International Journal of Geo-Information

MDPI

Article

Building Height Extraction Based on Spatial Clustering and a Random Forest Model

Jingxin Chang [1], Yonghua Jiang [1,*], Meilin Tan [2], Yunming Wang [1] and Shaodong Wei [3]

[1] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China; jxchang@whu.edu.cn (J.C.); wangyunming@whu.edu.cn (Y.W.)
[2] Inner Mongolia Autonomous Region Surveying and Mapping Geographic Information Center, Hohhot 010050, China; tanmeilin@whu.edu.cn
[3] State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China; weishaodong@whu.edu.cn
* Correspondence: jiangyh@whu.edu.cn

Abstract: Building height (BH) estimation is crucial for urban spatial planning and development. BH estimation using digital surface model data involves obtaining ground and roof elevations. However, vegetation and shadows around buildings affect the selection of the required elevation, resulting in large BH estimation errors. In highly urbanized areas, buildings of similar heights often have similar characteristics and spatial proximity, which have reference significance in BH estimation but are rarely utilized. Herein, we propose a BH estimation method based on BIRCH clustering and a random forest (RF) model. We obtain the initial BH results using a method based on the optimal ground search area and a multi-index evaluation. BIRCH clustering and an RF classification model are used to match buildings of similar heights based on their spatial distance and attribute characteristics. Finally, the BH is adjusted based on the ground elevation obtained from the secondary screening and the BH matching. The validation results from two areas with over 12,000 buildings show that the proposed method reduces the root-mean-square error of the final BH results compared with the initial results. Comparing the obtained height maps shows that the final results produce a relatively accurate BH in areas with high shading and vegetation coverage, as well as in areas with dense buildings. Thus, the proposed method has been validated for its effectiveness and reliability.

Keywords: digital surface model; building height; BIRCH; random forest; optimal height

check for updates

Citation: Chang, J.; Jiang, Y.; Tan, M.; Wang, Y.; Wei, S. Building Height Extraction Based on Spatial Clustering and a Random Forest Model. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 265. https://doi.org/10.3390/ijgi13080265

Academic Editors: Wolfgang Kainz and Maria Antonia Brovelli

Received: 3 June 2024
Revised: 17 July 2024
Accepted: 24 July 2024
Published: 26 July 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Building height (BH) data are used in various fields, such as pollutant diffusion [1], urban climate [2–4], land use [5], and socioeconomic development [6]. At present, both cities that have not yet obtained building basic data and those that have already obtained building basic data need to continuously automate the acquisition of building attribute information (such as BH). However, manual BH measurement methods cannot address this demand. Due to the swift advancement of aerospace technology, both airborne and satellite data now boast high spatial and temporal resolutions, meeting the requirements for obtaining BH data over extensive regions. Three main types of data sources have been used for BH extraction: light detection and ranging (LiDAR) data [7–11], synthetic aperture radar (SAR) images [12–14], and optical images [15–20].

LiDAR can provide three-dimensional coordinate information while exhibiting a certain penetrability into ground features [7]. Aviation LiDAR data have high accuracy and are used for 3D building reconstruction [21], 2D building contour (BC) extraction [22,23], and to generate high-precision digital surface models (DSMs) and digital terrain models (DTMs). The BH can then be determined from the difference between the DSM and DTM [11]. However, obtaining aviation LiDAR data incurs high costs and presents administrative regional limitations that hinder its wider application. In 2003, NASA launched

the Ice, Cloud, and land Elevation Satellite (ICESat-1), which enabled operation of the world's first space-borne laser altimetry system; its main payload is the geoscience laser altimeter system (GLAS) [24]. GLAS is not intended to generate building data; however, scholars have discovered its potential for obtaining BH data [8,25,26]. Gong et al. [26] proposed a method to obtain building height via GLAS data. This method can obtain a root mean square error (RMSE) of 6.4 m for building height within each GLAS footprint. However, because GLAS is recorded in the form of a complete waveform, more echoes are obtained in complex scenes, which leads to inaccurate matching of waveforms with ground objects and causes the problem of "same-wave foreign objects". The large footprint of GLAS (approximately 70 m) further limits its promotion. As a successor to ICESat-1, ICESat-2 realizes an improved ground trajectory density and footprint size. Therefore, the high-density photon distribution characteristics of ICESat-2 can be utilized to obtain building surface and top photons, and the BH can then be obtained from the height difference between the two [9,27]. Lao et al. [9] proposed an effective framework for estimating building height based on ICESAT-2 data. The RMSE obtained by this framework ranges from 0.35 to 0.45 m. However, because ICESat-2 data cannot fully cover the ground, the obtained BH data are discretely distributed, requiring further research to generate BH data at the grid scale. To obtain continuous BH data, researchers have proposed the integration of Global Ecosystem Dynamics Investigation (GEDI) derived relative height indicators, optical data (i.e., Landsat-8 and Sentinel-2), and radar data (Sentinel-1) to plot the large-scale distribution of BHs [10]; good results have been achieved at 30 m and 150 m grid sizes.

SAR can be used to obtain high-resolution, all-weather, light-independent data. Initially, researchers had to make assumptions regarding the properties of materials before the use of TerraSAR-X data [12] or high-resolution SAR data [13] to obtain the BH. However, the method has many limitations; primarily, it can only extract BH data from small areas. To obtain large-scale and high-precision BH data using SAR data, Li et al. [14] developed a method for estimating the BH using Sentinel-1 ground distance detection data. They proposed a VVH (i.e., VV and VH) index, and the BH model was established using this VVH index and a reference BH. The model has an RMSE of 1.5 m between the estimated height and the reference height in seven study cities in the United States. Additionally, Zhou et al. [28] developed a global urban BH atlas in 2015 using Sentinel-1 data at a resolution of 500 m, thus revealing the scale of infrastructure demand required by southern countries worldwide to achieve sustainable development goals. Optical images, digital elevation model (DEM) data, and night light remote sensing data have been integrated based on SAR data to improve the grid size and overcome the small area limitation; the random forest (RF) model could then be used to obtain more accurate, large-scale BH data [29,30]. Methods that obtain large-scale BH data through multi-source data fusion can reflect trends and reveal patterns. However, more detailed research will require the grid size to be reduced precisely to the height of each building.

Optical images have been used for BH estimation for several decades. Both single-lens and stereo optical images have been studied for their potential use in BH estimation applications [31]. With monocular images, the BH is mostly calculated based on shadows (SD). It is usually divided into two steps: the first step involves extracting the SD of the building, and the second step involves obtaining the BH by calculating the relative geometric relationship from the data among the building, the SD, and satellites [16,17,20,32–34]. Zhao et al. [34] proposed a method that combines using Google Earth Satellite images and ICESat-2 photons to extract building height, which can obtain an RMSE of 5.34 m. Methods based on stereo-optical imaging usually first perform matching and then generate a DSM. Subsequently, BC data can be obtained through methods such as deep learning [35], or the existing building footprint data can be used directly [31,36]. Finally, the BH can be obtained based on the difference in elevation between the roof and the ground of the building. Obtaining the BH by filtering the DSM also utilizes DSM data [35,37,38]. However, it is often impossible to completely filter out ground objects, and subsequent supplementary editing is required [39]. Both SD- and DSM-based methods can determine the height of

a single building; however, SD-based methods are susceptible to the influence of the SD extraction results. When SD extraction is inaccurate or incomplete, the BH calculation is prone to bias. DSM-based methods are easily affected by the quality of the DSM and the environment in which the building is located. It is difficult to obtain the ground elevation of a building surrounded by vegetation or SD, which leads to inaccurate calculation of the BH.

In summary, calculating the height of buildings using optical stereo images remains challenging. In particular, abnormal elevation values in the DSM are easily generated when a large amount of SD or vegetation is present, or when buildings are too densely packed, which results in inaccurate ground elevation selections for buildings. Existing research often fails to effectively address this type of building. We have found through research on the properties and spatial relationships of buildings that they can effectively handle such problems. Therefore, we propose a BH extraction method based on spatial clustering and an RF model. The proposed method obtains the initial BH results using the optimal ground search area and a multi-index evaluation mechanism (OGMI) [40]. Subsequently, the accuracy of the initial BH is determined based on the proposed method. The Balanced Iterative Reducing and Clustering Using Hierarchies (BIRCH) clustering algorithm and RF classification model are then used to match buildings with appropriate and accurate heights based on the constraint indicators for buildings with inaccurate heights. Finally, the optimal height for a building with an inaccurate height is selected from among multiple candidate building heights.

2. Data and Study Area

The experimental data were obtained from GF-7 stereo imaging data. The panchromatic image resolution captured by the two-line array camera on the GF-7 satellite was 0.8 m for the forward view image (FWD) and 0.65 m for the backward view image (BWD). We used the FWD and BWD images from GF-7 to generate DSM image data with a 1-m resolution using the more global matching (MGM) algorithm. Moreover, we extracted BC and SD data from the BWD using the GEOWAY and ENVI 5.3 commercial software applications [40], respectively. As shown in Figure 1a, the study areas were in the cities of Hohhot and Ordos in the Inner Mongolia Autonomous Region of China. The first study area is the eastern urban area in Dongsheng District, Ordos City (Figure 1b). The topography of the Dongsheng District is mainly characterized by hills and valleys in the east and undulating plateaus in the west. The buildings in the area are generally simple in shape, predominantly consisting of rectangular structures, with most buildings being under 40 m in height. The second study area is in the urban region of Hohhot City, as depicted in Figure 1c. Hohhot City is primarily characterized by two major geomorphic units, the northern Daqing Mountain and the southeastern Manhan Mountain, which feature mountainous terrain. The southern and southwestern regions are defined by the Tumochuan Plain, where the terrain gradually slopes from northeast to southwest. The urban area itself sits at an altitude of approximately 1040 m and is home to a mix of dense and varied buildings.

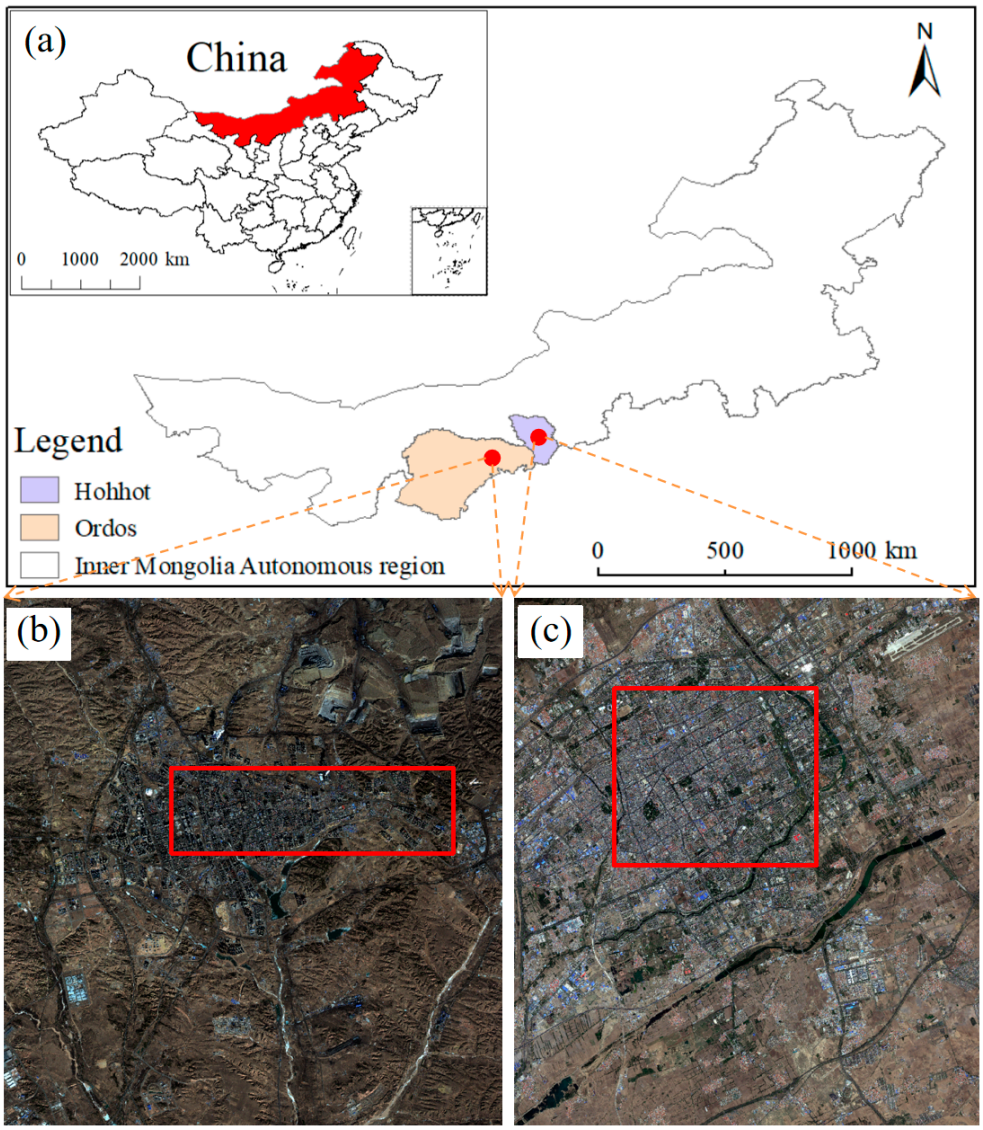


Figure 1. Study areas and corresponding validation areas. (**a**) Geographical location of the study area. (**b**) Main built-up area in Ordos City; the area within the red box is verification area S1. (**c**) Main built-up area in Hohhot City; the area within the red box is verification area S2.

3. Methodology

We propose a framework based on BIRCH clustering and RF classification models [41] for estimating the BH based on satellite stereo images. As shown in Figure 2, the BH estimation method can be divided into four parts:

(1) The initial BH estimation is performed based on OGMI. First, the middle elevation is used to obtain the roof elevation of the building. The process of selecting the best ground elevation is then employed to derive the building's ground elevation. Subsequently, the building's height is calculated by subtracting the ground elevation from the roof elevation. Additionally, the accuracy of the initial BH is evaluated based on the judgment criteria.

(2) Building clustering is performed based on the BIRCH clustering algorithm. The center of gravity of the BC represents the position of the building in the image, and the BIRCH clustering algorithm is used to cluster the buildings.

(3) Similar buildings are screened based on the RF classification model. The RF classification model is used to make predictions within the same class of buildings after clustering based on constraint indicators to determine whether different buildings have similar heights.

(4) A new BH (NBH) is selected. Upon obtaining an accurate BH (ABH) matched by the inaccurate BH (IBH), the optimal height value for the IBH is selected.

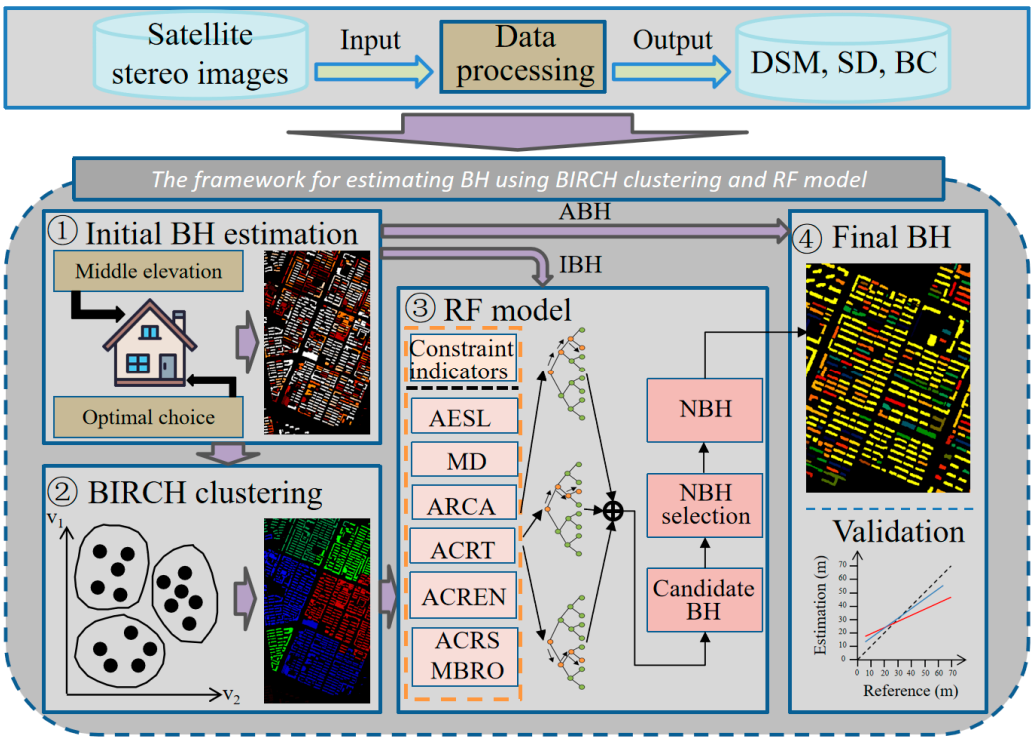


Figure 2. Flowchart of the proposed framework.

3.1. Initial BH Estimation

3.1.1. Obtaining the Initial BH Estimation Results

A BH estimation method based on the OGMI was used. OGMI determines the roof and ground elevations of a building and then uses the difference between these two elevations as the height of the building. The acquisition of the building roof elevation can be divided into two steps: (1) a set of building roof elevation values is obtained based on the DSM and BC, and (2) the set of roof elevation values is sorted, with the middle value assigned as the roof elevation of the building. The acquisition of the building ground elevation can be divided into three steps: (1) the optimal ground search area is obtained based on the k-nearest neighbor algorithm and visual building screening; (2) the points within the optimal ground search area are evaluated based on the elevation, distance, tortuosity, and directional folding indices; (3) weighted calculations are performed for each indicator to obtain a score for each ground point in the optimal ground search area. The elevation of the ground point with the highest score is taken as the ground elevation of the building. Chang et al. [40] provides a more detailed description.

3.1.2. Classification of Initial BH Estimation Results

Errors occur between the initial BH estimation results and the actual values owing to the influence of abnormal ground elevation values in the DSM and to limitations of the method. Therefore, it is necessary to recalculate the BH values with significant errors in the initial BH estimation results. To distinguish the initial BH estimation error, initial BH estimation results are divided into two categories: ABH with smaller errors, and IBH with larger errors. The OGMI primarily focuses on the selection of ground elevation, and the building ground elevation is determined based on index-weighted scores (score range $[-100, 100]$). The higher the ground point score, the more likely the point represents the building's ground. When a point on the ground is close to the edge of the BC, and the elevation value of the ground point is abnormal compared to the normal ground elevation

around the building, the OGMI may make an incorrect selection, leading to deviations in the BH. Therefore, to distinguish the accuracy of the initial BH estimation results, we propose the following standards: assuming that in the calculation of the ground points of a building, the ground points with weighted index scores greater than 60 points are $gs = \{gs_1, gs_2, \ldots gs_n\}$, the elevation value corresponding to gs is $gh = \{gh_1, gh_2, \ldots gh_n\}$. The middle value of gh that can be calculated is $gh_{\text{middle}} = \text{middle}(\text{sort}(gh))$. Here, $\text{sort}(\cdot)$ represents descending order, and $\text{middle}(\cdot)$ represents taking the middle value. Assuming that the initial ground elevation value of the BH is gh_{origin}, the absolute difference can be obtained as $gh_{\text{diff}} = \left|gh_{\text{middle}} - gh_{\text{origin}}\right|$, where $|\cdot|$ represents the absolute value. When the judgment threshold is gh_{thres}, if the condition $gh_{\text{diff}} > gh_{\text{thres}}$ is satisfied, the elevation of the building is judged as IBH; otherwise, it is judged as ABH.

3.2. BIRCH Clustering

BIRCH is a comprehensive hierarchical clustering algorithm that summarizes clustering descriptions using clustering features and clustering feature trees (CFs). BIRCH clustering requires three parameters: the branch factor B_r, threshold T_s, and cluster count C_k [42]. Upon input of data points into BIRCH, a well-balanced CF tree is built, where each node represents a cluster within the hierarchical structure. Intermediate nodes serve as supersets, while leaf nodes represent actual clusters. The branch factor, denoted as B_r, specifies the maximum number of child nodes a parent node can possess, serving as a global parameter. Each node holds essential information pertaining to its respective cluster. It is assumed that the cluster center of a sub-cluster is $x_0 = \sum_{i=1}^{N} x_i / N$, where N indicates the number of objects in the sub-cluster and x_i denotes the ith object in the sub-cluster. The clustering radius of each cluster is $R = \left(\sum_{i=1}^{N} (x_i - x_0)^2 / N\right)^{1/2}$. As each new point is introduced, it traverses the tree from the root, progressively moving towards the subset group with the closest center until it reaches a leaf node. Upon reaching a leaf node, the new point is incorporated into the leaf cluster assuming doing so does not surpass the specified threshold, T_s, for the cluster radius. If the threshold is exceeded, a new cluster is formed with the new point as its sole member. Therefore, the threshold parameter governs the size of the clusters. When a new cluster formation leads to more child nodes than the specified threshold B_r for the parent node, the parent node undergoes a split. To maintain tree balance, recursive splitting of nodes may be necessary. Following the submission of all points to BIRCH, the global clustering stage involves the center of the leaf cluster being fed into a clustering algorithm, such as cluster clustering or k-means clustering, with the cluster count C_k as a parameter. Subsequently, the quality of clustering is enhanced through the merging of neighboring clusters.

In urban areas with well-developed urban planning, buildings of the same height are generally located in similar spatial locations. Therefore, space-based clustering can better distinguish buildings in different regions, thereby significantly reducing the computational cost of matching buildings with similar heights. Here, we use the center of gravity of the BC to replace the building position, and then the center of gravity coordinates of all buildings is input into the BIRCH clustering algorithm. The clustering of buildings can then be achieved by clustering their centers of gravity. To reduce the impact of urban shape and layout, it is generally recommended to select fewer clustering groups, so that each group has enough buildings.

3.3. RF Classification Model

To match IBH buildings with ABH buildings of similar height, an RF classification model is used to match buildings of the same class after clustering. ABH buildings with similar attributes to those of the IBH buildings are selected for matching. We set six variables (absolute error of the shadow length (AESL), matching degree (MD), absolute rate of change in area (ARCA), absolute change rate of tightness (ACRT), absolute change rate of the edge number (ACREN), and absolute change rate in the orientation of the longest side of the minimum area bounding rectangle (ACRSMBRO)) as predictive variables. Two

crucial parameters in the RF model are the number of trees (N_{tree}) and the number of randomly selected features used for dividing each node (M_{try}) [30]. Increasing the value of N_{tree} can improve the RF performance. Due to the strong computational efficiency of the RF classifier and its capacity to prevent overfitting, it is advisable to set N_{tree} to the highest feasible value [43]. In the majority of prior research, the value of N_{tree} has been set to 500. Before reaching this number, the error in the classification tree tends to stabilize. Reducing M_{try} leads to decreased model variance but to increased bias due to the potential neglect of important features. Conversely, increasing M_{try} raises model variance while decreasing bias by considering more features. Prior studies have recommended using the square root of the number of variables as the M_{try} value. Considering the small number of variables and their importance, we set N_{tree} and M_{try} to 300 and all variables, respectively [41].

The six variables used in the RF classification model revolve around changes in the shadow and contour shapes of the two buildings, ensuring maximum similarity in their attributes. The data for the variables is generated by manually searching for samples of similar and dissimilar buildings that satisfy the criteria. A dataset that can be used to train the RF classification model is obtained by calculating the corresponding variable values and saving them. The calculation process for each variable is described below:

(1) AESL: we use maximum minimum linear stretching to obtain the SD of buildings in panchromatic images with higher spatial resolution in stereo images. To accurately calculate the length of the SD for buildings, we propose a calculation method that considers the shadow edge information. It is assumed that there exists a set of contour points $bc = \{cpt_1, cpt_2, \ldots, cpt_p\}$ for building BD (as shown in Figure 3a). Then, bc is expanded outward and overlapped with the shadow (as shown in Figure 3b). The point set $bso = \{opt_1, opt_2, \ldots, opt_q\}$ (as shown in Figure 3c) closest to the building contour in each direction of the overlapping part (the vector direction between the points in bc and the corresponding overlapping points) is obtained. The points in bso are then screened, and rays formed along the direction of illumination starting from themselves are used to determine whether the ray passes through BD (as shown in Figure 3d). Finally, as shown in Figure 3e, the point set $sbso = \{spt_1, spt_2, \ldots, spt_m\}$ describing the rays in bso that do not pass through BD is preserved. Starting from the point itself in the $sbso$, a pixel-by–pixel search is conducted along the direction of illumination, and the length of overlap with the shadow in the search direction is measured until the shadow cuts off or enters the contour of another building. When closely located buildings obstruct shadows, existing methods often only calculate the shadow length between buildings. This approach ignores the length information from the edges of the SD, resulting in a shorter calculated shadow length [18]. To consider the length of the shadow edges, we divide the statistical shadow length into two categories: the SD length obtained by the normal SD cutoff, and the SD length obtained by searching for other buildings. As shown in Figure 3f, to avoid interference from the length of extremely short shadows, we remove the SD lengths that are less than half of the maximum length of shadows of the same type. Meanwhile, if only a second type of shadow length appears in a building's statistics, the average length of the second type of SD is used as the final SD length of the building. Otherwise, the average length of the first type of SD is used as the final SD length of the building. The calculation formula for variable AESL is as follows:

$$\text{AESL} = |SL_A - SL_B| \tag{1}$$

where A and B represent two different buildings, and SL_A and SL_B represent the shadow lengths of buildings A and B, respectively.

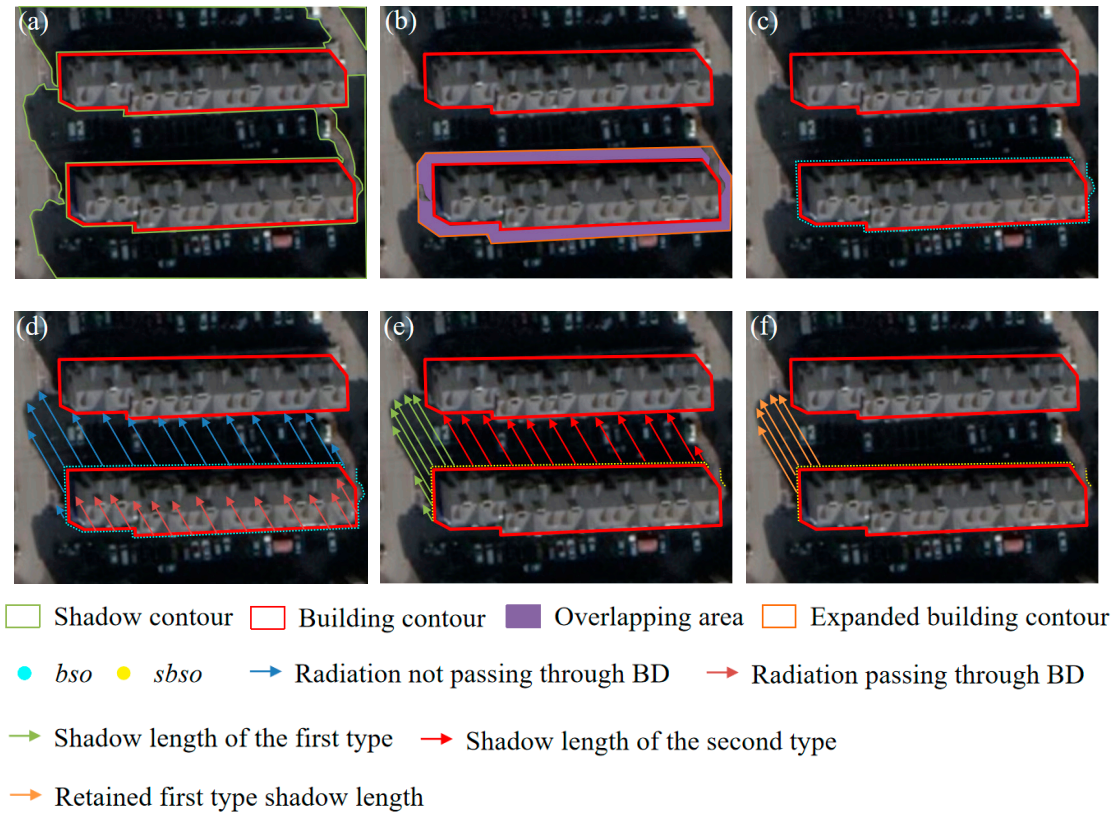


Figure 3. Calculation of the SD length of a building. (**a**) Point set and SD area of BC. (**b**) Area in which the building outline expands and overlaps with the SD. (**c**) Point set of the SD closest to the outline of the building. (**d**) Filtered point set of SD. (**e**) Classification of calculation methods for the SD length. (**f**) Final calculation result for the SD length of the example building.

(2) MD: MD describes the degree of matching between the contours of two buildings. Hu moments are used to calculate the similarity between the contours of two buildings. The Hu moment is a global invariant moment used to describe shapes, which can capture their rotation, scaling, and translation invariance [44]. The formula for calculating MD is as follows:

$$\mathrm{MD} = \sum_{i=1\ldots7} \left| \frac{1}{m_i^A} - \frac{1}{m_i^B} \right| \tag{2}$$

$$\begin{cases} m_i^A = \mathrm{sign}(h_i^A) \cdot \log h_i^A \\ m_i^B = \mathrm{sign}(h_i^B) \cdot \log h_i^B \end{cases} \tag{3}$$

where h_i^A and h_i^B represent the Hu moments of buildings A and B, respectively. The smaller the MD value, the more similar the contours of the two buildings will be.

(3) ARCA: the area of a building is a common comparison variable that can effectively reflect the size of the building. Combined with the matching degree, the shape and size of the building can be fixed more accurately based on the building area. The calculation formula for the ARCA is as follows:

$$\mathrm{ARCA} = \frac{|BA_A - BA_B|}{BA_A} \tag{4}$$

where BA_A and BA_B represent the areas of different buildings.

(4) ACRT: tightness is a feature used to describe shapes. It measures the roundness of a shape by calculating the ratio between its perimeter and area [45]. The formula for calculating ACRT is as follows:

$$\text{ACRT} = \frac{|CRT_A - CRT_B|}{CRT_A} \tag{5}$$

$$\begin{cases} CRT_A = \frac{4\pi S_A}{L_A^2} \\ CRT_B = \frac{4\pi S_B}{L_B^2} \end{cases} \tag{6}$$

where S_A and S_B represent the areas of buildings A and B, respectively, and L_A and L_B represent the perimeters of buildings A and B, respectively.

(5) ACREN: the number of sides of a building reflects its complexity [46], and the rate of change in the number of sides of buildings with the same shape should be small. ACREN is calculated as follows:

$$\text{ACREN} = \frac{|EN_A - EN_B|}{EN_A} \tag{7}$$

where EN_A and EN_B represent the number of sides of buildings A and B, respectively.

(6) ACRSMBRO. The angle between the longest side of the bounding rectangle with the minimum area of the building outline and the positive x-axis [45] reflects the main direction of the building. The calculation formula for ACRSMBRO is as follows:

$$\text{ACRSMBRO} = \frac{|SO_A - SO_B|}{SO_A} \tag{8}$$

where SO_A and SO_B represent the angles of buildings A and B, respectively.

3.4. Selection of NBH

After matching multiple ABHs for the IBH, the most suitable height should be selected as the NBH for IBH. We comprehensively consider gh_{middle} and the matched ABHs. When the IBH does not match the ABH, gh_{middle} is used as the new ground elevation for the IBH, and the roof elevation uses the initial result. The building height (bh_{ref}), which is obtained by subtracting the roof elevation from the ground elevation, is used as the NBH for the IBH. The ground elevation of the IBH typically contains a combination of a small amount of abnormal ground elevations and a large amount of normal ground elevations. At this point, bh_{ref} typically approaches the true height of the building. When the IBH matches the ABHs, the ABHs will also contain heights that are close to the true height of the IBH. We use bh_{ref} as a benchmark and calculate the minimum absolute difference between bh_{ref} and the ABHs to obtain the best ABH (BABH). When the absolute difference between bh_{ref} and the BABH is small, a slight deviation exists between bh_{ref} and the BABH. We believe that the IBH represents the approximate height of ABH. Therefore, the ground elevation of bh_{ref} is reliable, and bh_{ref} is used as the NBH for the IBH. Otherwise, we believe that the BABH is closer to the true height of the IBH, and BABH is used as the NBH of the IBH. The proposed method is summarized in Algorithm 1.

Algorithm 1: Selection of NBH

Data: IBH building roof elevation value, H_{roof}; gh_{middle}; ABHs.
Result: The NBH of the IBH building.
Begin
$bh_{\text{ref}} = H_{\text{roof}} - gh_{\text{middle}}$
if the number of ABHs is 0 then
  NBH = bh_{ref}
  end algorithm
BABH = argmin (| ABHs $- bh_{\text{ref}}$ |)
gap = | bh_{ref} − BABH |
if gap meets the condition, then
  NBH = bh_{ref}
else
  NBH = BABH
End

3.5. Accuracy Evaluation

As the proposed method focuses solely on BH estimation, we utilize relative accuracy for evaluation. For each building sample, the roof and ground elevations are manually acquired from the DSM. Subsequently, we calculate the difference between the roof and ground elevations to determine the reference height of the building.

To verify the accuracy of classifying the initial BH estimation results, we calculate the classification correctness (CR), as follows:

$$\text{CR} = \frac{V_{\text{i_i}} + V_{\text{a_a}}}{V_{\text{i_i}} + V_{\text{a_i}} + V_{\text{i_a}} + V_{\text{a_a}}} \tag{9}$$

where $V_{\text{i_i}}$ indicates that IBH buildings are classified as IBH buildings, $V_{\text{a_i}}$ indicates that ABH buildings are classified as IBH buildings, $V_{\text{i_a}}$ indicates that IBH buildings are classified as ABH buildings, and $V_{\text{a_a}}$ indicates that ABH buildings are classified as ABH buildings.

To quantitatively assess the accuracy of the building height estimation, we utilize the mean error (ME), mean absolute error (MAE), and root mean square error (RMSE) to evaluate the estimation results. The formulas are as follows:

$$\text{ME} = \frac{1}{n}\sum_{i=1}^{n}(H_{es_i} - H_{re_i}) \tag{10}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|H_{es_i} - H_{re_i}| \tag{11}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(H_{es_i} - H_{re_i})^2} \tag{12}$$

where n denotes the number of building samples, and H_{es_i} and H_{re_i} denote the estimated and reference values of the ith building, respectively.

4. Results and Discussion

4.1. Results

4.1.1. Reliability of the Classification

Our proposed method reselects the height for IBH buildings, whereas ABH buildings use the initial BH estimation results. Therefore, the accuracy of the initial BH estimation results determines the likelihood of IBH buildings being correctly identified and further processed. In most previous studies, the height of each floor of a building was assumed to be 3 m [14,19,47]. Therefore, 3 m is used as the error limit. If the initial BH estimation result is within 3 m of the reference height, it is considered accurate. Otherwise, the initial BH estimation result is considered inaccurate. Since the classification is mainly based

on the ground elevation of buildings, we conducted an accuracy analysis on the same ground elevation as the height of the buildings. We randomly selected 5240 buildings and 7202 buildings in the main built-up areas of validation areas S1 and S2, respectively, for the validation analysis. Figure 4 presents the four classification scenarios (V_{i_i}, V_{a_i}, V_{i_a}, and V_{a_a}). Among them, V_{i_i} and V_{a_a} represent correct classifications, whereas V_{a_i} and V_{i_a} represent incorrect classifications. Based on the classification results, most buildings were correctly classified. Buildings that were not correctly classified are often located in densely populated areas (with missing or incorrect ground elevations, making it difficult to select the ground elevation), or have a complex BC (with multiple changes in roof elevations and large differences in roof elevation errors). The classification accuracies are listed in Table 1. In both validation areas, the accuracy of the ground elevation classification is greater than 80% and both are better than the accuracy of the BH classification. This is because the evaluation based on ground elevation classification does not consider the roof elevation, resulting in higher classification accuracy. However, whether evaluating based on ground elevation or BH, the proposed method achieves high classification accuracy, thus verifying the reliability of the method.

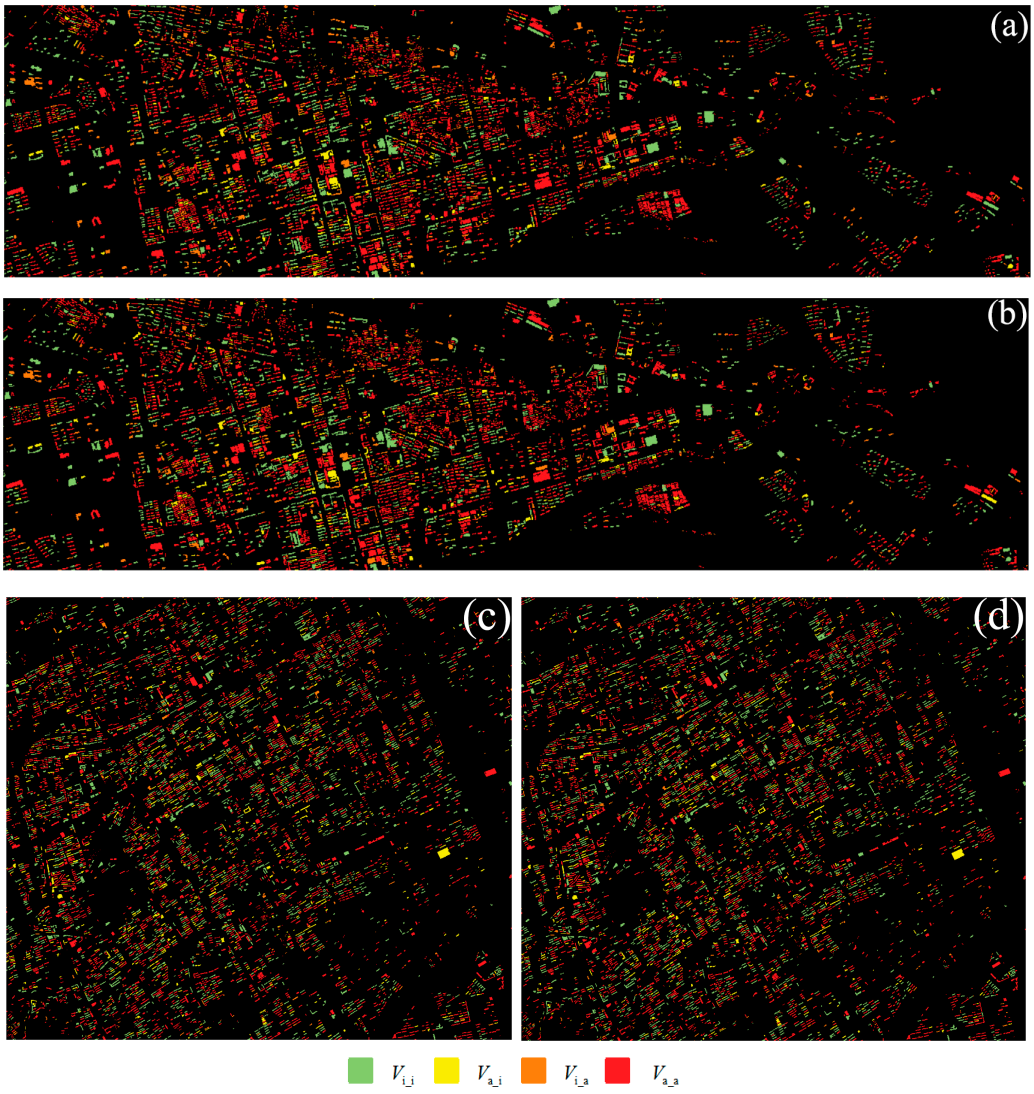


Figure 4. Classification results for different validation areas based on the building ground elevation and building height. (**a**,**b**) are the classification results of S1 based on BH and building ground elevation, respectively. (**c**,**d**) are the classification results of S2 based on BH and building ground elevation, respectively.

Table 1. Comparison of the classification accuracies for different validation areas based on the ground elevation and BH.

Study Area	CR	
	Based on BH/%	Based on Ground Elevation/%
S1	79.160	80.973
S2	82.422	83.588

4.1.2. Accuracy of the NBH Selection

To evaluate the accuracy of the NBH selection for IBH buildings based on the BIRCH clustering and RF classification model methods, we evaluated the accuracy of the initial BH and NBH for all buildings classified as IBH. Figure 5 shows the buildings classified as IBH in validation area S1. As shown in Figure 5a, the IBH buildings are scattered throughout the verification area, and the frequency of high-rise buildings (with heights greater than 50 m) is extremely high. This is because the roof elevation changes in high-rise buildings are more complex and the ground occlusion is more severe, resulting in larger errors. The incorrect selection of ground elevation for low-rise can be attributed to a small number of abnormal ground elevation values and the insufficient amount of ground elevation data caused by overly dense buildings. From the height map of the initial BH estimation results (Figure 5b) and the height map of the NBH for the IBH buildings (Figure 5c), the NBH is closer to the reference building height and has higher consistency. The same situation occurs in validation area S2 (as shown in Figure 6). Compared with validation area S1, the problem of BH estimation errors caused by dense buildings is more pronounced in validation area S2 (as shown in Figure 6b). Therefore, more evenly distributed IBH buildings exist in the highly urbanized validation area S2. Highly urbanized areas are more advantageous for the proposed NBH selection method because more buildings of the same height are clustered together, thus providing more choices when selecting ABH buildings of an approximate height for IBH buildings. Therefore, the height map of the NBH of the IBH buildings (as shown in Figure 6c) shows that its height is generally closer to the height map of the reference building than to the NBH height map for validation area S1.

We conducted a quantitative evaluation to more accurately evaluate the NBH improvement of the IBH compared to the initial BH. As shown in Figure 7a, in validation area S1, 1402 buildings were identified as IBH buildings, and most of their initial BHs were higher than the reference height (ME = 5.107 m). Similarly, 2982 buildings in validation area S2 were classified as IBH buildings, and their initial BHs (as shown in Figure 7c) were also higher than the reference height (ME = 4.172 m). The initial BH extraction method tends to favor the selection of low-elevation values in the ground elevation. Therefore, abnormally low ground elevations are easily selected as the ground elevation, which leads to a higher BH. However, the NBH selection method combines the matching selection of buildings with approximate heights and the secondary screening of ground elevation. This method can refer to precise building heights and prevents extreme ground elevation values, thereby obtaining more accurate BH values. As shown in Figure 7b, the RMSE of the NBH in validation region S1 is increased by 4.266 m compared with the initial BH. As shown in Figure 7d, the RMSE of the NBH in validation region S2 is increased by 1.484 m compared with the initial BH. These accuracy improvements demonstrate the effectiveness of the proposed height selection method.

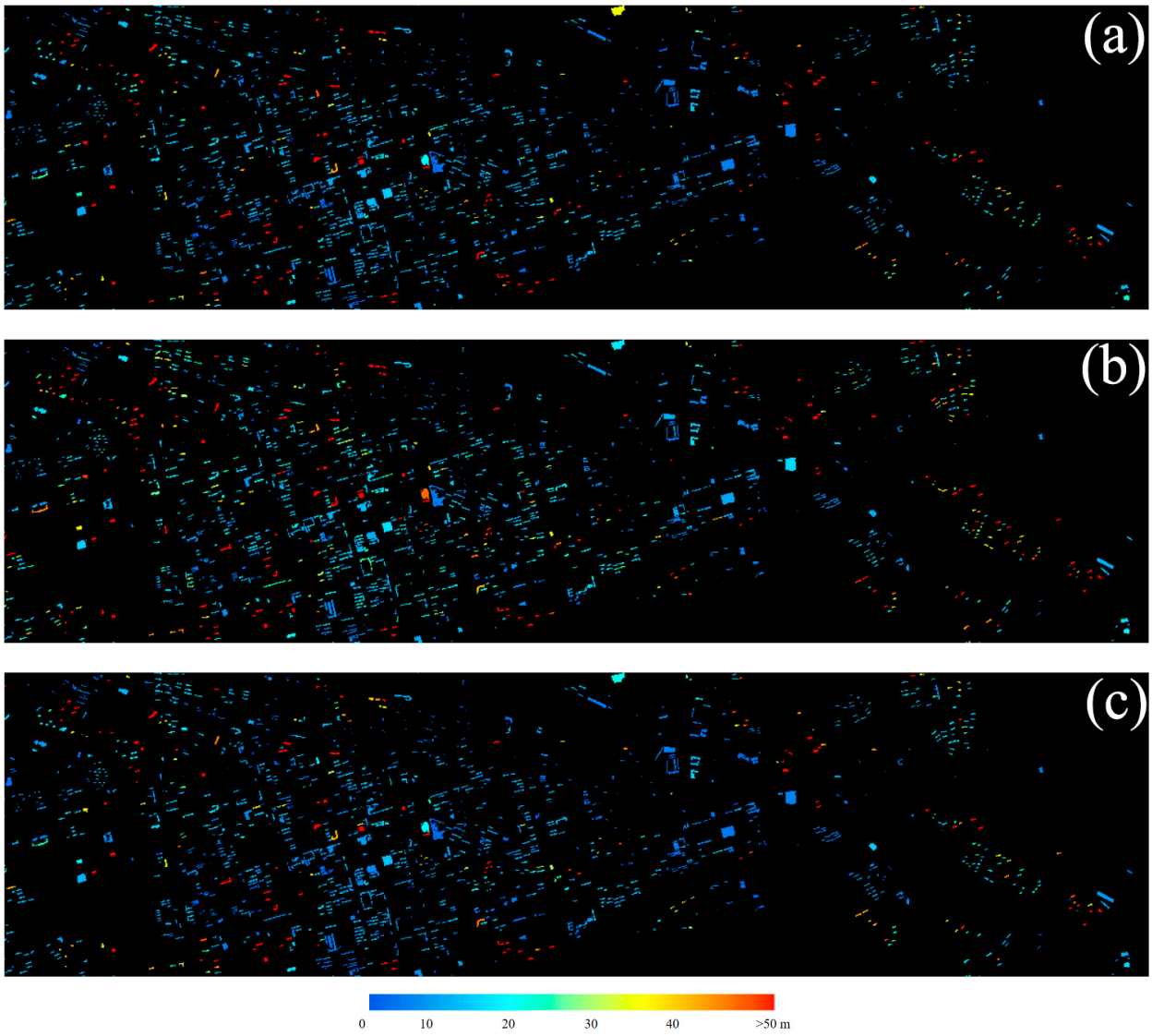


Figure 5. Comparison of the initial BH and NBH height maps of the IBH in validation region S1. (**a**) Reference height map of IBH. (**b**) Height map of IBH. (**c**) Height map of the NBH for IBH.

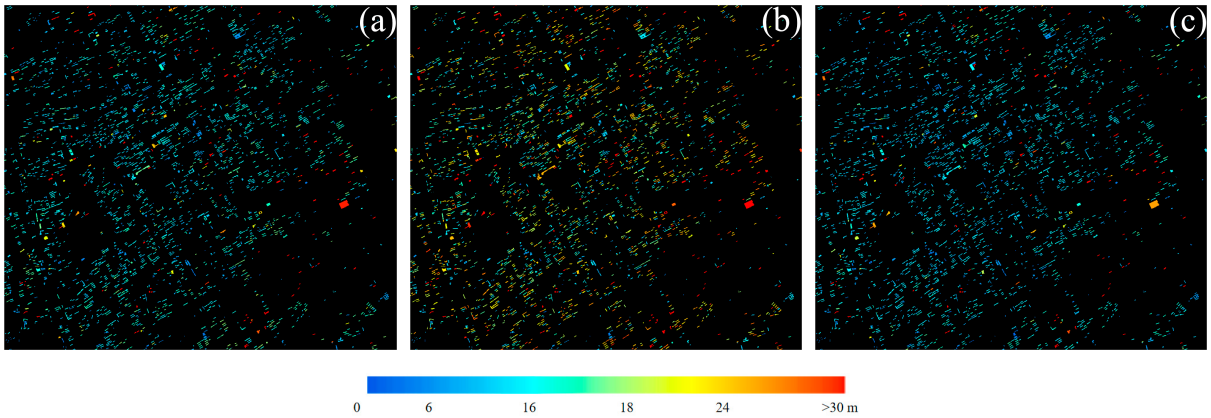


Figure 6. Comparison of the initial BH and NBH height maps of the IBH in validation region S2. (**a**) Reference height map of the IBH. (**b**) Height map of the IBH. (**c**) Height map of the NBH for the IBH.

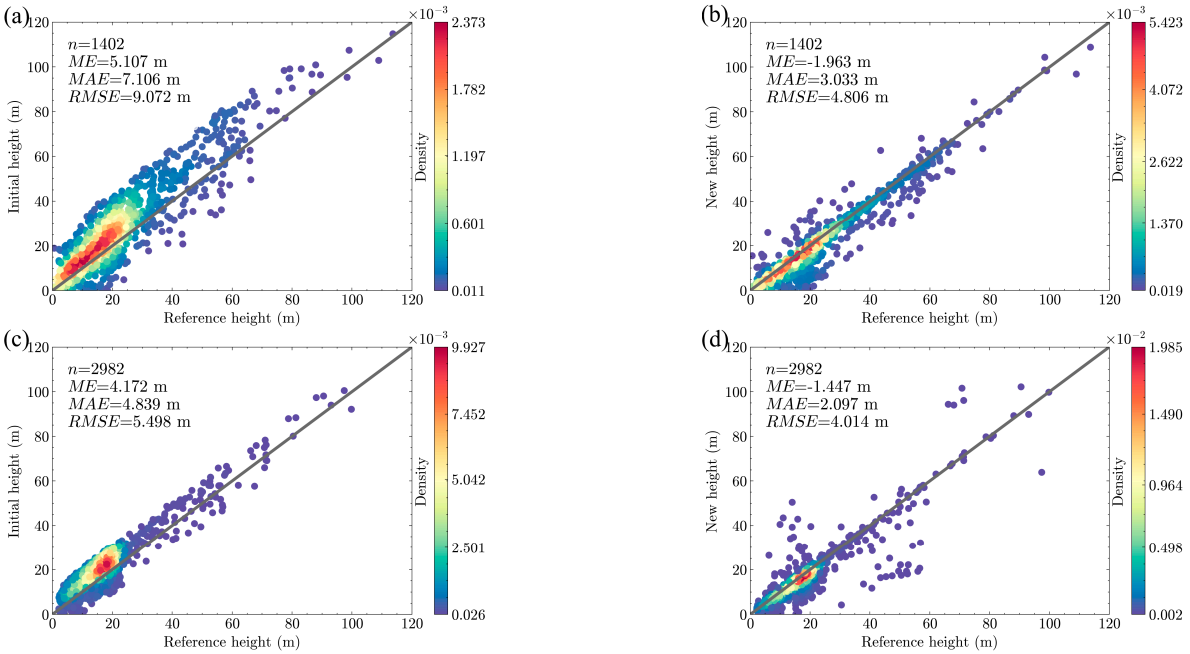


Figure 7. Evaluation of the accuracy of the initial BH and NBH for IBH buildings in validation regions S1 and S2. (**a**,**b**) are the accuracy evaluations of the initial BH and NBH of IBH buildings in the validation area S1, respectively. (**c**,**d**) are the accuracy evaluations of the initial BH and NBH of IBH buildings in the validation area S2, respectively.

4.2. Discussion

4.2.1. Accuracy of the Overall Building Height

All the buildings used for validation were evaluated to assess the proposed method's performance. To compare the reliability of the methods, we chose the method of elevation difference model (EDM) based on DSM [31] as the comparative method. As shown in Figure 8a, the buildings in validation area S1 are densely arranged, and the shadows are longer. In particular, in the central built-up area with tall buildings (on the left of validation area S1), many shadows are present between the buildings, and the ground is less exposed. The ground elevation steadily becomes more complex toward the right side of verification area S1. As shown in Figure 8b, the buildings in validation area S2 are arranged in a more orderly and more dense manner, with more buildings of similar heights. Although the impact of building shadows in S2 is not as significant as in S1, numerous non-ground disturbances are present, such as vegetation, vehicles, and small man-made objects between buildings. Six small regions from validation regions S1 and S2 were extracted to compare the initial BH estimation results with the final results (as shown in Figure 9). The small area map clearly shows that the three small areas (A1–A3) in S1 have more shadows on the buildings, covering much of the ground, and the shapes and heights of the buildings are also diverse. The buildings in the three small areas (A4–A6) in S2 are often disturbed by vegetation or non-ground objects, but more buildings of the same height are present. A comparison of the height maps of different small regions shows that the initial BH estimation results perform better in S2 than in S1 because S2 has more ground exposure, making ground elevation selection simpler. Therefore, a small number of buildings with more ground interference have a larger height error than the reference building. This is due to the EDM method requiring a large and stable amount of ground elevation to obtain accurate ground elevation. Therefore, when shadows and dense buildings result in a lack of ground elevation, the accuracy of ground elevation selection decreases. Similarly, although the initial BH estimation method does not require many ground elevation values, vegetation and shadows can cause outliers, affecting the selection of ground elevation. After processing with the proposed method, the correction

of the BH is obvious, particularly in A1, where many buildings with significant height errors are corrected. This proves that the height replacement of buildings with similar attributes is reliable, and the proposed method can effectively handle the problem of BH estimation errors caused by a large amount of occlusion of the building's ground. The maps of each small area show that the closer the area is to the concentration of buildings with similar heights, the closer the building height will be to the reference building height after correction.



Figure 8. Multispectral images of validation regions S1 and S2 and the locations of small regions. (**a**) Multispectral image of validation region S1 and the locations of small regions A1–A3. (**b**) Multispectral image of validation region S2 and the locations of small regions A4–A6.

The quantified validation results are listed in Table 2. Owing to the complexity of the ground elevation and shadow interference in S1, the initial RMSE of the BH is 5.444 m. Although the buildings in S2 are dense, the ground visibility is good, and the selection of the ground elevation is much simpler. However, vegetation presents interference; therefore, the RMSE is 3.894 m, which is better than that of S1. Following height correction via the proposed method, the accuracy of the IBH is improved, resulting in a significant overall improvement in the RMSE. The RMSEs for S1 and S2 are increased by 1.729 m and 0.842 m, respectively. The improvement in accuracy indicates that the proposed method can effectively address the poor accuracy in BH extraction caused by vegetation and shadow occlusion. Both the accuracy of the initial and final results are superior to the comparison method EDM. The main reason is that the EDM method requires higher accuracy in ground elevation.

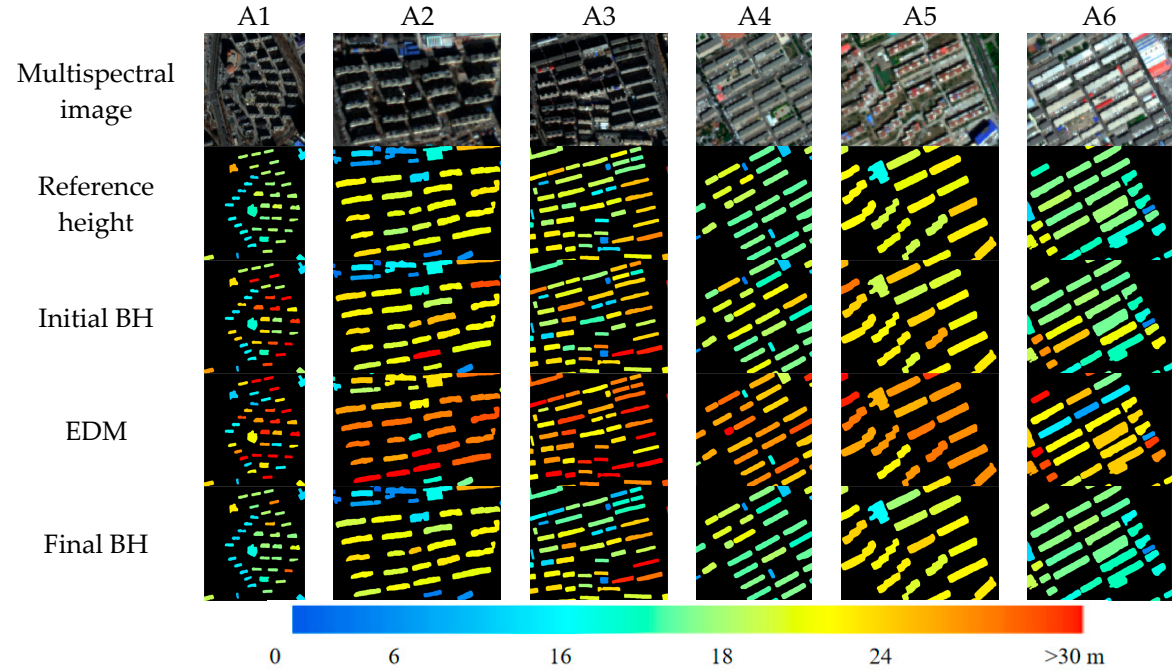


Figure 9. Comparison of the results for small areas in validation regions S1 and S2.

Table 2. Comparison of the accuracy of the initial and final heights of all validated buildings in the validation areas.

Study Area	Method	ME/m	MAE/m	RMSE/m
S1	Initial BH	0.908	3.374	5.444
	EDM	2.483	5.954	11.209
	Final BH	−0.984	2.284	3.715
S2	Initial BH	1.733	2.859	3.894
	EDM	2.511	5.422	9.594
	Final BH	−0.594	1.723	3.052

4.2.2. Threshold for Height Error in Building Classification

The choice of threshold often affects the result. The main objective of the method proposed in this article is to reselect the appropriate height for IBH. Therefore, the threshold (gh_{thres}) for determining whether BH is IBH has a certain impact on the results of the proposed method. To explore this impact, we took validation area S2 as an example and conducted experiments while only changing gh_{thres}, and evaluated the accuracy of the experimental results (as shown in Table 3). From Table 3, as gh_{thres} increases, the limiting effect of the threshold weakens, the number of IBHs continues to decrease, and the accuracy of classification correspondingly increases. Meanwhile, as the number of IBHs decreases, the retained IBHs will be a part of the larger height error. Therefore, as gh_{thres} increases, the RMSE of IBH gradually increases. The RMSE of NBH decreased as gh_{thres} increased from 2 m to 3 m. We believe that when gh_{thres} is 2 m, the quantity of ABH is relatively small, which affects the process of selecting NBH for IBH. However, when gh_{thres} is 3 m, there is enough ABH available. In fact, when ABH is sufficient, as gh_{thres} increases, the RMSE of NBH should slowly increase, and the gap between it and the RMSE of IBH should become larger and larger. Finally, the overall accuracy should show a trend similar to the NBH accuracy. However, when gh_{thres} is 5 m, the RMSE of overall accuracy is 2.989 m, which is better than the RMSE of overall accuracy when gh_{thres} is 3 m. We believe that when gh_{thres} is 3 m, there are some errors in selecting NBH for IBH, resulting in an increase in error. When gh_{thres} is 5 m, these IBHs are classified as ABHs, resulting in an abnormal decrease in the RMSE of overall accuracy. Considering the above analysis and actual data, a value

of 3 m for gh_{thres} is more in line with the application needs of the method proposed in this article.

Table 3. Comparison of experimental results accuracy for different gh_{thres} values.

gh_{thres}/m	Result	4.1.1. Reliability of the Classification				4.1.2. Accuracy of the NBH Selection			4.2.1. Accuracy of the Overall Building Height		
		IBH	ABH	CR (BH)/%	CR (Ground Elevation)/%	ME/m	MAE/m	RMSE/m	ME/m	MAE/m	RMSE/m
2	Initial	3912	3290	76.618	78.534	3.401	4.112	4.921	1.733	2.859	3.894
	Final					−1.247	2.159	4.923	−0.792	1.798	3.895
3	Initial	2982	4220	82.422	83.588	4.172	4.839	5.498	1.733	2.859	3.894
	Final					−1.447	2.097	4.014	−0.594	1.723	3.052
4	Initial	2347	4855	85.601	86.323	4.821	5.448	5.994	1.733	2.859	3.894
	Final					−1.503	2.196	4.248	−0.328	1.799	3.055
5	Initial	1887	5315	88.364	89.336	5.353	5.976	6.444	1.733	2.859	3.894
	Final					−1.541	2.231	4.212	−0.073	1.877	2.989
6	Initial	1475	5727	90.225	91.016	5.793	6.469	6.888	1.733	2.859	3.894
	Final					−1.628	2.338	4.426	0.213	2.012	3.076

4.2.3. Algorithm Complexity

In this part, we will mainly discuss the running environment of the experiment and the time complexity of the algorithm. The experimental setup consisted of Windows 11 as the operating system, an Intel Core i7-12700H processor, 16 GB of RAM, and Python version 3.7 as the programming language. The proposed method is mainly divided into three parts besides obtaining initial results: building classification, BIRCH clustering, and new height selection. The time complexity of both the first and second parts is approximately $O(n)$, while the time complexity of the third part is approximately $O\left(\sum_{i=1}^{C_k} N_{IBH}^i N_{ABH}^i\right)$. Among them, n represents the total number of buildings, C_k represents the number of clusters, and N_{IBH}^i and N_{ABH}^i represent the number of IBH and ABH in the ith class, respectively. Therefore, the overall time complexity is $O\left(n + n + \sum_{i=1}^{C_k} N_{IBH}^i N_{ABH}^i\right) \approx O\left(n + \sum_{i=1}^{C_k} N_{IBH}^i N_{ABH}^i\right)$. The memory requirements for the method depend on the size of the image and the number of buildings. To meet the needs of large-scale applications, batch reading and calculation can be performed to reduce memory usage.

5. Conclusions

This paper proposed a method for estimating building heights from the DSM. First, we used the optimal ground search area and multi-index evaluation mechanism to obtain the initial BH. The initial estimation results were then classified, and spatial clustering and matching were performed. Finally, an appropriate height was selected and reassigned to buildings with significant height errors. Our method was tested in Hohhot and Ordos, Inner Mongolia Autonomous Region. We evaluated over 12,000 buildings and found that the accuracy of building classification was better than 79% when the threshold was 3 m. In addition, we evaluated over 4000 new heights selected for IBH and increased RMSE by 4.266 m and 1.484 m in Ordos and Hohhot, respectively. Finally, our building height estimation method was compared with similar types of methods. Our method estimated building heights that are more consistent with the reference building height data, with an accuracy of 3.052 m, which is better than the comparison methods of 3.894 m and 9.594 m.

The findings demonstrate that the proposed method is effective in mitigating the issue of substantial errors in building height estimation resulting from vegetation and ground shading. Additionally, we demonstrated that the more densely populated the areas with similar buildings or building heights, the more significant the improvement effect of the proposed method. The RMSE accuracy of the final results is significantly better than that of the initial BH extraction results. These findings and the accuracy verification results highlight that the proposed method has enormous potential for BH estimation. However, this method still has biases when selecting building heights in areas with a small number of

buildings and diverse forms of buildings. In future work, it is possible to further integrate data from other sources or improve height selection schemes to obtain more accurate height selection.

Author Contributions: Conceptualization, Jingxin Chang and Yonghua Jiang; methodology, Jingxin Chang; software, Jingxin Chang; validation, Yunming Wang and Shaodong Wei; resources, Meilin Tan; data curation, Meilin Tan; writing—original draft preparation, Jingxin Chang; writing—review and editing, Yonghua Jiang; supervision, Yonghua Jiang; funding acquisition, Yonghua Jiang. 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 No. 42171341, No. 42301424), and the Chinese Universities Scientific Fund (NO. 2042019kf0209).

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

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

References

1. Huang, Y.-D.; Song, Y.; Xu, X.; He, W.-R.; Kim, C.-N. Impacts of Upstream Building Height and Stack Location on Pollutant Dispersion from a Rooftop Stack. *Aerosol. Air Qual. Res.* **2017**, *17*, 1837–1855. [CrossRef]
2. Xi, C.; Ren, C.; Wang, J.; Feng, Z.; Cao, S.-J. Impacts of Urban-Scale Building Height Diversity on Urban Climates: A Case Study of Nanjing, China. *Energy Build.* **2021**, *251*, 111350. [CrossRef]
3. Yoshida, T.; Takemi, T.; Horiguchi, M. Large-Eddy-Simulation Study of the Effects of Building-Height Variability on Turbulent Flows over an Actual Urban Area. *Bound.-Layer Meteorol.* **2018**, *168*, 127–153. [CrossRef]
4. Wang, M.; Xu, H. The Impact of Building Height on Urban Thermal Environment in Summer: A Case Study of Chinese Megacities. *PLoS ONE* **2021**, *16*, e0247786. [CrossRef] [PubMed]
5. Chen, Y.; Chen, J.; Zhao, S.; Xu, X.; Liu, X.; Zhang, X.; Zhang, H. Inferring the Heterogeneous Effect of Urban Land Use on Building Height with Causal Machine Learning. *GIScience Remote Sens.* **2024**, *61*, 2321695. [CrossRef]
6. Yu, G.; Xie, Z.; Li, X.; Wang, Y.; Huang, J.; Yao, X. The Potential of 3-D Building Height Data to Characterize Socioeconomic Activities: A Case Study from 38 Cities in China. *Remote Sens.* **2022**, *14*, 2087. [CrossRef]
7. Aravinth, J.; Lavenya, R.; Shanmukha, K.; Vaishnavi, K. Evaluation and Analysis of Building Height with LiDAR Data. In Proceedings of the 2018 3rd International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 15–16 October 2018; pp. 397–402. [CrossRef]
8. Yang, X.; Wang, C.; Xi, X.; Wang, P.; Lei, Z.; Ma, W.; Nie, S. Extraction of Multiple Building Heights Using ICESat/GLAS Full-Waveform Data Assisted by Optical Imagery. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1914–1918. [CrossRef]
9. Lao, J.; Wang, C.; Zhu, X.; Xi, X.; Nie, S.; Wang, J.; Cheng, F.; Zhou, G. Retrieving Building Height in Urban Areas Using ICESat-2 Photon-Counting LiDAR Data. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *104*, 102596. [CrossRef]
10. Ma, X.; Zheng, G.; Chi, X.; Yang, L.; Geng, Q.; Li, J.; Qiao, Y. Mapping Fine-Scale Building Heights in Urban Agglomeration with Spaceborne Lidar. *Remote Sens. Environ.* **2023**, *285*, 113392. [CrossRef]
11. Kaplan, G.; Comert, R.; Kaplan, O.; Matci, D.K.; Avdan, U. Using Machine Learning to Extract Building Inventory Information Based on LiDAR Data. *IJGI* **2022**, *11*, 517. [CrossRef]
12. Liu, K.; Balz, T.; Liao, M. Building Height Extraction via a Deterministic Approach Using a TerraSAR-X Data Stack. In Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium, Honolulu, HI, USA, 23–25 July 2010; pp. 2920–2923. [CrossRef]
13. Guida, R.; Iodice, A.; Riccio, D. Height Retrieval of Isolated Buildings From Single High-Resolution SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2967–2979. [CrossRef]
14. Li, X.; Zhou, Y.; Gong, P.; Seto, K.C.; Clinton, N. Developing a Method to Estimate Building Height from Sentinel-1 Data. *Remote Sens. Environ.* **2020**, *240*, 111705. [CrossRef]
15. Li, Z.; Ji, S.; Fan, D.; Yan, Z.; Wang, F.; Wang, R. Reconstruction of 3D Information of Buildings from Single-View Images Based on Shadow Information. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 62. [CrossRef]
16. Xu, W.; Feng, Z.; Wan, Q.; Xie, Y.; Feng, D.; Zhu, J.; Liu, Y. Building Height Extraction From High-Resolution Single-View Remote Sensing Images Using Shadow and Side Information. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 6514–6528. [CrossRef]
17. Kadhim, N.; Mourshed, M. A Shadow-Overlapping Algorithm for Estimating Building Heights from VHR Satellite Images. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 8–12. [CrossRef]
18. Xie, Y.; Feng, D.; Xiong, S.; Zhu, J.; Liu, Y. Multi-Scene Building Height Estimation Method Based on Shadow in High Resolution Imagery. *Remote Sens.* **2021**, *13*, 2862. [CrossRef]
19. Cao, Y.; Huang, X. A Deep Learning Method for Building Height Estimation Using High-Resolution Multi-View Imagery over Urban Areas: A Case Study of 42 Chinese Cities. *Remote Sens. Environ.* **2021**, *264*, 112590. [CrossRef]

20. Zhou, X.; Myint, S.W. Shadow Pattern-Enhanced Building Height Extraction Using Very-High-Resolution Image. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 180–190. [CrossRef]
21. Albano, R. Investigation on Roof Segmentation for 3D Building Reconstruction from Aerial LIDAR Point Clouds. *Appl. Sci.* **2019**, *9*, 4674. [CrossRef]
22. Wang, X.; Li, P. Extraction of Urban Building Damage Using Spectral, Height and Corner Information from VHR Satellite Images and Airborne LiDAR Data. *ISPRS J. Photogramm. Remote Sens.* **2020**, *159*, 322–336. [CrossRef]
23. Nguyen, T.H.; Daniel, S.; Guériot, D.; Sintès, C.; Le Caillec, J.M. Super-Resolution-Based Snake Model—An Unsupervised Method for Large-Scale Building Extraction Using Airborne LiDAR Data and Optical Image. *Remote Sens.* **2020**, *12*, 1702. [CrossRef]
24. Zwally, H.J.; Schutz, B.; Abdalati, W.; Abshire, J.; Bentley, C.; Brenner, A.; Bufton, J.; Dezio, J.; Hancock, D.; Harding, D.; et al. ICESat's Laser Measurements of Polar Ice, Atmosphere, Ocean, and Land. *J. Geodyn.* **2002**, *34*, 405–445. [CrossRef]
25. Tan, J.; Cheng, C.; Yang, S.; Xu, Z.; Cheng, G. Building Height Extraction with Full Waveform Decomposition of Spaceborne Laser Altimetry. *Sci. Surv. Mapp.* **2020**, *45*, 102–109. [CrossRef]
26. Gong, P.; Li, Z.; Huang, H.; Sun, G.; Wang, L. ICESat GLAS Data for Urban Environment Monitoring. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 1158–1172. [CrossRef]
27. Dandabathula, G.; Sitiraju, S.R.; Jha, C.S. Retrieval of Building Heights from ICESat-2 Photon Data and Evaluation with Field Measurements. *Environ. Res. Infrastruct. Sustain.* **2021**, *1*, 011003. [CrossRef]
28. Zhou, Y.; Li, X.; Chen, W.; Meng, L.; Wu, Q.; Gong, P.; Seto, K.C. Satellite Mapping of Urban Built-up Heights Reveals Extreme Infrastructure Gaps and Inequalities in the Global South. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2214813119. [CrossRef]
29. Li, M.; Koks, E.; Taubenböck, H.; Van Vliet, J. Continental-Scale Mapping and Analysis of 3D Building Structure. *Remote Sens. Environ.* **2020**, *245*, 111859. [CrossRef]
30. Wu, W.-B.; Ma, J.; Banzhaf, E.; Meadows, M.E.; Yu, Z.-W.; Guo, F.-X.; Sengupta, D.; Cai, X.-X.; Zhao, B. A First Chinese Building Height Estimate at 10 m Resolution (CNBH-10 m) Using Multi-Source Earth Observations and Machine Learning. *Remote Sens. Environ.* **2023**, *291*, 113578. [CrossRef]
31. Zeng, C.; Wang, J.; Zhan, W.; Shi, P.; Gambles, A. An Elevation Difference Model for Building Height Extraction from Stereo-Image-Derived DSMs. *Int. J. Remote Sens.* **2014**, *35*, 7614–7630. [CrossRef]
32. Glinka, S.; Bajer, J.; Wierzbicki, D.; Karwowska, K.; Kedzierski, M. The Use of Deep Learning Methods for Object Height Estimation in High Resolution Satellite Images. *Sensors* **2023**, *23*, 8162. [CrossRef]
33. Wu, B.; Huang, H.; Zhao, Y. Utilizing Building Offset and Shadow to Retrieve Urban Building Heights with ICESat-2 Photons. *Remote Sens.* **2023**, *15*, 3786. [CrossRef]
34. Zhao, Y.; Wu, B.; Li, Q.; Yang, L.; Fan, H.; Wu, J.; Yu, B. Combining ICESat-2 photons and Google Earth Satellite images for building height extraction. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *117*, 103213. [CrossRef]
35. Wang, J.; Hu, X.; Meng, Q.; Zhang, L.; Wang, C.; Liu, X.; Zhao, M. Developing a Method to Extract Building 3D Information from GF-7 Data. *Remote Sens.* **2021**, *13*, 4532. [CrossRef]
36. Zhang, C.; Cui, Y.; Zhu, Z.; Jiang, S.; Jiang, W. Building Height Extraction from GF-7 Satellite Images Based on Roof Contour Constrained Stereo Matching. *Remote Sens.* **2022**, *14*, 1566. [CrossRef]
37. Casella, V.; Zampori, B.; Dell'Acqua, F.; Gamba, P.; Mainardi, A. DTM Extraction in Urban Areas: A Detailed Comparison of Methodologies, Algorithms and Results. In Proceedings of the IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No.01CH37217), Sydney, Australia, 9–13 July 2001; Volume 3, pp. 976–978. [CrossRef]
38. Li, Y.; Zhu, L.; Gong, P.; Shimamura, H. A Refined Marker Controlled Watershed for Building Extraction from DSM and Imagery. *Int. J. Remote Sens.* **2010**, *31*, 1441–1452. [CrossRef]
39. Hu, H.; Ding, Y.; Zhu, Q.; Jiang, J.; Wen, X.; Zhang, L.; Tang, W.; Yang, J.; Zhong, R. Precision global DEM generation based on adaptive surface filter and Poisson terrain editing. *Acta Geod. Et Cartogr. Sin.* **2019**, *48*, 374–383.
40. Chang, J.; Jiang, Y.; Li, J.; Tan, M.; Wang, Y.; Wei, S. Building Height Extraction Based on Joint Optimal Selection of Regions and Multi-Index Evaluation Mechanism. *IEEE Trans. Geosci. Remote Sens.* **2023**, *62*, 5603113. [CrossRef]
41. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
42. Lorbeer, B.; Kosareva, A.; Deva, B.; Softić, D.; Ruppel, P.; Küpper, A. Variations on the Clustering Algorithm BIRCH. *Big Data Res.* **2018**, *11*, 44–53. [CrossRef]
43. Guan, H.; Li, J.; Chapman, M.; Deng, F.; Ji, Z.; Yang, X. Integration of Orthoimagery and Lidar Data for Object-Based Urban Thematic Mapping Using Random Forests. *Int. J. Remote Sens* **2013**, *34*, 5166–5186. [CrossRef]
44. Hu, M.-K. Visual Pattern Recognition by Moment Invariants. *IEEE Trans. Inform. Theory* **1962**, *8*, 179–187. [CrossRef]
45. Meng, N.; Wang, Z.; Gao, C.; Li, J. A vector building clustering algorithm based on local outlier factor. *Geomat. Inf. Sci. Wuhan Univ.* **2024**, *49*, 562–571. [CrossRef]

46. Steiniger, S.; Lange, T.; Burghardt, D.; Weibel, R. An Approach for the Classification of Urban Building Structures Based on Discriminant Analysis Techniques. *Trans. GIS* **2008**, *12*, 31–59. [CrossRef]
47. Zhou, C.; Wang, Z.; Chen, Q.; Jiang, Y.; Pei, J. Design Optimization and Field Demonstration of Natural Ventilation for High-Rise Residential Buildings. *Energy Build.* **2014**, *82*, 457–465. [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.