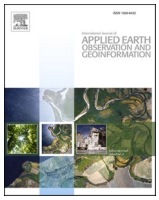
International Journal of Applied Earth Observations and Geoinformation 117 (2023) 103213

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Combining ICESat-2 photons and Google Earth Satellite images for building height extraction

Yi Zhao [a,b,d], Bin Wu [a,c,*], Qiaoxuan Li [a,b], Lei Yang [a,b], Hongchao Fan [d], Jianping Wu [a,b], Bailang Yu [a,b,*]

[a] Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai 200241, China
[b] School of Geographic Sciences, East China Normal University, Shanghai 200241, China
[c] School of Geospatial Engineering and Science, Sun Yat-sen University, Zhuhai 519082, China
[d] Department of Civil and Environmental Engineering, Norwegian University of Science and Technology, Trondheim 7491, Norway

ARTICLE INFO

Keywords:
Building height
ICESat-2
Google Earth images
Building shadow
U-Net

ABSTRACT

Building heights are one of the crucial data for comprehending the functions of urban systems. Employing optical remote sensing imagery, the shadow-based method is one of the most promising methods which have been proposed for estimating building height. However, the existing shadow-based studies for building height estimation are restricted to a small area due to the lack of building height annotations and ignorance of building azimuth variations. The Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) allows large-scale building height retrieval in the along-track direction and thus can be taken as ground truth building height annotations to support the shadow-based algorithms for large-scale building height extraction. Here, we proposed an approach for extracting building height by combining Google Earth Satellite (GES) images and ICESat-2 photons. Building and shadow instances were first extracted using a U-Net deep learning framework. Based on the building height annotations retrieved from ICESat-2 photons, an improved shadow-based building height estimation model by minimizing the global error across all sample buildings was developed. A typical urban area located in the city center of Shanghai, China with an area of around 90 km^2 was selected to validate the proposed method. In total 15,966 buildings were successfully extracted and the results indicated that the estimated building heights have high accuracy with an absolute mean error of 4.08 m. Moreover, the proposed method shows a better performance compared to the existing shadow-based method and existing building height datasets. The method holds great potential for large-scale building-level height retrieval which contributes to further studies of urban morphologies.

1. Introduction

As a primary parameter of three-dimensional (3D) urban morphologies (Ren et al., 2020; Zhao et al., 2020a), the building height is essential in understanding the impacts of the vertical characteristics of urban development and evolution. Various studies, such as energy consumption estimation (Güneralp et al., 2017; Pérez-Lombard et al., 2008), urban microclimate simulation (Huang and Wang, 2019; Xu et al., 2017; Yu et al., 2020), 3D reconstruction (Fan et al., 2021; Haala and Kada, 2010; Wu et al., 2017), urban population estimation (Chen et al., 2021; Langford et al., 2008; Wu et al., 2023), and urban planning (Wu et al., 2021; Wu et al., 2018; Yu et al., 2016; Zhao et al., 2020b), have mentioned the importance of building heights. Building height estimation has thus been a primary subject in studies related to the urban environment and human well-being (Yu et al., 2016).

The existing approach which acquires building heights from field surveys is usually time-consuming and labor-intensive, thus could only be conducted on a limited number of buildings and is hard to apply on a large scale. Remote sensing technologies can provide simultaneous observation on a large scale in a cost-effective way and were thus utilized to estimate the building height. Up to date, many regional or global scale building height datasets (e.g., World Settlement Footprint 3D (Esch et al., 2022), Global 3D Building Structure Data (Li et al., 2022), Continental-scale 3D Building Structure Data (Li et al., 2020a), and Building Height Map of Germany dataset (Frantz et al., 2021)) have been established with remote sensing technologies. However, most of the

* Corresponding authors at: Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai 200241, China.
E-mail addresses: wubin65@mail.sysu.edu.cn (B. Wu), blyu@geo.ecnu.edu.cn (B. Yu).

https://doi.org/10.1016/j.jag.2023.103213
Received 18 August 2022; Received in revised form 29 December 2022; Accepted 24 January 2023
Available online 30 January 2023

existing large building height datasets remain constrained at the grid level. Existing height estimation methods for individual buildings are mainly based on Light Detection and Ranging (LiDAR), Synthetic Aperture Radar (SAR) images, and high-resolution optical imagery (Frantz et al., 2021). LiDAR data provides precise building height measurements (Baltsavias, 1999; Yu et al., 2010) but requires high-cost data collection at the same time (Gong et al., 2010). Alternatively, the SAR (Soergel et al., 2009; Torres et al., 2012) and Interferometric Synthetic Aperture Radar (InSAR) images are of great potential in extracting building height (Li et al., 2020b; Wegner et al., 2013), however, individual building interpretation from SAR images is highly challenging due to the side-looking imaging geometry and complex backscattering mechanism (Sun et al., 2022). High-resolution optical imagery has fine spatial information which can be used for building height analysis. Furthermore, optical imagery does not have signal interference problems that occurred in the SAR technique due to the different data collection methods (Cao and Huang, 2021). Besides, compared with LiDAR data, it has a lower cost of data collection (Xie et al., 2021). As one of the optical images, stereo/multi-view images are also a common data source used for building height estimation (Jung, 2004; Li et al., 2020b). However, the multi-view stereo algorithms require at least two images from different views and involve complex arithmetic processes (Cao and Huang, 2021).

Using a single optical image, studies (Izadi and Saeedi, 2011; Nagao et al., 1979; Qi et al., 2016; Shao et al., 2011) have also demonstrated that building height can be calculated according to a building's shadow length and its geometric relationship between the sun and sensors. For example, Liasis and Stavrou (2016) proposed an automatic shadow detection method and obtained the shadow lengths using the median parallel line lengths of each shadow. They implemented the proposed method in 198 buildings and reported a 4.13% overall variance. Xie et al. (2021) proposed a cutline and fishnet method for shadow length extraction and applied their method to 131 buildings' height estimations, most of the building estimations show an absolute error of lower than two meters. However, most existing shadowed-based studies require satellite metadata or building height annotation data. Moreover, shadow-based building height estimation is usually limited to small areas based on the assumption that building azimuths in the study area are consistent, which ignores the influence of the building azimuths on the building height estimation accuracy. Many shadow-based building height estimation studies (Raju et al., 2014; Shao et al., 2011) have demonstrated that building azimuth is a significant factor in large-scale building height estimation of shadow-based methods, especially under the situation when the difference in azimuths between the sun and the sensor is smaller than 180 degrees.

The photons collected by ICESat-2 can provide a new data source of building height annotation to supplement the shadow-based building height estimation methods. As a level 2 product of ICESat-2, ATL03 provides the longitude, latitude, and elevation of each photon event reference to the World Geodetic System 1984 ellipsoid. The ATL03 product makes it possible to obtain building heights on a global scale. Particularly, the accuracy of ATL03 in measuring building heights in urban area has been verified by previous studies (Dandabathula et al., 2021; Zhao et al., 2021). More recently, Lao et al. (2021) proposed a random sample consensus method to retrieve urban building height from ATL03 photons and a root mean square error of 0.45 m was reported. Although high accuracy of building height estimation can be obtained from ICESat-2 ATL03 photons, ICESat-2 can only provide the height of a limited number of sample buildings along its tracks.

In view of this, a new method by combining Google Earth Satellite (GES) images and ICESat-2 ATL03 photons was proposed for extracting building height. Utilizing the shadow and building boundary information extracted from the GES images and the sample building height annotations provided by ATL03 photons, a building height estimation model by minimizing the global error across all sample buildings was developed. The remainder of this paper is organized as follows. The study area and data are described in Section 2. Section 3 introduces the method in detail. Section 4 reports the experimental results. Section 5 displays the method comparisons and discussions; Finally, conclusions of this study are drawn in Section 6.

2. Study area and data

2.1. Study area

The study area, as shown in Fig. 1(a), is located in the city center of Shanghai, China adjacent to the Huangpu River, and covers an area of about 90 km^2. The study area is a typical urban area with characteristics of high building density, diverse building types, and intricate building distributions, which are ideal for testing the performance of our method.

2.2. Data

2.2.1. GES image data

Fig. 1(b) shows the GES image of the study area, which was taken on February 7, 2021, with a high spatial resolution of 0.3 m per pixel. The GES image was downloaded with Python scripts through the Google Earth application program interface (API) provided by Google.

2.2.2. ICESat-2 data

The ATL03 data from October 2018 to November 2021 were collected from the National Snow and Ice Data Center (NSIDC) (https://nsidc.org/data/atl03, accessed in November 2021). The green dots shown in Fig. 1(b) represent the collected ATL03 photons in the study area. In total, the ATL03 data covers 11 granules and 32 tracks, with a total of 1,014,974 photons acquired.

2.2.3. Reference data for building height

The Digital Surface Model (DSM) data provided by Shanghai Surveying and Mapping Institute was employed as the reference data for the building height estimation accuracy evaluation. The DSM data was produced using the stereo pair method based on the high-resolution aerial imagery taken in January 2016. It covers the inner area of the outer ring road of Shanghai with a fine spatial resolution of 0.5 m/pixel. The precision of the DSM data was verified with a bias of under 0.58 m. Considering the DSM contains the terrain and non-terrain objects, it must be further normalized based on the bottom height of the buildings to acquire the absolute height of the buildings, and the specific method is presented in Section 3.4. All the datasets used in our study were finally processed and projected in Universal Transverse Mercator Zone 51N, World Geodetic System 1984 datum in unit of meters.

3. Method

The high-resolution GES images can be used to extract building boundaries and shadows, while the ATL03 photons have the potential to provide the height of a number of sample buildings along the tracks. Therefore, we combine GES images and ATL03 photons data to develop an optimal shadow-based model for building height estimation. Fig. 2 illustrates the workflow of our method. The method contains four steps: 1) building boundaries and shadows extraction from GES image; 2) sample buildings' height extraction from ATL03 photons; 3) large-scale building height determination; 4) accuracy validation.

3.1. Building boundaries and shadows extraction from GES image

3.1.1. Building boundary extraction

As one of the state-of-the-art semantic segmentation frameworks of convolutional neural networks (CNN), U-Net (Ronneberger et al., 2015) has been successfully and widely used for feature extraction and semantic segmentation tasks (e.g., building boundaries and roads extraction) from satellite images (Falk et al., 2019; Zhang et al., 2018). To its

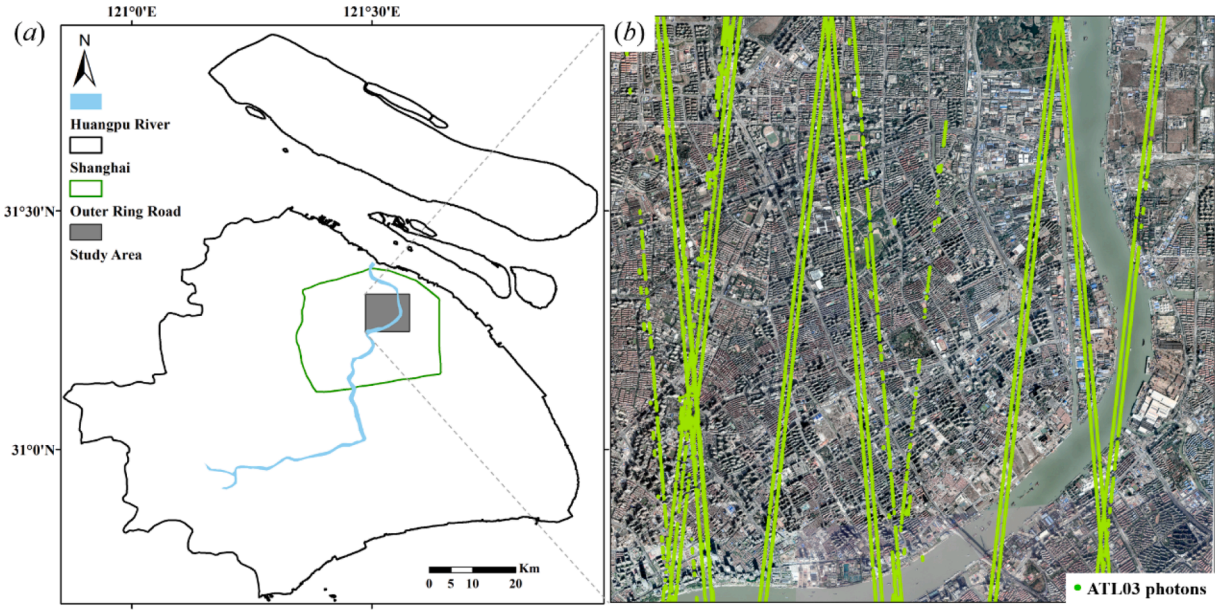


Fig. 1. (*a*) The geolocation of the study area; (*b*) The GES image of the study area.

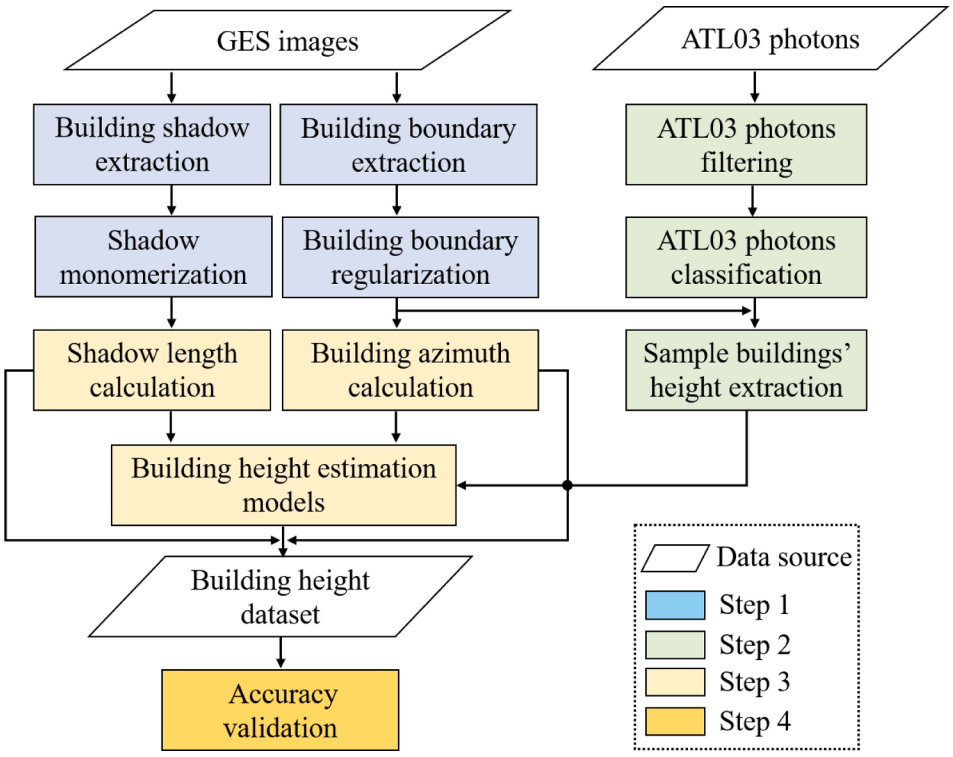


Fig. 2. The flowchart of our method.

superior performance in image semantic segmentation, the U-Net framework was employed for extracting building boundaries in our study. In this study, we have annotated 1,057 images with a size of 256 × 256 pixels from GES images to train the U-Net model for building boundary extraction. The building boundaries in the 1,057 images were labeled with the help of the DSM data. The building boundary labeling from DSM data follows a semi-automatic process. We first employed a mathematical morphological filter (Arefi et al., 2011) to create a digital terrain model (DTM). By subtracting the DTM from the DSM, a normalized DSM (nDSM) was generated. The building boundaries were then automatically identified through a recursive connected component identification and indexing algorithm (Yu et al., 2010) from the nDSM data. Finally, a carefully manual examination was conducted to eliminate spurious boundaries and refine boundaries. Fig. 3(a) and (e) show the DSM data of two sample areas, and Fig. 3(b) and (f) display the corresponding GES images of the two sample areas. The labeled building boundaries are shown in Fig. 3(c) and (g) with green frames. As can be seen from Fig. 3(c) and (g), building boundaries were successfully identified and matched well with the DSM data. By overlaying the annotated building boundaries onto the GES images (Fig. 3(d) and (h)), the position accuracy of the annotated building boundaries was further verified as an obvious offset can be found between the building

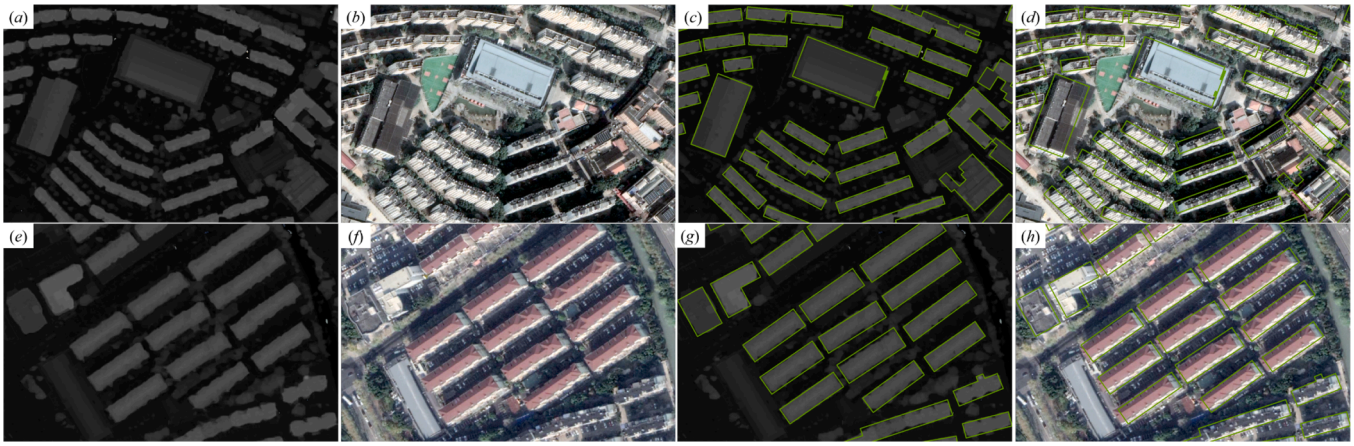


Fig. 3. The building boundaries labeled from the DSM data. (*a*) and (*e*) show the DSM data of two sample areas; (*b*) and (*f*) display the GES images; (*c*) and (*g*) show the labeled building boundaries overlay onto the DSM data; (*d*) and (*h*) display the labeled building boundaries overlay onto the GES images.

boundaries and building rooftops.

Before network training, data augmentation was used for enlarging the training samples. Then the GES images were randomly categorized into training, testing, and validation sets by a ratio of 7:2:1. The U-Net model for building boundary extraction was trained on the training sets for 300 epochs with an initial learning rate of 0.001 and a batch size of 16. The U-Net model was implemented using TensorFlow and Keras on a consumer-level PC with an Intel Core CPU i7-8700 at 3.20 GHz and Nvidia GeForce GTX 1080 8 G graphics card.

As shown in Fig. 4(a), the initial extracted building boundaries by the U-Net model are usually irregular and incomplete, while many buildings have regular boundaries in reality. To refine the building outlines, the building boundaries were further regularized using the "Regularize Building Footprint" tool in Esri ArcGIS Desktop software. Fig. 4(b) shows the regularized building boundary corresponding to Fig. 4(a).

3.1.2. Shadow extraction

The previous study (Luo et al., 2020) proved the U-Net has an excellent performance in shadow extraction. Using the same 1,057 GES images for building boundary extraction model training, we labeled the shadows manually and then built the U-Net model for building shadow extraction. Similarly, the GES images are randomly categorized into training, testing, and validation sets by a ratio of 7:2:1 after data augmentation. In the training process, the U-Net model for shadow extraction was trained on the training sets for 300 epochs with an initial learning rate of 0.001 and a batch size of 16 using the same computer configuration as the building boundary extraction model.

Using the shadow extraction model, the shadows of non-building objects such as vegetation (see Fig. 4(c)) are also detected due to the similar spectral characteristics between those objects' shadows and building shadows. To mitigate the influence of non-building shadows, the shadows were filtered according to the location and area. Firstly, only the shadows which have intersected or the nearest spatial relationship with the building boundaries remained; Then, the area of the shadows lower than 30 m^2 was further removed. Fig. 4(d) shows the filtered shadow extraction results.

3.2. Sample buildings' height extraction from ATL03 photons

Here we used ATL03 photons to retrieve the building height of the

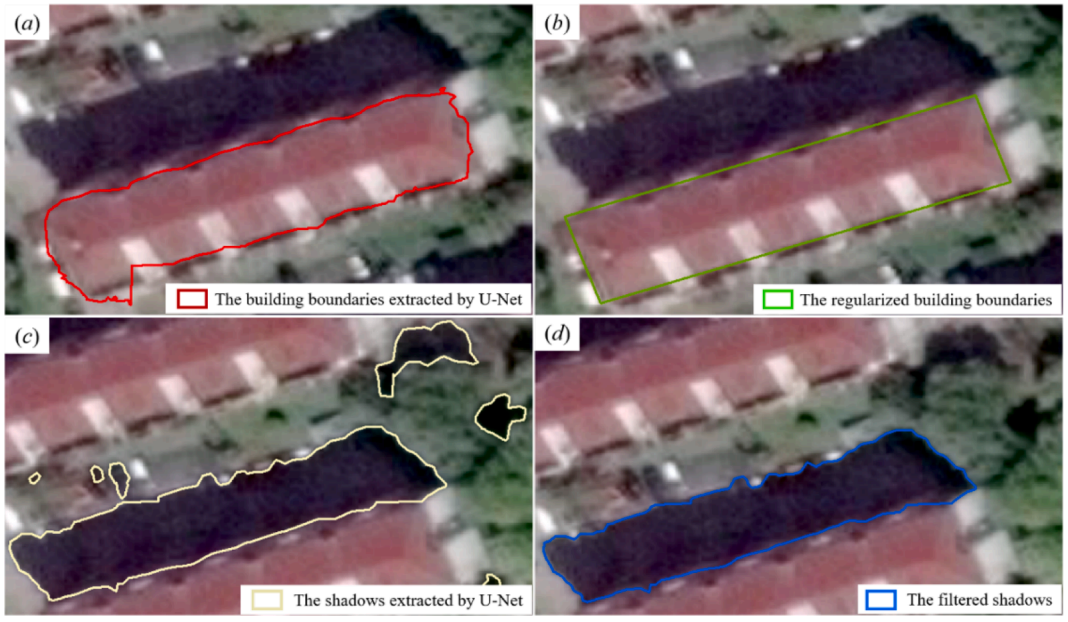


Fig. 4. Building boundaries and shadows extraction. (*a*) shows the initial building boundaries extracted by the U-Net; (*b*) shows the corresponding regularized building boundaries; (*c*) displays the initial shadows extracted by the U-Net; (*d*) shows the filtered shadows.

sample buildings located along the ATL03 tracks. The sample building height extraction contains three steps. Firstly, we conducted a noise removal step to filter the noise photons. The parameter of *signal_conf_ph* in the ATL03 product indicates the signal photon confidence of each photon event. Referring to the previous studies (Lao et al., 2021; Neuenschwander and Magruder, 2019; Zhao et al., 2021), only photons with medium/high (*signal_conf_ph* = 3/4) confidence levels remained. Secondly, the ATL03 photons were classified into ground and off-ground photons using the cloth simulation filter (CSF) method (Zhang et al., 2016). The building photons were further distinguished from the off-ground photons by extracting the photons within the building boundaries (Fig. 5(a)). Finally, the height of the buildings was acquired according to the elevation difference between the building photons and the nearest ground photon to the building boundary (as shown in Fig. 5 (b)).

3.3. Large-scale building height determination

3.3.1. Shadow length calculation

The building shadows often overlap with each other (Fig. 6(a)) affected by the geometric relationship between the sun, the sensor, and the buildings. To calculate the corresponding shadow length of each building from the overlapping shadows, the method proposed by Xie et al. (2021) was employed for shadow monomerization and length calculation. Firstly, we erased the shadows with the extracted building boundaries to separate the shadows from each other along the sun's azimuth direction. Secondly, the cutlines were generated for monomerizing the shadows (Fig. 6(b)). The corner coordinates of the building boundaries were taken as the origins of the cutlines, and the direction of the cutlines was determined according to the sun azimuth angle. Thirdly, the calculation of shadow length is a critical precondition for shadow-based building height calculation. As one of the most efficient and accurate approaches, the parallel line method has been used in many studies (Liasis and Stavrou, 2016; Xie et al., 2021) for shadow length calculation. As shown in Fig. 6(c), a series of parallel lines with small intervals (0.2 m) was created based on the azimuth angle of the sun for measuring the shadow length. Finally, to avoid the random errors caused by the spots and holes in the shadow area on the shadow length calculation, the parallel lines were further filtered by the Pauta criterion, and the average of the remained parallel lines was taken as the length of each shadow. The calculation function is shown as follows:

$$L = \frac{\sum_{j=0}^{n} L_j}{n}, L_j \in [\mu - 3\sigma, \mu + 3\sigma] \tag{1}$$

where L indicates the length of the shadow, L_j represents the length of each parallel line, n is the quantity of the parallel lines; μ and σ represent the mean and standard deviation of the parallel lines' length in each shadow.

3.3.2. Building height estimation

The shadow-based building height estimation models are usually divided into two cases based on the geometrical relationship between the sun and the sensor. Fig. 7 shows the two typical geometric relationships. Fig. 7(a) displays the first type in which the relative azimuth angle between the sun and sensor is in the range from 0 to 180 degrees. As shown in Fig. 7(a), if the relative azimuth angle difference between the sun and sensor is smaller than 180 degrees, part of the shadows are blocked by the building. Based on the azimuths of the sun (A_{su}) and sensor (A_{sa}), the elevation of the sun (E_{su}) sensor (E_{sa}), and the azimuth of buildings (A_{bu}), the algebra relationship between building height (H) and shadow length (L) can be mathematically expressed as follows (Xie et al., 2021):

$$H = L \times K_1$$

$$K_1 = \frac{tanE_{su}tanE_{sa}\sin(A_{bu} - A_{su})}{tanE_{sa}\sin(A_{bu} - A_{su}) - tanE_{su}\sin(A_{bu} - A_{sa})} \tag{2}$$

where K_1 is the proportionality coefficient between H and L.

The second type is the relative azimuth angle between the sun and the sensor is greater than 180 degrees (Fig. 7(b)), all the shadow areas of the building can be captured by the sensor, and the geometrical relationship between H and L is only related to the sun elevation, the calculation function is shown in Equation (3) (Irvin and McKeown, 1989; Shettigara and Sumerling, 1998):

$$H = L \times K_2$$

$$K_2 = tanE_{su} \tag{3}$$

where K_2 is defined as the proportionality coefficient between H and L.

As shown in Equation (2), when the difference in azimuth angle between the sun and the sensor is in the interval of 0 to 180 degrees, the proportionality coefficient K_1 varies with the building azimuths as the other variables (e.g. azimuths of the sun and sensor, the elevation of the sun and sensor) are consistent in an image. Therefore, different buildings have different K_1 values due to the various building azimuths. It is practically not possible for us to calculate the K_1 value for every building, so we classified the buildings into several categories based on their building azimuths to diminish the influence of building azimuths on the building height estimation. While considering each category must have a sufficient number of building samples, we classified the buildings which have a building azimuth angle in the range between 0 and 180 degrees into six categories with an equal interval of 30 degrees. Each category contains a certain number of sample buildings and each sample building corresponds to a proportionality coefficient. To determine the best proportionality coefficient for each category, an optimized shadow-based model by minimizing the global error across all sample buildings

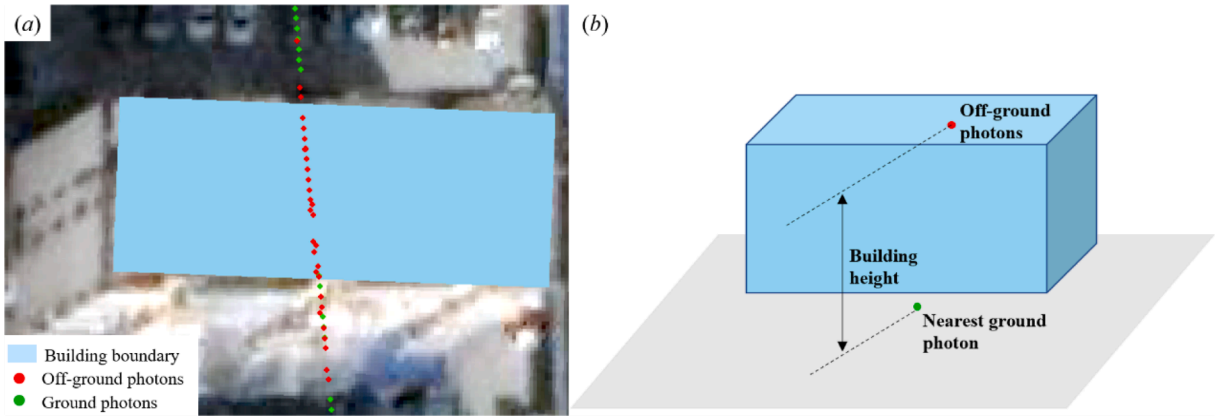


Fig. 5. Building height calculation using ATL03 photons.

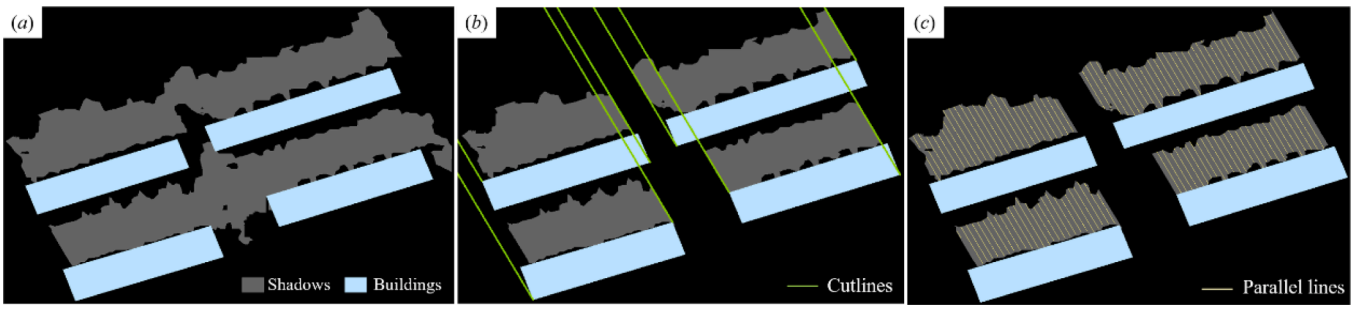
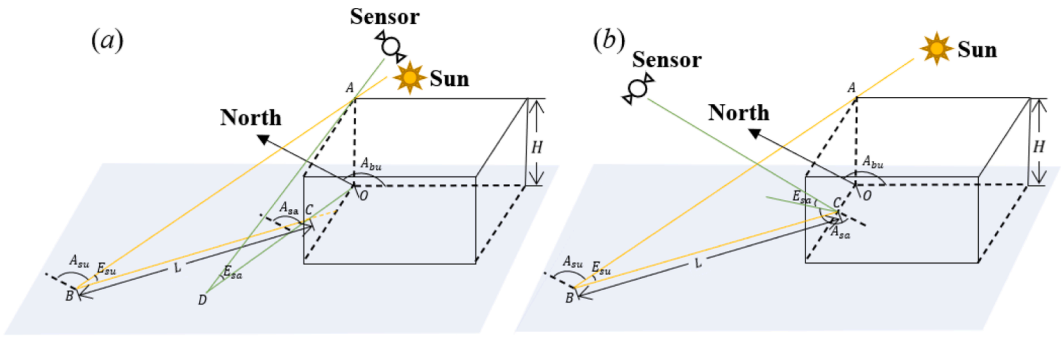


Fig. 6. Shadow length calculation. (*a*) the overlapping shadows; (*b*) the monomerized shadows; (*c*) the parallel lines.



***A*: the building vertices nearest to the sun; *O*: the orthographic projection point of *A*; *B*: the oblique projection point of *A* along A_{su} direction; *D*: the oblique projection point of *A* along A_{sa} direction; *C*: the intersection of the ray along A_{su} direction starting at B and the ray along A_{sa} direction starting at D.**

Fig. 7. The difference in azimuths between the sun and the sensor is less than 180 degrees (*a*) and larger than 180 degrees (*b*).

in each category was developed:

$$H_{es,i} = L_i \times K_o + b_o \tag{4}$$

where $H_{es,i}$ represents the optimal estimated building height for the *i*-th building ($i = 1, \cdots, n$, n is the number of sample buildings); L_i is the shadow length of the *i*-th building; K_o and b_o are the optimal proportionality coefficient and intercept; $H_{ATL,i}$ represents the height of *i*-th building obtained from ATL03 photons.

As shown in Equation (5), we employed the least squares method to obtain the optimal K_o by minimizing the sum of the squares of the residuals. As the proportionality coefficient K is calculated from building samples, the optimal K_o should be limited in a range between the minimum and maximum values. Thus, we limited the optimal K_o in the range from the minimum K to the maximum K to constrain the optimal solution of the least squares method. With the optimal K_o and b_o, we can use Equation (4) to estimate the heights of the rest buildings extracted

Fig. 8. (*a*) The nDSM data; (*b*) The reference building height data.

from the entire images.

$$f(K, b) = \sum_{i=1}^{n} \left(H_{ATL,i} - H_{es}(L_i)\right)^2 = \sum_{i=1}^{n} \left(H_{ATL,i} - b - K \times L_i\right)^2$$
$$\text{minimize} \quad f(K, b) \tag{5}$$
$$\text{subject to}$$
$$\min\left(\frac{H_{ATL,i}}{L_i}\right) \le K \le \max\left(\frac{H_{ATL,i}}{L_i}\right)$$

3.4. Accuracy validation

As mentioned in Section 2.2.3, the DSM data (Fig. 8(a)) needs to be further normalized by terrain height to retrieve the reference building height. The method proposed by Wu et al. (2022) was first employed for generating a normalized DSM (nDSM), as shown in Fig. 8(a). For each building, the reference building height was then calculated as the maximum value of the nDSM pixels within its corresponding building boundary. Fig. 8(b) displays the reference building height distribution.

To validate the accuracy of building height estimation quantitatively, the mean absolute error (MAE), and root mean square error (RMSE) was used for evaluating the estimation results:

$$MAE = \frac{\sum_{i=1}^{n} \left|H_{re,i} - H_{es,i}\right|}{n} \tag{6}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(H_{es,i} - H_{re,i}\right)^2}{n}} \tag{7}$$

where $H_{re,i}$ represents the reference building height, and $H_{es,i}$ is the building heights estimated from the proposed method, n is the count of building samples.

4. Results

4.1. Building boundaries and shadows extraction

Fig. 9(a) shows the building boundaries extracted from the building boundary extraction model and Fig. 9(b) displays the shadows extracted by the shadow extraction model. We selected the area marked by the black box shown in Fig. 9(a) to have a closer look at the extracted results. Fig. 9(c) shows the original GES image, and Fig. 9(d) and (e) show the extracted building boundaries and shadows, respectively. It can be seen that both the building boundaries and shadows were successfully extracted. Moreover, compared with the reference data, the large overlapping area between the extracted results with reference data indicates the good performance of the selected U-Net models. In the study area, 17,088 building boundaries with an area of 13.79 km^2 were initially extracted, which is about 90 percent of the reference data compared to the 19,099 reference building boundaries extracted from

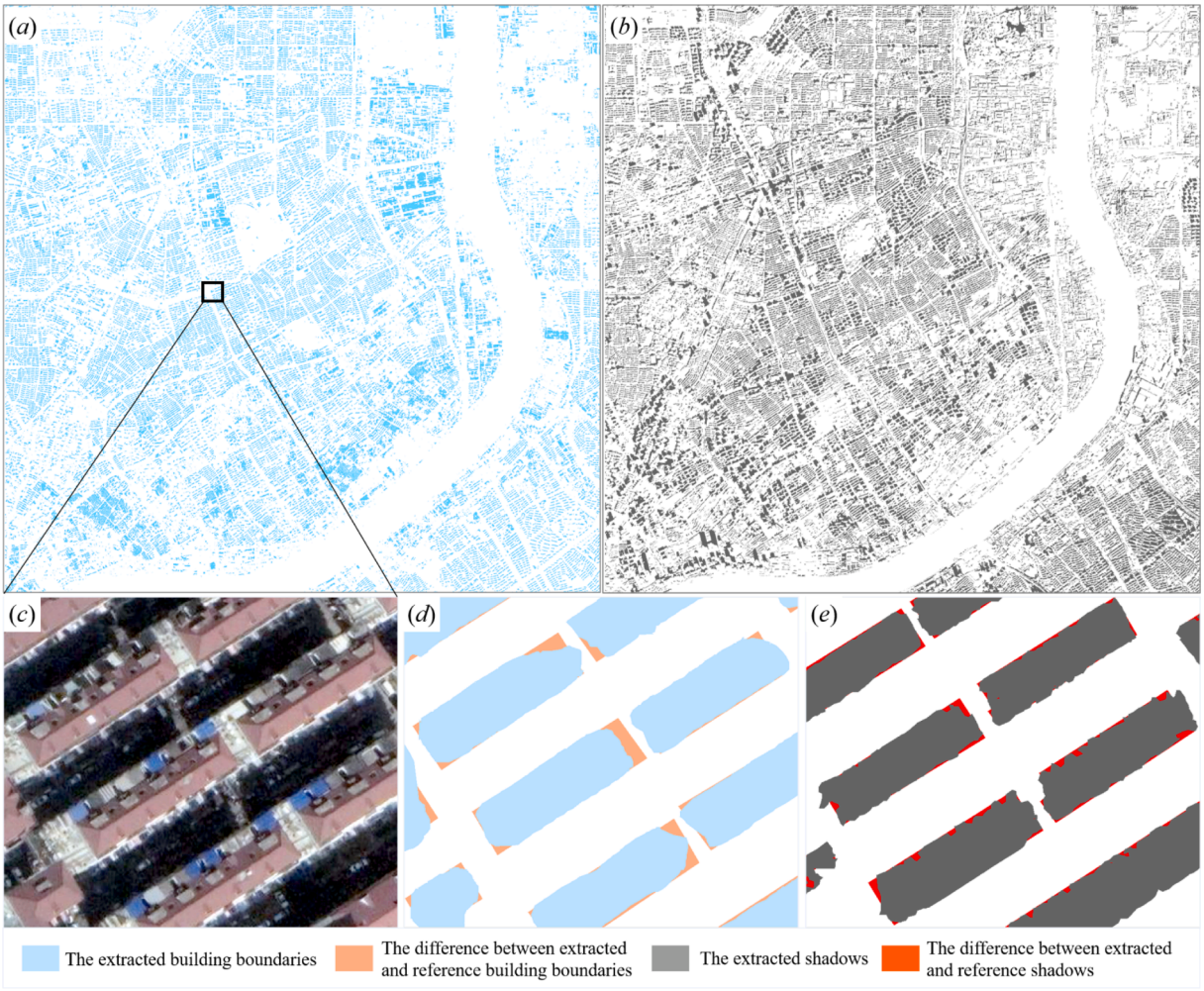


Fig. 9. (*a*) shows the building boundaries extraction results; (*b*) displays the shadows extraction results; (*c*) shows the GES image of the zoomed region (black box); (*d*) displays the building boundaries extraction results corresponding to (*c*); (*e*) shows the shadows extraction results corresponding to (*c*).

the nDSM data. As we mentioned in Section 3.3, only buildings with valid shadows were retained. As shown in Fig. 9(a), 15,966 building boundaries with a total area of 13.01 km^2 remained.

Accuracy assessment for building boundaries and shadows extraction includes not only the evaluation of how many building boundaries and shadows are extracted but also the evaluation of how well the building boundaries and shadows match the reference data. To quantitatively evaluate the building boundaries and shadows extraction results, the indexes of overall accuracy (OA), precision (P), recall (R), F-score (F), and the intersection over union (IOU) were used. Table 1 lists the accuracy of building boundaries and shadows extraction for the test dataset. Overall, our method obtained an average OA of 90.82% and 94.21%, respectively for building boundaries and shadows extraction. It also can be seen that the building shadows extraction has a better performance as the values of OA, P, R, F, and IOU were higher than that of building boundaries. This is probably because the spectral characteristics of buildings are more complex than that of shadows, resulting in a relatively weak generalization of the building boundaries extraction model (Elbakary and Iftekharuddin, 2013).

4.2. Building height estimation

A total of 189,883 ATL03 photons with medium and high signal confidence levels were obtained in our study area. Among them, 86,642 photons are classified as ground photons and the rest 103,241 photons are identified as off-ground photons. Besides, 1,176 sample buildings along the ICESat-2 tracks and their corresponding heights were also obtained according to the method proposed in Section 3.2. The 1,176 sample buildings are validated with an accuracy of MAE = 2.33 m and RMSE = 3.23 m. These sample buildings were then classified into six categories based on their building azimuths to calculate the optimal parameters for building height estimation. Table 2 lists the optimal K_o and b_o value of each category.

The extracted heights were firstly associated with the extracted building boundary polygons and 3D building models at the Level of Detail-1 (LoD-1) level were then generated. Fig. 10 shows the 3D view of the extracted building heights with a total of 9,612 buildings that have a height range between 10 and 20 m. There are 900 buildings have a height higher than 35 m (marked in red in Fig. 10) and these high-rise buildings are mostly distributed in the center of the study area. It can be seen from Fig. 10 that the low-rise buildings (marked in blue) are mainly located adjacent to the Huangpu River, and most of them are factories and industrial buildings.

Fig. 11 displays the overall accuracies of the building height estimation results. Fig. 11(a) shows the frequency histogram of the building height estimation residuals. As shown in Fig. 11(a), the majority (69%) of the residuals fall in the interval from −5 to 5 m, and a slight overestimation was found with an average bias of 0.68 m. Fig. 11(b) shows the scatterplot between the estimation and reference building height. The estimated building heights are in good agreement with the reference building heights, and most of the estimations are distributed around the $y = x$ line. In total, 15,966 buildings' heights were well estimated with an MAE of 4.08 m and an RMSE of 5.34 m. Fig. 11(c) shows the residual distribution of the buildings. It can be seen that the distribution of the residuals is correlated with the types of the buildings, e.g., the buildings with underestimated building heights are mostly located adjacent to the Huangpu River, and most of them are factories and industrial buildings.

To further evaluated the performance of our method, we used the Essential Urban Land Use Categories in China (EULUC-China) land use data (Gong et al., 2020), which is generated with the random forest classifier using the features extracted from Sentinel-2 images, OpenStreetMap, nighttime lights, Point of Interests, and Tencent social big data, to categorize the buildings into four types, i.e., industrial, residential, commercial, and other (including schools, hospitals, and offices). Fig. 12(a) to (d) summarizes the building height estimation accuracies for different building types. In each of these figures, the blue dotted line represents the $y = x$, the solid red line represents the fitting trend line. Fig. 12(e) to (h) displays the representative buildings belonging to each category. Fig. 12(e) to (h) displays the representative buildings belonging to each category. Residential buildings have the best accuracy with an MAE of 3.86 m and RMSE of 5.08 m while commercial buildings own the lowest accuracy (MAE = 4.51 m, RMSE = 5.87 m) in building height estimation. As shown in Fig. 12(f) and (g), this is mainly because residential buildings which have regular shapes are usually neatly arranged, which improves the extraction accuracy of building boundaries and shadow lengths. Conversely, the irregular shapes and concentrated distribution of commercial buildings increase the uncertainties of the building height estimations. For industrial buildings, the shadow length extraction uncertainties caused by the compact layout and similar color with shadows result in the biases of building height estimation (see Fig. 12(e)).

To illustrate the performance of our method on buildings with different geometric shapes, four sites (Fig. 13) with various building geometric shapes were selected to show the accuracy of building boundary, shadow, and building height extraction. A visual analysis of Fig. 13 indicated that the building boundaries and shadows were well extracted by our method in the selected four sites. Obviously, regularly shaped buildings (Fig. 13(*c*, *d*)) were found to have better building boundaries and shadows extraction results. A quantitative comparison with the reference building boundaries and shadows suggested that regularly shaped buildings yield a boundary and shadow extraction accuracy of 5.87% and 3.02% better than that of irregularly shaped buildings (Fig. 13(*a*, *b*)) in the four sites. For building height estimation, we also found that regularly shaped buildings have a better height accuracy (0.94 m) than that irregularly shaped buildings (2.01 m). The reasons are mainly twofold. First, regularly shaped buildings usually have relatively regular shadows, the extracted shadow lengths tend to be more accurate. Second, the regularized building boundaries for irregularly shaped buildings have a relatively larger bias than that of regularly shaped buildings. These biases will undoubtedly influence the accuracy of the extracted shadow lengths.

Table 1
The accuracies of building boundaries and shadows extraction.

Task	OA (%)	P (%)	R (%)	F (%)	IOU (%)
Building boundaries	90.82	79.13	74.85	76.84	66.42
Building shadows	94.21	89.35	87.75	88.55	76.73

Table 2
The optimal parameters of the building height estimation model for each category.

Category	Building azimuth	Sample building count	K_o	b_o
1	$A_{bu} \leq 30^{\circ}$	31	1.66	4.42
2	$30^{\circ} < A_{bu} \leq 60^{\circ}$	54	1.70	0.47
3	$60^{\circ} < A_{bu} \leq 90^{\circ}$	436	1.54	−3.64
4	$90^{\circ} < A_{bu} \leq 120^{\circ}$	321	1.34	5.93
5	$120^{\circ} < A_{bu} \leq 150^{\circ}$	168	1.28	4.72
6	$A_{bu} > 150^{\circ}$	166	1.20	2.47

5. Method comparison and discussion

To comprehensively evaluate the performance of the optimized building height estimation model, we compared it with the existing shadow-based building height estimation model (Section 5.1) and two existing building height estimation studies (Section 5.2). The impacts of the building categories' number on the building height estimation results are reported in Section 5.3. At the end of this section, we reported the limitations of the proposed method and planned future work.

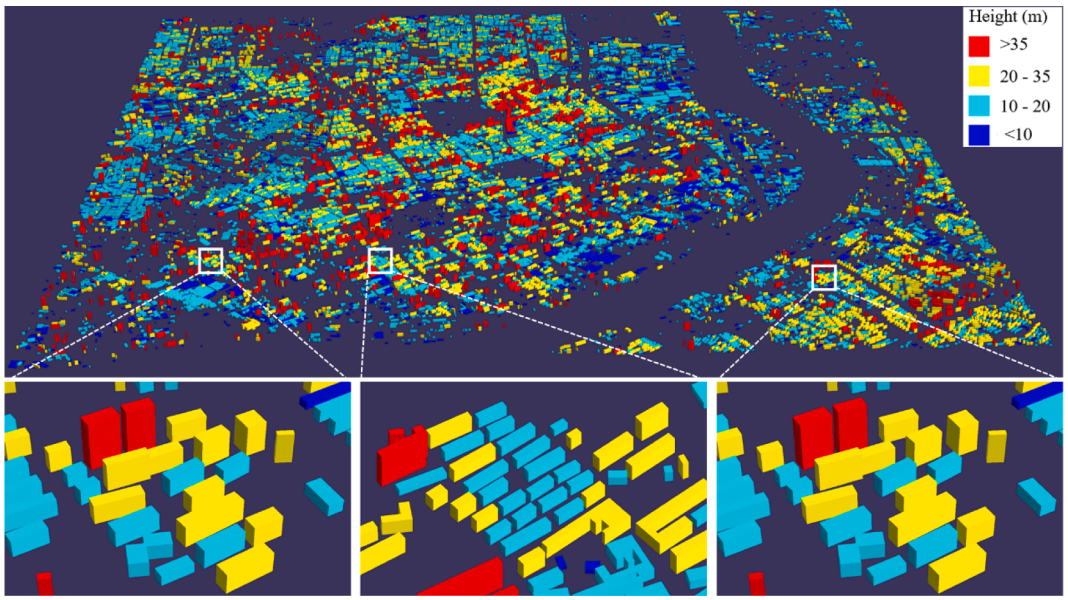


Fig. 10. A 3D view of the extracted building heights.

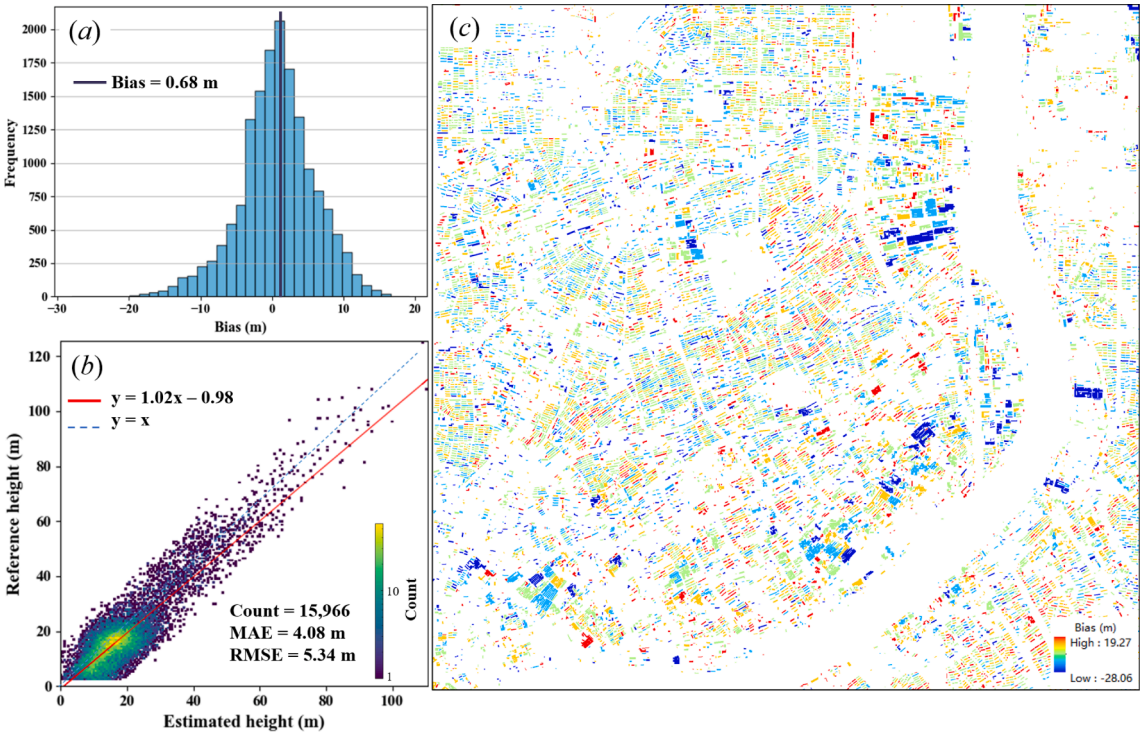


Fig. 11. The overall accuracy of the building height estimations. (*a*) the frequency histogram of the building height estimation residuals; (*b*) the scatterplot of the estimation and reference building heights; (*c*) the residual distribution of the buildings.

5.1. Comparison with the existing shadow-based algorithm

For comparison, we implemented the existing shadow-based algorithm (Liasis and Stavrou, 2016; Xie et al., 2021). The existing shadow-based algorithm uses the same K value calculated from all the sample buildings for building height estimation regardless of the variations of building azimuths. While our method divides all the buildings into different categories based on their building azimuths, and calculates the optimal K_o for each building category using an optimization approach. In this study, the K_o and b_o values for the existing shadow-based algorithm are calculated as 1.29 and 3.01, respectively using 1,176 sample buildings' heights. Table 3 lists the building height estimation results of the two algorithms.

Table 3 denotes that our model has a better overall performance than the existing model in building height estimation. This is mainly because we classified the buildings according to the building azimuths and developed the optimal model for each building azimuths category. For the existing model, due to a lack of consideration of the building azimuths, the existing model represents the global optimal solution obtained from all the sample buildings' heights instead of the optimal solution of each building azimuths category.

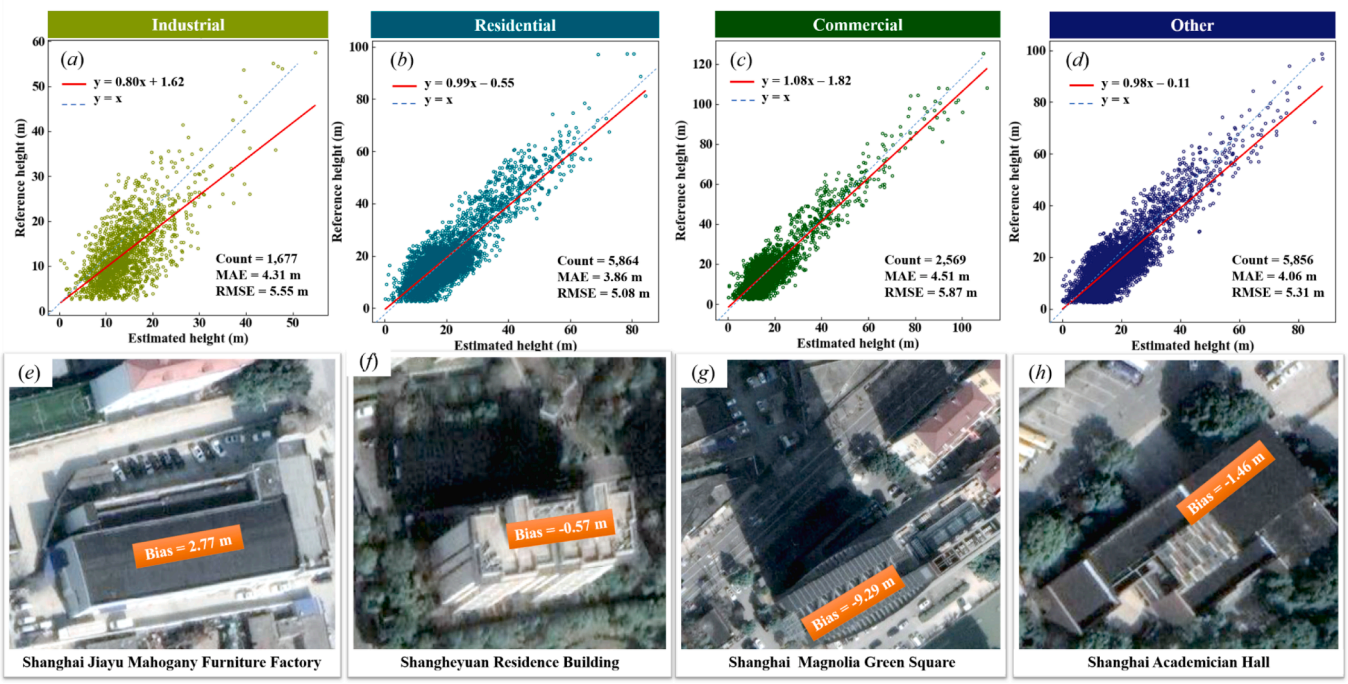


Fig. 12. The height estimation results for industrial buildings (*a*), residential buildings (*b*), commercial buildings (*c*), and other buildings (*d*). (*e*) to (*h*) display the representative buildings of the four types.

Fig. 13. The boundary and shadow extraction results of buildings with different geometric shapes.

5.2. Comparing to existing building height datasets

To further assess the performance of our method, we contrasted our results with two state-of-the-art datasets for building height estimation, including the multi-view images-based building height estimations at a 2.5 m spatial resolution (Cao and Huang, 2021) and the Sentinel-1 and Sentinel-2 radar images-based building height dataset with 10 m spatial resolution (Frantz et al., 2021). Fig. 14 displays the building height distribution maps of Cao and Huang (2021)'s and our results. As the selected dataset for comparison is in raster format, we hence rasterized our dataset into grids by using an average aggregation method (see Fig. 14(a)).

Table 3
The accuracy of building height estimation between existing and our models.

Building azimuth	Building count	Models	MAE (m)	RMSE (m)
$A_{bu} \leq 30^{\circ}$	425	Ours	2.76	3.59
		Existing	5.08	6.53
$30^{\circ} < A_{bu} \leq 60^{\circ}$	420	Ours	4.50	5.71
		Existing	4.79	6.21
$60^{\circ} < A_{bu} \leq 90^{\circ}$	5,717	Ours	4.55	5.85
		Existing	4.54	5.86
$90^{\circ} < A_{bu} \leq 120^{\circ}$	4,695	Ours	4.02	5.22
		Existing	4.64	5.75
$120^{\circ} < A_{bu} \leq 150^{\circ}$	1,861	Ours	4.15	5.37
		Existing	4.57	5.96
$A_{bu} > 150^{\circ}$	2,848	Ours	3.35	4.58
		Existing	4.84	5.93
Total	15,966	Ours	4.08	5.34
		Existing	4.64	5.88

We first implemented the M^{3}Net method proposed by Cao and Huang (2021) and generated the building height estimation dataset which has a spatial resolution of 2.5 m/pixel for the study area, as shown in Fig. 14 (b). The reference building height datasets and our estimated datasets were then rasterized into grids with the same size of 2.5 m. A total of 1,019,126 grids were acquired. As shown in Table 4, our model obtained a lower MAE and RMSE than that of the M^{3}Net model. This is mostly thanks to the precise building height provided by ICESat-2 photons and the optimized building height estimation model, which decreases the biases of building height estimation. By contrast, the M^{3}Net model holds higher uncertainties in building height estimation.

To further compare with the existing building height dataset and evaluate the transferability of our method, we applied our method to a new study area located in Hamburg, Germany, and compared our results with the 10 m-grided building height dataset proposed by Frantz et al. (2021). Fig. 15(a) shows the new study area with an area of 31.38 km^2 in Hamburg, and Fig. 15(b) displays the GES image and ATL03 photons we collected in the study area. In total, the ATL03 data covers 12 granules and 30 tracks, with a total of 155,623 medium/high confidence photons acquired. A total of 521 sample buildings with a height accuracy of MAE = 2.09 m and RMSE = 2.85 m were obtained based on the ATL03 photons. We have annotated 465 images with a size of 256 × 256 pixels from GES images to fine-tune the U-Net models. The building boundaries and building shadows in the 465 images were labeled manually with the help of the reference building height data. The reference building height data is obtained from the open LoD-1 data, which is available at https://metaver.de/startseite.

Fig. 15(c) and (d) show the building heights extracted by our method and Frantz et al. (2021), respectively. We conducted a building level comparison by assigning height from the height dataset estimated by Frantz et al. (2021). Table 5 gives a summary of the comparison results. As listed in Table 5, our method shows a better performance with a lower MAE and RMSE than that of Frantz et al. (2021) on the 9,424 buildings. This is mainly because the unstable backscatter of SAR results in the uncertainties of building height estimation, as shown in Fig. 15 (d), many non-building objects such as roads are mistaken as buildings.

5.3. Effects of the building categories' number

To analyze the impact of the number of building categories on the building height estimation, a range of building azimuth intervals (from 10 to 90 degrees) were used to categorize buildings, and the results are shown in Fig. 16.

The results suggested that the estimated building heights have a relatively low MAE and RMSE when the building azimuth interval ranges from 10 to 30 degrees (the difference of MAE is less than 10 cm). Generally, a small building azimuth interval will increase building categories, leading to a small size of sample buildings to get optimal K_o values; On the contrary, a large building azimuth interval will lead to large biases in building height estimation as the variations of the building azimuths are ignored. Therefore, in this study, we choose a building azimuth interval of 30 degrees to categorize the buildings (i.e., categorized all the buildings into six categories).

5.4. Limitations and further work

There are several limitations to our proposed approach. Firstly, the

Table 4
The accuracies of building heights estimated by Cao and Huang (2021) and our method.

Grid size	Dataset	Count	MAE (m)	RMSE (m)
2.5 m	Ours	1,019,126	4.97	5.96
	Cao and Huang (2021)	1,019,126	5.60	6.64

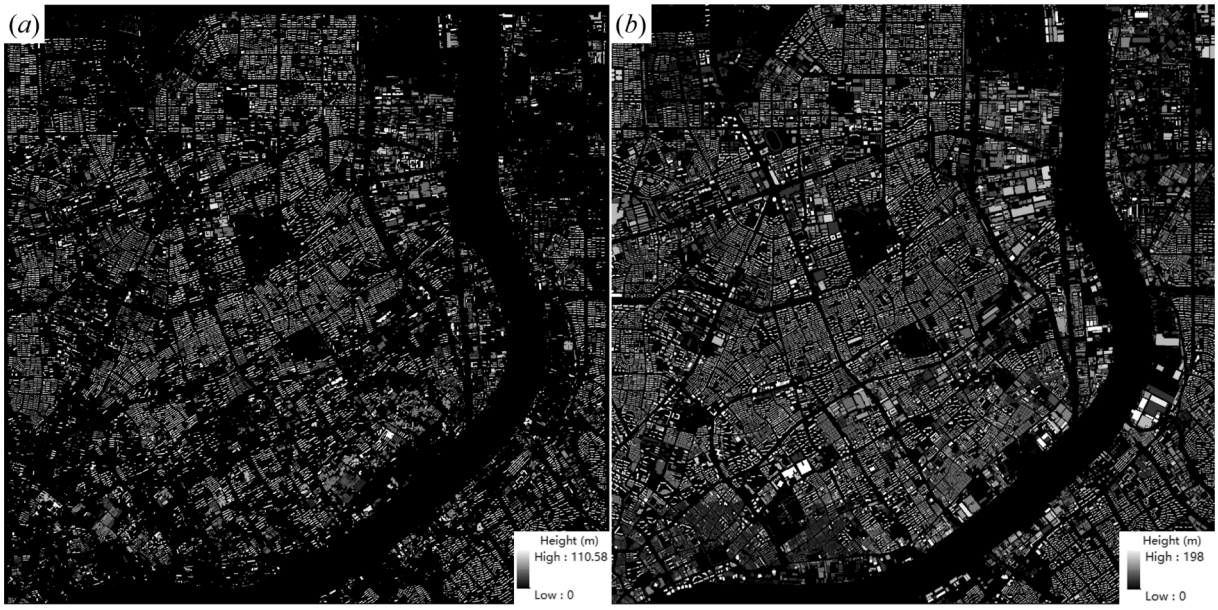


Fig. 14. The comparison of three building height datasets. (*a*) ours; (*b*) the building height dataset estimated by Cao and Huang (2021).

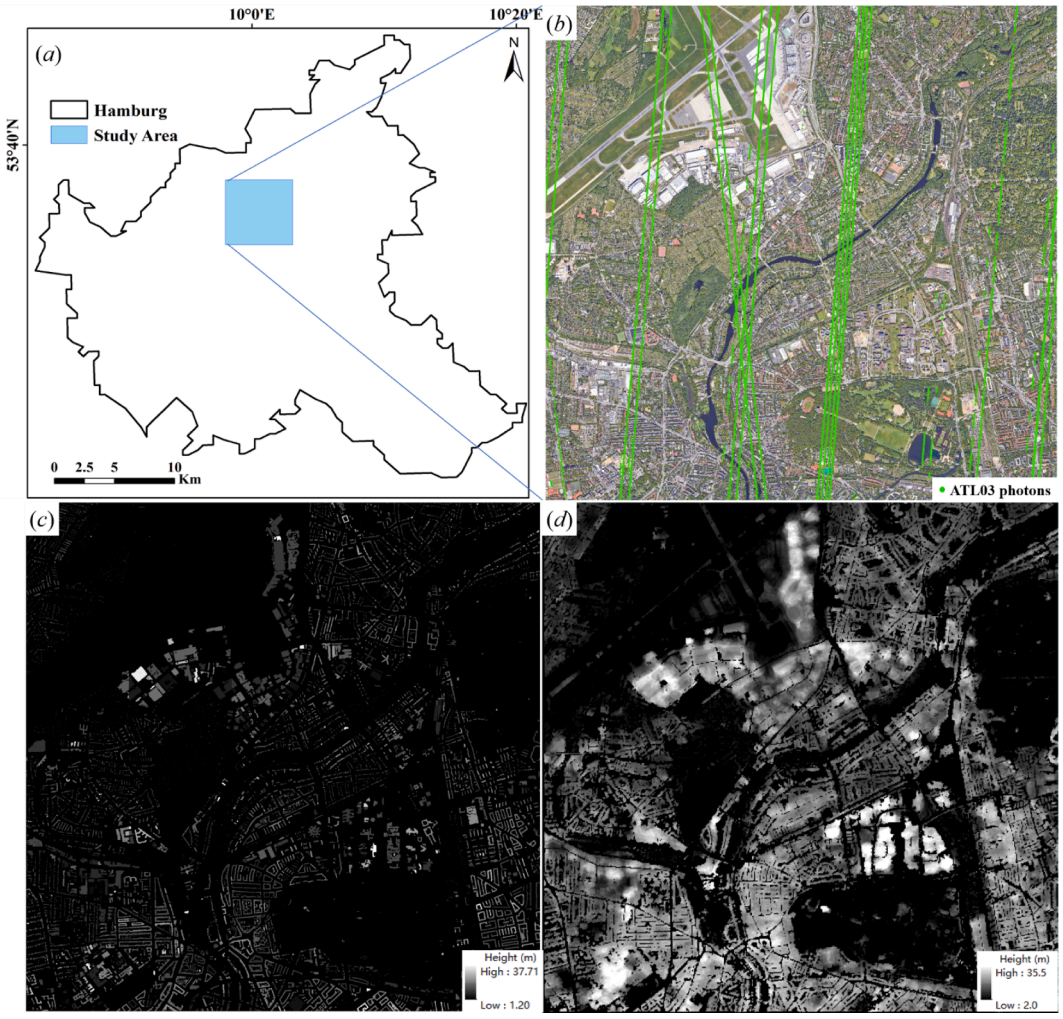


Fig. 15. (*a*) shows the study area in Hamburg; (*b*) shows the GES images of the study area; (*c*) our building height estimation results; (*d*) shows the building height dataset published by Frantz et al. (2021).

Table 5
The accuracies of building heights estimated by Frantz et al. (2021) and our method.

Dataset	Count	MAE (m)	RMSE (m)
Ours	9,424	3.87	5.11
Frantz et al. (2021)	9,424	4.37	5.54

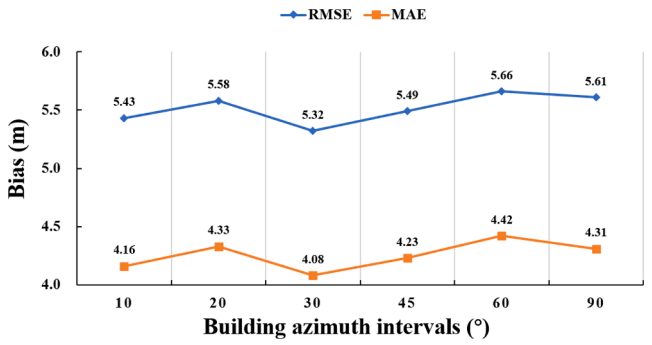


Fig. 16. The impact of building azimuth intervals on building height estimation.

shadow-based building height estimation methods rely on the shadow information of the input images. Commonly, the high resolution of images and high quality of shadows produce precise estimation results, on the contrary, the unclear shadows caused by factors such as the low resolution of images and the mutual occlusion between buildings lead to uncertainties in building height estimation results. As shown in Fig. 17, we enlarged the two representative buildings in Fig. 12 (residential building A and commercial building B) to have a better understanding of the differences. In Fig. 17, the yellow lines represent the extracted building shadow lines. As can be seen from Fig. 17(a), the building shadow of building A was fully examined by our method; while for building B in Fig. 17(b), only part of the shadow was extracted as part of the building shadow is blocked by building C. Therefore, the extracted building shadow length of building B is less reliable than that of building A. The underestimated building shadow length makes building B a large height bias. In this situation, the shadow length is typically underestimated and will lead to an underestimation of building height according to Equation (4).

Overall, the method we proposed in this study has a good performance in the areas where the buildings are regularly shaped and the distance between buildings is relatively large. As we mentioned in Section 4, our method tends to extract more accurate shadow lengths for regularly shaped buildings with relatively large building distances to avoid the problem of shadow occlusion. As shown in Fig. 18(a), the buildings are regularly shaped and located within an equal building

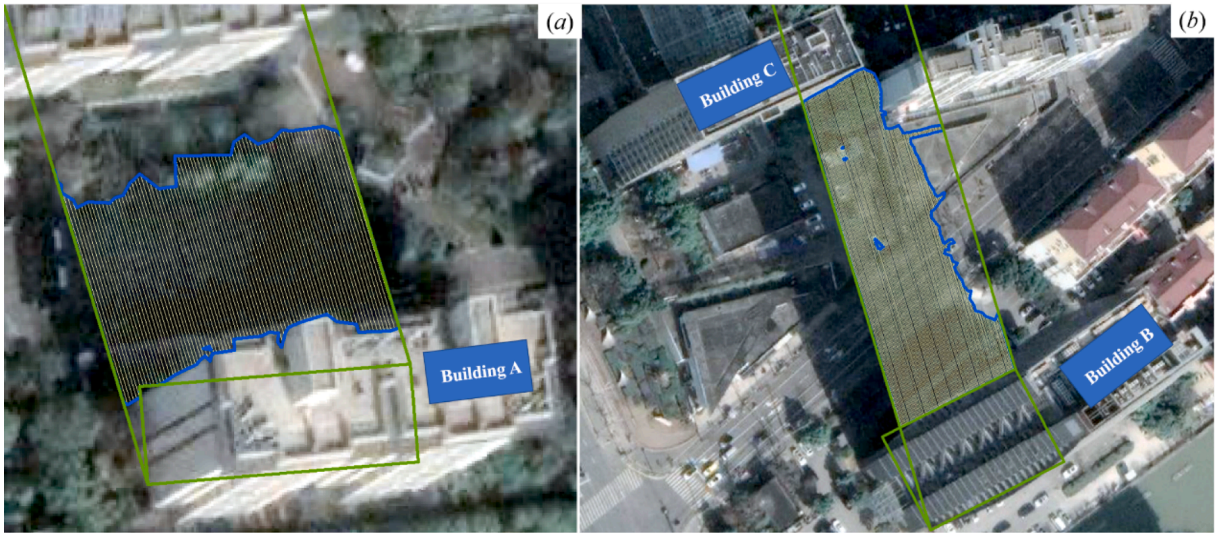


Fig. 17. (*a*) shows the scenario of the residential building; (*b*) displays the scenario of the commercial building.

distance, our method shows a good performance with an absolute height error of 1.39 m. For the region shown in Fig. 18(b), the buildings are irregularly shaped and located, our method shows a moderate performance with an absolute height error of 7.43 m. In the future, the information on building side length (vertical side length) can be incorporated to cope with the underestimation of building height. As shown in Fig. 18(b), the building side length is also a good indicator of building height (Qi et al., 2016) and thus can be combined to improve the height accuracy when the building shadows were not well extracted. Secondly, even though Google Earth provides high-resolution and large-scale satellite images openly, the image metadata such as the sun azimuth angle may not identical due to different data sources such as QucikBird and Worldview being fused, which increased the uncertainties of the shadow-based models' proportionality coefficient. Thirdly, the complex structures of building roofs will also lead to uncertainties during the building height estimation using ATL03 photons. As shown in Fig. 18(c), different building height bias are found due to the different distributions of ATL03 photons on the building roofs.

6. Conclusion

In this study, we proposed a new framework that retrieves building heights in large-scale urban areas from ICESat-2 photons and GES images. ICESat-2 has its advantages in providing high-accuracy building heights along its tracks, while GES images can provide the building and shadow information in detail. The proposed method makes use of their respective advantages to strengthen the retrieval of building heights. Moreover, we proposed an optimized algorithm to enhance the shadow-based building height estimation model by minimizing the global error across all sample buildings in each building's azimuth category. A site covering 90 km^2 located in the city center of Shanghai, China was used to test our method. The estimated heights agree with the reference building height data up to an accuracy of 4.08 m. The results are significant, given that they are achieved from a single GES image and 32 ICESat-2 tracks. Therefore, our method is effective and comprehensive for building height estimation at a large scale and the reported results are more dependable in contrast with existing shadow-based algorithms. Besides, the comparisons with two existing building height datasets also indicate that our method has a better performance. However, the building heights estimated by the proposed method are still biased in retrieving the buildings that are located in high-density areas and the areas where the shadows of the buildings are occluded by other objects. In future work, the method can be further improved by employing other environmental factors, and thus building-level height datasets using GES images and ICESat-2 data on regional or global scales can be produced. It should be noticed that the suggested method may be modified and optimized in different geographic regions. Therefore, adaptation techniques for improving the transferability of our method across different

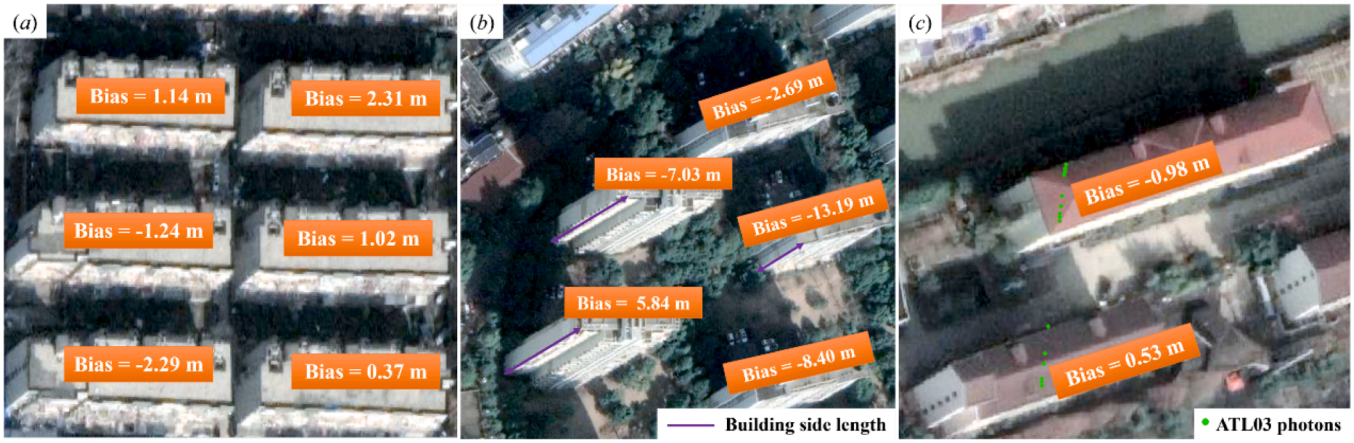


Fig. 18. (*a*) and (*b*) shows the building height estimation results for different building arrangements; (*c*) displays the bias of ATL03 photons-measured building height.

geographic regions are expected.

CRediT authorship contribution statement

Yi Zhao: Validation, Software, Data curation, Writing – original draft. **Bin Wu:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Qiaoxuan Li:** Validation, Resources, Data curation. **Lei Yang:** Data curation, Formal analysis. **Hongchao Fan:** Writing – review & editing. **Jianping Wu:** Software, Resources, Supervision. **Bailang Yu:** Conceptualization, Writing – review & editing, Supervision.

Declaration of Competing 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.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 42001357 and No. 41871331), the Open Fund of Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University (Grant No. KLGIS2022A04), the China Scholarship Council (Grant No. 202105530007) and the Norwegian-Chinese Government Scholarship Program 2021/2022 (No. NGSC-CN-2021/00007) and Diku-Norwegian Agency for International Cooperation and Quality Enhancement in Higher Education.

References

Arefi, H., et al., 2011. Iterative approach for efficient digital terrain model production from CARTOSAT-1 stereo images. J. Appl. Remote Sens. 5, 053527 https://doi.org/10.1117/1.3595265.

Baltsavias, E.P., 1999. A comparison between photogrammetry and laser scanning. ISPRS J. Photogramm. Remote Sens. 54, 83–94. https://doi.org/10.1016/S0924-2716(99)00014-3.

Cao, Y., Huang, X., 2021. 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. 264, 112590 https://doi.org/10.1016/j.rse.2021.112590.

Chen, H., et al., 2021. A new method for building-level population estimation by integrating LiDAR, nighttime light, and POI data. J. Remote Sens. 2021, 9803796. https://doi.org/10.34133/2021/9803796.

Dandabathula, G., et al., 2021. Retrieval of building heights from ICESat-2 photon data and evaluation with field measurements. Environ. Res.: Infrastruct. Sustain. 1, 011003 https://doi.org/10.1088/2634-4505/abf820.

Elbakary, M.I., Iftekharuddin, K.M., 2013. Shadow detection of man-made buildings in high-resolution panchromatic satellite images. IEEE Trans. Geosci. Remote Sens. 52, 5374–5386. https://doi.org/10.1109/TGRS.2013.2288500.

Esch, T., et al., 2022. World Settlement Footprint 3D - A first three-dimensional survey of the global building stock. Remote Sens. Environ. 270, 112877 https://doi.org/10.1016/j.rse.2021.112877.

Falk, T., et al., 2019. U-Net: deep learning for cell counting, detection, and morphometry. Nat. Methods 16, 67–70. https://doi.org/10.1038/s41592-018-0261-2.

Fan, H., et al., 2021. An interactive platform for low-cost 3D building modeling from VGI data using convolutional neural network. Big Earth Data 5, 49–65. https://doi.org/10.1080/20964471.2021.1886391.

Frantz, D., et al., 2021. National-scale mapping of building height using Sentinel-1 and Sentinel-2 time series. Remote Sens. Environ. 252, 112128 https://doi.org/10.1016/j.rse.2020.112128.

Gong, P., et al., 2020. Mapping essential urban land use categories in China (EULUC-China): preliminary results for 2018. Sci. Bull. 65, 182–187. https://doi.org/10.1016/j.scib.2019.12.007.

Gong, P., et al., 2010. ICESat GLAS data for urban environment monitoring. IEEE Trans. Geosci. Remote Sens. 49, 1158–1172. https://doi.org/10.1109/TGRS.2010.2070514.

Güneralp, B., et al., 2017. Global scenarios of urban density and its impacts on building energy use through 2050. Proc. Natl. Acad. Sci. 114, 8945–8950. https://doi.org/10.1073/pnas.1606035114.

Haala, N., Kada, M., 2010. An update on automatic 3D building reconstruction. ISPRS J. Photogramm. Remote Sens. 65, 570–580. https://doi.org/10.1016/j.isprsjprs.2010.09.006.

Huang, X., Wang, Y., 2019. Investigating the effects of 3D urban morphology on the surface urban heat island effect in urban functional zones by using high-resolution remote sensing data: A case study of Wuhan, Central China. ISPRS J. Photogramm. Remote Sens. 152, 119–131. https://doi.org/10.1016/j.isprsjprs.2019.04.010.

Irvin, R.B., McKeown, D.M., 1989. Methods for exploiting the relationship between buildings and their shadows in aerial imagery. IEEE Trans. Syst. Man. Cybernet 1076, 156–164. https://doi.org/10.1109/21.44071.

Izadi, M., Saeedi, P., 2011. Three-dimensional polygonal building model estimation from single satellite images. IEEE Trans. Geosci. Remote Sens. 50, 2254–2272. https://doi.org/10.1109/TGRS.2011.2172995.

Jung, F., 2004. Detecting building changes from multitemporal aerial stereopairs. ISPRS J. Photogramm. Remote Sens. 58, 187–201. https://doi.org/10.1016/j.isprsjprs.2003.09.005.

Langford, M., et al., 2008. Urban population distribution models and service accessibility estimation. Comput. Environ. Urban Syst. 32, 66–80. https://doi.org/10.1016/j.compenvurbsys.2007.06.001.

Lao, J., et al., 2021. Retrieving building height in urban areas using ICESat-2 photon-counting LiDAR data. Int. J. Appl. Earth Obs. Geoinf. 104, 102596 https://doi.org/10.1016/j.jag.2021.102596.

Li, M., et al., 2020a. Continental-scale mapping and analysis of 3D building structure. Remote Sens. Environ. 245, 111859 https://doi.org/10.1016/j.rse.2020.111859.

Li, M., et al., 2022. Global maps of 3D built-up patterns for urban morphological analysis. Int. J. Appl. Earth Obs. Geoinf. 114, 103048 https://doi.org/10.1016/j.jag.2022.103048.

Li, X., et al., 2020b. Developing a method to estimate building height from Sentinel-1 data. Remote Sens. Environ. 240, 111705 https://doi.org/10.1016/j.rse.2020.111705.

Liasis, G., Stavrou, S., 2016. Satellite images analysis for shadow detection and building height estimation. ISPRS J. Photogramm. Remote Sens. 119, 437–450. https://doi.org/10.1016/j.isprsjprs.2016.07.006.

Luo, S., et al., 2020. Deeply supervised convolutional neural network for shadow detection based on a novel aerial shadow imagery dataset. ISPRS J. Photogramm. Remote Sens. 167, 443–457. https://doi.org/10.1016/j.isprsjprs.2020.07.016.

Nagao, M., et al., 1979. Region extraction and shape analysis in aerial photographs. Comput.Gr.Image Process. 10, 195–223. https://doi.org/10.1016/0146-664X(79)90001-7.

Neuenschwander, A.L., Magruder, L.A., 2019. Canopy and terrain height retrievals with ICESat-2: A first look. Remote Sens. 11, 1721. https://doi.org/10.3390/rs11141721.

Pérez-Lombard, L., et al., 2008. A review on buildings energy consumption information. Energy Build. 40, 394–398. https://doi.org/10.1016/j.enbuild.2007.03.007.

Qi, F., et al., 2016. Building height estimation using Google Earth. Energy Build. 118, 123–132. https://doi.org/10.1016/j.enbuild.2016.02.044.

Raju, P.L.N., et al., 2014. Shadow analysis technique for extraction of building height using high resolution satellite single image and accuracy assessment. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XL-8, 1185–1192. https://doi.org/10.5194/isprsarchives-XL-8-1185-2014.

Ren, C., et al., 2020. Developing a rapid method for 3-dimensional urban morphology extraction using open-source data. Sustain. Cities Soc. 53, 101962 https://doi.org/10.1016/j.scs.2019.101962.

Ronneberger, O., et al., 2015. U-net: Convolutional networks for biomedical image segmentation. In: Proc. Int. Conf. Medical Image Comput. Comput.-Assisted Intervention. Springer, pp. 234–241.

Shao, Y., et al., 2011. Shadow detection and building-height estimation using IKONOS data. Int. J. Remote Sens. 32, 6929–6944. https://doi.org/10.1080/01431161.2010.517226.

Shettigara, V.K., Sumerling, G.M., 1998. Height determination of extended objects using shadows in SPOT images. Photogramm. Eng. Remote Sens. 64, 35–44.

Soergel, U., et al., 2009. Stereo analysis of high-resolution SAR images for building height estimation in cases of orthogonal aspect directions. ISPRS J. Photogramm. Remote Sens. 64, 490–500. https://doi.org/10.1016/j.isprsjprs.2008.10.007.

Sun, Y., et al., 2022. Large-scale building height retrieval from single SAR imagery based on bounding box regression networks. ISPRS J. Photogramm. Remote Sens. 184, 79–95. https://doi.org/10.1016/j.isprsjprs.2021.11.024.

Torres, R., et al., 2012. GMES Sentinel-1 mission. Remote Sens. Environ. 120, 9–24. https://doi.org/10.1016/j.rse.2011.05.028.

Wegner, J.D., et al., 2013. Combining high-resolution optical and InSAR features for height estimation of buildings with flat roofs. IEEE Trans. Geosci. Remote Sens. 52, 5840–5854. https://doi.org/10.1109/TGRS.2013.2293513.

Wu, B., et al., 2022. Automatic building rooftop extraction using a digital surface model derived from aerial stereo images. J. Spatial Sci. 67, 21–40. https://doi.org/10.1080/14498596.2020.1720836.

Wu, B., et al., 2023. A building volume adjusted nighttime light index for characterizing the relationship between urban population and nighttime light intensity. Comput. Environ. Urban Syst. 99, 101911 https://doi.org/10.1016/j.compenvurbsys.2022.101911.

Wu, B., et al., 2021. Mapping fine-scale visual quality distribution inside urban streets using mobile LiDAR data. Build. Environ. 206, 108323 https://doi.org/10.1016/j.buildenv.2021.108323.

Wu, B., et al., 2018. An Extended Minimum Spanning Tree method for characterizing local urban patterns. Int. J. Geog. Inf. Sci. 32, 450–475. https://doi.org/10.1080/13658816.2017.1384830.

Wu, B., et al., 2017. A graph-based approach for 3D building model reconstruction from airborne LiDAR point clouds. Remote Sens. 9, 92. https://doi.org/10.3390/rs9010092.

Xie, Y., et al., 2021. Multi-Scene building height estimation method based on shadow in high resolution imagery. Remote Sens. 13, 2862. https://doi.org/10.3390/rs13152862.

Xu, Y., et al., 2017. Urban morphology detection and computation for urban climate research. Landsc. Urban Plan. 167, 212–224. https://doi.org/10.1016/j.landurbplan.2017.06.018.

Yu, B., et al., 2010. Automated derivation of urban building density information using airborne LiDAR data and object-based method. Landsc. Urban Plan. 98, 210–219. https://doi.org/10.1016/j.landurbplan.2010.08.004.

Yu, S., et al., 2020. Exploring the relationship between 2D/3D landscape pattern and land surface temperature based on explainable eXtreme Gradient Boosting tree: A case study of Shanghai. China. Sci. Total Environ. 725, 138229 https://doi.org/10.1016/j.scitotenv.2020.138229.

Yu, S., et al., 2016. View-based greenery: A three-dimensional assessment of city buildings' green visibility using Floor Green View Index. Landsc. Urban Plan. 152, 13–26. https://doi.org/10.1016/j.landurbplan.2016.04.004.

Zhang, W., et al., 2016. An easy-to-use airborne LiDAR data filtering method based on cloth simulation. Remote Sens. 8, 501. https://doi.org/10.3390/rs8060501.

Zhang, Z., et al., 2018. Road extraction by deep residual U-Net. IEEE Geosci. Remote Sens. Lett. 15, 749–753. https://doi.org/10.1109/LGRS.2018.2802944.

Zhao, C., et al., 2020a. Characterizing the 3-D urban morphology transformation to understand urban-form dynamics: A case study of Austin, Texas, USA. Landsc. Urban Plan. 203, 103881 https://doi.org/10.1016/j.landurbplan.2020.103881.

Zhao, Y., et al., 2021. Evaluation of ICESat-2 ATL03/08 surface heights in urban environments using airborne LiDAR point cloud data. IEEE Geosci. Remote Sens. Lett. 19, 1–5. https://doi.org/10.1109/LGRS.2021.3127540.

Zhao, Y., et al., 2020b. Mapping 3D visibility in an urban street environment from mobile LiDAR point clouds. GISci. Remote Sens. 57, 797–812. https://doi.org/10.1080/15481603.2020.1804248.