, Site 2, and Site 3), with different colors representing different individual trees. The visualization clearly illustrates that most tree point clouds are fully and distinctly separated, featuring well-defined boundaries between trees and effectively preventing the adhesion between point cloud clusters.

https://doi.org/10.3390/f17010027




**Figure 10.** Results of street tree segmentation in three experimental areas using our method. Randomly assign colors to represent different segmented individual trees.

In Site 1, the proposed method effectively segments continuous roadside trees on both sides of the road. In Site 2, despite localized dense tree distribution, the method still accurately distinguishes individual trees, demonstrating its strong ability to segment individual trees. In Site 3, even with curved roads and irregularly distributed roadside trees, the proposed method maintains good segmentation results, highlighting its adaptability and robustness in complex scenarios. The segmentation visualization results confirm that the proposed method can effectively segment roadside trees in different road scenarios, exhibiting high accuracy and robustness.

### *3.3. Comparative Analysis*

To objectively evaluate the performance of the proposed method, this paper selects two other classic tree segmentation methods for comparative testing. Wang et al. [34] proposed an unsupervised method based on the superpoint graph structure single tree classification (SSSC). This method first clusters the point cloud through geometric partitioning, dividing the point cloud into several super-points. Each super-point is regarded as a primitive in the graph structure and carries a fixed-size feature of its partition. Then, an undirected graph is constructed based on these super-points, and node segmentation is performed using the features of the super-points and shortest path analysis. Finally, by segmenting the nodes in the super-point graph, each super-point cluster corresponds to a single tree.

Peng et al. [35] proposed a single-tree segmentation method based on Nyström spectral clustering (NSC). First, the point cloud is scaled along the *Z*-axis to avoid the influence of height differences. Then, the point cloud is voxelized and divided into blocks using MeanShift clustering to obtain preliminary spatial cluster centers. Next, it is further divided using Nyström spectral clustering to determine the correspondence between each cluster center and the single-tree point cloud. Finally, each point is reassigned with the corresponding tree label, and parameters such as crown width and tree height are calculated to achieve refined segmentation and attribute extraction of single trees.

Figure 11 provides a detailed comparison of the segmentation results from the three methods. The comparison reveals that the proposed method outperforms both the NSC and SSSC methods in terms of the completeness and accuracy of boundary handling.

https://doi.org/10.3390/f17010027

Specifically, both methods exhibit significant limitations in addressing the canopy adjacency problem. The NSC method (Figure 11d,h,l) cannot solve the problem of canopy point cloud adhesion. In densely wooded areas, multiple individuals are often misclassified as the same tree, resulting in widespread under-segmentation errors. The SSSC method (Figure 11c,g,k) is limited by its difficulty in handling dense canopies and background noise, leading to severe under-segmentation in Site 2 and Site 3, and erroneously segmenting streetlight poles into individual trees in Site 1.

In contrast, the proposed method (Figure 11b,f,j) effectively overcomes these challenges in individual tree segmentation. Even in complex situations with severely intersecting canopies, it can accurately delineate the boundaries of adjacent trees, thereby avoiding under-segmentation and ensuring the accuracy of the segmentation results. Furthermore, considering the details of the trunk and canopy, the proposed method not only maintains the continuity of the trunk point cloud well but also achieves relatively accurate segmentation in areas with intersecting canopies. This demonstrates that the proposed method can balance segmentation accuracy and integrity when dealing with complex urban road scenes, exhibiting higher applicability and stability.

Table 1 presents a comparative analysis of the accuracy metrics point clouds contained in the ground truth instance of a single tree. the two methods mentioned above and the method proposed in this paper, focusing on street tree extraction across three experimental regions. As can be seen from Table 1, the method proposed in this paper performs better overall in most cases across the five segmentation metrics of street trees in the three experimental regions. Specifically, the F1-score is optimal in all three experimental regions when the IoU is greater than the threshold of 0.75, with scores of 0.772, 0.767, and 0.765, respectively.

In the Site 1 scenario, the proposed method significantly outperforms SSSC and NSC in terms of precision, Recall, and F1-score. Compared to the second-best performing SSSC method, it achieves an improvement of approximately 24% in F1-score. Furthermore, it also demonstrates the best performance in coverage and Weighted Coverage, achieving the highest Overall Accuracy (OA) of 0.6194, which is a substantial gain of 36.0% over SSSC (0.4519) and 89.3% over NSC (0.3277). This marked superiority in OA fully underscores the overall high precision and completeness of our segmentation results in this scenario.

**Table 1.** Comparison of accuracy indicators of three methods.

| Sites | Methods | Precision | Recall | F1-Score | Coverage | Weighted Coverage | Overall Accuracy |
|---|---|---|---|---|---|---|---|
| Site 1 | NSC | 0.483 | 0.62 | 0.542 | 0.702 | 0.752 | 0.3727 |
| | SSSC | 0.522 | 0.771 | 0.623 | 0.786 | 0.745 | 0.4519 |
| | Ours | 0.762 | 0.782 | 0.772 | 0.855 | 0.860 | 0.6285 |
| Site 2 | NSC | 0.558 | 0.504 | 0.53 | 0.584 | 0.701 | 0.3602 |
| | SSSC | 0.584 | 0.609 | 0.596 | 0.752 | 0.763 | 0.4247 |
| | Ours | 0.735 | 0.802 | 0.767 | 0.822 | 0.842 | 0.6221 |
| Site 3 | NSC | 0.539 | 0.539 | 0.539 | 0.748 | 0.771 | 0.3689 |
| | SSSC | 0.646 | 0.875 | 0.743 | 0.778 | 0.833 | 0.5914 |
| | Ours | 0.771 | 0.759 | 0.765 | 0.731 | 0.713 | 0.6194 |

https://doi.org/10.3390/f17010027

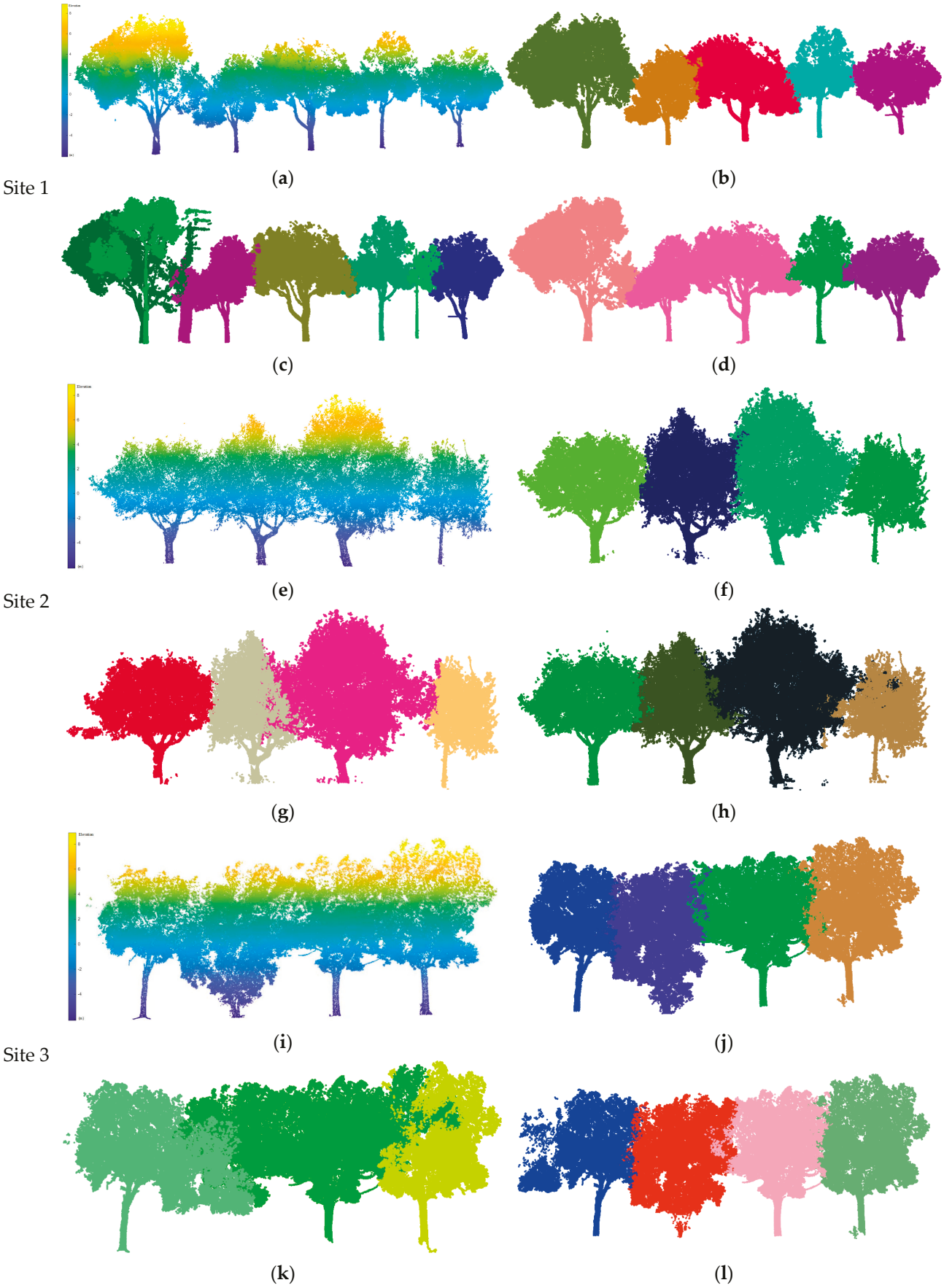



**Figure 11.** Segmentation details comparison. (**a**,**e**,**i**) are the details of the three experimental areas, (**b**,**f**,**j**) are the segmentation results of this paper, (**c**,**g**,**k**) are the NSC segmentation results, and (**d**,**h**,**l**) are the SSSC segmentation results.

https://doi.org/10.3390/f17010027

For the Site 2 scenario, our proposed method outperforms the other two methods in F1-score, improving upon NSC and SSSC by 0.237 and 0.171, respectively. Simultaneously, our method also outperforms the other two methods in Recall, coverage, and Weighted Coverage and achieves the highest Overall Accuracy of 0.6221, surpassing SSSC by 29.9% and NSC by 73.7%. These results indicate that, in this sample site, the proposed method delivers superior segmentation accuracy while maintaining high Recall, thereby demonstrating excellent comprehensiveness and balance.

In the Site 3 scenario, although SSSC performed exceptionally well in Recall, its relatively low precision led to a high false positive rate. In contrast, the proposed method outperformed both SSSC and NSC in terms of precision and F1-score, and attained the highest OA of 0.6194, exceeding SSSC (0.5914) and NSC (0.3689) by 4.7% and 67.9%, respectively. However, the Weighted Coverage at Site 3 was lower, primarily due to the increased structural complexity of individual trees and the greater heterogeneity of the surrounding environment. In contrast to Sites 1 and 2, this site contains a substantially higher proportion of trees with heavily overlapping crowns, irregular and asymmetric branching patterns, and pronounced occlusion caused by adjacent street infrastructure. Because the Weighted Coverage metric assigns larger weights to trees represented by denser point clouds, any omission or misclassification affecting these large-crown trees produces a disproportionate influence on the final score. Consequently, despite the model's stable performance on smaller or less occluded trees, the compounded effects of crown overlap, structural irregularity, and environmental occlusion introduce more segmentation errors in high weight trees, ultimately leading to the reduction in Weighted Coverage. Nevertheless, the superior OA achieved by our method demonstrates that it effectively balances accuracy and completeness of identification, avoids excessive false positives, and maintains a more robust overall segmentation performance even under challenging urban conditions.

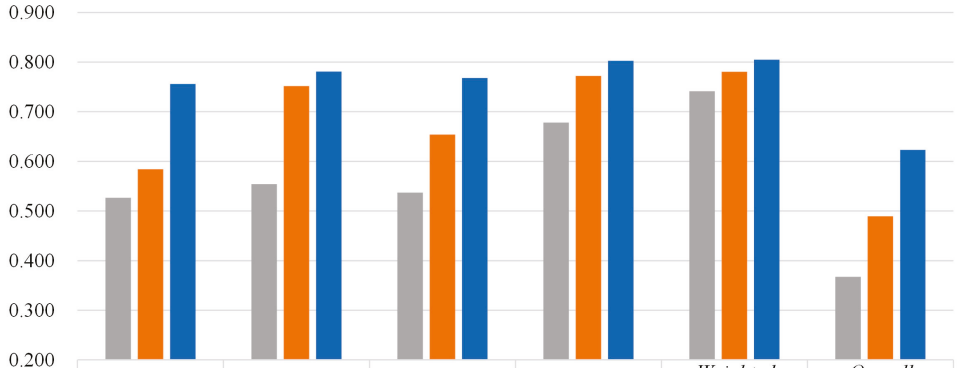
The elevated Recall of SSSC in several sites primarily derives from its tendency to over-aggregate superpoints. SSSC first partitions the point cloud into geometric superpoints and then merges them using graph-connectivity cues and shortest-path similarity. In environments characterized by dense canopy overlap or weak trunk-crown separation, this merging scheme frequently fuses neighboring crowns or elongated linear structures into a single entity. Consequently, SSSC captures a large portion of true tree points, yielding high Recall, but at the expense of incorporating substantial non-tree points or collapsing multiple trees into the same cluster, which in turn degrades precision and the overall F1-score. This issue is particularly evident in Site 3, where severe crown occlusion drives SSSC's Recall to 0.875 while substantially lowering its precision to 0.646. NSC, in contrast, exhibits an opposite pattern of failure. As it relies on spatial clustering of voxelized blocks followed by Nyström-based spectral refinement, NSC is more susceptible to height discontinuities and block boundary artifacts. It therefore tends to under-segment trees with irregular crown morphology or leaning stems, leading to reduced Recall but comparatively stable precision. Our method achieves a more balanced segmentation performance. By integrating multi-constraint graph partitioning with dual-stage optimization, it suppresses both over- and under-segmentation. Although minor omissions persist under extreme occlusion, the proposed approach attains consistently higher F1-score and coverage by preserving inter-tree boundaries while maintaining structural completeness. Across all three evaluation areas, it surpasses competing methods in average accuracy, F1-score, and coverage. In particular, the method demonstrates strong stability and adaptability across diverse roadway conditions, validating its effectiveness and robustness.

Our method achieves a more balanced segmentation performance. By integrating multi-constraint graph partitioning with dual-stage optimization, it suppresses both over- and under-segmentation. Although minor omissions persist under extreme occlusion,

https://doi.org/10.3390/f17010027

the proposed approach attains consistently higher F1-score and Coverage by preserving inter-tree boundaries while maintaining structural completeness. Across all three evaluation areas, it surpasses competing methods in average accuracy, F1-score, and Coverage. In particular, the method demonstrates greater stability and adaptability across diverse roadway conditions, validating its effectiveness and robustness.

Figure 12 compares the mean values of various precision metrics for the three methods. The results show that our proposed method achieves the best performance across all five evaluation metrics. Specifically, *Precision_mean*, *Recall_mean*, and *F1-score_mean* reach 0.756, 0.781, and 0.768, respectively, significantly outperforming NSC (0.527, 0.554, 0.537) and SSSC (0.584, 0.752, 0.654). This indicates that our proposed method maintains high *precision* while also considering *Recall*, resulting in a more balanced and robust segmentation outcome. Our proposed method also performs excellently in *Coverage_mean* and *Weighted Coverage_mean*, reaching 0.803 and 0.805, respectively, representing improvements of approximately 18.5% and 8.6% over NSC, and approximately 4.0% and 3.2% over SSSC. In addition, our method attains the highest *Overall Accuracy_mean*, achieving 0.623 compared with 0.367 for NSC and 0.489 for SSSC. This substantial improvement further confirms the superior and more reliable segmentation performance of our approach across different evaluation criteria.




| | *Precision_mean* | *Recall_mean* | *F1_score_mean* | *Coverage_mean* | *Weighted Coverage_mean* | *Overall Accuracy_mean* |
|---|---|---|---|---|---|---|
| NSC | 0.527 | 0.554 | 0.537 | 0.678 | 0.741 | 0.367 |
| SSSC | 0.584 | 0.752 | 0.654 | 0.772 | 0.780 | 0.489 |
| Ours | 0.756 | 0.781 | 0.768 | 0.803 | 0.805 | 0.623 |

**Figure 12.** Comparison of the mean values of each accuracy indicators of the three methods.

Overall, in the three experimental regions with an IoU greater than 0.75, our method outperforms both comparative methods in key metrics such as accuracy, F1-score, and coverage. The experimental results prove that our method not only demonstrates strong competitiveness in overall segmentation accuracy, but also shows greater stability and adaptability when dealing with road scenarios featuring diverse data characteristics. This validates its effectiveness and robustness in complex urban environments.

### *3.4. Visual Analysis of Segmentation Errors*

Although the shortest-path and PCA-based trunk axis estimation techniques employed in this paper can accurately segment individual trees in numerous instances, both methods still encounter separation errors in complex street environments. To better elucidate this issue, Figure 13 showcases four representative cases, encompassing the ground truth distribution, two types of incorrect segmentations, and the correct result achieved by the proposed method. This comparison aids in clarifying the origins of segmentation errors and highlights the advantages of the proposed approach.

https://doi.org/10.3390/f17010027

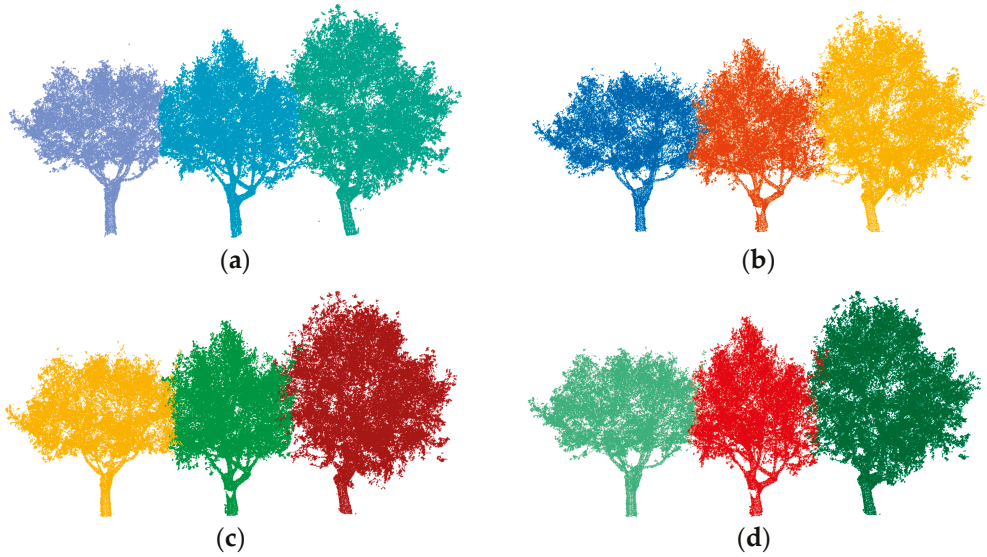



**Figure 13.** Comparison of individual tree segmentation results by three different methods. (**a**) Ground-truth segmentation result. (**b**) Shortest-path segmentation result. (**c**) PCA-based trunk-axis–estimation segmentation result. (**d**) The proposed method segmentation result.

Figure 13a displays the actual tree distribution, featuring clearly and well-separated tree crowns. It functions as the visual reference standard for evaluating the accuracy of subsequent segmentation results.

Figure 13b underscores the typical failure mode of the shortest-path method. The algorithm wrongly connects points that belong to two adjacent trees, causing the stem skeleton to deviate from the true trunk path. The results in partial merging of crown regions or their assignment to the wrong tree, creating a noticeable discrepancy compared to the ground truth are shown in Figure 13a. Such errors mainly arise when two trunks are in close proximity, branches intersect, or the point density is uneven. Given that the shortest-path method solely depends on geodesic distance, it is highly susceptible to local occlusion and irregular sampling, which makes it likely to choose an incorrect path.

Figure 13c illustrates the typical error generated by PCA-based trunk axis estimation. The PCA axis incorrectly inclines towards high-density branches or off-axis structures, particularly when the point cloud around the trunk is uneven or the trunk has curvature. The resulting splitting plane deviates from the true stem orientation, leading to boundary shifts and mis-segmentation. When compared to the ground truth in Figure 13a, it is clear that adjacent crowns are not properly separated, and several segments extend into neighboring trees. In real street tree datasets, leaning stems, asymmetric cross-sections, and uneven point sampling can easily distort the principal components, resulting in unstable segmentation outcomes.

Figure 13d demonstrates that the proposed method generates segmentation results that closely resemble the ground truth. Each tree is clearly demarcated, and the crown boundaries align well with the actual tree structures. By incorporating neighborhood constraints, angular filtering, and structural cues, the method maintains the continuity of the trunk and the natural transition between the trunk and the crown. The subsequent refinement using shortest-path optimization and stem-axis distance analysis further eliminates most of the remaining misalignments. As a consequence, cross-assignments between neighboring trees are significantly reduced, even in areas with dense foliage and overlapping crowns.

https://doi.org/10.3390/f17010027

## 4. Discussion

### 4.1. The Limitations of the Proposed Method

The proposed multi-constraint and shortest-path segmentation method exhibits commendable robustness and accuracy across a variety of urban environments. In contrast to classical geometry-based methods, this study enhances structural integrity by integrating multiple spatial constraints and path-based optimization strategies. These combined efforts contribute to a more precise delineation of overlapping crowns and ensure the continuity of trunks.

Despite the fact that the proposed multi-constraint and shortest-path segmentation method attains robust performance in diverse urban settings, several limitations need to be recognized and critically analyzed.

i. Sensitivity to Point Density

The performance of the proposed method is influenced by the density of the point cloud data. Low point density can lead to incomplete delineation of tree structures, particularly for large or irregular canopies. When the point density is low or when data contain severe occlusions and noise, the adaptive neighborhood radius may fail to capture enough local geometric information to preserve crown continuity. This can result in incomplete delineation of large or irregular canopies, especially in cases where trees are partially obscured by urban structures or vehicles. Pre-filtering and data normalization can mitigate this issue but cannot fully compensate for structural gaps in the source data.

ii. Influence of Scanning Trajectory

The scanning trajectory of the LiDAR system can significantly impact the completeness and accuracy of the point cloud data, especially concerning tree crowns. Given that the performance of the proposed method is contingent upon the quality of the LiDAR data, as previously discussed, it follows that the scanning trajectory will also influence the final outcomes of individual tree segmentation. To optimize the extraction of individual trees, the scanning trajectory should be carefully planned to minimize occlusions as much as possible. Additionally, maintaining a slow and uniform scanning speed is crucial. By adhering to these guidelines, the acquired point clouds will be of high quality, with occlusions minimized to the greatest extent possible.

iii. Dependence on the Quality of Stem Detection

In this study, the quality of stem detection directly affects the accuracy of individual tree segmentation. When stem points are correctly identified and localized, the subsequent steps in the segmentation pipeline, such as connectivity analysis, shortest path optimization, and point-to-stem axis distance calculation can proceed more effectively. This leads to more precise delineation of tree boundaries and reduces the likelihood of over-segmentation or under-segmentation. Conversely, if stem detection is inaccurate due to factors such as noise, missing points, occlusions, or complex tree architectures, the segmentation results can be severely compromised. For instance, missing stem points can lead to incomplete tree structures, while falsely detected stem points can introduce erroneous associations between crowns and trunks, resulting in mis-segmentation.

### 4.2. Discussion on Parameter Selection and Sensitivity

All parameters used in this paper are summarized in Table 2. As shown in Table 2, the quality of geometric primitive acquisition is jointly controlled by the normal vector threshold, verticality threshold, and neighborhood radius. An adaptive strategy is adopted for the neighborhood radius, which is dynamically determined based on the local point density estimated from each query point and its ten nearest neighbors. In the structural

https://doi.org/10.3390/f17010027

extraction stage, the height threshold is used to distinguish the basal trunk section from the ground. In addition, morphological constraints, including aspect ratio, linearity, and matching radius, collectively influence the accuracy of trunk extraction. The crown radius plays a key role in determining the accuracy of the final individual tree segmentation and canopy delineation.

**Table 2.** Parameter settings used in this paper.

| Parameters | Values |
|---|---|
| Normal Angle Threshold | 0.8 |
| Verticality Threshold | 0.8 |
| Neighborhood Radius | 10 |
| Aspect Ratio Threshold | 0.85 |
| Height Threshold | 2 (m) |
| Linearity Threshold | 0.65 |
| Matching Radius | 3 (m) |
| Crown Radius | 10 (m) |

To systematically evaluate the dependency of the proposed method on its key parameters, specifically the threshold $H$ and crown radius $R$, a sensitivity analysis was conducted. As quantified in Table 3, the algorithm exhibits a high degree of robustness, characterized by marginal fluctuations in F1-score, Weighted Coverage, and Overall Accuracy across the tested parameter space $H \in \{1.5, 2, 2.5\}$, $R \in \{5, 10, 15\}$. This observed insensitivity to parameter variations suggests that the method possesses strong generalization capabilities and is not prone to instability caused by minor parameter perturbations. However, to ascertain the optimal configuration for maximizing segmentation accuracy, a comparative analysis of the local maxima was performed. The empirical results demonstrate that the combination of $H = 2$ m and $R = 10$ m yields the superior balance between feature coverage and precision. Under this configuration, the model achieves peak performance across all indicators, recording an F1-score of 0.772, a Weighted Coverage of 0.86, and an Overall Accuracy of 0.6285. Consequently, the parameters selected here is identified as the empirically optimal setting and is adopted as the default configuration for the subsequent experiments.

**Table 3.** Parameter sensitivity analysis.

| $H$ (m) | $R$ (m) | F1-score | Weighted Coverage | Overall Accuracy |
|---|---|---|---|---|
| 1.5 | 5 | 0.738 | 0.797 | 0.550 |
| 2 | 5 | 0.751 | 0.819 | 0.605 |
| 2.5 | 5 | 0.714 | 0.781 | 0.530 |
| 1.5 | 10 | 0.762 | 0.847 | 0.609 |
| 2 | 10 | 0.772 | 0.860 | 0.629 |
| 2.5 | 10 | 0.711 | 0.761 | 0.529 |
| 1.5 | 15 | 0.742 | 0.757 | 0.574 |
| 2 | 15 | 0.758 | 0.810 | 0.609 |
| 2.5 | 15 | 0.725 | 0.771 | 0.532 |

https://doi.org/10.3390/f17010027

### 4.3. Evaluation of Processing Efficiency and Algorithmic Complexity

To conduct a more in-depth assessment of the computational performance of the proposed framework, we compared its runtime with that of two classical methods, namely SSSC and NSC, across the three experimental sites within the WHU-STree dataset. Site 1 comprises 178 individual trees, with a total of 22,064,482 points; Site 2 encompasses 64 trees, containing 12,291,874 points; and Site 3 consists of 137 trees, with 20,914,084 points.

All experiments were carried out on a workstation configured with a 13th-Gen Intel (R) Core (TM) i7-13700K CPU operating at 3.40 GHz, 64 GB of RAM, and an NVIDIA RTX 3060 GPU, running under the Windows 11 operating system. All the codes were implemented using MATLAB R2023b.

As shown in Table 4, the proposed method requires a longer processing time (approximately 7.5 h per scene on average) compared with the SSSC and NSC methods. This increased computational cost is mainly attributed to the multi-constraint graph construction and shortest-path optimization procedures, which ensure accurate tree delineation even under dense canopy overlap and occlusion. From a theoretical perspective, the method's dominant computational cost arises from two stages: (1) the KD-tree-based k-nearest neighbor search used for graph construction, and (2) the Dijkstra shortest-path computation for tree point assignment.

**Table 4.** Comparison of average runtime among different methods (unit: hours).

| Methods | Site 1 | Site 2 | Site 3 | Time_Mean |
|---|---|---|---|---|
| NSC | 3.3 | 1.7 | 3.1 | 2.7 |
| SSSC | 2.3 | 1.1 | 2.2 | 1.9 |
| Ours | 8.5 | 4.6 | 9.3 | 7.5 |

Despite the relatively long processing time, the proposed method consistently achieves higher segmentation accuracy, with mean F1-score and coverage values surpassing those of the baseline methods. Specifically, the mean F1-score of our method (0.768) is significantly higher than that of NSC (0.537) and SSSC (0.654). These results indicate that the trade-off between computational cost and accuracy is acceptable, particularly for applications requiring high-precision mapping of urban street trees.

## 5. Conclusions

Individual tree segmentation in urban environments is a pivotal step for resource inventorying and vegetation parameter estimation. Nevertheless, current methods for extracting urban street trees continue to encounter challenges stemming from point cloud quality issues, occlusions, and environmental noise.

To tackle these problems, this paper introduces an individual tree segmentation method specifically designed for urban street trees, which is based on the integration of multi-constraint and shortest-path optimization techniques. Initially, the method employs radius constraints, vertical angle constraints, and normal vector constraints for primitive object segmentation. Subsequently, it extracts root nodes by combining the aspect ratio of objects with linear feature detection, and further refines trunk point cloud segmentation through tree vertex optimization. Finally, precise segmentation of tree point clouds is accomplished by integrating the shortest path with the point-to-stem axis distance, thereby effectively enhancing both the completeness and accuracy of individual tree delineation.

An evaluation conducted on three representative regions of the WHU-STree dataset reveals that the proposed method consistently demonstrates improved performance compared to existing approaches. The method achieves F1-score of 0.772, 0.767, and 0.765, respectively, outperforming the NSC and SSSC methods. In comparison with SSSC, the

https://doi.org/10.3390/f17010027

proposed method shows increases in both precision and F1-score while maintaining competitive coverage and Weighted Coverage metrics. These results suggest that the method offers a well-balanced segmentation outcome, improving delineation accuracy without increasing the likelihood of false detections, and showcases reliable adaptability across a variety of urban road conditions.

Future research will concentrate on expanding the framework to encompass a wider range of urban forest compositions, investigating its performance under different point cloud densities and acquisition conditions, and incorporating deep learning-based or hybrid strategies to enhance the delineation of complex tree structures. These advancements are anticipated to facilitate more accurate, scalable, and automated assessments of urban vegetation resources.

**Author Contributions:** Conceptualization, Methodology, Writing–original draft preparation, S.Y.; Investigation, Formal Analysis, Supervision, D.L.; Funding acquisition, Writing–review & editing, Supervision, Z.H.; Formal Analysis, X.X.; Formal Analysis, H.L.; Writing—review & editing, X.C.; Writing—review & editing, F.H.; Resources, Supervision, L.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Outstanding Young Talents Funding of Jiangxi Province (20232ACB213017), Double Thousand Plan of Jiangxi Province (DHSQT42023002), the Natural Science Foundation of Jiangxi Province (20242BAB25176, 20192BAB217010), Funding of National Key Laboratory of Uranium Resources Exploration-Mining and Nuclear Remote Sensing (2025QZ-YZZ-08-1, 2024QZ-TD-26) and the National Natural Science Foundation of China (NSF) (42161060, 41801325) for their financial support.

**Data Availability Statement:** The WHU-STree dataset used in this paper is accessible for download from the following link: https://github.com/WHU-USI3DV/WHU-STree. (accessed on 1 July 2025). The corresponding codes can be acquired after a request.

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

## References

1. Xie, F.; Yang, F.; Wei, H. Urban Tree Extraction Method Based on LiDAR Data and Orthophoto. *Laser Optoelectron. Prog.* **2022**, *59*, 435–442.
2. Yu, Z.; Li, L.; Deng, L. Extraction of Individual Tree DBH and Tree Height by Fusing Terrestrial and UAV Laser Data. *J. Northeast. For. Univ.* **2025**, *53*, 33–40.
3. Morais, S.M.F.; Pereira, A.A.; Oliveira, U.F.D. Inventário Florestal Urbano do município de Botelhos, MG. *Ciência Florest* **2024**, *34*, e71628. [CrossRef]
4. Xu, M.; Zhong, X.; Zhong, R. A method for automatic extraction and individual segmentation of urban street trees from laser point clouds. *Opt. Laser Technol.* **2025**, *180*, 111431. [CrossRef]
5. Pataki, D.E.; Alberti, M.; Cadenasso, M.L. The benefits and limits of urban tree planting for environmental and human health. *Front. Ecol. Evol.* **2021**, *9*, 603757. [CrossRef]
6. Liu, M.; Han, Z.; Chen, Y. Tree Species Classification of LiDAR Data based on 3D Deep Learning. *Measurement* **2021**, *177*, 109301. [CrossRef]
7. Fassnacht, F.E.; Latifi, H.; Stereńczak, K. Review of studies on tree species classification from remotely sensed data. *Remote Sens. Environ.* **2016**, *186*, 64–87. [CrossRef]
8. Liu, Y.; Liang, W. Application of vehicle-mounted laser point clouds in extraction of geometric attributes of urban street tree survey. *Beijing Surv. Mapp.* **2024**, *38*, 1599–1603.
9. Luo, Y.; He, L.; Ma, S. Multimodal LiDAR Enhancement Algorithm Based on Multiscale Features. *Laser Optoelectron. Prog.* **2024**, *61*, 238–246.
10. Shang, G.; Xie, R.; Lei, Z. LiDAR Ground-Segmentation Algorithm Based on Slope Threshold and Convolution Filtering Processing. *Laser Optoelectron. Prog.* **2023**, *60*, 308–317.
11. Hui, Z.; Li, N.; Cheng, P. Single Tree Segmentation Method for Terrestrial LiDAR Point Cloud Based on Connectivity Marker Optimization. *Chin. J. Lasers* **2023**, *50*, 0610002.

12. Xu, D.; Wang, H.; Xu, W.; Luan, Z.; Xu, X. LiDAR Applications to Estimate Forest Biomass at Individual Tree Scale: Opportunities, Challenges and Future Perspectives. *Forests* **2021**, *12*, 550. [CrossRef]
13. Liu, D.; Jiang, Y.; Wang, R. Establishing a citywide street tree inventory with street view images and computer vision techniques. *Comput. Environ. Urban Syst.* **2023**, *100*, 101924. [CrossRef]
14. Luo, H.; Khoshelham, K.; Chen, C. Individual tree extraction from urban mobile laser scanning point clouds using deep pointwise direction embedding. *ISPRS J. Photogramm.* **2021**, *175*, 326–339. [CrossRef]
15. Revelli, R.; Porporato, A. Ecohydrological model for the quantification of ecosystem services provided by urban street trees. *Urban Ecosyst.* **2018**, *21*, 489–504. [CrossRef]
16. Yun, T.; Jiang, K.; Li, G. Individual tree crown segmentation from airborne LiDAR data using a novel Gaussian filter and energy function minimization-based approach. *Remote Sens. Environ.* **2021**, *256*, 112307. [CrossRef]
17. Dubrovin, I.; Fortin, C.; Kedrov, A. An open dataset for individual tree detection in UAV LiDAR point clouds and RGB orthophotos in dense mixed forests. *Sci. Rep.* **2024**, *14*, 21938. [CrossRef]
18. Chang, N.; Liao, Z. Street tree extraction method based on vehicle-borne laser point cloud data. *Beijing Surv. Mapp.* **2023**, *37*, 1617–1622.
19. Yan, Y.; Li, Q.; Li, W. A single tree segmentation method for street trees facing side-looking MLS point clouds. *J. Nanjing For. Univ. (Nat. Sci. Ed.)* **2024**, *48*, 166–174.
20. Mongus, D.; Žalik, B. An efficient approach to 3D single tree-crown delineation in LiDAR data. *ISPRS J. Photogramm.* **2015**, *108*, 219–233. [CrossRef]
21. Cabo, C.; Ordóñez, C.; López-Sánchez, C.A. Automatic dendrometry: Tree detection, tree height and diameter estimation using terrestrial laser scanning. *Int. J. Appl. Earth Obs.* **2018**, *69*, 164–174. [CrossRef]
22. Zhang, K.; Zhang, Y. Street tree extraction method based on vehicle laser scanning technology. *Geomat. Technol. Equip.* **2022**, *024*, 81–85.
23. Li, J.; Cheng, X.; Xiao, Z. A branch-trunk-constrained hierarchical clustering method for street trees individual extraction from mobile laser scanning point clouds. *Measurement* **2022**, *189*, 110440. [CrossRef]
24. Chen, X.; Wang, R.; Shi, W. An Individual Tree Segmentation Method That Combines LiDAR Data and Spectral Imagery. *Forests* **2023**, *14*, 1009. [CrossRef]
25. De Oliveira, P.A.; Conti, L.A.; Neto, F.C.N. Mangrove individual tree detection based on the uncrewed aerial vehicle multispectral imagery. *Remote Sens. Appl. Soc. Environ.* **2024**, *33*, 101100. [CrossRef]
26. Zhen, Z.; Li, X.; Xiu, S. Individual Tree Crown Delineation Using Maker-controlled Region Growing Method. *J. Northeast. For. Univ.* **2016**, *44*, 22–29.
27. Ullah, S.; Adler, P.; Dees, M. Comparing image-based point clouds and airborne laser scanning data for estimating forest heights. *iForest* **2017**, *10*, 273. [CrossRef]
28. Hui, Z.; Cheng, P.; Yang, B. Multi-level self-adaptive individual tree detection for coniferous forest using airborne LiDAR. *Int. J. Appl. Earth Obs.* **2022**, *114*, 103028. [CrossRef]
29. Hua, Z.; Xu, S.; Liu, Y. Individual Tree Segmentation from Side-View LiDAR Point Clouds of Street Trees Using Shadow-Cut. *Remote Sens.* **2022**, *14*, 5742. [CrossRef]
30. Zhang, W.; Qi, J.; Wan, P.; Wang, H.; Xie, D.; Wang, X.; Yan, G. An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. *Remote Sens.* **2016**, *8*, 501. [CrossRef]
31. Tao, S.; Wu, F.; Guo, Q. Segmenting tree crowns from terrestrial and mobile LiDAR data by exploring ecological theories. *ISPRS J. Photogramm.* **2015**, *110*, 66–76. [CrossRef]
32. Becker, C.; Rosinskaya, E.; Häni, N. Classification of aerial photogrammetric 3D point clouds. *Photogramm. Eng. Remote Sens.* **2018**, *84*, 287–295. [CrossRef]
33. Ding, R.; Chen, Z.; Fan, W. WHU-STree: A Multi-modal Benchmark Dataset for Street Tree Inventory. *arXiv* **2025**, arXiv:2509.13172.
34. Wang, D. Unsupervised semantic and instance segmentation of forest point clouds. *ISPRS J. Photogramm.* **2020**, *165*, 86–97. [CrossRef]
35. Pang, Y.; Wang, W.; Du, L. Nyström-based spectral clustering using airborne LiDAR point cloud data for individual tree segmentation. *Int. J. Digit. Earth* **2021**, *14*, 1452–1476. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

*forests*

MDPI

*Article*

# Estimation of Forest Canopy Height from Spaceborne Full-Waveform LiDAR Data Using a Bisection Approximation Decomposition Method

**Song Chen [1,2,3], Ming Gong [1,2,3], Hua Sun [1,2,3,*], Ming Chen [1,2,3] and Binbin Wang [1,2,3]**

[1] Research Center of Forestry Remote Sensing & Information Engineering, Central South University of Forestry & Technology, Changsha 410004, China; 20210100003@csuft.edu.cn (S.C.); 20211100036@csuft.edu.cn (M.G.); 20241100070@csuft.edu.cn (M.C.); 20231100050@csuft.edu.cn (B.W.)
[2] Hunan Provincial Key Laboratory of Forestry Remote Sensing Based Big Data & Ecological Security, Changsha 410004, China
[3] Key Laboratory of National Forestry and Grassland Administration on Forest Resources Management and Monitoring in Southern China, Changsha 410004, China
* Correspondence: sunhua@csuft.edu.cn; Tel.: +86-138-7588-2184

**Abstract:** Forest canopy height (FCH) is a vital indicator for assessing forest health and ecosystem service capacity. Over the past two decades, full-waveform (FW) LiDAR has been widely employed for estimating forest biophysical variables due to its high precision in measuring vertical forest structures. However, the impact of terrain undulations on forest parameter estimation remains challenging. To address this issue, this study proposes a bisection approximation decomposition (BAD) method for processing GEDI L1B data and FCH estimation. The BAD method analyzes the energy composition of simplified echo signals and determines the fitting parameters by integrating overall signal energy, the differences in unresolved signals, and the similarity of inter-forest signal characteristics. FCH is subsequently estimated based on waveform peak positions. By dynamically adjusting segmentation points and Gaussian fitting parameters, the BAD method achieved precise separation of mixed canopy and ground signals, substantially enhancing the physical realism and applicability of decomposition results. The effectiveness and robustness of the BAD method for FCH estimation were evaluated using 2049 footprints across varying slope conditions in the Harvard Forest region of Petersham, Massachusetts. The results demonstrated that digital terrain models (DTMs) extracted using the GEDI data and the BAD method exhibited high consistency with the DTMs derived using airborne laser scanning (ALS) data (coefficient of determination $R^2 > 0.99$). Compared with traditional Gaussian decomposition (GD), wavelet decomposition (WD), and deconvolution decomposition (DD) methods, the BAD method showed significant advantages in FCH estimation, achieved the smallest relative root mean square error (rRMSE) of 17.19% and greatest mean estimation accuracy of 84.57%, and reduced the rRMSE by 10.74%, 21.49%, and 28.93% compared to GD, WD, and DD methods, respectively. Moreover, the BAD method exhibited a significantly stronger correlation with ALS-derived canopy height mode data than the relative height metrics from GEDI L2A products ($r = 0.84$, $p < 0.01$). The robustness and adaptability of the BAD method to complex terrain conditions provide great potential for forest parameters using GEDI data.

**Keywords:** full-waveform LiDAR; GEDI; waveform decomposition; bisection approximation decomposition; forest canopy height; slope

## 1. Introduction

Forests play a crucial role in terrestrial ecosystems as the largest carbon sink on land. Forest canopy height (FCH), a key indicator of forest growth, is closely linked to forest health and directly affects ecosystem service capacities [1,2]. Rapid and accurate acquisition of FCH is essential for assessing forest carbon sequestration and carbon emission, providing a solid scientific basis for formulating effective climate change mitigation and adaptation measures [3,4].

Compared with traditional ground-based measurements, remote sensing technologies greatly improve the efficiency and spatial details of information in forest monitoring [5]. Based on different data sources, remote sensing methods are categorized into optical, microwave, and Light Detection and Ranging (LiDAR) remote sensing. Optical remote sensing, however, is limited by spectral saturation, making it challenging to accurately capture the three-dimensional (3D) structures of forests [6,7]. Microwave remote sensing offers penetration capability, but its accuracy in estimating forest parameters remains limited [8,9]. LiDAR, on the other hand, emits laser pulses and measures the time difference of returned signals, enabling high-precision acquisition of 3D structural information of terrestrial features. Compared with optical and microwave remote sensing, LiDAR can penetrate canopy gaps and is unaffected by weather or cloud cover, providing detailed 3D structural information about forests [10,11]. Among LiDAR platforms, spaceborne LiDAR, with its high orbital altitude and extensive observation coverage, surpasses ground-based and airborne systems in monitoring forest parameters on a regional scale [12,13].

The Global Ecosystem Dynamics Investigation (GEDI) mission, launched in April 2019, is mounted on the International Space Station (ISS), covering latitudes from 51.6° N to 51.6° S. GEDI's primary goal is to collect global 3D forest ecosystem information for studying forest growth, species diversity, tree structure, and vegetation dynamics [14,15]. GEDI data feature high resolution and dense sampling, capturing continuous vertical forest structures, making them uniquely advantageous for forest parameter estimation [16,17]. As the latest generation of full-waveform (FW) LiDAR platforms, GEDI provides valuable resources for studying vertical structural parameters such as FCH.

In FW LiDAR data, the first peak often represents the canopy top, while the last peak corresponds to the ground. The relative distance between these two peaks is used to estimate FCH [18,19]. The waveform decomposition method enables more accurate identification of peak positions, facilitating the precise extraction of forest canopy height [19,20]. Unlike optical and SAR data sources, which often rely on regression models and ground-truth data [12], waveform decomposition directly estimates canopy height using only LiDAR data, offering significant advantages for this application. Existing waveform decomposition methods primarily include Gaussian decomposition (GD), wavelet decomposition (WD), and deconvolution decomposition (DD). GD fits the waveform with multiple Gaussian components, offering simplicity and computational efficiency [20]. WD leverages wavelet functions for local feature extraction and multi-scale analysis, suitable for complex signal decomposition [21]. DD removes system noise, yielding waveform features closer to the actual vertical distribution of forest structures [22].

Although these methods perform well in flat terrains, challenges arise in non-flat areas where complex forest structures and topographic variations broaden and mix waveform signals. In rugged terrains, echoes from the ground, shrubs, and trees of the lower layer are mixed, increasing uncertainty in FCH estimation [23,24]. Notably, slopes significantly affect FCH estimation accuracy, with root mean square error (RMSE) increasing by 14% when the slope exceeds 20° compared to slopes between 10° and 20° [25].

Slope data can be used to correct FCH estimation. Lefsky et al. [24] corrected waveform decomposition results using slope data from a digital elevation model (DEM), achieving

explainable tree heights with an RMSE of 5 m. Similarly, Nie et al. [26] improved terrain correction methods with DEM data, enhancing FCH estimation accuracy ($R^2$ = 0.77).

Machine learning models can be used to estimate FCH by combining GEDI-extracted relative height metrics with field data. Zhu et al. [27] used GEDI L2A relative height metrics and Landsat imagery to model FCH in the Harvard Forest, achieving high accuracy with an RMSE of 3.11 m. Similarly, Chen et al. [28] extracted GEDI waveform and terrain features to establish regression models for estimating forest canopy height in the mountainous regions of northern Henan, obtaining satisfactory accuracy with an RMSE of 2.23 m. However, this approach relies on external data sources, increasing model complexity and uncertainty. Moreover, errors in intermediate FCH products are difficult to quantify during subsequent modeling processes, affecting overall accuracy.

To address these challenges, we developed a bisection approximation decomposition (BAD) method for processing spaceborne LiDAR full-waveform (FW) data to estimate FCH. This method analyzed the internal energy composition of waveform signals, determined fitting coefficients by integrating overall energy, differences in unresolved signals, and canopy signal characteristics, and ultimately estimated FCH. GEDI L1B waveform data were decomposed using this method, and the results were validated against airborne-laser-scanning-derived canopy height model (ALS-CHM) data from the National Ecological Observatory Network (NEON) Airborne Observation Platform (AOP) [29]. The proposed method was compared with traditional methods including GD, WD, and DD based on the FCH values derived from ALS-CHM. This study aimed to provide an accurate FCH estimation method using spaceborne FW LiDAR while mitigating the influence of terrain slopes on FCH estimates.

## 2. Materials and Methods

### *2.1. Study Area*

The study area is Harvard Forest (Figure 1), located in Petersham, MA, USA. Harvard Forest has an area of approximately 1500 hectares and features a hilly terrain with elevations ranging from 90 to 410 m above sea level. The region is characterized by a temperate continental climate, with cold winters, warm and humid summers, and an average annual precipitation of approximately 1000 mm, evenly distributed throughout the year. The forest vegetation represents a typical New England mixed forest ecosystem. Dominant coniferous species include white pine (*Pinus strobus*) and eastern hemlock (*Tsuga canadensis*). Broadleaf species are represented by red maple (*Acer rubrum*), sugar maple (*Acer saccharum*), paper birch (*Betula papyrifera*), and American beech (*Fagus grandifolia*).

### *2.2. GEDI Data and Processing*

The GEDI tracks of the study area are shown in Figure 1b. We obtained the GEDI L1B footprint data from the Land Processes Distributed Active Archive Center (LP DAAC, https://cmr.earthdata.nasa.gov/search (accessed on 7 November 2024)). The GEDI data products are categorized into four levels: Level 1 provides geographic location waveform data, Level 2 offers footprint-level canopy height and profile metrics data, Level 3 includes gridded canopy height metrics and variability data, and Level 4 comprises footprint and gridded aboveground carbon estimation data [30]. Our dataset consists of version 2.0 GEDI L1B footprint data collected between February and December 2022, covering the Harvard Forest area. These data record geolocated ground return waveforms, essential for studying forest canopy structures and dynamics.

The GEDI L1B footprint data were filtered based on the geographical boundaries of the study area to exclude footprints falling outside the target region. To ensure data quality, footprints were selected based on the criteria of "degrade_flag = 0" and "solar_elevation < 0".

The "degrade_flag" set to 0 indicates that the data were not degraded, preserving their original quality and precision. Additionally, a "solar_elevation" value of less than 0 corresponds to nighttime acquisitions, which reduce interference between solar radiation and laser pulses [30,31].

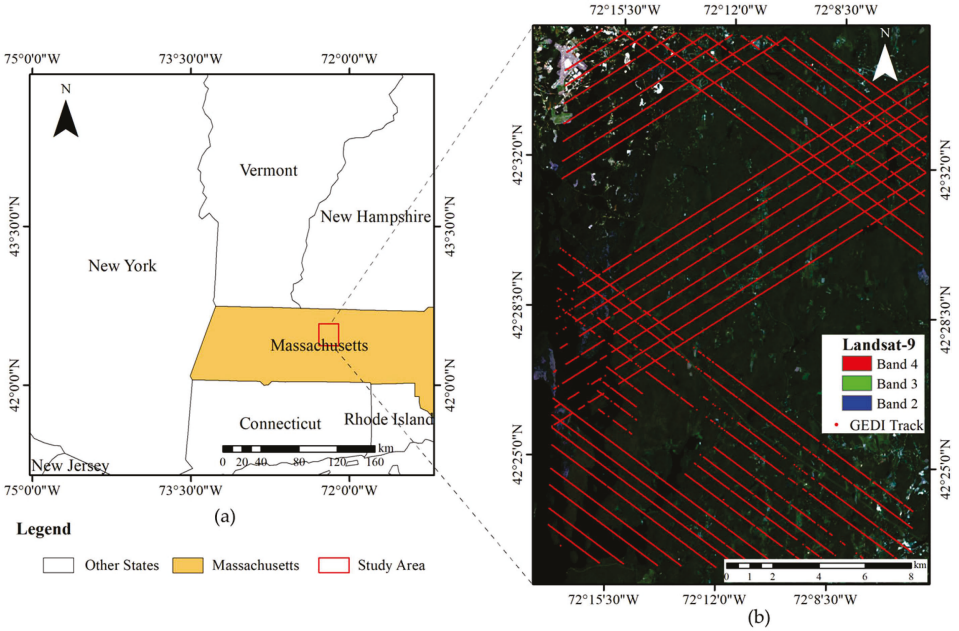



**Figure 1.** The study area location: (**a**) Harvard Forest, Massachusetts, and (**b**) the positions of GEDI tracks.

Waveform signals during propagation generally approximate a Gaussian distribution. However, atmospheric factors inevitably introduce background noise, while interactions with foliage and atmospheric water vapor may result in sharp spikes in the waveform [32,33]. Background noise was estimated using Equation (1), where the constant $v$ was set to 4 [32]. Figure 2a presents a comparison of the echo signals before and after denoising. The waveform amplitude reflects the intensity of the LiDAR signal, with the green curve representing the average background noise. The main oscillatory component (blue curve) is clearly emphasized. To mitigate sudden disturbances and high-frequency components, a Gaussian smoothing filter was applied (Figure 2b), which effectively extracted the underlying trends and fundamental features of the signals [34]. According to the Algorithm Theoretical Basis Document (ATBD), a smoothing width of 6.5 ns was chosen to ensure the integrity of waveform characteristics [35,36].

$$RXwave = RXwave - (noise_{mean} + v \times noise_{\theta}) \quad (1)$$

where $RXwave$ is the receiver echo waveform, $noise_{mean}$ is the mean background noise, $noise_{\theta}$ is the standard deviation of background noise, and $v$ is a constant.

Although the majority of the Harvard Forest area is covered by forest, certain GEDI footprints inevitably fell on non-forest areas. After denoising and smoothing the waveform signals, the footprints clearly represent the echo characteristics of ground objects at their respective locations. Non-forest footprints were excluded based on the following criteria: (a) chaotic signals lacking distinct main components or recognizable features, indicating interference at the footprint location and rendering the result unreliable [24,37]; (b) a single peak observed, whereas waveforms in forested areas typically exhibit multiple peaks corresponding to different reflective surfaces, such as the ground, vegetation, and canopy [24,37]; and (c) the heights of the first and last peaks being less than 4 m, as these peaks reflect canopy height, which is generally above this threshold in our study area.

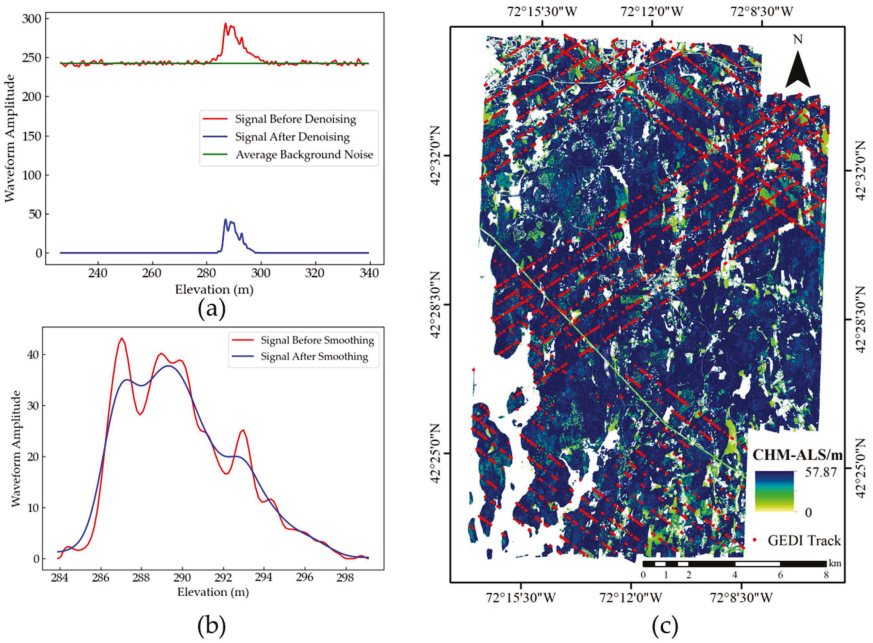



**Figure 2.** Waveform data pre-processing results: (**a**) the effect of noise reduction, (**b**) the effect after smoothing, and (**c**) ultimately retaining the GEDI's shot.

Ultimately, 2049 high-quality GEDI footprints were retained (Figure 2c) for subsequent waveform decomposition and analysis.

### *2.3. ALS Data and Processing*

The ALS data used in this study were acquired from the NEON field observatory (https://data.neonscience.org/data-products/ (accessed on 7 November 2024)) in August 2022. Funded by the U.S. National Science Foundation and operated by Battelle, NEON is a continental-scale observation facility designed to collect long-term, open access ecological data to better understand ecosystem changes across the United States. NEON employs a multi-scale sampling strategy that integrates a ground-based sensor network and high-resolution AOP to link terrestrial and spaceborne remote sensing scales [28].

The LiDAR instrument used in the NEON AOP is the Optech Gemini system, which has a laser footprint diameter of 0.25 m (0.8 m in wide beam divergence mode) and an elevation accuracy of <5–35 cm (1σ) [29,38]. For this study, we collected CHM, digital terrain model (DTM), and slope raster data provided by NEON AOP, all with a spatial resolution of 1.0 m. The ALS-CHM served as the reference dataset for evaluating the FCH extracted from waveform decomposition.

To account for the GEDI footprint diameter of 25 m, the NEON AOP raster data (ALS-CHM, ALS-derived DTM (ALS-DTM), and ALS-derived slope (ALS-slope)) were resampled to a resolution of 25 m using an averaging method. Forest canopy height (ALS-CHM) values were extracted for the 2049 retained footprints. The overall ALS-CHM distribution across all footprint locations ranged from 4.45 m to 31.15 m, with an average value of 19.50 m. The statistical distribution of heights within specific intervals is shown in Table 1.

### *2.4. Bisection Approximation Decomposition*

During the transmission of forest echo pulses, the pulse signal initially interacts with the top of the forest canopy, followed by the ground, as illustrated in Figure 3a. The signal composition of the echo pulses during transmission is characterized by

$$Signal_{total} = Signal_{canopy} + Signal_{ground} \quad (2)$$

where $Signal_{total}$ is the total echo signal, $Signal_{canpoy}$ is the canopy echo signal, and $Signal_{ground}$ is the ground echo signal.

**Table 1.** Statistical distribution of forest canopy heights for footprint locations.

| Height Range (m) | Number of Footprints | Minimum Height (m) | Maximum Height (m) | Average Height (m) | Coefficient of Variation (%) |
|---|---|---|---|---|---|
| <10 | 92 | 4.45 | 9.97 | 7.76 | 20.85 |
| 10–15 | 233 | 10.01 | 14.99 | 12.92 | 11.00 |
| 15–20 | 682 | 15.06 | 19.99 | 17.98 | 7.50 |
| 20–25 | 855 | 20.00 | 24.98 | 22.24 | 6.32 |
| >25 | 187 | 25.02 | 31.15 | 26.50 | 4.61 |

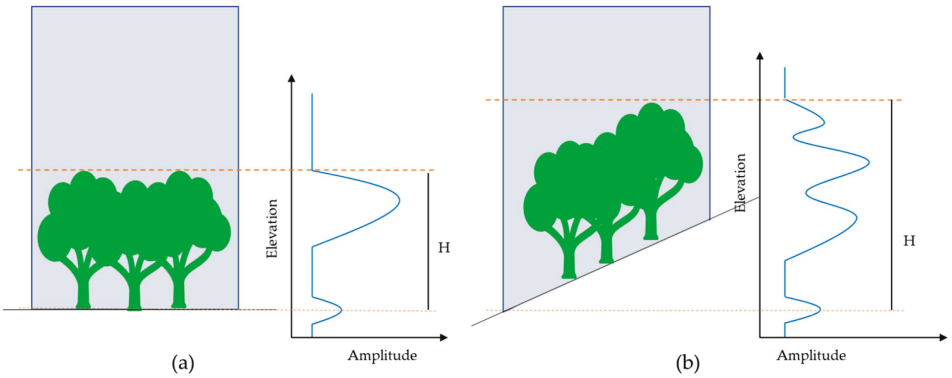



**Figure 3.** Comparison of waveforms under different slopes: (**a**) flat areas and (**b**) slope areas.

However, in areas with steep terrain slopes and complex forest stands, mixed waveforms can occur at the same horizontal elevation position. Moreover, the terminal peak in the waveform within the footprint represents the lowest point on the ground, leading to an overestimation of FCH estimates, as depicted in Figure 3b. In such cases, the signal composition of the echo pulses during transmission is influenced by

$$Signal_{total} = Signal_{canopy} + \sum Signal_{c\&t} + \sum Signal_{trunk} + \sum Signal_{t\&g} + Signal_{ground} \quad (3)$$

where $Signal_{c\&t}$ is canopy–trunk mixed echoes and $Signal_{t\&g}$ is trunk–ground mixed echoes.

However, when FCH is estimated, only the positions of the first and last wave peaks are required. Although these positions may shift due to slope effects and deviate from the actual locations, the broadening of the waveform ensures that the true first and last wave peaks are located within the range of the original peaks. To simplify the waveform decomposition process, we divided the signal into three components: ground signal, canopy signal, and intermediate mixed signal. We assume that the total energy of the entire echo signal is the sum of these three components, even though slope-induced effects may cause mixing between sub-canopy signals and the ground or canopy signals.

$$Signal_{total} = Signal_{canpy} + \sum Signal_{midst} + Signal_{ground} \quad (4)$$

where $Signal_{midst}$ is intermediate mixed echoes.

To accurately determine the locations of these three signal components, a bisection search-based decomposition method was adopted. This approach iteratively approximates the positions of the ground, canopy, and mixed signals within the waveform. The specific steps of the proposed BAD method are described as follows:

(1) The number of peaks in the pre-processed echo signal is first determined, and the signal is divided into two segments based on the median peak count. This segmentation

ensures that one segment predominantly contains canopy-mixed signals and the other contains ground-mixed signals, providing a foundation for subsequent decomposition.

For the canopy-mixed signal segment, the peak count ($k$) is determined:

- If $k$ = 1, the entire segment is considered a canopy signal and is directly fitted using a Gaussian model.
- Otherwise, the peak position ($M_1$) nearest to the split point ($O$) is identified, and a tangent point ($T$) is defined at $M_1$. A tangent ($L_t$) is drawn perpendicular to the x-axis, centered at $T$. The width is set to $L_t$ $O$, and the average height of the peaks is used as the amplitude to fit the sub-canopy signal (Figure 4a).

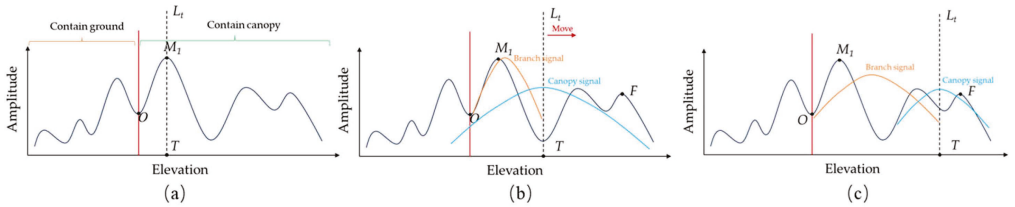



**Figure 4.** Bisection approximation decomposition process: (**a**) is the signal segmentation through the split point ($O$), and (**b**,**c**) are the decomposition results of the tree crown region.

(2) For the remaining portion of the signal, half the horizontal distance between the tangent point ($T$) and the first peak ($F$) is used as the width, with the midpoint as the center and the average height of the peaks as the amplitude to fit the canopy signal (Figure 4b,c).

Throughout this process, key features are recorded, including the energy ratio of the canopy-mixed signals and the amplitude-to-width ratio of sub-canopy signals. If the distance ($M_1F - TF$) exceeds a predefined threshold (Threshold), the tangent point ($T$) is adjusted by moving it toward $F$ by a distance proportional to $k/(M_1F - TF)$. This adjustment is repeated until the distance falls below the threshold.

(3) The ground-mixed signal is processed using the same method as the canopy-mixed signal to extract features for both sub-canopy and ground signals.

(4) After completing all iterations, the decomposition results are evaluated. The optimal result is selected based on a combination of the overall energy difference between the decomposed and raw signals and the similarity of sub-canopy signal features. As shown in Equation (5), the evaluation $E_{score}$ is used to assess the performance of waveform decomposition. When $E_{score}$ approaches 1, it indicates accurate decomposition results.

$$E_{score} = \frac{|E_C - E_c|}{E_C} + \frac{|E_G - E_g|}{E_G} + \frac{d_2 a_1 w_1}{d_1 a_2 w_2} \quad (5)$$

where $E_{score}$ is the decomposition score; $E_C$ and $E_c$ represent the total energy in the canopy region before and after decomposition, respectively; $E_G$ and $E_g$ represent the total energy in the ground region before and after decomposition, respectively; $d_1$ and $d_2$ denote the total distance in the canopy and ground regions, respectively; $a_1$ and $w_1$ represent the amplitude and width of the inter-layer signal in the canopy region, respectively; and $a_2$ and $w_2$ represent the amplitude and width of the inter-layer signal in the ground region, respectively.

### *2.5. Forest Canopy Height Estimation and Accuracy Evaluation*

We utilized the proposed BAD method to analyze waveform signals obtained within forest footprints and estimated FCH based on the relative positions of decomposed sub-wave signals. Additionally, we compared this method with traditional methods including GD, WD, and DD. The FCH values derived from ALS were used as references for accuracy validation

using metrics including coefficient of determination ($R^2$), root mean square error (RMSE), mean absolute error (MAE), relative RMSE (rRMSE), and estimation accuracy (EA) [13,39].

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} \quad (6)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}} \quad (7)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |\hat{y}_i - y_i| \quad (8)$$

$$\text{rRMSE} = \frac{\sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}}}{\overline{y}} \times 100\% \quad (9)$$

$$\text{EA} = 1 - \frac{1}{n} \sum_{i=1}^{n} \frac{|\hat{y}_i - y_i|}{y_i} \times 100\% \quad (10)$$

where $y_i$ is the ALS-derived FCH values, $\hat{y}_i$ is the estimated FCH using GEDI L1B footprint data, and $n$ is the number of samples. In addition, we conducted the tests of significant differences of mean residuals from zero by comparing the FCH estimates with the airborne-LiDAR-derived FCH reference values and the tests of significant differences of mean absolute residuals from the paired methods based on Student-$t$ distribution and the significance level of smaller than 0.05.

## 3. Results

### 3.1. Waveform Decomposition

Figure 5 illustrates the decomposition results of four waveform decomposition methods applied to forest footprint waveforms under varying slope conditions. The figure highlights notable differences in the ability of each method to handle slope-induced waveform distortions. The GD method maintains the original positions of waveform peaks and relies on Gaussian fitting based on the identified peak locations. While effective in scenarios with minimal slope influence, this method's performance heavily depends on the accuracy of background noise estimation and signal smoothing. In highly variable signal conditions, GD often introduces spurious peaks. Additionally, the method does not account for slope-induced waveform broadening, making it particularly susceptible to significant errors in peak identification under steep slope conditions. The DD and WD methods exhibit robust smoothing capabilities, akin to secondary smoothing of the raw signal. This process effectively reduces spurious peaks and partially mitigates slope-induced waveform broadening. However, the secondary processing of the signal may lead to shifts in peak positions, particularly under steep slope conditions. The limited controllability of these shifts can compromise the stability and accuracy of the decomposition results. The BAD method redefines peak positions by decomposing the waveform signal into three components: ground, canopy, and branches (upper and lower). By integrating overall energy distribution, residual differences from undecomposed signals, and intra-forest signal similarities, the BAD method achieves a more precise and stable decomposition. Notably, the BAD method excels in addressing slope-induced waveform broadening while minimally altering the initial and final peak positions in flat terrain. This dual capability enhances the physical realism of the decomposition results and extends their applicability to diverse terrain conditions.

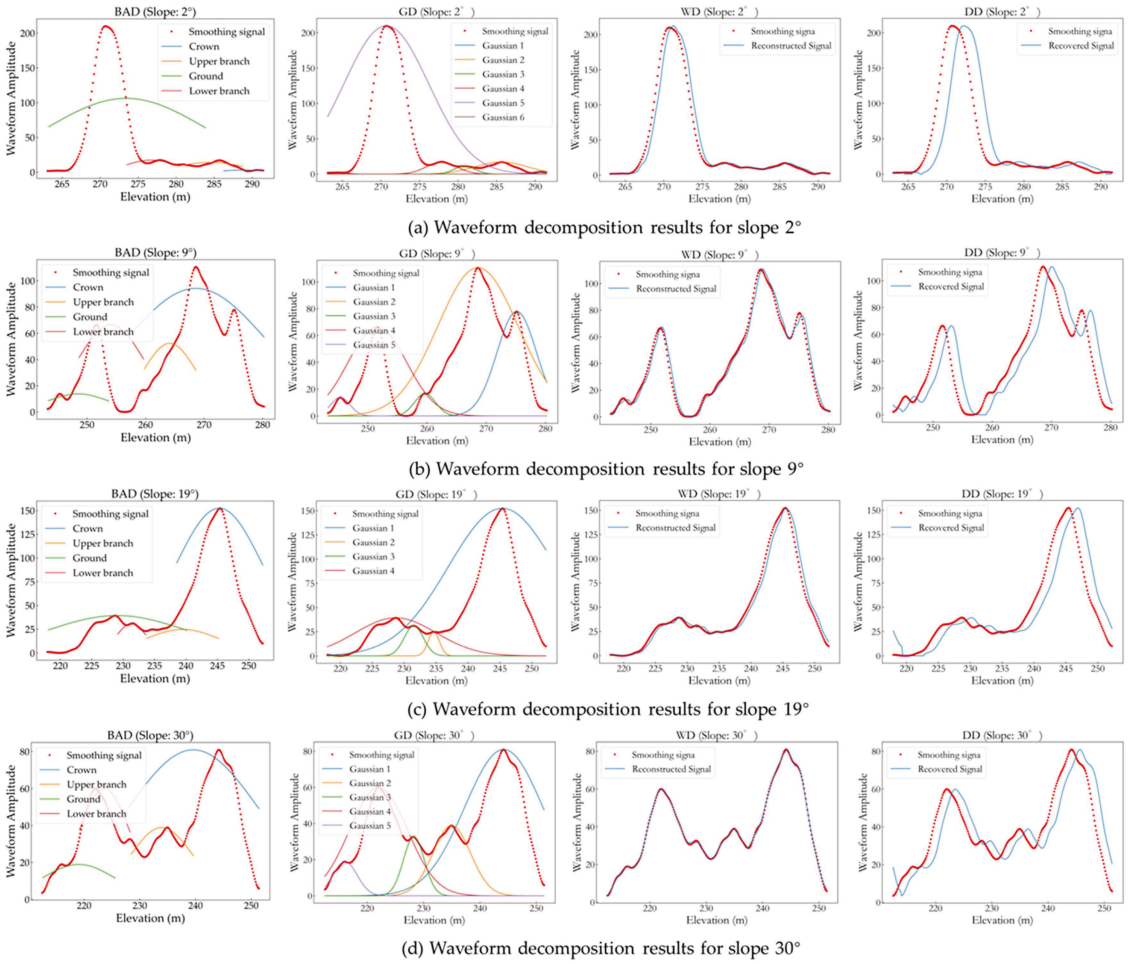



**Figure 5.** Comparison of the proposed BAD method with three traditional approaches for waveform decomposition effects.

### *3.2. Ground Elevation Estimation with Waveform Decomposition*

The estimation of ground elevation beneath each GEDI footprint is a critical component of waveform decomposition, as its accuracy directly reflects the effectiveness of the decomposition algorithm [26,40]. In this study, ground elevation data were extracted from 2049 pre-processed footprints within the study area using four different waveform decomposition methods. The extracted digital terrain models (DTMs) were then compared with the ALS-DTM to assess the accuracy of the decomposition methods. The results are shown in Figure 6.

DTMs obtained from all four decomposition methods exhibited high consistency with the ALS-DTM, as indicated by strong coefficients of determination ($R^2 > 0.99$). This consistency demonstrates the reliability of ground elevation results regardless of the chosen decomposition method, confirming the robustness of waveform decomposition for this application. Despite the overall high accuracy, differences in error levels were observed among the methods. Although the GD method shows the best performance in MAE (2.93 m) and EA (98.83%), it is evident from Figure 6b that the majority of data points fall below the Y=X line, indicating that the GD method tends to underestimate the DTM. This is primarily due to the waveform broadening effect, which causes the GD method to misclassify lower elevation points as ground level. In contrast, the DTM estimated using the BAD method is more evenly distributed around the Y=X line, with lower RMSE and rRMSE values of 4.33 m and 1.61%, respectively, demonstrating better overall consistency. These results are closely aligned with those of Nie et al. [26] and Wang et al. [41], but our method stands out

by not requiring the introduction of additional data sources. These findings demonstrate the superiority of the BAD method in accurately estimating ground elevation.

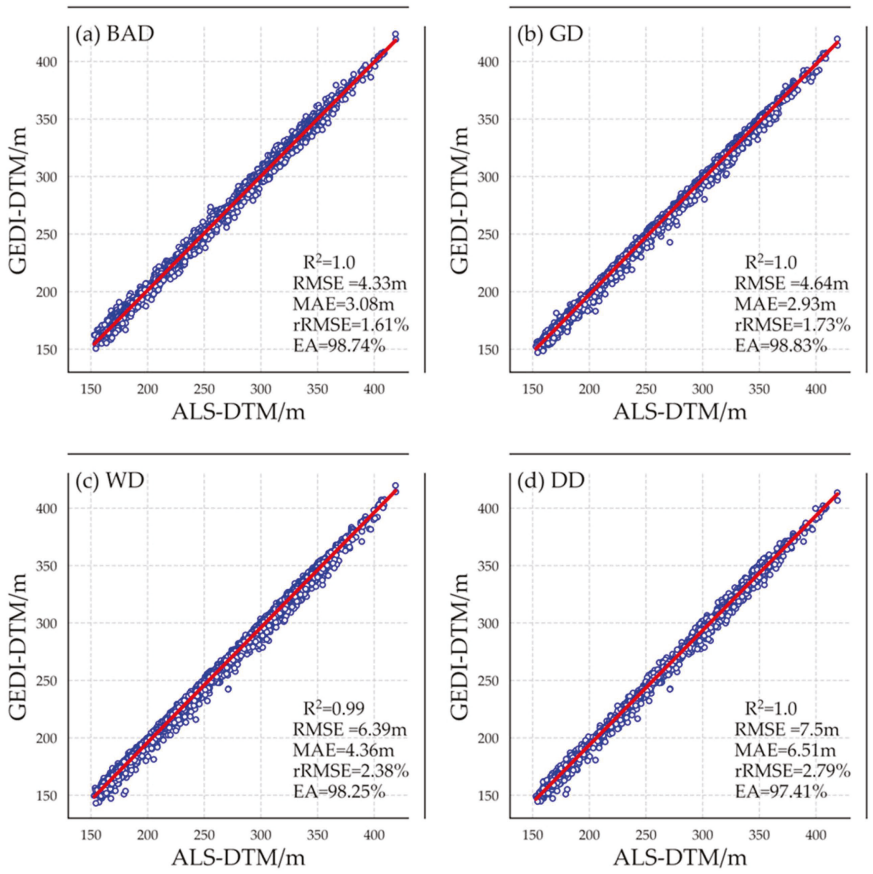



**Figure 6.** Comparison of estimation results of DTM based on the graphs of the predicted values against the referenced values among four decomposition methods: (**a**) BAD; (**b**) GD; (**c**) WD; (**d**) DD.

### *3.3. Forest Canopy Height Estimation with Waveform Decomposition*

In Table 2, the results of FCH estimation using four distinct methods are compared. The proposed BAD demonstrated the highest estimation accuracy by achieving a greatest $R^2$ of 0.67 and EA value of 84.57%, with the smallest RMSE of 3.33 m, MAE of 2.83 m, and rRMSE of 17.19%. Based on error and accuracy metrics, the best performance was achieved by BAD, followed by GD and DD, and WD yielded the poorest performance. Compared to GD, DD, and WD, BAD reduced the rRMSE by 10.74%, 21.49%, and 28.93%, respectively. Moreover, the fitting plots illustrated that BAD closely matches the measured values, demonstrating superior estimation accuracy with a close alignment along the Y=X line (Figure 7). Although all the methods led to approximate normal distributions of the predicted values similar to those of the reference values, BAD resulted in estimates close to the line of Y=X, overall, while the three traditional methods led to the overestimation being obvious.

**Table 2.** Accuracy comparison of FCH estimates among four decomposition methods.

| Methods | $R^2$ | RMSE/m | MAE/m | rRMSE/% | EA/% |
|---|---|---|---|---|---|
| BAD | 0.67 | 3.33 | 2.83 | 17.19 | 84.57 |
| GD | 0.37 | 5.66 | 4.20 | 29.21 | 75.82 |
| WD | 0.25 | 7.49 | 5.60 | 38.68 | 67.47 |
| DD | 0.23 | 8.94 | 7.42 | 46.12 | 57.19 |

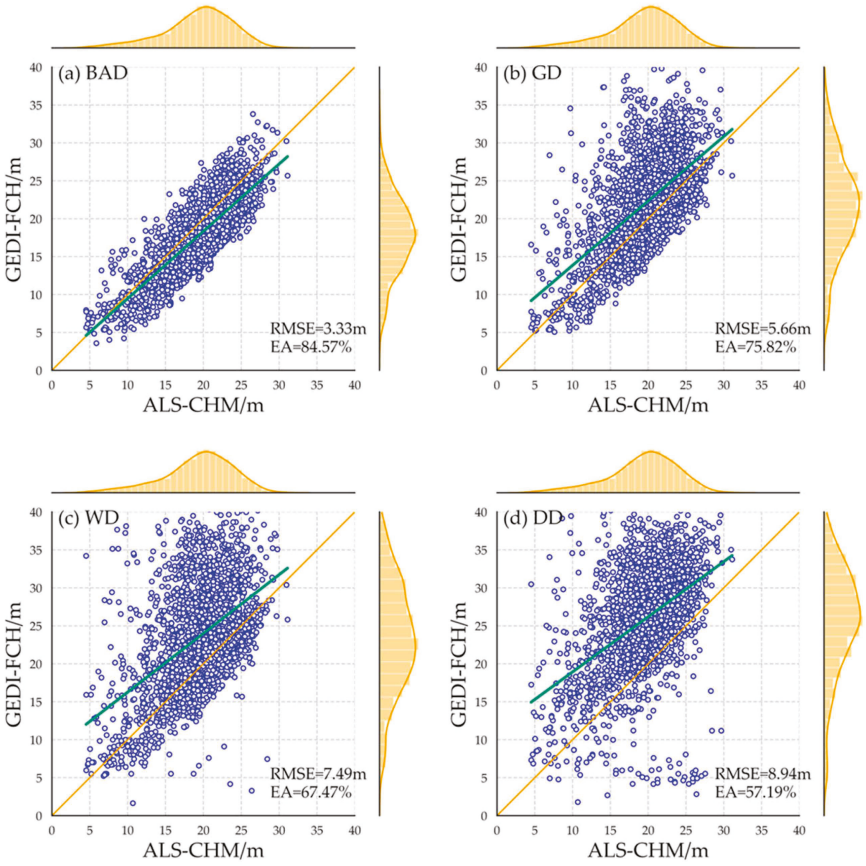



**Figure 7.** Comparison of estimation results of FCH based on the graphs of the predicted values against the referenced values among four decomposition methods: (**a**) BAD; (**b**) GD; (**c**) WD; (**d**) DD.

In Figure 8, the distributions of residual errors from four methods are compared. The residuals of the predicted values of FCH from BAD were predominantly within ±6 m and evenly spread around the zero line, indicating robust performance. Conversely, GD, DD, and WD exhibited residuals mainly falling in the range of 10 m to −10 m, suggesting less accurate and notable overestimations.

Additionally, as shown in Table 3, a comparison of the FCH values estimated by the four methods with the ALS-CHM reference values revealed significant correlations ($p < 0.01$) for all methods. Among them, the BAD-FCH method exhibited the highest correlation (0.82) with ALS-CHM, while the correlations for GD-FCH, WD-FCH, and DD-FCH with ALS-CHM were 0.61, 0.50, and 0.48, respectively. Furthermore, a comparison of the absolute residuals of FCH extraction results among the four methods and the obtained t-values showed greater than the critical t-value at the significance level of less than 0.01, indicating that the mean absolute residuals of the methods were statistically significantly different from each other.

Compared to the study by Chen et al. [28], which utilized GEDI waveform feature data and machine learning models to estimate forest canopy height in the mountainous regions of northern Henan, our approach achieved comparable estimation accuracy (rRMSE of 17.19% and 17%, respectively). Similarly, when compared to the work by Zhu et al. [27], which estimated forest canopy height in the Harvard Forest, the accuracy of the BAD method also proves to be comparable, with RMSE values of 3.33 m and 3.11 m, respectively. However, the key difference is that we did not employ a regression model; instead, we directly analyzed the waveform data, which enhanced the interpretability, practicality, and physical realism of the results.

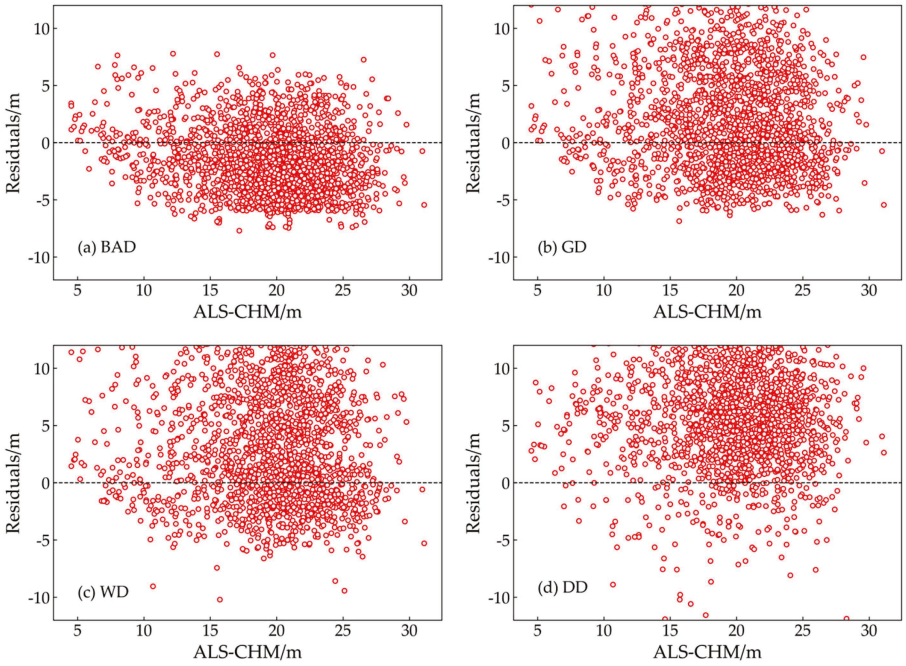



**Figure 8.** Comparison of residuals of FCH estimates among four decomposition methods: (**a**) AGD; (**b**) GD; (**c**) WD; (**d**) DD.

**Table 3.** Comparison of the four decomposition methods based on their correlation with ALS-CHM and the significance difference test of the absolute residual differences between methods (significance level $\leq 0.01$).

| Methods | Correlation Difference (R-Value) Between ALS-CHM and Estimated Results | Significant Difference (t-Value) Among Methods of Mean Absolute Residuals | | | |
|---|---|---|---|---|---|
| | FCH | BAD | GD | WD | DD |
| BAD | 0.82(0.00) | - | 39.62 | 47.11 | 77.95 |
| GD | 0.61(0.00) | 39.62 | - | 26.40 | 64.26 |
| WD | 0.50(0.00) | 47.11 | 26.40 | - | 55.82 |
| DD | 0.48(0.00) | 77.95 | 64.26 | 55.82 | - |

## 4. Discussion

### *4.1. The Impact of Terrain Slope*

After the laser pulse from spaceborne LiDAR reaches the ground, it forms a large footprint, with diameters ranging from several tens of meters to hundreds of meters. Terrain slope variations significantly affect the shape of these footprints, thereby influencing the estimation of forest parameters [39]. Currently, there is no universally effective method to mitigate the impact of slope on full-FW LiDAR data. To further investigate the effect of slope on forest canopy height (FCH) extraction, we extracted ALS-slope data from NEON AOP and obtained the corresponding slope values for each footprint based on its location. We classified the slope values into six intervals: I [0–5°], II [5–10°], III [10–15°], IV [15–20°], V [20–25°], and VI [>25°], and analyzed the robustness of the four methods within each slope interval. The slope gradations were determined based on natural breaks observed in the topographic slope distribution of the study area, ensuring that the classification reflects the terrain's inherent variability. This approach is consistent with the classification thresholds adopted by Fayad et al. [25]. Moreover, the use of 5° intervals provides a balanced resolution to analyze the effect of slope on FCH extraction results, offering clear insights into the trends within each range. Figure 9 shows the trends of RMSE and MAE for FCH estimation using the four decomposition methods across different slope intervals.

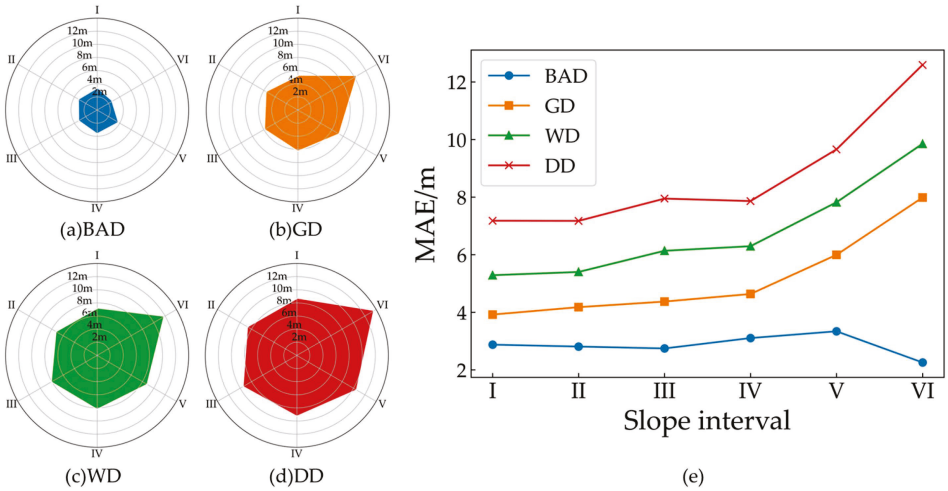



**Figure 9.** Comparison of errors of FCH estimates in different slope intervals among four decomposition methods: (**a**–**c**) and (**d**) represent the root mean square errors for BAD, GD, WD, and DD methods, respectively, while (**e**) shows the evaluation of the mean absolute error for all four methods.

The results showed that for the GD, WD, and DD methods, RMSE gradually increased with the slope interval, a trend consistent with the findings of Lee et al. [42] and Liu et al. [29], indicating that the accuracy of FCH extraction decreases as the slope increases. In contrast, the RMSE for the BAD method remained relatively stable across the slope intervals, with no significant correlation to slope changes. In all slope intervals, the RMSE for the BAD method was consistently lower than that of the other three methods. Specifically, the order of RMSE from lowest to highest was BAD < GD < WD < DD. For each slope interval, the RMSE for the BAD method was always lower than the RMSE for any of the other three methods across all slope intervals. The trend for MAE (Figure 9e) followed a similar pattern.

The BAD method we proposed effectively redefines the sub-wave positions within the valid signal range by integrating the overall energy of the signal, the difference from the undecomposed signal, and the similarity of the in-between signals. This process adjusts the peak positions and reduces the broadening effect caused by the slope. Overall, the BAD method significantly outperforms the GD, WD, and DD methods in terms of FCH extraction accuracy across different slope intervals, showing stronger slope adaptability. In contrast, the GD, WD, and DD methods are more sensitive to slope changes. Compared to the studies of Lefsky et al., Liu et al., and others, the BAD method demonstrates stronger slope adaptability and achieves the smallest estimation error (rRMSE = 17.19%) without relying on additional data sources.

### *4.2. The Impact of ALS Canopy Height Resampling Method*

This study evaluated the effects of different sampling methods on the accuracy of FCH estimation by comparing waveform-decomposition-estimated FCH with CHM data generated from airborne LiDAR. During the ALS-CHM generation process, point cloud data were gridded to match the GEDI footprint diameter (25 m), and window mean resampling was applied to obtain the average stand height within each footprint. However, the window mean resampling does not fully align waveform-decomposition-extracted canopy height with the average FCH across the stand.

To explore the suitability of various sampling methods, this study further resampled the ALS-CHM data at 1 m resolution using three distinct approaches (Figure 10):

(1) Maximum Value Sampling (ALS-CHM(max)): extracts the maximum canopy height within the footprint, representing the vertical distance from the highest treetop to the

lowest ground point. This metric corresponds to the distance between the start of the first echo and the peak of the last echo, providing an approximation of the total vertical range of the canopy.

(2) 95% Relative Height Sampling (ALS-CHM(95%)): extracts the 95% relative canopy height within the footprint. This approach accounts for potential deviations of the waveform's first peak due to canopy structural complexity or environmental factors such as wind or water vapor, correcting for these influences.

(3) Interval Average Sampling (ALS-CHM(sec)): calculates the average height of pixels within the 45%–95% relative height interval. This approach better reflects the upper canopy height while mitigating the effects of gaps and low ground-level values.

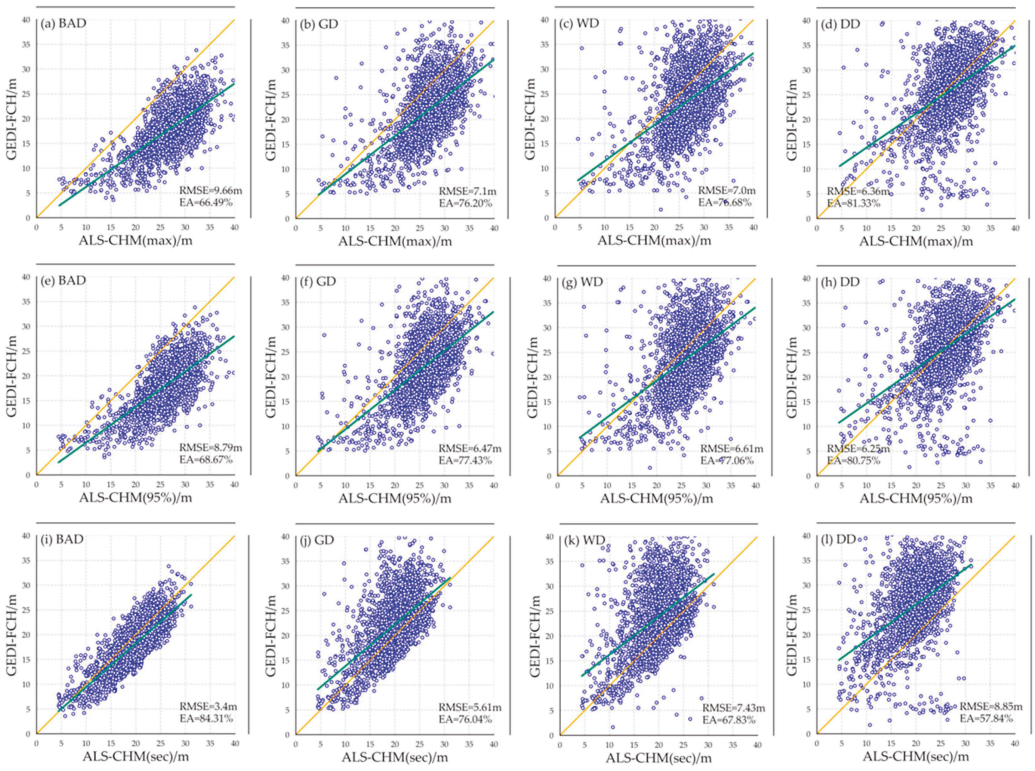



**Figure 10.** Comparison of FCH extraction results among four decomposition methods and different sampling methods: (**a**–**d**) represent BAD, GD, WD, and DD compared to ALS-CHM(max); (**e**–**h**) represent BAD, GD, WD, and DD compared to ALS-CHM(95%); (**i**–**l**) represent BAD, GD, WD, and DD compared to ALS-CHM(sec).

When using ALS-CHM(max) and ALS-CHM(95%) as reference variables, the DD method achieved the highest accuracy, with estimation accuracies of 81.33% (ALS-CHM(max)) and 80.75% (ALS-CHM(95%)). Both WD and GD methods yielded comparable results, with accuracies of 76.88% and 77.06% (ALS-CHM(max)) and 76.20% and 77.43% (ALS-CHM(95%)), respectively. However, both methods showed underestimation in some regions. The BAD method demonstrated the lowest accuracy, with extraction accuracies of 66.49% (ALS-CHM(max)) and 68.67% (ALS-CHM(95%)).

When using ALS-CHM(sec) as the reference variable, the BAD method outperformed other methods, achieving an accuracy of 84.31%. The GD method ranked second, with an accuracy of 76.04%. The WD and DD methods showed lower accuracies of 67.83% and 57.84%, respectively. Furthermore, the GD, WD, and DD methods generally overestimated canopy height compared to reference values.

Across different CHM sampling strategies, the WD and DD methods are more suited for extracting maximum canopy height but are prone to underestimating high values in

ALS-CHM(max) and ALS-CHM(95%), while overestimating low values in ALS-CHM(95%) and ALS-CHM(sec). The BAD method consistently demonstrated superior performance in ALS-CHM(sec) and ALS-CHM, achieving significantly higher accuracy, particularly for estimating average canopy height.

The results highlighted the distinct advantage of the BAD method in estimating average FCH, which is a critical metric for assessing forest quality and estimating carbon sequestration potential. Average FCH serves as a key indicator in forestry applications, providing essential insights for forest ecosystem management and decision-making.

### *4.3. Comparison with GEDI L2A Products*

The GEDI Level 2A (L2A) dataset, a Level 2 product released by NASA, includes height metric products that extend beyond the waveform data provided in GEDI Level 1B (L1B). GEDI L2A consists of 100 relative height (RH) metrics, which capture vertical height variations at footprint locations. These RH metrics have been widely used for estimating forest parameters, including FCH [43,44].

In this study, we obtained corresponding GEDI L2A data together with the downloaded GEDI L1B waveform data. By matching footprint identifiers, we extracted 11 default relative height metrics (RH90–RH100) for the same footprints. GEDI L2A (version 2.0) includes RH metrics extracted using six different algorithms, with the officially recommended results set as default [17,45]. To evaluate the effectiveness of the BAD method, we analyzed the correlation between FCH extracted from these 11 relative height metrics and the BAD-extracted canopy heights, using four sampling strategies for CHM data: ALS-CHM, ALS-CHM(sec), ALS-CHM(95%), and ALS-CHM(max). The results are presented in Figure 11.

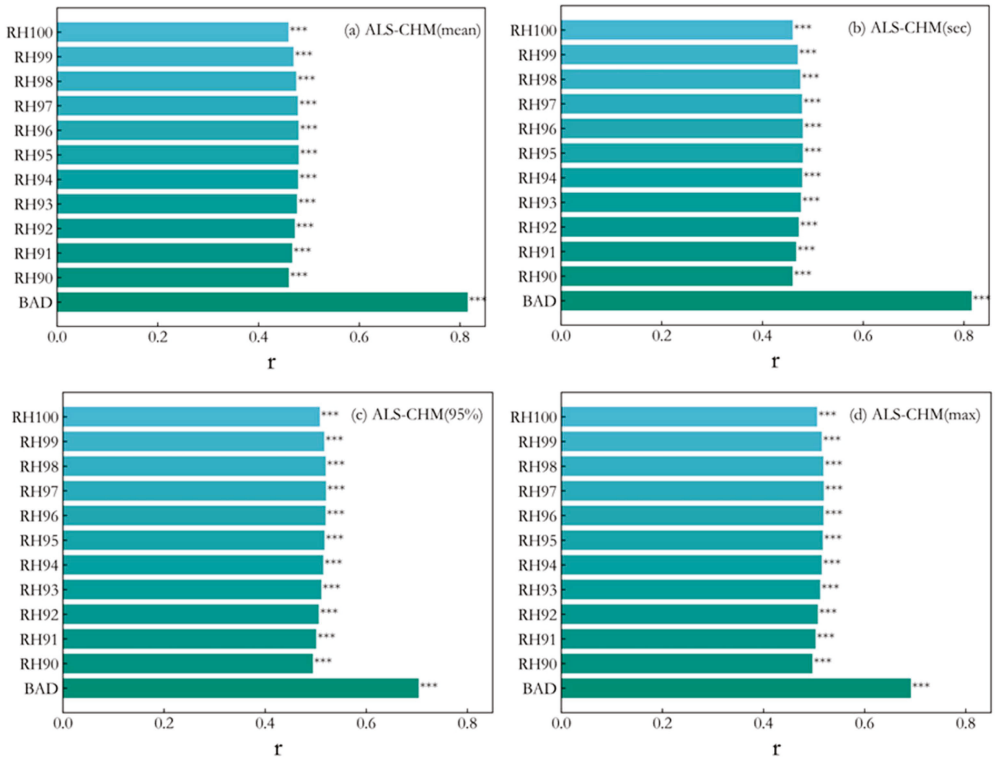



**Figure 11.** Comparison of correlation between the BAD extraction results, 11 relative GEDI 2A height metrics with ALS-CHM: (**a**) ALS-CHM(mean), (**b**) ALS-CHM(sec), (**c**) ALS-CHM(95%), (**d**) ALS-CHM(max). *** indicates a significance level of 0.001.

The overall variation in RH90–RH100 metrics was small, with RH95, RH96, RH97, and RH98 showing the highest correlations, which is consistent with the findings of Zhu

et al. [27] and Potapov et al. [46]. However, the FCH extracted from the BAD method demonstrated significantly stronger correlations with the four CHM sampling strategies compared to GEDI L2A metrics. The correlation coefficients between BAD-extracted FCH and ALS-CHM were 0.816 (ALS-CHM(mean)), 0.816 (ALS-CHM(sec)), 0.704 (ALS-CHM(95%)), and 0.692 (ALS-CHM(max)). In contrast, GEDI L2A RH90–RH100 metrics did not achieve comparable levels of correlation in the same comparisons.

The results highlighted the superior performance of the BAD method in FCH estimation. Its extracted canopy heights demonstrated significantly higher correlations with CHM-extracted canopy heights compared to the default RH90–RH100 metrics in GEDI L2A. This indicated that the BAD method offered greater adaptability and potential for accurately estimating FCH.

## 5. Conclusions

This study utilized the GEDI full-waveform spaceborne LiDAR data to estimate forest canopy height (FCH) in the Harvard Forest region of Petersham, Massachusetts, and validated the results using airborne-laser-scanning-derived canopy height model (ALS-CHM) data from the National Ecological Observatory Network. A novel bisection approximation decomposition (BAD) method was developed, which outperformed Gaussian decomposition (GD), wavelet decomposition (WD), and deconvolution decomposition (DD) approaches based on estimation accuracy (rRMSE: 17.19%, 29.21%, 38.68%, and 46.12%; EA: 84.57%, 75.82%, 67.47%, and 57.19%, respectively) and adaptability across varying slope conditions. At the same time, the digital terrain model (DTM) obtained using the BAD method showed good consistency with the airborne-laser-scanning-derived digital terrain model (ALS-DTM) data ($R^2 > 0.99$), further confirming the robustness of its waveform decomposition. The BAD method demonstrated a significantly stronger correlation with ALS-CHM data compared to the relative height metrics from GEDI L2A products ($r = 0.84$, $p < 0.01$). The BAD method's ability to dynamically adjust segmentation points and Gaussian fitting parameters enables precise separation of mixed signals from the canopy and ground, enhancing the physical realism and applicability of decomposition results. These results highlighted the BAD method's potential for fine-scale LiDAR data analysis and its value in providing robust technical and theoretical support for forest parameter estimation, particularly under complex terrain conditions.

**Author Contributions:** Conceptualization and methodology, S.C. and M.G.; visualization and writing—original draft, S.C.; formal analysis, S.C. and M.C.; resources, M.G. and M.C.; data curation, M.C. and B.W.; writing—review and editing, S.C. and H.S.; supervision, project administration and funding acquisition, H.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Key Research and Development Program of China, grant number 2023YFD2201703; the Science and Technology Innovation Program of Hunan Province, grant number 2023RC1065; the Hunan Provincial Natural Science Foundation of China, grant number 2022JJ30078; and the Hunan Province Graduate Research Innovation Program of China, grant number CX20230777.

**Data Availability Statement:** GEDI data were downloaded free of charge from the Land Processes Distributed Active Archive Center [https://cmr.earthdata.nasa.gov/search (accessed on 7 November 2024)]. ALS data were downloaded free of charge from the National Ecological Observatory Network [https://data.neonscience.org/data-products/ (accessed on 7 November 2024)].

**Acknowledgments:** The authors would like to thank the GEDI team and the NASA LPDAAC (Land Processes Distributed Active Archive Center) for providing the GEDI data. The authors also extend their gratitude to the NEON AOP team for their support of the NEON AOP project and for providing the ALS datasets.

**Conflicts of Interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References


1. Lefsky, M.A. A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophys. Res. Lett.* **2010**, *37*, L15401. [CrossRef]
2. Dubayah, R.O.; Sheldon, S.L.; Clark, D.B.; Hofton, M.A.; Blair, J.B.; Hurtt, G.C.; Chazdon, R.L. Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica. *J. Geophys. Res.* **2010**, *115*, G00E09. [CrossRef]
3. Yu, Q.; Ryan, M.G.; Ji, W.; Prihodko, L.; Anchang, J.Y.; Kahiu, N.; Nazir, A.; Dai, J.; Hanan, N.P. Assessing canopy height measurements from ICESat-2 and GEDI orbiting LiDAR across six different biomes with G-LiHT LiDAR. *Environ. Res. Ecol.* **2024**, *3*, 025001. [CrossRef]
4. Bruening, J.M.; May, P.; Armston, J.; Dubayah, R. Precise and unbiased biomass estimation from GEDI data and the US Forest Inventory. *Front. For. Global Chang.* **2023**, *7*, 1149153. [CrossRef]
5. Abshire, J.B.; Sun, X.; Riris, H.; Sirota, J.M.; McGarry, J.F.; Palm, S.; Yi, D.; Liiva, P. Geoscience Laser Altimeter System (GLAS) on the ICESat Mission: On-orbit measurement performance. *Geophys. Res. Lett.* **2003**, *32*, L21S02. [CrossRef]
6. Neumann, C.; Forster, M.; Kleinschmit, B.; Itzerott, S. Utilizing a PLSR-Based Band-Selection Procedure for Spectral Feature Characterization of Floristic Gradients. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 3982–3996. [CrossRef]
7. Soomro, B.N.; Xiao, L.; Huang, L.; Soomro, S.H.; Molaei, M. Bilayer Elastic Net Regression Model for Supervised Spectral-Spatial Hyperspectral Image Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *9*, 4102–4116. [CrossRef]
8. Hichem, M.; Mounira, O. Volume Height Estimation based on Fusion of Discrete Fourier Transform (DFT) and Least Square (LS) in a Tomographic SAR Application. *J. Indian Soc. Remote Sens.* **2016**, *45*, 217–228. [CrossRef]
9. Zribi, M.; Guyon, D.; Motte, E.; Dayau, S.; Wigneron, J.P.; Baghdadi, N.; Pierdicca, N. Performance of GNSS-R GLORI Data for Biomass Estimation over the Landes Forest. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *74*, 150–158. [CrossRef]
10. Chen, Q. Assessment of terrain elevation derived from satellite laser altimetry over mountainous forest areas using airborne lidar data. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 111–122. [CrossRef]
11. Silva, C.A.; Duncanson, L.D.; Hancock, L.; Neuenschwander, A.; Thomas, N.; Hofton, M.; Fatoyinbo, L.; Simard, M.; Marshak, C.Z.; Armston, J.; et al. Fusing simulated GEDI, ICESat-2 and NISAR data for regional aboveground biomass mapping. *Remote Sens. Environ.* **2021**, *253*, 112234. [CrossRef]
12. 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]
13. Cui, L.; Jiao, Z.; Zhao, K.; Sun, M.; Dong, Y.; Yin, S.; Zhang, X.; Guo, J.; Xie, R.; Zhu, Z.; et al. Retrieving Forest Canopy Elements Clumping Index Using ICESat GLAS Lidar Data. *Remote Sens.* **2021**, *13*, 948. [CrossRef]
14. Torresani, M.; Rocchini, D.; Alberti, A.; Moudry, V.; Heym, M.; Thouverai, E.; Kacic, P.; Tomelleri, P. LiDAR GEDI derived tree canopy height heterogeneity reveals patterns of biodiversity in forest ecosystems. *Ecol. Inf.* **2023**, *76*, 102082. [CrossRef] [PubMed]
15. Padalia, H.; Prakash, A.; Watham, T. Modelling aboveground biomass of a multistage managed forest through synergistic use of Landsat-OLI, ALOS-2 L-band SAR and GEDI metrics. *Ecol. Inf.* **2023**, *77*, 102234. [CrossRef]
16. Pang, Y.; Lefsky, M.; Sun, G.; Ranson, J. Impact of footprint diameter and off-nadir pointing on the precision of canopy height estimates from spaceborne lidar. *Remote Sens. Environ.* **2011**, *115*, 2798–2809. [CrossRef]
17. Musthafa, M.; Singh, G.; Kumar, P. Comparison of forest stand height interpolation of GEDI and ICESat-2 LiDAR measurements over tropical and sub-tropical forests in India. *Environ. Monit. Assess.* **2023**, *195*, 71. [CrossRef]
18. Lefsky, M.A.; Harding, D.J.; Keller, M.; Cohen, W.B.; Carabajal, C.C.; Espirito-Santo, F.D.B.; Hunter, M.O.; Oliveira, R.D. Estimates of forest canopy height and aboveground biomass using ICESat. *Geophys. Res. Lett.* **2005**, *32*, L22S02. [CrossRef]
19. Shen, X.; Li, Q.Q.; Wu, G.; Zhu, J. Decomposition of LiDAR waveforms by B-spline-based modeling. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 182–191. [CrossRef]
20. Gu, Z.; Lai, J.; Wang, C.; Yan, W.; Li, Z. Generalized Gaussian decomposition for full waveform LiDAR processing. *Meas. Sci. Technol.* **2022**, *33*, 065201. [CrossRef]
21. Wu, H.; Xing, Y. Wavelet transform and its application to ICEat-GLAS full waveform data. In Proceedings of the 2010 International Symposium on Intelligence Information Processing and Trusted Computing (IPTC), Huanggang, China, 28–29 October 2010; IEEE Computer Society: Washington, DC, USA, 2010.
22. Wang, Y.; Zhang, J.; Roncat, A.; Wagner, W. Regularizing method for the determination of the backscatter cross section in lidar data. *J. Opt. Soc. Am. A.* **2009**, *26*, 1071–1079. [CrossRef] [PubMed]
23. Magnussen, S.; Nord-Larsen, T.; Riis-Nielsen, T. LiDAR supported estimators of wood volume and aboveground biomass from the Danish national forest inventory (2012–2016). *Remote Sens. Environ.* **2018**, *211*, 146–153. [CrossRef]
24. Lefsky, M.A.; Keller, M.; Pang, Y.; Camargo, P.B.; Hunter, M.O. Revised method for forest canopy height estimation from Geoscience Laser Altimeter System waveforms. *J. Appl. Remote Sens.* **2007**, *1*, 6656–6659. [CrossRef]

25. Fayad, I.; Baghdadi, N.; Alcarde, A.C.; Stape, J.L.; Bailly, J.S.; Scolforo, H.F.; Cegatta, I.R.; Zribi, M.; Maire, G.L. Terrain Slope Effect on Forest Height and Wood Volume Estimation from GEDI Data. *Remote Sens.* **2021**, *13*, 2136. [CrossRef]
26. Nie, S.; Wang, C.; Zeng, H.; Xi, X.; Xia, S. A revised terrain correction method for forest canopy height estimation using ICESat/GLAS data. *ISPRS J. Photogramm. Remote Sens.* **2015**, *108*, 183–190. [CrossRef]
27. Zhu, W.; Yang, F.; Qiu, Z.; He, N.; Zhu, X.; Li, Y.; Xu, Y.; Lu, Z. Enhancing Forest Canopy Height Retrieval: Insights from Integrated GEDI and Landsat Data Analysis. *Sustainability* **2023**, *15*, 10434. [CrossRef]
28. Chen, R.; Wang, X.; Liu, X.; Wang, S. Optimizing GEDI Canopy Height Estimation and Analyzing Error Impact Factors Under Highly Complex Terrain and High-Density Vegetation Conditions. *Forests* **2024**, *15*, 2024. [CrossRef]
29. 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*, 12571. [CrossRef]
30. Adrah, E.; Jaafar, W.S.W.M.; Bajaj, S.; Omar, H.; Leite, R.V.; Silva, C.A.; Mohan, M. Analyzing canopy height variations in secondary tropical forests of Malaysia using NASA GEDI. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *880*, 012031. [CrossRef]
31. Li, X.; Li, L.; Ni, W.; Mu, X.; Wu, X.; Laurin, G.V.; Vangi, E.; Sterenczak, K.; Chirici, G.; Yu, S.; et al. Validating GEDI tree canopy cover product across forest types using co-registered aerial LiDAR data. *ISPRS J. Photogramm. Remote Sens.* **2024**, *207*, 326–337. [CrossRef]
32. Fayad, I.; Baghdadi, N.; Alcarde, C.; Stape, J.L.; Maire, G.L. Assessment of GEDI's LiDAR Data for the Estimation of Canopy Heights and Wood Volume of Eucalyptus plantations in Brazil. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7095–7110. [CrossRef]
33. Thomas, A.; Peter, B.; Christopher, P. Extracting More Data from LiDAR in Forested Areas by Analyzing Waveform Shape. *Remote Sens.* **2012**, *4*, 682–702. [CrossRef]
34. Gao, S.; Niu, Z.; Sun, G.; Zhao, D.; Kun, J.; Qin, Y. Height Extraction of Maize Using Airborne Full-Waveform LIDAR Data and a Deconvolution Algorithm. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1978–1982. [CrossRef]
35. Dubayah, S.L.R. GEDI L1B Geolocated Waveform Data Global Footprint Level V001. 2020. Available online: https://lpdaac.usgs.gov/products/gedi01_bv001/ (accessed on 8 November 2024).
36. Dubayah, S.L.R. GEDI L2A Elevation and Height Metrics Data Global Footprint Level V001. 2020. Available online: https://lpdaac.usgs.gov/products/gedi02_av001/ (accessed on 8 November 2024).
37. Anderson, J.; Martin, M.E.; Smith, M.-L.; Dubayah, R.O.; Hofton, M.A.; Hyde, P.; Peterson, B.E.; Blair, J.B.; Knox, R.G. The use of waveform lidar to measure northern temperate mixed conifer and deciduous forest structure in New Hampshire. *Remote Sens. Environ.* **2006**, *105*, 248–261. [CrossRef]
38. Scholl, V.M.; Cattau, M.E.; Joseph, M.B.; Balch, J.K. Integrating national ecological observatory network (NEON) airborne remote sensing and in-situ data for optimal tree species classification. *Remote Sens.* **2020**, *12*, 1414. [CrossRef]
39. Song, S.; Wang, B.; Gong, W.; Chen, Z.; Lin, X.; Sun, J.; Shi, S. A new waveform decomposition method for multispectral LiDAR. *ISPRS J. Photogramm. Remote Sens.* **2019**, *149*, 40–49. [CrossRef]
40. Allouis, T.; Durrieu, S.; Couteron, P. A new method for incorporating Hillslope effects to improve Canopy-height estimates from large-footprint LIDAR waveforms. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 730–734. [CrossRef]
41. Wang, Y.; Ni, W.; Sun, G.; Chi, H.; Zhang, Z.; Guo, Z. Slope-adaptive waveform metrics of large footprint lidar for estimation of forest aboveground biomass. *Remote Sens. Environ.* **2019**, *224*, 386–400. [CrossRef]
42. Lee, S.; Ni-Meister, W.; Yang, W.; Chen, Q. Physically based vertical vegetation structure retrieval from ICESat data: Validation using LVIS in White Mountain National Forest, New Hampshire, USA. *Remote Sens. Environ.* **2011**, *115*, 2776–2785. [CrossRef]
43. Adam, M.; Urbazaev, M.; Dubois, C.; Schmullius, C. Accuracy Assessment of GEDI Terrain Elevation and Canopy Height Estimates in European Temperate Forests: Influence of Environmental and Acquisition Parameters. *Remote Sens.* **2020**, *12*, 3948. [CrossRef]
44. Vatandaslar, C.; Narin, N.O.; Abdikan, S. Retrieval of forest height information using spaceborne LiDAR data: A comparison of GEDI and ICESat-2 missions for Crimean pine (*Pinus nigra*) stands. *Trees-Struct. Funct.* **2023**, *37*, 717–731. [CrossRef]
45. Li, H.; Li, X.; Kato, T.; Hayashi, M.; Fu, J.; Hironshima, T. Accuracy Assessment of GEDI Terrain Elevation, Canopy Height, and Aboveground Biomass Density Estimates in Japanese Artificial Forests. *Sci. Remote Sens.* **2024**, *10*, 100144. [CrossRef]
46. 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]


**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.

*forests*

MDPI

*Article*

# Multi-Criteria Filtration and Extraction Strategy for Understory Elevation Control Points Using ICESat-2 ATL08 Product

**Jiapeng Huang *, Yunqiu Wang and Yang Yu**

School of Geomatics, Liaoning Technical University, Fuxin 123000, China; wangyq202410@163.com (Y.W.); 472220873@stu.lntu.edu.cn (Y.Y.)

* Correspondence: huangjp@nefu.edu.cn; Tel.: +86-18941821626

**Abstract:** Understory terrain plays a multi-faceted role in ecosystems, biodiversity, and productivity in forests by influencing different major factors, such as hydrological processes, soils, climate, and light conditions. Strong illuminants (e.g., sunlight) from ground surfaces and atmosphere can introduce additional photons into the ATLAS system. These photons can, consequently, be mistakenly identified as laser photons reflected from ground surfaces. The presence of such ambient light, particularly under low-photon-count conditions, can significantly increase elevation measurement errors. In this context, this study aims to propose a method for extracting reliable understory elevation control points under varying forest conditions, based on the parameter attributes of ICESat-2/ATLAS data. The overall filtered data resulted in a coefficient of determination ($R^2$), root mean square error (*RMSE*), and standard deviation (*STD*) of 0.99, 2.77 m, and 2.42 m, respectively. The greatest accuracy improvement was found in the Puerto Rico study area, showing decreases in the *RMSE* and *STD* values by 2.68 and 2.67 m, respectively. On the other hand, canopy heights and slopes exhibited relatively large impacts on noise interferences. In addition, there were decreases in the *RMSE* and *STD* values by 4.57 and 4.64 m, respectively, under the very tall canopy category, whereas under steep slope conditions, the *RMSE* and *STD* values of the filtering results decreased by 4.59 and 4.34 m, respectively. The proposed method can enhance the overall accuracy of elevation data, allowing for the significant extraction of understory elevation control points, ultimately optimizing forest management practices and improving ecological assessments.

**Keywords:** ICESat-2_ATLAS; understory topography; elevation control points; multi-criteria filtration

## 1. Introduction

Global surface temperatures are rising at an alarming rate [1] due to the ongoing rapid climate changes, exposing forest ecosystems to significant challenges [2]. There is an urgent need to accurately detect actual disturbances and predict future impacts on forest ecosystems [3]. Understory terrain typically refers to the topography, vegetation structures, and ground features beneath forest canopies in densely wooded areas [4]. They play a multi-level role in ecosystems, shaping biodiversity and productivity in forests by influencing multiple important influencing factors, such as hydrological processes, soils, climate, and light availability [5]. Therefore, estimating understory terrain parameters is also crucial.

Traditional forest resource surveys primarily rely on field-based observations. However, this approach only provides measurements at specific individual points. On the other hand, advanced remote sensing technology provides a more comprehensive perspective on forest management and ecological monitoring [6]. For example, Light Detection and Ranging (LiDAR) can accurately measure forest canopy height, tree density, and biomass, with a horizontal error of only 5% to 10% and a vertical error of less than 10 cm. This can not only enhance the efficiency of surveys, but also effectively monitor dynamic changes

in forest resources, providing a scientific basis for sustainable forest management. Unlike optical remote sensing, synthetic aperture radar (SAR) signals can be saturated through biomass increases and canopy closure. The SAR signals are, in fact, often affected by terrain variations [7,8], limiting the effective applications of the SAR-based remote sensing method for monitoring forest resources to some extent. Indeed, optical remote sensing is highly affected by atmospheric conditions and cloud covers, while microwave remote sensing has relatively low spatial resolutions despite its strong penetration capability.

LiDAR has rapidly developed as an active remote sensing technology in recent years. The LiDAR method can provide high-resolution vertical measurements. Its laser pulses can penetrate vegetation covers to directly reach soil surfaces [9]. Compared with airborne LiDAR, spaceborne LiDAR can cover larger surface areas, making it uniquely advantageous for large-scale forest height inversion and mapping [10]. It plays an increasingly important role in environmental monitoring studies [11], digital cities [12], forestry surveys [13], and global change studies [14]. LiDAR has become a critical tool for high-spatiotemporal-resolution–3D Earth observations [15]. NASA's ICESat-2 (Ice, Cloud, and Land Elevation Satellite-2) can effectively provide data for high spatiotemporal resolution 3D Earth observations [16]. The Advanced Topographic Laser Altimeter System (ATLAS) onboard ICESat-2 can provide high-resolution and dense surface elevation data, supporting various environmental monitoring and management applications, thanks to its high frequency and sensitivity [17]. ATLAS is capable of collecting dense photon cloud data and is often used in multiple applications, such as drought stress monitoring [18,19], lake water level monitoring [20,21], global forest biomass monitoring [22,23], and forest tree species mapping [24].

Light sources, such as sunlight, moonlight, and other strong ground and atmospheric light sources, can generate noise photons in the ATLAS system, which can be mistakenly identified as laser photons reflected from ground surfaces and, consequently, increase altimetry errors, especially under low photon counts. Furthermore, aerosols and particles in the atmosphere can scatter laser photons during their propagation, creating extra noise signals, which further affects the performance of the altimetry system. In forest environments, the occurrence of noise photons is mainly due to the influence of vegetation density, solar radiation, terrain complexity, and atmospheric conditions. It is crucial to consider these factors in ATLAS altimetry data processing. Noise photons can be effectively minimized and removed through investigations and quality screening to enhance data accuracy in forested environments and minimize ambient light- and scattering-induced errors. Research shows that in some cases, the *RMSE* (root mean square error) may decrease by 20% to 50%. The data retention rate may reach 70% to 90%, depending on the data quality and filtering criteria.

Existing research focuses on improving the noise removal and elevation estimation accuracy of ATL08 data products, proposing methods such as photon classification, filtering threshold setting, photon count standard expansion, and machine learning algorithms (e.g., XGBoost) to enhance data quality and maintain a high data retention rate. Dong et al. [25] analyzed the reliability of photon classification in ATL08 data products and their accuracies in forest height estimation by comparing the ATL03 and ATL08 data with reference to airborne LiDAR data to effectively remove noises, thus providing a basis for data filtering. They revealed the greater performance of the ATL08 data in representing forest canopy structures in temperate forests, with coefficients of determination ($R^2$) of 0.54 and 0.61 for average and maximum canopy heights, respectively. Wang et al. [26] significantly improved the accuracy of elevation data and data retention rates by setting filtering thresholds based on elevation and ATL08 product attributes, thereby providing a more reliable foundation for ecosystem research that aids in the better monitoring of forest health and supporting sustainable management. According to their results, the root mean square error (*RMSE*) of the elevation control points after filtering ranged from 0.57 to 0.69 m, with a data retention rate range of 60.00%–61.38%, providing critical data support for the high-resolution mapping and product quality verification of high-resolution

satellites in China. Zheng et al. [27] expanded the filtering criteria by including the number of ground photons, further improving the accuracy of the extracted elevation control points, resulting in an average absolute error range of laser footprint points after filtering of 0.13–0.82 m, and consequently making this method applicable under various terrain scenarios. Wang et al. [28] proposed an ICESat-2 ATL08 terrain data filtering method using the XGBoost algorithm, taking into account error-related factors to effectively remove unreliable data. After comparing with ground measurement data, this method improved the accuracy of terrain data by over 40%, with a data retention rate exceeding 60%. However, there is still a lack of comprehensive studies on understory environments. Therefore, this study utilizes ATLAS data parameter attributes to select understory elevation control points. Moreover, there are spatial differences in the effectiveness of elevation point filtering between different regions, particularly in complex terrains, where it is challenging to filter out elevation control points that conform to actual true values. Accuracy analysis is based on different landforms, slopes, aspects, surface cover types, and canopy height classes.

In this context, the present study aims to explore the applicability of ICESat-2 ATL08 in extracting elevation control points from spaceborne photon-counting LiDAR data in forested areas. ① Specifically, multiple parameter constraints were applied for fine filtering using an ATL08-based multi-criteria approach to extract elevation control points. ② The accuracy of the ALTAS data in forested environments was further evaluated in this study to further validate the feasibility of the proposed method. ③ The impacts of different landforms, slopes, aspects, land cover types, and canopy heights on noise interferences were also quantitatively assessed in this study by comparing their related *RMSE* and standard deviation (*STD*) values. This provides an applicable solution for extracting elevation points in forest areas under different conditions, thereby enhancing forest management practices and improving ecological assessments.

## 2. Material and Methods

### *2.1. Study Area*

(1) California, USA

California (USA) is located on the western coast of the United States, between the longitude and latitude ranges of 32.5–42° N and 114–124° W, respectively. The elevation difference in the study area is significant, ranging from the baseline of 0 m at sea level to over 4400 m. California's natural landscapes host diverse forest types, including evergreen coniferous, deciduous broadleaf, montane mixed, and desert forests. The forest cover is characterized by a significant regional pattern. Specifically, 20% of this cover is located in the Sierra Nevada and coastal areas, while less than 5% is found in the Central Valley and eastern deserts. Satellite remote sensing data and GIS analyses indicate that approximately 33% of California's land area is forested. The dominant tree species in these forest areas include coast redwood (*Sequoia sempervirens*) and Douglas fir (*Pseudotsuga menziesii*), with canopy heights typically ranging from 40 to 100 m. Because of its rich ecological diversity and complex topographical features, there are significant impacts of climate change in this area, and advanced applications of remote sensing technology and sustainable forest management are thus made difficult.

(2) Quintana Roo, Mexico

Quintana Roo (Mexico) is located on the eastern Yucatán Peninsula of Mexico, between the longitude and latitude ranges of 19.6–21.3° N and 86.8–88.3° W, respectively, with an average elevation range of 0–100 m. This region has abundant forest resources, distributed across coastal areas, wetlands, and inland regions. The dense vegetation in this area can create complex understory terrain, especially during the rainy season, where the leaf litter and low vegetation layers can further hinder the vegetation layer. Quintana Roo has different forest types, including tropical rainforests, dry forests, and mangroves. The forest cover occupies about 50% of the total area of Quintana Roo, and deforestation has affected this ecosystem along coastal areas due to the development of tourism. The heights of the trees in the tropical forests typically range from 20 to 40 m, while the larger emergent trees

in the rainforests reach over 30 m. The dominant species in this area include red mangrove (*Rhizophora mangle*), neem (*Azadirachta indica*), and Puerto Rican oak (*Quercus prinoides*).

(3) Puerto Rico, USA

Puerto Rico (USA) is located in the Caribbean, with geographic coordinates between 17.9–18.5° N and 65.2–67.3° W. The island's terrain ranges from coastal plains to inland mountains and rainforests, creating a diverse landscape. The majority of the coastal areas are located at low elevations, between 0 and 100 m. The forests in this area are primarily distributed in the Central Cordillera, inland lowlands, hills, coastal areas, and estuaries. Puerto Rico's forest cover is about 55% of the total area of Puerto Rico, with tropical rainforests, dry forests, mangroves, and secondary forests being the main forest types. In the coastal areas and lowlands, the understory terrain is generally flat, with dense vegetation in the mangroves and wetlands. In contrast, the understory terrain in the central mountain ranges and tropical rainforests is more complex, with dense vegetation, thick leaf litter layers, and high densities of low plants, making it more difficult for photons to penetrate the vegetation layer. The heights of the canopies in the tropical rainforests and wet forests of this study area range from 20 to 30 m, with some large trees exceeding 35 m. In the dry forests and coastal mangroves, on the other hand, the tree heights are typically between 5 and 15 m. The major tree species in these forests include West Indian mahogany (*Swietenia mahagoni*), Puerto Rican oak (*Quercus prinoides*), red mangrove (*Rhizophora mangle*), and bullet tree (*Bucida buceras*).

(4) South Carolina, USA

South Carolina (USA) is located in the southeastern part of the United States, with geographic coordinates between 32–35° N and 78.5–83.5° W. The understory terrain in this state varies depending on the topography and vegetation types. In the coastal plain, the terrain is generally flat, with an elevation range of 0–30 m, particularly in the wetlands and floodplains, where the vegetation is dense. In contrast, the Piedmont Plateau has more rolling hills and valleys, with elevations of about 1085 m. The slope in South Carolina gradually increases from the northwestern to the southeastern part. Approximately 66% of South Carolina's land area is forested, including pine forests, mixed forests, hardwood forests, and wetland forests. The dominant tree species include longleaf pine (*Pinus palustris*), swamp chestnut oak (*Quercus michauxii*), shortleaf pine (*Pinus echinata*), and hickory (*Carya* spp.). The forest types in the four study areas are shown in Figure 1.

### *2.2. Study Data*

In this study, ICESat-2/ATLAS data were utilized, and G-LiHT DTM data were used for validation to assess the accuracy of the filtered ICESat-2/ATLAS data. In addition, this study also utilized slope and aspect information derived from the Digital Terrain Model (DTM) provided by G-LiHT, forest cover type data provided by Global Land Cover (GLC), and canopy height data provided by GEDI as classification criteria to assess the accuracy and reliability of the filtered data. These are used to evaluate the performance of the filtered ICESat-2/ATLAS ground control points under different terrains, slopes, aspects, forest types, and canopy heights, thus ensuring the precision and validity of the research results.

#### 2.2.1. ICESat-2/ATLAS Data Product

ICESat-2 was launched in 2018, employing photon-counting multi-beam LiDAR technology to collect elevation data through an ATLAS laser altimeter system [29,30]. ATLAS uses the photon-counting principle to mark and locate each photon [31,32]. The laser beam is divided into six beams (three pairs), with each pair consisting of a strong and a weak beam at an energy ratio of 4:1 [33]. The laser footprint has a diameter and cross-track distance of 17 m and 3.3 km, respectively [34]. The satellite orbits within an altitude range of 496–500 km, covering a large area from 88° S to 88° N [35]. In addition, it emits 10,000 laser pulses per second, with each pulse containing about 200 trillion photons. However, only a small number of photons return to the receiver and reach the detector via fiber optics [36]. ICESat-2 provides up to 20 standard data products, ranging from level 0 to level 3 and

covering different data from raw telemetry to processed elevation, ice sheet, and vegetation canopy information [37]. Although the horizontal and vertical accuracies of ICESat-2 are within meters and sub-meters, respectively, some elevation data have low precision due to atmospheric dispersion, solar background noise, and instrument noise, requiring further filtering for precise measurements [38,39].

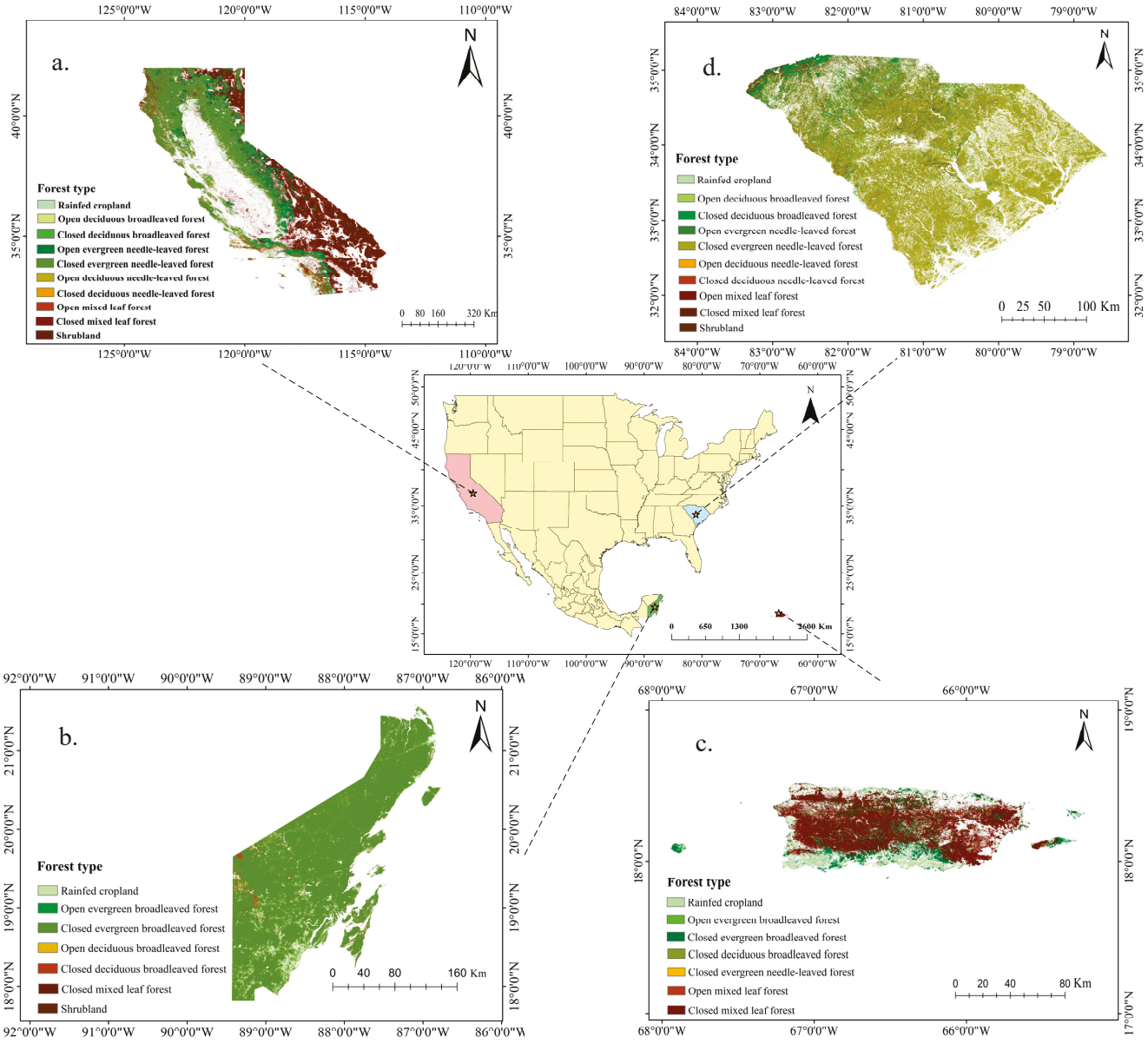



**Figure 1.** Forest types in the study areas: (**a**) California research area. (**b**) Quintana Roo area. (**c**) Puerto Rico area. (**d**) South Carolina area.

Among these data products, the ATL08 (Land and Vegetation Height) data are generated based on the ATL03 data through a series of processing steps and algorithm applications. The processing begins by filtering useful photons from the ATL03 photon data. Subsequently, the data are classified into different surface types (e.g., forest and bare soils) based on the remaining photon data before calculating the heights of each surface type. The height of the surface and vegetation is calculated using the laser reflection time, converting the photon signals into actual ground and vegetation height data. ATL08 is a more application-oriented product obtained by processing and analyzing the data from ATL03. The ATL03 and ATL08 products contain a large amount of parameter information.

By matching the parameters of the two products based on latitude and longitude and filtering the parameters, ground control points were extracted to improve the accuracy of subsequent understory terrain inversion. The ICESat-2/ATLAS parameters used in this experiment are shown in Table 1.

**Table 1.** Introduction to ICESat-2/ATLAS product parameters.

| Data | Parameter | Brief Description |
|---|---|---|
| ATL03 | lat_ph | The latitude of the photon |
| ATL03 | lon_ph | The longitude of the photon |
| ATL03 | segment_id | A seven-digit number that uniquely identifies each track segment |
| ATL03 | Geoid | Add to the geodetic ground to get the geodetic ground height in the measurement system |
| ATL03 | signal_conf_ph | The confidence level is associated with each photon event selected as the signal for each surface type |
| ATL03 | ph_index_beg | The index of the first photon in a given segment |
| ATL08 | latitude | 100 m segmented latitude |
| ATL08 | longitude | 100 m segmented longitude |
| ATL08 | h_te_best_fit_20m | The best topographic elevation for the midpoint of the 20 m lot |
| ATL08 | h_te_uncertainty | Uncertainty in ground height estimates |
| ATL08 | centroid_height | The absolute height above the reference spheroid associated with the centroid of all signal photons |
| ATL08 | classed_PC_flag | The classification markers for each photon are noise, ground, top of the crown, or top of the crown |
| ATL08 | classed_PC_indx | The photon index is tracked to ATL03, and surface discovery software is identified and used in the creation of data products |
| ATL08 | segment_landcover | Reference land cover for the world's best land cover products |
| ATL08 | night_flag | A flag indicating whether the measured value was obtained in nighttime conditions |
| ATL08 | h_te_skew | The slope of the topographic height of the line segment |
| ATL08 | msw_flag | Multiple scattering warning products |
| ATL08 | Cloud_flag_atm | Cloud flags |
| ATL08 | snr | Background letter to noise level |
| ATL08 | n_seg_ph | The number of photons in each land segment |
| ATL08 | n_te_photons | The number of ground photons in a segment |

Specifically, the ATL03 data were employed for generating ATL08, while the ATL08 data were used to convert ATL03 photon data into actual environmental and surface information. Both datasets were freely downloaded from the National Snow and Ice Data Center (NSIDC) website (https://nsidc.org/data/, accessed on 16 November 2024) [40].

#### 2.2.2. G-LiHT Data

The G-LiHT system (Goddard's LiDAR, Hyperspectral, and Thermal Imager) is a portable airborne imaging system that combines optical and laser technologies to acquire

surface elevation data. NASA's official team designed G-LiHT as a multi-sensor airborne imaging system, enabling it to simultaneously map the composition, structure, and function of terrestrial ecosystems [41]. The G-LiHT system includes imaging spectrometers for species composition and biophysical parameter distinctions, thermal payload for surface temperature analysis, and LiDAR payload for accurate 3D observations of land ecosystems [42]. G-LiHT is equipped with the VQ-480 LiDAR to provide vertical and horizontal canopy distribution information [43]. Detailed parameters of the VQ-480 airborne laser are reported in Table 2.

**Table 2.** Hardware parameters of the VQ-480.

| Parameters | Values |
|---|---|
| Beam divergence/mrad | 0.3 |
| Wavelength/nm | 1550 |
| Pulse frequency/kHz | 300 |
| Maximum number of pulse returns | 8 |
| Maximum scanning angle/(°) | 36 |
| Point Density/$m^2$ | 29.6 |

The G-LiHT data products will be used to provide information on the photosynthetic function and vegetation productivity of forest vegetation in the four study areas: California research area, Quintana Roo area, Puerto Rico area, and South Carolina area. These products include spatially explicit remote sensing indicators of dynamic forest vegetation. Additionally, the products were used to describe fine-scale spatiotemporal heterogeneity in the ecosystem structure and function under different environmental and climatic conditions [44]. However, new methods of data fusion could be created to monitor ecosystem health and the impacts of climate and human-induced changes on these ecosystems [45]. The data were freely downloaded in this study from the G-LiHT official website (https://glihtdata.gsfc.nasa.gov, 16 November 2024).

#### 2.2.3. Land Cover Data

GLC_FCS30-2020 is a high-resolution global land cover classification product provided by the Institute of Remote Sensing and Digital Earth (RADI) of the Chinese Academy of Sciences and other collaborative institutions. The data had a resolution of 30 m and included 29 land cover types. The forest cover type was subdivided into 12 different forest subtypes. This dataset provides a fundamental data source for the fine-scale analysis of land cover changes.

In this study, 12 forest types were considered as classification standards to monitor the understory terrains of the study areas. These forest types include tropical rainforests, shrublands, open evergreen broadleaf forests, closed evergreen broadleaf forests, open deciduous broadleaf forests, closed deciduous broadleaf forests, open evergreen needleleaf forests, closed evergreen needleleaf forests, open deciduous needleleaf forests, closed deciduous needleleaf forests, open mixed forests, and closed mixed forests. The data were downloaded from the website (https://zenodo.org/records/3986872, 16 November 2024).

#### 2.2.4. Canopy Height Data

The canopy height data provided by Global Ecosystem Dynamics Investigation (GEDI) were obtained using LiDAR technology. These data were used mainly to measure and monitor the vertical structure of global forests. They are, in fact, characterized by high sampling frequency and a 30 m spatial resolution, with a four-day revisit cycle, covering the 51.6° S–51.6° N area. The data are typically provided in a raster format. The footprint diameter of 25 m can effectively reduce the effects of terrain slopes to some extent [46]. The sensor measures the height of the target by emitting laser pulses and receiving reflected signals. The laser pulses penetrate the forest canopy before reflecting back from the tops of trees and ground surfaces. The time difference of these reflected signals is used to calculate

heights of vegetation from ground surfaces to the tops of trees, as well as other information on the vertical structure of forests [47].

GEDI's canopy height data provide height information at each pixel, thereby supporting forest biomass assessments, health monitoring, ecosystem management, and climate change impact studies. These data were downloaded from NASA's public data platform and the GEDI data portal at (https://gedi.umd.edu/, 16 November 2024).

## 3. Methods

ICESat-2/ATLAS ATL08 can continuously provide terrain height data at 100 m intervals, along with associated descriptive parameters. Due to the extremely high sensitivity of the ATLAS, the data can be significantly affected by atmospheric dispersion and solar background noise, resulting in a substantial amount of randomly distributed noisy photons in the provided ATL08 products. It cannot be directly applied to 30 m-resolution understory terrain estimation. Therefore, it is crucial to enhance the accuracy of the ATL08 product.

In this study, the ATL08 terrain products were filtered based on evaluation labels to obtain high-precision elevation control points for subsequent understory terrain inversion. The specific process used in this study is illustrated in Figure 2.

### 3.1. Preliminary Filtering of the ICESat-2/ATLAS Elevation Data

The Geoid parameter values from ATL03 were associated with the label "h_te_best_fit _20m" parameter from ATL08 using latitude and longitude information from the ICESat-2/ATLAS data. In addition, elevation calibration was performed based on the "h_te_best_fit _20m" label using the Geoid. This process resulted in corrected elevation values consistent with the actual elevation coordinate system. The detailed formulas are as follows:

$$H_{ATL08} = h\_te\_best\_fit\_20m - Geoid \quad (1)$$

In the formula, $H_{ATL08}$ represents the elevation correction value, *h_te_best_fit_20m* is the best fit elevation value for the midpoint position of each 20 m segment provided by ATL08, and *Geoid* refers to the photon height from ATL03 that can be converted to the height above the reference ellipsoid.

The obtained corrected elevation values obtained were subsequently filtered. First, the segment_id parameter from ATL03 was matched with the ph_segment_id parameter from ATL08 to obtain the starting photon index (ph_index_beg) for this segment. The transmission time index of the photon in the ATL03 file was determined by adding the relative photon index (classed_pc_indx) to the starting photon index for that segment. The ground photons in ATL03 were determined based on the classed_PC_flag parameter, retaining only those with classed_PC_flag values of 1, while discarding others. The signal_conf_ph parameter from ATL03 was subsequently extracted, retaining only high-confidence photons with signal_conf_ph values 4, mitigating the influences of noisy photons on the data accuracy to some extent. These filtering steps provided final photons with actual surface information.

### 3.2. Extraction of the Understory Terrain Control Points

Wang et al. [26] and Zheng et al. [27] filtered the elevation data of ICESat-2 ATL08 based on reference elevations and certain attribute parameters, successfully improving the accuracy of the extracted ATL08 elevation values to some extent and providing high-precision points as elevation control points. However, relying solely on a single or a few attribute parameters makes it difficult to significantly enhance the precision. Moreover, terrain changes, especially in areas with complex topography and land cover, make it challenging to effectively filter out suitable elevation control points. Therefore, this experiment further expands the filtering criteria, incorporating more parameters for filtering and downloading a large amount of ICESat-2/ATLAS data for a more comprehensive selection, thereby providing a sufficient number of high-precision elevation control points.

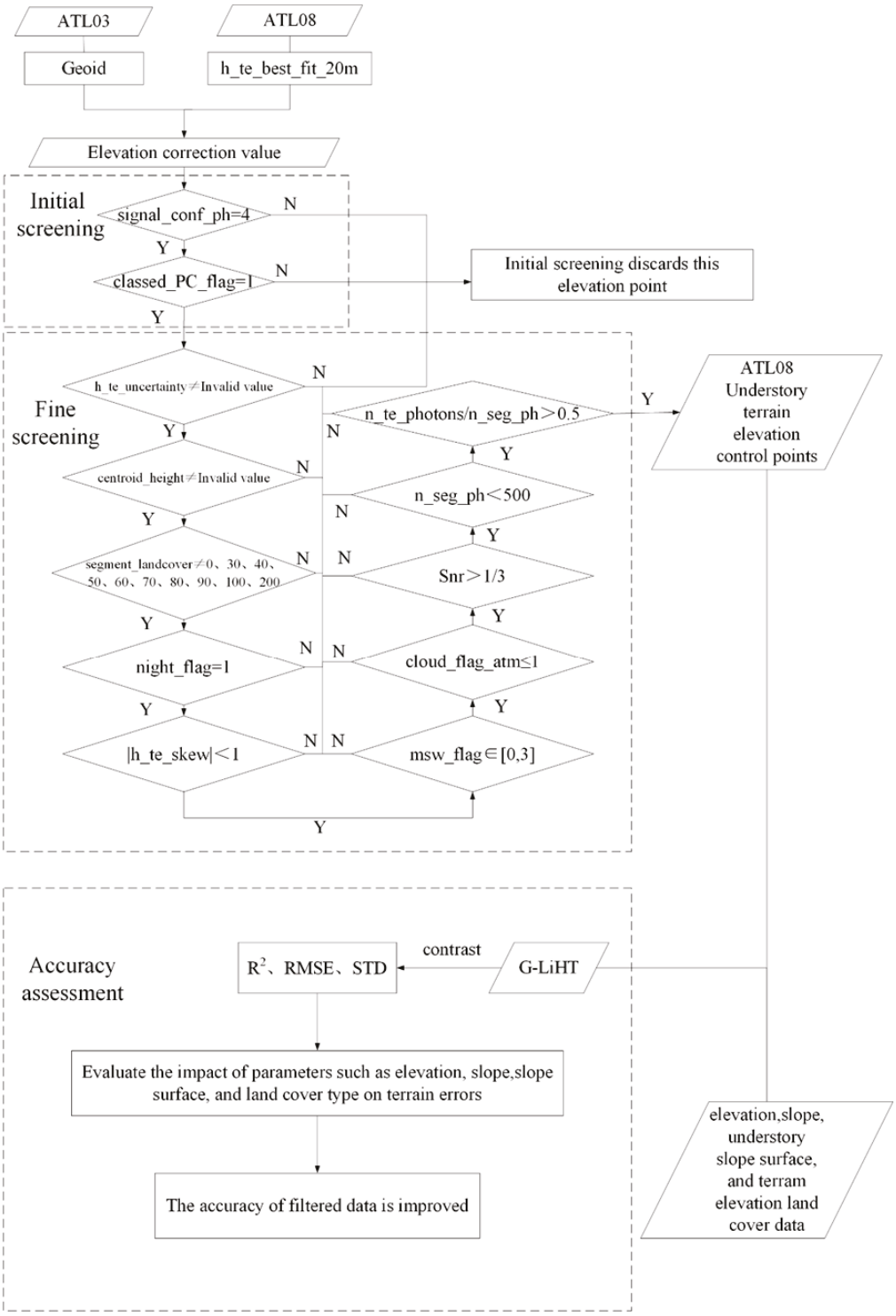



**Figure 2.** Flowchart of the methodology adopted in this study.

In this study, we proposed a comprehensive process for extracting understory elevation control points from the ICESat-2/ATLAS data, taking into account the different parameter characteristics of the ICESat-2/ATLAS data. This process meticulously considers the impacts of various factors, such as cloud covers, slopes, and signal-to-noise ratios (SNR), on the ground elevation accuracy. We performed fine filtering using the parameters provided by the ATL08 product and designed a detailed process that includes nine filtering

steps to effectively remove low-precision photons caused by multiple factors (e.g., the atmosphere, terrain, and sunlight). These steps are as follows:

(1) Filtration based on the h_te_uncertainty and centroid_height parameter. The h_te_uncertainty parameter was assigned an invalid value ($3.4028235 \times 10^{38}$) under yhe condition of a serious shortage of terrain photons, making it unable to reflect the variability of elevation measurements in specific segments. Under these conditions, the 3.4 parameter can be discarded. If the number of returned photons in a segment is too low or if the photons cannot be correctly classified as canopy or ground points, the centroid_height parameter was given an invalid value ($3.4028235 \times 10^{38}$). Hence, it cannot provide accurate terrain or vegetation height information, necessitating its removal;

(2) Filtration based on the segment_landcover parameter. Since the present experiment focuses mainly on the understory terrain, the photon points were filtered based on the land cover type criteria provided by the segment_landcover parameter, retaining those with segment_landcover values of 20, 111–116, and 121–126 as understory terrain photons;

(3) Filtration based on the night_flag parameter. The presence of sunlight can generate a large number of noisy photons, decreasing the SNR values and, consequently, affecting the accuracy of the ground elevation data. To mitigate the impact of sunlight on the ground elevation accuracy, we only retain nighttime photons (night_flag = 1);

(4) Filtration based on the h_te_skew parameter. The terrain skewness provided by h_te_skew can significantly influence the accuracy of the understory elevation data. We filter segments with $|\text{h\_te\_skew}| < 1$ to ensure an even distribution of the elevation values and reduce the occurrence of potential outliers;

(5) Filtration based on the msw_flag parameter. A higher msw_flag value indicates a more severe influence of multiscattering, which may negatively affect the accuracy of the extracted data. Therefore, the msw_flag value was within the 0–3 range to filter out unreliable data and improve the data accuracy;

(6) Filtration based on the cloud_flag_atm parameter. Photon points with a high number of cloud or aerosol layers were removed through the filtration process. The cloud_flag_atm values were at $\leq$1 to reduce the impact of cloud or aerosol-induced photon signal attenuation and scattering on the ground elevation measurement accuracy;

(7) Filtration based on the SNR. Higher SNR values indicate that signal strengths are far above noise levels, typically suggesting better data quality and more reliable measurement results. Therefore, only elevation data with SNR values greater than 1:3 were retained in this study to ensure good measurement accuracy;

(8) Filtration based on the n_seg_ph parameter. Generally, a higher n_seg_ph value indicates a greater photon density in a specific segment, providing a more accurate description of surface features and improved elevation data accuracies. However, when the photon count exceeds a certain threshold, errors may be introduced. Therefore, the n_seg_ph values were set below 500 to ensure more accurate and reliable elevation data;

(9) Filtration based on the n_te_photons parameter. A higher n_te_photons value often indicates the presence of abundant ground photons in a specific segment, which can typically improve the accuracy of ground elevation measurements. However, an excessive number of photons may introduce noise from non-ground signals, thereby reducing the accuracy of elevation data. Hence, it is crucial to accurately set the n_te_photons range to provide more accurate ATL08 data. It is recommended to retain segments where ground photons account for over 50% to optimize the accuracy of elevation data, according to the following formula:

$$T = \frac{N_{terrain}}{N_{Total}} > 50\% \quad (2)$$

where $T$ denotes the coefficient of proportionality, $N_{terrain}$ denotes the number of terrain photons within a specific segment, and $N_{Total}$ denotes the total number of photons within a specific segment.

### 3.3. Influences of Landform, Slope, Aspect, Land Cover Type, and Canopy Height on the Photon Noise Generation

The ATLAS-based generation of noise photons can be influenced not only by solar background noises, but also by multiple factors, including topography [48], slope [49], aspect [50], land cover [51], and canopy height [52]. These influences can subsequently affect the estimated vegetation elevations. Complex terrains, such as mountains, valleys, and hills, can cause variations in laser pulse echo timing, affecting the interaction between the laser and the ground. This can result in more irregular distributions of noise photons and increase the uncertainty of echo signals.

The reflection angles of laser pulses on the slope become more negatively affected with increasing slope degrees, leading to unclear signals, increased noise photons, and greater errors. The slope aspect parameter can also influence reflection angles. For instance, slopes facing the sun may produce more background noise photons due to increased solar radiation when compared with shaded slopes. Land cover types (e.g., bare soil, vegetation, and water bodies) can also affect the reflection characteristics of lasers. For example, densely vegetated areas can absorb or scatter more photons, while water bodies and snow have high reflectivity, increasing noise photons. Moreover, the height of the canopy in forests can negatively affect signals through reflections, particularly from the canopy, further increasing the number of noise photons. Therefore, it is crucial to consider these influencing factors through elevation data processing to effectively improve the accuracy and reliability of the extracted data.

### 3.4. Accuracy Assessment

We extract the understory terrain photon data using latitude and longitude information, along with topography, slope, aspect, forest cover type, and canopy height categories, and compare the overlapping points with the corresponding G-LiHT DTM validation data.

To evaluate the accuracy of the filtered photons, the ATLAS-derived understory terrain control points were compared with the G-LiHT DTM reference data using three statistical metrics, namely, $R^2$, *RMSE*, and *STD*, according to the following equations [53]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (e_i - r_i)^2}{\sum_{i=1}^{n} (r_i - \bar{r})^2} \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (e_i - r_i)^2} \tag{4}$$

$$STD = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2} \tag{5}$$

where $n$ denotes the number of ground photons; $e_i$ denotes the elevation of the $i$-th ground point or the elevation of the photon minus the corresponding interpolated ground elevation from the ground photons; $r_i$ denotes the DTM elevation, corresponding to the $i$-th photon; $x_i$ denotes the $i$-th data point; and $\overline{x}$ denotes the mean of all data points.

## 4. Results and Discussion

### 4.1. Discussion on the Extraction Results of Understory Terrain Control Points and the Influence of External Factors on Filtering Accuracy

In this study, the airborne laser scanning point cloud data of California, South Carolina, Puerto Rico, and Quintana, Mexico, were filtered to extract the ground elevation control points using the ICESat-2/ATLAS data-based method, thereby effectively describing the understory terrain in these areas. In this experiment, the G-LiHT DTM data were used as the validation dataset. The obtained analysis results are reported in Table 3.

**Table 3.** Results of the stepwise analysis.

| Method | Evaluation Indicators | | | |
|---|---|---|---|---|
| | N | $R^2$ | RMSE/m | STD/m |
| Unfiltered results | 15,101 | 0.99 | 3.22 | 3.18 |
| Primary screening | 14,686 | 0.99 | 3.19 | 3.15 |
| Fine Screening—(1) | 10,913 | 0.99 | 3.15 | 3.10 |
| Fine Screening—(2) | 8766 | 0.99 | 3.05 | 3.00 |
| Fine Screening—(3) | 4888 | 0.99 | 3.04 | 3.00 |
| Fine Screening—(4) | 3848 | 0.99 | 3.03 | 3.00 |
| Fine Screening—(5) | 3837 | 0.99 | 2.97 | 2.94 |
| Fine Screening—(6) | 3383 | 0.99 | 2.95 | 2.92 |
| Fine Screening—(7) | 3371 | 0.99 | 2.90 | 2.87 |
| Fine Screening—(8) | 3331 | 0.99 | 2.87 | 2.84 |
| Fine Screening—(9) | 1132 | 0.99 | 2.77 | 2.47 |

In this study (Table 3), a total of 1132 photon point cloud data points were retained as understory elevation control points through multiple filtering steps, including initial screening and ATL08-based refined selection. The effectiveness of this filtration process was validated using the different accuracy evaluation metrics. The results reveal a continuous decrease in the *RMSE* and *STD* values through the different filtering steps, reaching decrease rates of the *RMSE* and *STD* of 0.45 and 0.71 m, respectively. This not only indicates that the proposed filtering method had a strong optimization capability for photon point cloud data, but also demonstrates its outstanding performance in noise removal and accuracy enhancement. In addition, the results reveal improved filtering data accuracies for the different study areas, demonstrating the broad applicability of the filtration process. The specific evaluation results are reported in Table 4, further demonstrating the effectiveness and reliability of our proposed method. The error histogram of the study area is shown in Figure 3.

**Table 4.** Comparison between the filtered and unfiltered results in the different study areas.

| Study Areas | Unfiltered | | | | Filtered | | | |
|---|---|---|---|---|---|---|---|---|
| | N | $R^2$ | RMSE/m | STD/m | N | $R^2$ | RMSE/m | STD/m |
| California | 3688 | 0.99 | 4.30 | 3.87 | 596 | 1 | 3.70 | 3.09 |
| Quintana | 3659 | 0.99 | 1.89 | 1.88 | 100 | 1 | 0.82 | 0.75 |
| Puerto Rico | 3662 | 0.99 | 4.26 | 4.26 | 69 | 1 | 1.58 | 1.59 |
| South carolina | 4092 | 0.99 | 1.50 | 1.43 | 367 | 1 | 0.86 | 0.74 |
| All | 15,101 | 0.99 | 3.22 | 3.18 | 1132 | 1 | 2.77 | 2.42 |

The obtained results demonstrate the effectiveness of the filtration process in improving the accuracy of the photon point cloud data in the different study areas (Table 4 and Figure 3) with varying improvement effects. The most significant improvement in the accuracy was found in the Puerto Rico study area, with reductions in the *RMSE* and *STD* values of 2.68 and 2.67 m, respectively. This could be attributed to the complex terrain and larger initial errors in this study area, strengthening the effectiveness of the filtering method in removing noises and improving the accuracy. In contrast, the *RMSE* and *STD* values in the Quintana study area decreased from 1.89 to 0.82 m and from 1.88 to 0.75 m, respectively, while those of South Carolina decreased to 0.86 and 0.74 m. These findings further demonstrate the effectiveness of the filtering method in enhancing the data accuracy in these areas, particularly by separating photon signals from noises. However, the lowest accuracy in this study was found in the California study area, showing decreasing rates of the *RMSE* and *STD* values by 0.6 and 0.78 m, respectively. This may be related to the terrain and forest structure characteristics of this area. The validation data in this area were, in fact, concentrated in the flatter and more sparsely vegetated central region, limiting the

potential of the filtration process to improve the data accuracy. It is worth noting that the filtering method could exhibit a weaker enhancement effect on the data accuracy under relatively stable photon signals with less noise. In summary, the proposed filtering method in this study demonstrated good applicability in different study areas.

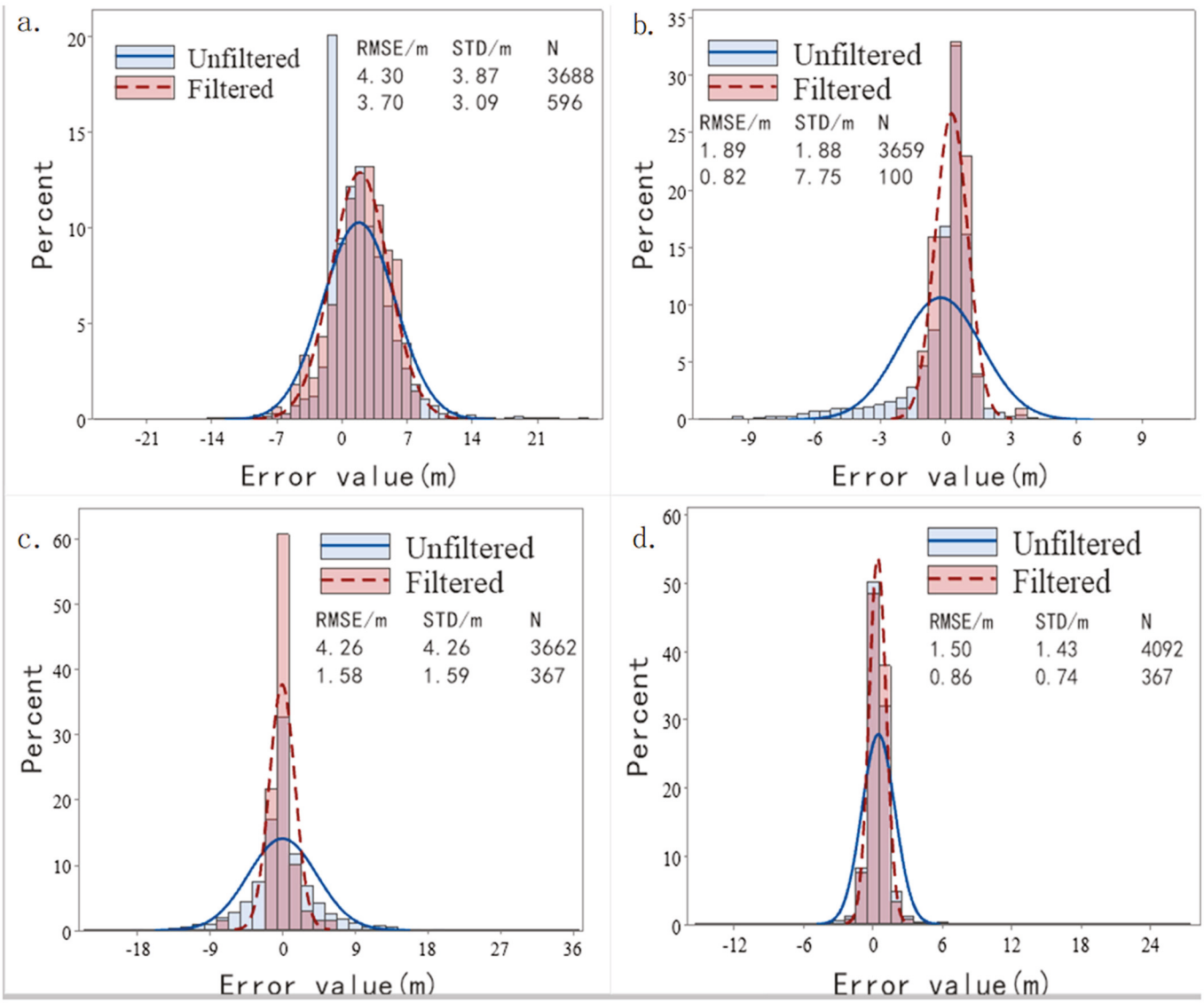



**Figure 3.** Error histogram of the study area: (**a**) California research area. (**b**) Quintana Roo area. (**c**) Puerto Rico area. (**d**) South Carolina area.

### *4.2. Impacts of Topography on the Filtering Results*

The filtering results covered five types of landforms, namely, lowlands, hills, highlands, plateaus, and mountains. In this study, a detailed assessment of the photon point cloud data accuracy was conducted under these typical topographic features before and after the filtration process based on the G-LiHT DTM validation dataset. The results reveal differences in the *RMSE* and *STD* values between the filtered and unfiltered data (Table 5) with the different topographic features. The specific filtering effects are shown in Figure 4.

**Table 5.** Comparative analysis between the filtered and unfiltered results under the different geomorphological conditions.

| Landforms | Unfiltered | | | Filtered | | |
|---|---|---|---|---|---|---|
| | $R^2$ | RMSE/m | STD/m | $R^2$ | RMSE/m | STD/m |
| Lowland | 0.99 | 2.13 | 2.13 | 1 | 1.18 | 1.13 |
| Hills | 0.99 | 4.92 | 4.93 | 0.99 | 3.45 | 3.13 |
| Highlands and plateaus | 0.99 | 5.44 | 5.37 | 0.99 | 3.65 | 3.09 |
| Mountains | 0.99 | 3.33 | 2.00 | 0.99 | 3.32 | 1.50 |
| Plateaus | 0.99 | 6.08 | 5.11 | 0.99 | 4.51 | 2.95 |

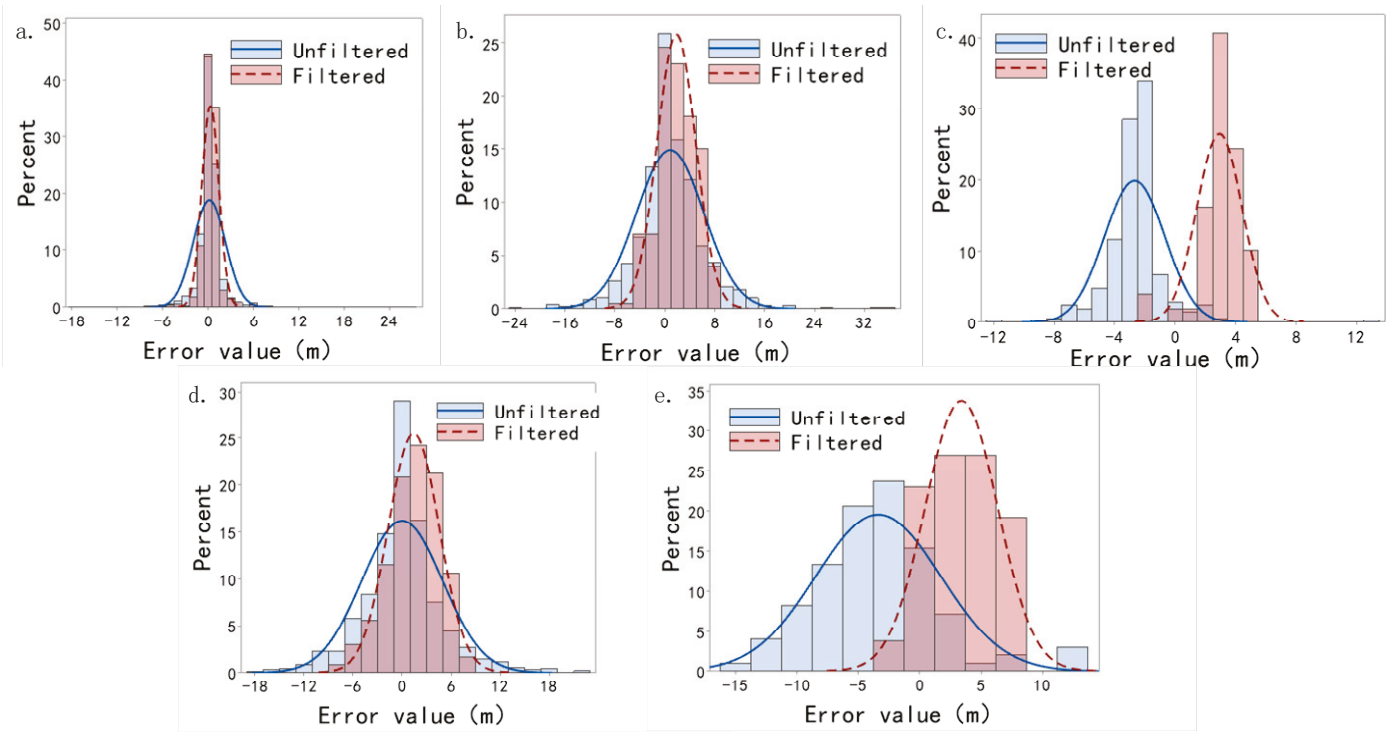



**Figure 4.** Histogram of error distribution for landform types: (**a**) Lowlands. (**b**) Highlands. (**c**) Plateau. (**d**) Mountains. (**e**) Hills.

The results show an improvement in the filtered photon point cloud data under the five different topographic features (Table 5 and Figure 4). However, the extent of the improvement was different between the topographic features. In the mountainous regions, the differences in the *RMSE* and *STD* between the filtered and unfiltered data were only 0.01 m and 0.5 m, respectively. This might be due to the great effect of undulating mountainous terrain, making the separation of photon signals from noise challenging, and consequently limiting the effectiveness of the filtering method. In the lowland areas, the *RMSE* and *STD* errors decreased by 0.95 and 1 m, respectively. This improvement might be attributed to the relatively flat terrain in the lowland areas, making the differences between photon point cloud signals and noise more apparent. This can further enhance the effectiveness of the filtering method in removing noises under such simple topographic conditions.

In the hilly areas, the *RMSE* and *STD* increased to 3.45 and 3.13 m, respectively. In addition, the differences in the *RMSE* and *STD* values between the filtered and unfiltered data were 1.47 and 1.8 m, respectively. This demonstrates the strong adaptability of the filtering method under more complex topographic conditions. In the plateau regions, *RMSE* and *STD* decreased by 1.57 and 2.16 m, even though this terrain condition has significant undulations, showing good overall filtering stability. The most accurate filtering

results were found in the highland and plateau areas, showing decreasing rates of the *RMSE* and *STD* values by 1.79 and 2.28 m, respectively. This might be due to the relatively stable terrain and sparse vegetation in these conditions, resulting in clearer photon signals.

Overall, the filtering method showed strong generalizability, and can enhance the accuracy of photon point cloud data across different topographic conditions, making it suitable for optimizing the ATL08 data under different topographic features.

### *4.3. Impacts of the Slope on the Filtering Results*

This study utilizes the GLC_FCS30-2020 product as a reference for the forest cover data. Three vegetation cover types were identified based on the distribution of the G-LiHT reference data in the selected study areas, namely, closed evergreen needleleaf, closed evergreen broadleaf, and closed mixed forests. Figure 5 shows the *RMSE* and *STD* values for four Digital Elevation Model (DEM) datasets under the different forest types.

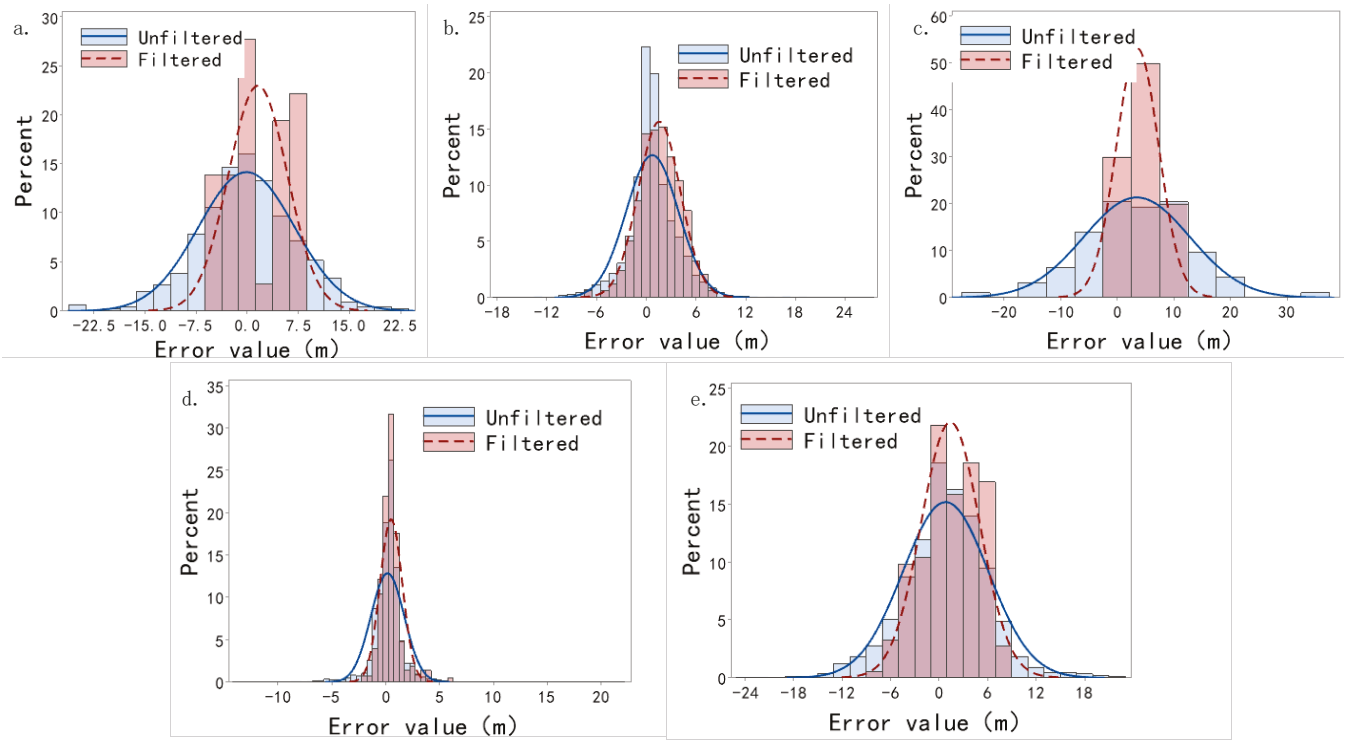



**Figure 5.** Histogram of error distribution for slope types: (**a**) Steep slopes. (**b**) Gentle slopes. (**c**) Abrupt slopes. (**d**) Flat slopes. (**e**) Inclined slopes.

The filtering results cover various slope types, including flat (0°–5°), gentle (5°–15°), moderate (15°–25°), steep (25°–35°), and very steep (35°–45°) slopes. The impacts of these slope conditions on the data accuracy were different. Table 6 compares the unfiltered and filtered ground elevations between the different slope types.

**Table 6.** Comparison between the filtered and unfiltered results under the different slope conditions.

| Slope Types | Unfiltered | | | Filtered | | |
|---|---|---|---|---|---|---|
| | $R^2$ | RMSE/m | STD/m | $R^2$ | RMSE/m | STD/m |
| Flat slope | 1 | 1.56 | 1.55 | 1 | 1.17 | 1.04 |
| Gentle slope | 1 | 3.24 | 3.15 | 1 | 3.03 | 2.55 |
| Slope | 0.99 | 5.33 | 5.27 | 1 | 3.90 | 3.62 |
| Steep slope | 0.99 | 7.04 | 7.04 | 0.99 | 4.59 | 4.34 |
| Precipitous | 0.99 | 9.92 | 9.36 | 0.99 | 5.11 | 3.76 |

According to Table 6, the improvement of the filtered accuracy was relatively small on the flat and gentle slopes. In the range of moderate slopes, on the other hand, the differences in the *RMSE* and *STD* between the filtered and unfiltered data were 1.43 and 1.65 m, respectively, demonstrating the effectiveness of the filtering method in removing some noises under these slope conditions. The *RMSE* and *STD* values of the filtered data were 4.59 and 4.34 m, respectively. This improvement might be due to the complexity of the steep slope terrain, making noises in the photon signals more pronounced and, consequently, enhancing the effectiveness of the filtering method under these slope conditions.

In Figure 5, the *RMSE* and *STD* values of the filtered data in the presence of very steep slopes were 5.11 and 3.76 m, corresponding to reduction rates of 4.81 and 5.6 m, respectively. This result might be attributed to the dramatic changes in the terrain conditions under very steep conditions, increasing noises in the photon signals. The filtration process can, indeed, effectively remove these noises, significantly enhancing the accuracy of the ground elevation data. This indicates that the filtration process has greater advantages in terms of improving data accuracy under steep slope conditions.

### *4.4. Impacts of the Slope Aspect on the Filtering Results*

In this study, we further evaluated the accuracy of the photon point cloud data under eight different G-LiHT DTM-derived aspect conditions. These slope aspect conditions included north-facing (0°–22°, 338°–360°), northeast-facing (23°–67°), east-facing (68°–112°), southeast-facing (113°–157°), south-facing (158°–202°), southwest-facing (203°–247°), west-facing (248°–292°), and northwest-facing (293°–337°) slopes (Table 7). The specific filtering effects are shown in Figure 6.

**Table 7.** Comparison between the filtered and unfiltered results under the different aspect conditions.

| Slope Aspect | Unfiltered | | | Filtered | | |
|---|---|---|---|---|---|---|
| | $R^2$ | RMSE/m | STD/m | $R^2$ | RMSE/m | STD/m |
| North | 0.99 | 3.23 | 3.19 | 0.99 | 3.19 | 2.92 |
| Northeast | 0.99 | 3.14 | 3.11 | 0.99 | 3.02 | 2.74 |
| East | 0.99 | 3.27 | 3.19 | 0.99 | 2.91 | 2.58 |
| Southeast | 0.99 | 3.20 | 3.01 | 0.99 | 2.27 | 2.07 |
| South | 0.99 | 2.64 | 2.63 | 0.99 | 2.32 | 2.19 |
| Southwest | 0.99 | 3.12 | 3.09 | 0.99 | 2.37 | 2.21 |
| West | 0.99 | 3.04 | 3.01 | 0.99 | 2.69 | 2.48 |
| Northwest | 0.99 | 3.55 | 3.52 | 0.99 | 3.05 | 2.50 |

The results reveal different improvement degrees of the accuracy of the ARL08-derived ground elevation between the eight different slope aspects (Table 7 and Figure 6). The southeast-facing aspect exhibited the most significant improvement, with *RMSE* and *STD* values decreased by 0.93 and 0.94 m, respectively, whereas the southwest-facing aspect resulted in the second best improvement in the accuracy of the filtered data, with *RMSE* and *STD* values decreased by 0.75 and 0.88 m, respectively. On the other hand, the lowest improvement in the data accuracy was found under the north-facing aspect, with differences in the *RMSE* and *STD* values between the filtered and unfiltered results by only 0.04 and 0.27 m, respectively. Although there were differences in the improvement of the accuracy of filtered data between the different slope aspects, the overall impact of the slope aspect factor on the filtering results was not substantial. Indeed, some factors, such as terrain uniformity, lighting conditions, and noise characteristics, might collectively limit the overall impact of the slope aspects on filtering results. This suggests the applicability of the filtration method under various aspect conditions, effectively enhancing the accuracy of the ground elevation data. Improvements in accuracy under moderate slope conditions can optimize management practices, enhance ecological assessments, improve estimates of carbon sequestration potential, and deepen the understanding of forest ecosystem dynamics.

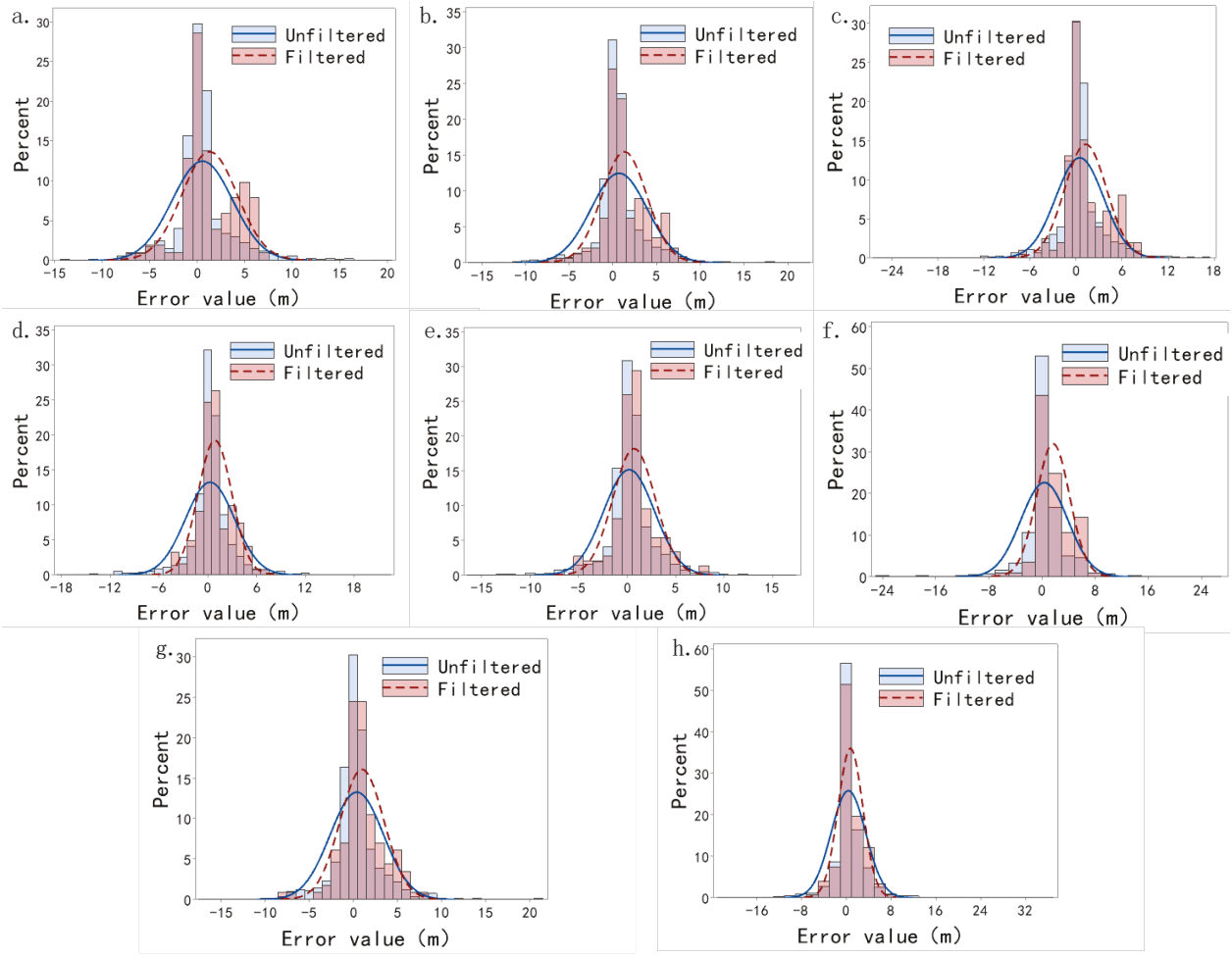



**Figure 6.** Histogram of error distribution for aspect: (**a**) North; (**b**) east; (**c**) northeast; (**d**) southeast; (**e**) south; (**f**) northwest; (**g**) west; (**h**) southwest.

### *4.5. Impacts of the Land Cover Types on the Filtering Results*

The results show the presence of multiple forest types following the filtration process, including tropical rainforest, closed evergreen broadleaf forest, closed deciduous broadleaf forest, open evergreen coniferous forest, closed evergreen coniferous forest, closed deciduous coniferous forest, closed mixed forest, and shrubland. The filtered and unfiltered data were further compared in this study under these eight different forest types (Table 8). The specific filtering effects are shown in Figure 7.

**Table 8.** Comparison between the filtered and unfiltered results under the different land cover types.

| Forest Types | Unfiltered | | | Filtered | | |
|---|---|---|---|---|---|---|
| | $R^2$ | RMSE/m | STD/m | $R^2$ | RMSE/m | STD/m |
| Tropical rainforest | 0.99 | 0.87 | 0.73 | 0.99 | 0.73 | 0.67 |
| Closed evergreen broadleaved forest | 0.99 | 3.14 | 3.11 | 0.99 | 3.02 | 2.74 |
| Closed deciduous broadleaved forest | 0.99 | 3.27 | 3.19 | 0.99 | 2.91 | 2.58 |
| Open evergreen coniferous forest | 0.99 | 3.20 | 3.01 | 0.99 | 2.27 | 2.07 |
| Closed evergreen coniferous forest | 0.99 | 2.64 | 2.63 | 0.99 | 2.32 | 2.19 |
| Closed deciduous coniferous forest | 0.99 | 3.12 | 3.09 | 0.99 | 2.37 | 2.21 |
| Closed mixed forest | 0.99 | 3.04 | 3.01 | 0.99 | 2.69 | 2.48 |
| Shrubland | 0.99 | 3.55 | 3.52 | 0.99 | 3.05 | 2.50 |

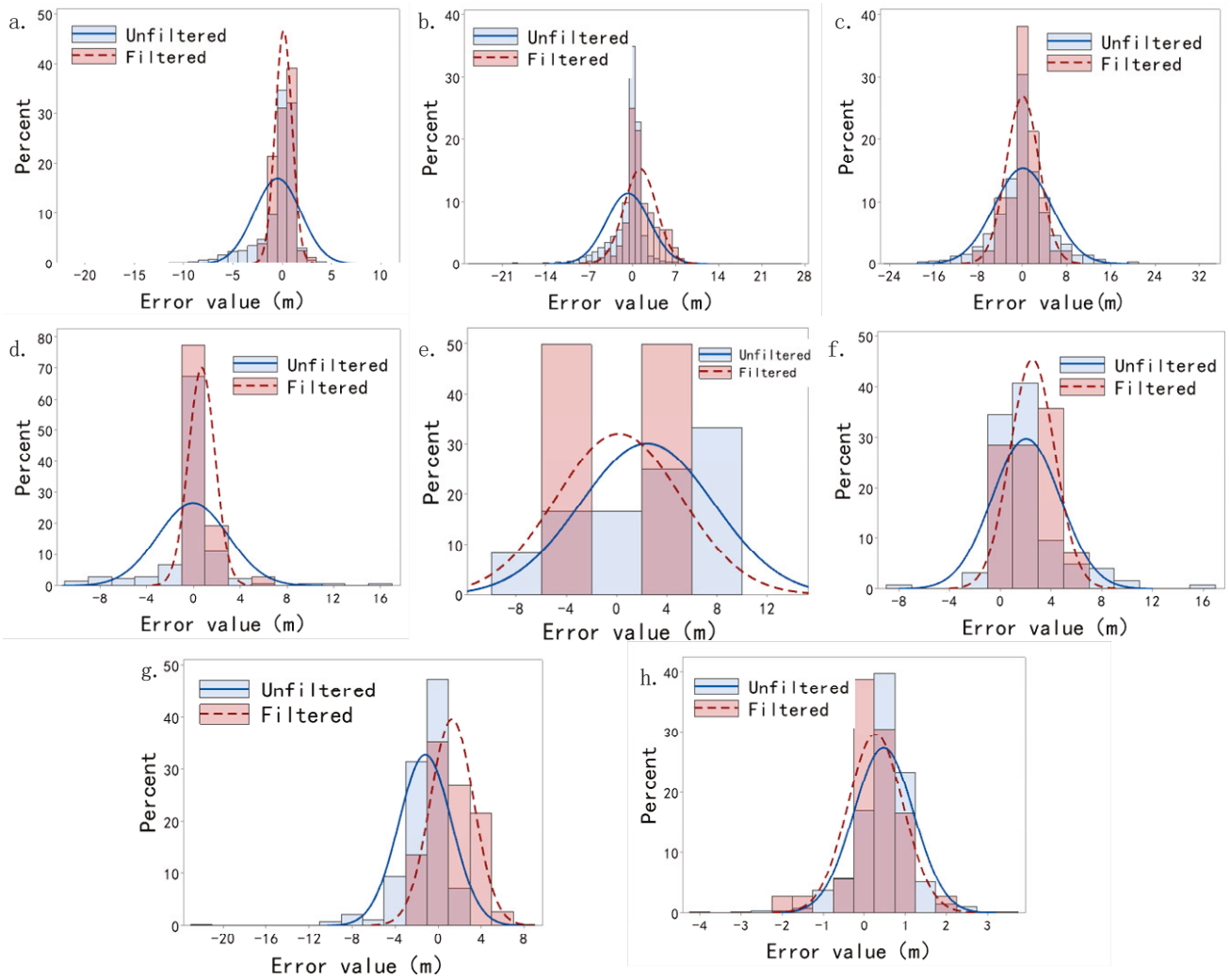



**Figure 7.** Histogram of error distribution for surface cover types: (**a**) closed evergreen broadleaf forest; (**b**) closed evergreen coniferous forest; (**c**) closed mixed-leaf forest; (**d**) closed deciduous broadleaf forest; (**e**) closed deciduous coniferous forest; (**f**) shrubland; (**g**) open evergreen coniferous forest; (**h**) tropical rainforest.

The accuracy of the filtered photon point cloud data showed significant differences between the different forest types (Table 8 and Figure 7). For instance, the *RMSE* improved by only 0.14 and 0.12 m under the presence of the tropical rainforests and closed evergreen broadleaf forests, respectively. This might be related to the high densities and complex structures of these two forest types, making it difficult to effectively remove noises from the photon signals.

The improvement effects on the filtered data accuracy under the closed deciduous broadleaf, closed evergreen coniferous, and closed mixed forest conditions were relatively similar, resulting in decreases in the *RMSE* and *STD* values by average values of 0.34 and 0.53 m, respectively. This indicates that these forest types were structurally more uniform, effectively removing noises from the photon signals through the filtration process. The *RMSE* and *STD* values under the shrubland were increased by 0.5 and 1.02 m following the filtration process, which might be related to the low and sparse characteristics of the shrubland. Although the separation of the photon signals from noises was relatively simple, the overall accuracy still needs improvement. There was a noticeable improvement in the accuracy of the filtration process under the closed deciduous coniferous forest condition, showing *RMSE* and *STD* values that were decreased by 0.75 and 0.88 m, respectively. Under the open evergreen coniferous forests, on the other hand, the *RMSE* and *STD* values were

decreased by 0.93 and 0.94 m, respectively. This improvement might be due to the openness of this forest type, improving the photon signals and thus enhancing the effectiveness of the filtration process in reducing noises and significantly increasing the data accuracy.

In summary, the structural characteristics and lighting conditions of the different forest types significantly influenced the effectiveness of the filtration process in this study. Specifically, the most accurate filtration was obtained under the open and structurally simple forest types. The filtering accuracy is better in sparsely vegetated areas compared to densely vegetated areas. Lower coverage density and larger canopy gaps provide a clearer path for signal photons emitted by ICESat-2, allowing them to penetrate the ground more effectively, thereby significantly improving estimation accuracy.

### *4.6. Impacts of the Canopy Heights on the Filtering Results*

In this study, five canopy height categories were retained following the filtration process, namely, shrubs (below 6 m), small canopy (6–10 m), medium canopy (11–20 m), tall canopy (21–30 m), and very tall canopy (above 30 m). The unfiltered and filtered photon point cloud data were compared under these five different canopy height categories (Table 9). The error histograms under different canopy height are shown in Figure 8.

**Table 9.** Comparison between the filtered and unfiltered results under the different canopy categories.

| Canopy Heights | Unfiltered | | | Filtered | | |
|---|---|---|---|---|---|---|
| | $R^2$ | RMSE/m | STD/m | $R^2$ | RMSE/m | STD/m |
| Shrubs | 0.99 | 3.14 | 3.04 | 0.99 | 2.68 | 2.24 |
| Small Canopy | 0.99 | 3.92 | 3.87 | 0.99 | 3.54 | 3.07 |
| Medium Canopy | 0.99 | 3.65 | 3.65 | 0.99 | 2.68 | 2.39 |
| Large Canopy | 0.99 | 4.29 | 2.18 | 0.99 | 3.25 | 2.95 |
| Giant Canopy | 0.99 | 8.07 | 8.10 | 0.99 | 3.5 | 3.46 |

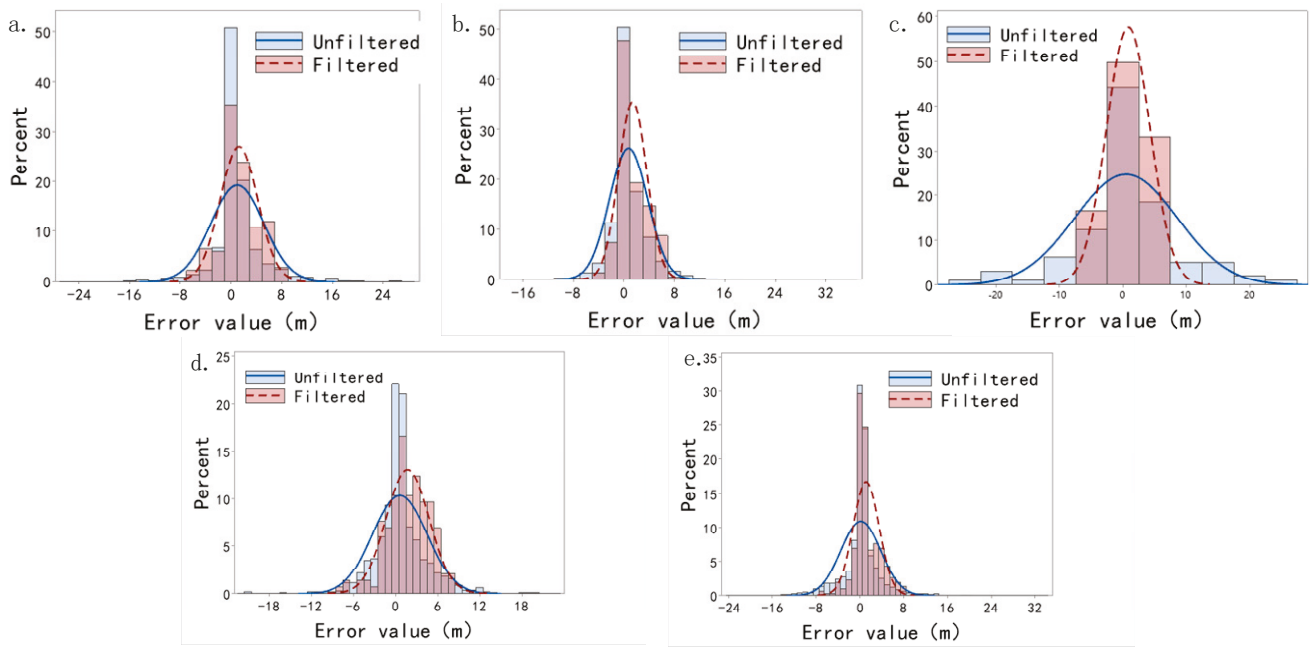



**Figure 8.** Error histogram of tree height: (**a**) large trees; (**b**) shrubs; (**c**) error histogram of tall trees; (**d**) small trees; (**e**) medium trees.

The accuracy of the filtered photon point cloud data shows significant differences between the canopy height categories (Table 9 and Figure 8). Under the presence of the shrub and small canopy categories, the *RMSE* and *STD* values were decreased by average

values of 0.42 and 0.8 m, respectively, whereas under the medium canopy category, the *RMSE* and *STD* values decreased by 0.96 and 1.26 m, respectively. In addition, there was a difference in the *RMSE* value between the filtered and unfiltered data in the presence of the large canopy category of 1.04 m. The *RMSE* and *STD* values decreased by 4.57 and 4.64 m, respectively, in the very large canopy category. This might be attributed to the strong effects of background noises on the photon signals in the presence of short canopy species, negatively influencing the effectiveness of the proposed filtration method. In contrast, the characteristics of the canopy can easily extract photon signals with increasing canopy heights, reducing the interference from noises and thus improving the effectiveness of the filtration process.

In summary, the characteristics of the different canopy height categories and their influences on the photon signals led to variations in the improvement degrees of the filtration accuracy. This indicates that the filtration method has strong adaptability and effectiveness in processing photon data in the presence of canopy species of different heights. The results show that the $R^2$ values are the same, likely due to the high accuracy of ICESat-2 data, resulting in a minimal impact on $R^2$. However, when using ICESat-2 data for denoising in different forest regions, the *RMSE* varies significantly.

## 5. Conclusions

This study proposes a method for extracting understory elevation control points from forested areas based on the ICESat-2/ATLAS data. The parameters of the ATL03 and ATL08 products were combined through the filtration process to ensure that almost all corrected elevation values were provided by actual signal photons. In addition, the use of additional ALT08-derived parameters enhanced the effectiveness of the filtration process in reducing the impacts of influencing factors, such as sunlight and clouds, on the photon data accuracy, resulting in the extraction of reliable understory elevation control points.

1. The experimental results indicate that the filtration process proposed in this study significantly improved the accuracy of the photon point cloud data, reducing the *RMSE* and *STD* to 2.33 and 2.18 m, respectively. The most notable improvement was observed in the Puerto Rico study area, where the *RMSE* and *STD* decreased by 2.68 and 2.67 m, respectively.
2. Among these factors, the canopy height and slope had relatively significant impacts on the noise interferences. The *RMSE* and *STD* values of the filtering results under the canopy height classes were decreased by 4.57 and 4.64 m, respectively, whereas under the steep slope conditions, the *RMSE* and *STD* values were 4.59 and 4.34 m, respectively.

Our results reveal enhanced filtering results under different conditions with varying degrees. Moreover, in this experiment, the filtering parameters are universal. In the future, the more precise classification of parameters such as h_te_skew, segment_landcover, and h_te_uncertainty can be conducted to improve the filtering accuracy of ground photons under the influence of slope, surface cover, and terrain. In future experiments, I will study more research areas to enhance the generalizability of the method. This finding demonstrates the applicability of the proposed method under various conditions and multiple study areas, providing a more effective approach for exploring terrain features in forested environments.

**Author Contributions:** J.H. and Y.W.: methodology. Y.Y.: formal analysis. J.H.: investigation. J.H.: resources. Y.Y.: data curation. Y.W.: writing and original draft preparation. J.H.: writing, review and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (Grant Nos. 42401488 and 42071351), the National Key Research and Development Program of China (Grant Nos. 2020YFA0608501 and 2017YFB0504204), the Liaoning Province Doctoral Research Initiation Fund Program (2023-BS-202), and Basic Research Projects of Liaoning Department of Education (JYTQN2023202).

**Data Availability Statement:** Data are contained within the article.

**Acknowledgments:** We acknowledge NASA and GEE for providing the datasets used in this study.

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

## References

1. Beloiu, M. Forest Response to Climate Warming and Drought in Europe, Bayreuth. Ph. D. Thesis, Faculty of Biology, Chemistry and Earth Sciences, University of Bayreuth, Bayreuth, Germany, 2022.
2. Petritan, A.M.; Schwenke, M.B. Forest Functioning under Climate Warming and Future Perspectives on Forest Disturbances. *Forests* **2023**, *14*, 2302. [CrossRef]
3. Coops, N.C.; Tompalski, P.; Goodbody, T.R.H.; Achim, A.; Mulverhill, C. Framework for near real-time forest inventory using multi source remote sensing data. *For. Int. J. For. Res.* **2023**, *96*, 1–19. [CrossRef]
4. Yang, Z.; Zhou, W.; Gu, C.; Xu, L. Application of ICESat-2/ATLAS Radar Data in Forestry. *Mod. Agric. Sci. Technol.* **2024**, 176–181. [CrossRef]
5. Food and Agriculture Organization of the United Nations. *Global Forest Land-Use Change from 1990 to 2010: An Updata to a Global Remote Sensing Survey of Forests*; Food and Agriculture Organization of the United Nations: Rome, Italy, 2014.
6. Muzalevskiy, K.; Zeyliger, A. Application of Sentinel-1B Polarimetric Observations to Soil Moisture Retrieval Using Neural Networks: Case Study for Bare Siberian Chernozem Soil. *Remote Sens.* **2021**, *13*, 3480. [CrossRef]
7. Pandit, A.; Ramsankaran, R.A.A.J.; Rao, Y.S. Generation and Validation of the Interferometric SAR DEMs from TanDEM-X data for Gangotri and Hamtah Glaciers of Indian Himalayas. *Procedia Technol.* **2014**, *16*, 793–805. [CrossRef]
8. Sun, G.; Ranson, K.; Kimes, D.; Blair, J.; Kovacs, K. Forest vertical structure from GLAS: An evaluation using LVIS and SRTM data. *Remote Sens. Environ.* **2008**, *112*, 107–117. [CrossRef]
9. 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]
10. Cao, W.; Chen, D.; Shi, Y.; Cao, Z.; Xia, S. Progress and Prospect of LiDAR Point Clouds to 3D Tree Models. *Geomat. Inf. Sci. Wuhan Univ.* **2021**, *46*, 203–220.
11. Tao, F. 3D Building Model Reconstruction Method Based on Airborne LiDAR Point Cloud Data. *J. Surv. Mapp. Sci. Technol.* **2023**, *11*, 200–206.
12. Acebes, P.; Lillo, P.; Jaime-González, C. Disentangling LiDAR Contribution in Modelling Species–Habitat Structure Relationships in Terrestrial Ecosystems Worldwide. A Systematic Review and Future Directions. *Remote Sens.* **2021**, *13*, 3447. [CrossRef]
13. Vosselman, G.; Maas, H. *Airborne and Terrestrial Laser Scanning*; Whittles Publishing: Dunbeath, UK; CRC Press: Scotland, UK, 2010; pp. 204–205.
14. Gong, W.; Shi, S.; Chen, B.; Song, S.; Niu, Z.; Wang, C.; Guan, H.; Li, W.; Gao, S.; Lin, Y.; et al. Development of Hyperspectral LiDAR for Earth Observation and Prospects. *J. Remote Sens.* **2021**, *25*, 501–513. [CrossRef]
15. Kui, M.; Xu, Y.; Wang, J.; Cheng, F. Research on the Adaptability of Typical Denoising Algorithms Based on ICESat-2 Data. *Remote Sens.* **2023**, *15*, 3884. [CrossRef]
16. Fayad, I.; Baghdadi, N.; Alvares, C.A.; Stape, J.L.; Bailly, J.S.; Scolforo, H.F.; Cegatta, I.R.; Zribi, M.; Le Maire, G. Terrain Slope Effect on Forest Height and Wood Volume Estimation from GEDI Data. *Remote Sens.* **2021**, *13*, 2136. [CrossRef]
17. Cheng, Y.; Liu, X.; Tan, Z.; Wang, S.; Wei, S. Research and development of spaceborne solid state laser technology for laser altimeter. *Infrared Laser Eng.* **2020**, *49*, 55–64.
18. Linares, J.C.; Taïqui, L.; Camarero, J.J. Increasing Drought Sensitivity and Decline of Atlas Cedar (*Cedrus atlantica*) in the Moroccan Middle Atlas Forests. *Forests* **2011**, *2*, 777–796. [CrossRef]
19. Linares, J.C.; Taïqui, L.; Sangüesa-Barreda, G.; Seco, J.I.; Camarero, J.J. Age-related drought sensitivity of Atlas cedar (*Cedrus atlantica*) in the Moroccan Middle Atlas forests. *Dendrochronologia* **2013**, *31*, 88–96. [CrossRef]
20. Zhu, X.; Wang, C.; Xi, X.; Nie, S.; Yang, X.; Li, D. Research Progress of ICESat-2 Satellite Photon Counting LiDAR Data Processing and Applications. *Infrared Laser Eng.* **2020**, *49*, 76–85.
21. Lao, J.; Wang, C.; Nie, S.; Xi, X.; Wang, J. Monitoring and Analysis of Water Level Changes in Mekong River from ICESat-2 Spaceborne Laser Altimetry. *Water* **2022**, *14*, 1613. [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*, 112234. [CrossRef]
23. Zhang, J.; Tian, J.; Li, X.; Wang, L.; Chen, B.; Gong, H.; Ni, R.; Zhou, B.; Yang, C. Leaf Area Index Retrieval with ICESat-2 Photon Counting LiDAR. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *103*, 102488. [CrossRef]
24. Mohajane, M.; Essahlaoui, A.; Oudija, F.; El Hafyani, M.; Teodoro, A.C. Mapping forest species in the central middle atlas of Morocco (Azrou Forest) through remote sensing techniques. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 275. [CrossRef]
25. Dong, J.; Ni, W.; Zhang, Z.; Sun, G. Effectiveness Evaluation of ICESat-2 Vegetation Canopy Height and Surface Elevation Data Products for Forest Height Extraction. *J. Remote Sens.* **2021**, *25*, 1294–1307.

26. Wang, M.; Wei, Y.; Yang, B.; Zhou, X. Extraction and Analysis of Global Elevation Control Points Using ICESat-2/ATLAS. *J. Wuhan Univ. Inf. Sci. Ed.* **2021**, *46*, 184–192.
27. Zheng, Y.; Zhang, Y.; Wang, T.; Zhao, X.; Zhang, K.; Wang, L. Extraction and Accuracy Validation of Elevation Control Points Based on ICESat-2 Data. *J. Geo-Inf. Sci.* **2022**, *24*, 1234–1244.
28. Wang, L. Filtering ICESat-2 ATL08 Terrain Data in Vegetation Areas and Its Application in DEM Error Correction. Ph.D. Thesis, Central South University, Changsha, China, 2023.
29. Popescu, S.; Zhou, T.; Nelson, R.; Neuenschwander, A.; Sheridan, R.; Narine, L.; Walsh, K. Photon counting LiDAR: An adaptive ground and canopy height retrieval algorithm for ICESat-2 data. *Remote Sens. Environ.* **2018**, *208*, 154–170. [CrossRef]
30. Jin, L.; Huang, H. Receiver performance and detection statistics of single photon lidar. *IET Radar Sonar Navig.* **2020**, *14*, 105–109. [CrossRef]
31. Narin, O.G.; Abdikan, S.; Gullu, M.; Lindenbergh, R.; Sanli, F.B.; Yilmaz, I. Improving global digital elevation models using space-borne GEDI and ICESat-2 LiDAR altimetry data. *Int. J. Digit. Earth* **2024**, *17*. [CrossRef]
32. Ranndal, H.; Sigaard Christiansen, P.; Kliving, P.; Baltazar Andersen, O.; Nielsen, K. Evaluation of a Statistical Approach for Extracting Shallow Water Bathymetry Signals from ICESat-2 ATL03 Photon Data. *Remote Sens.* **2021**, *13*, 3548. [CrossRef]
33. Wang, Y.; Fang, H. Derivation and Evaluation of LAI from the ICESat-2 Data over the NEON Sites: The Impact of Segment Size and Beam Type. *Remote Sens.* **2024**, *16*, 3078. [CrossRef]
34. Sun, W.; Jin, J.; Li, G.; Yao, J. Accuracy Evaluation of ICESat-2 Laser Altimetry Satellite for Monitoring Water Levels in Taihu Lake. *Sci. Surv. Mapp.* **2021**, *46*, 6–11.
35. Wu, H.; Wang, N.; Guo, Z. Application of ICESat-2/ATLAS Altimetry Data in Estimating Lake Water Levels of Qinghai Lake. *J. Water Resour. Water Eng.* **2021**, *32*, 11–18+26.
36. Li, Y.; Zhou, X.; Li, G.; Guo, J.; Ma, Y.; Chen, Y. Research Progress and Prospects of Spaceborne Single Photon Lidar for Shallow Water Bathymetry. *Infrared Laser Eng.* **2024**, 1–13. [CrossRef]
37. Wang, Z.; Nie, S.; Xi, X.; Wang, C.; Lao, J.; Yang, Z. A methodological framework for specular return removal from photon-counting LiDAR data. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *122*, 103387. [CrossRef]
38. E, X.; Dai, G.; Wu, S. Preprocessing and Correction Methods for ICESat-2 ATL03 Data. *Infrared Laser Eng.* **2021**, *50*, 92–102. [CrossRef]
39. Zhu, J.; Yang, P.-F.; Li, Y.; Xie, Y.-Z.; Fu, H.-Q. Accuracy assessment of ICESat-2 ATL08 terrain estimates: A case study in Spain. *J. Cent. S. Univ.* **2022**, *29*, 226–238. [CrossRef]
40. Tian, X.X.; Shan, J. Comprehensive evaluation of the ICE-Sat-2 ATL08 terrain product. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 8195–8209. [CrossRef]
41. Zhu, X.; Wang, C.; Xi, X. Research Progress on the Processing and Application of ICESat-2 Spaceborne Photon-Counting Lidar Data. *Infrared Laser Eng.* **2020**, *49*, 68–77.
42. Yu, Q.; Ryan, M.G.; Ji, W.; Prihodko, L.; Anchang, J.Y.; Kahiu, N.; Nazir, A.; Dai, J.; Hanan, N.P. Assessing canopy height measurements from ICESat-2 and GEDI orbiting LiDAR across six different biomes with G-LiHT LiDAR. *Environ. Res. Ecol.* **2024**, *3*, 025001. [CrossRef]
43. Wang, Z.; Liu, J.; Sheng, Y.; Yang, X. Intercomparison of the DART model and GEDI simulator for simulating GEDI waveforms in forests. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *134*, 104148. [CrossRef]
44. Roy, A.D.; Karpowicz, D.A.; Hendy, I.; Rog, S.M.; Watt, M.S.; Reef, R.; Broadbent, E.N.; Asbridge, E.F.; Gebrie, A.; Ali, T.; et al. Current Status of Remote Sensing for Studying the Impacts of Hurricanes on Mangrove Forests in the Coastal United States. *Remote Sens.* **2024**, *16*, 3596.
45. Huang, J.; Yang, Y.; Yu, Y.; Zhang, Y. Vertical Accuracy of Open-Source Remote Sensing Data (AW3D30, TanDEM-X, ATLAS) for Understory Terrain Estimation. *Geocarto Int.* **2024**, *39*, 2356855. [CrossRef]
46. Wang, C.; Jia, D.; Lei, S.; Numata, I.; Tian, L. Accuracy Assessment and Impact Factor Analysis of GEDI Leaf Area Index Product in Temperate Forest. *Remote Sens.* **2023**, *15*, 1535. [CrossRef]
47. Bruening, J.; May, P.; Armston, J.; Dubayah, R. Precise and unbiased biomass estimation from GEDI data and the US forest inventory. *Front. For. Glob. Chang.* **2023**, *6*, 1149153. [CrossRef]
48. Li, B.; Zhao, T.; Su, X.; Fan, G.; Zhang, W.; Deng, Z.; Yu, Y. Correction of Terrain Effects on Forest Canopy Height Estimation Using ICESat-2 and High Spatial Resolution Images. *Remote Sens.* **2022**, *14*, 4453. [CrossRef]
49. Neuenschwander, A.; Pitts, K. The ATL08 land and vegetation product for the ICESat-2 Mission. *Remote Sens. Environ.* **2019**, *221*, 247–259. [CrossRef]
50. You, H.; Li, Y.; Qin, Z.; Lei, P.; Chen, J.; Shi, X. Research on Multilevel Filtering Algorithm Used for Denoising Strong and Weak Beams of Daytime Photon Cloud Data with High Background Noise. *Remote Sens.* **2023**, *15*, 4260. [CrossRef]
51. Huang, J.; Xia, T. Multi-level adaptive photon cloud noise filtering algorithm for different observation time scenes in forest environments. *IEEE Trans. Geosci. Remote Sens.* **2023**, *62*, 5700816.

52. Huang, X.; Cheng, F.; Wang, J.; Duan, P.; Wang, J. Forest canopy height extraction method based on ICESat-2/ATLAS data. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–14. [CrossRef]
53. Li, B.; Fan, G.; Zhao, T.; Deng, Z.; Yu, Y. Retrieval of DTM under complex forest stand based on spaceborne LiDAR fusion photon correction. *Remote Sens.* **2022**, *14*, 218. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

*Forests* Editorial Office
E-mail: forests@mdpi.com
www.mdpi.com/journal/forests




Disclaimer/Publisher's Note: The title and front matter of this reprint are at the discretion of the Guest Editors. The publisher is not responsible for their content or any associated concerns. The statements, opinions and data contained in all individual articles are solely those of the individual Editors and contributors and not of MDPI. MDPI disclaims responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.





Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-8808-5