

Article

Automated Estimation of Building Heights with ICESat-2 and GEDI LiDAR Altimeter and Building Footprints: The Case of New York City and Los Angeles

Yunus Kaya 

Department of Geomatics Engineering, Harran University, Şanlıurfa 63050, Türkiye; yunuskaya@harran.edu.tr

Abstract: Accurate estimation of building height is crucial for urban aesthetics and urban planning as it enables an accurate calculation of the shadow period, the effective management of urban energy consumption, and thorough investigation of regional climatic patterns and human-environment interactions. Although three-dimensional (3D) cadastral data, ground measurements (total station, Global Positioning System (GPS), ground laser scanning) and air-based (such as Unmanned Aerial Vehicle—UAV) measurement methods are used to determine building heights, more comprehensive and advanced techniques need to be used in large-scale studies, such as in cities or countries. Although satellite-based altimetry data, such as Ice, Cloud and land Elevation Satellite (ICESat-2) and Global Ecosystem Dynamics Investigation (GEDI), provide important information on building heights due to their high vertical accuracy, it is often difficult to distinguish between building photons and other objects. To overcome this challenge, a self-adaptive method with minimal data is proposed. Using building photons from ICESat-2 and GEDI data and building footprints from the New York City (NYC) and Los Angeles (LA) open data platform, the heights of 50,654 buildings in NYC and 84,045 buildings in LA were estimated. As a result of the study, root mean square error (RMSE) 8.28 m and mean absolute error (MAE) 6.24 m were obtained for NYC. In addition, 46% of the buildings had an RMSE of less than 5 m and 7% less than 1 m. In LA data, the RMSE and MAE were 6.42 m and 4.66 m, respectively. It was less than 5 m in 67% of the buildings and less than 1 m in 7%. However, ICESat-2 data had a better RMSE than GEDI data. Nevertheless, combining the two data provided the advantage of detecting more building heights. This study highlights the importance of using minimum data for determining urban-scale building heights. Moreover, continuous monitoring of urban alterations using satellite altimetry data would provide more effective energy consumption assessment and management.

Keywords: building height estimation; GEDI; ICESat-2; building footprint; New York City; Los Angeles; RMSE; MAE



Citation: Kaya, Y. Automated Estimation of Building Heights with ICESat-2 and GEDI LiDAR Altimeter and Building Footprints: The Case of New York City and Los Angeles. *Buildings* **2024**, *14*, 3571. <https://doi.org/10.3390/buildings14113571>

Academic Editors: Li Yang and Haifeng Liao

Received: 8 October 2024

Revised: 1 November 2024

Accepted: 7 November 2024

Published: 9 November 2024



Copyright: © 2024 by the author. 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

Urbanization is one of the most important human impacts on the planet [1]. Access to education and health services, easy access to infrastructure and technology, and cultural and social opportunities are among the factors that encourage people to live in cities, and 68% of the world's population is expected to live in cities by 2050 [2,3]. The rapid increase in the urban population causes people's living spaces to change and cities to become more crowded and complex. Rapid urban development and the increasing population require a continuous supply of energy. Energy is an indispensable resource for building a modern society [4]. Many aspects of human society, such as transportation, building conditioning, and manufacturing, require energy, either directly or indirectly [5]. Non-agricultural areas such as the building industry and factories have the greatest impact on energy consumption [6]. Building morphology also directly affects energy use. Building height, building shape, building surface area, and distance between buildings directly or

indirectly affect energy consumption in buildings [7,8]. It is seen that two-dimensional (2D) building features (such as building footprint, building type, built year, floor area ratio) do not provide sufficient information about energy use [9]. Therefore, determining three-dimensional (3D) building features (such as building height, building volume) is crucial.

In recent years, there have been numerous applications involving the footprints of human settlements in urban areas, and many high-resolution products have been obtained [10]. Nevertheless, the accurate measurement of building height is important to understand the impacts of urbanization on the urban environment [11]. Dense and tall buildings resulting from increasing urbanization are responsible for 39% of global energy-related carbon emissions [2,12–14]. Accurately estimating building energy consumption is the best way to reduce greenhouse gas emissions. Knowing the physical properties of a building is of great practical importance to model building energy consumption [15,16]. Building heights are also associated with urban energy use [17] and human well-being [18]. Building height is also useful in analyzing earthquake hazards [19]. Building height is one of the important parameters used to convert a 2D footprint into a 3D model [20]. Although building heights are important in managing urban planning and urban energy consumption, automatically determining building heights is still a challenging task.

Currently, various data sources and methods are used to estimate building heights. For example, 3D models provided by open data platforms such as 3D CityGML (3D City Geography Markup Language), CityJSON, and OpenStreetMap provide information about building heights [12,21,22]. In addition, 3D cadastral data and population-based calculations are used to estimate building heights [22]. However, determining building height from cadastral data may not be possible in every situation or for every region. For example, when it comes to energy demand and need, city, country, or world-wide analyses need to be conducted when necessary. Fortunately, similar records are kept regularly in all cities and countries. On the other hand, updating the records can be difficult, depending on the situation. Therefore, in addition to these open-source data, satellite-based data are also frequently used in building height estimation [20].

Advances in remote sensing technology are facilitating the automatic estimation of building heights through the increased availability of remote sensing data. The most used data for determining building heights include free high-resolution platforms (such as Google Earth), remote sensing data, global Digital Elevation Models (DEMs)/Digital Surface Models (DSMs), and aerial LiDAR data. Kadhim and Mourshed [20] developed an automatic estimation model of building heights from shadows using very-high-resolution (VHR) multispectral pan-sharpened satellite imagery. Liu et al. [23] estimated building heights with shadow correction from Gaofen-7 images. Lee and Kim [24] used building footprints and shadow lengths to determine building heights from single image data. They determined building heights with RMSE twice the ground sampling distance (GSD) of the image. Li et al. [23] proposed a Region Based Convolutional Neural Networks (RCNN) algorithm based on the geometric relationship between building and building shadow. The results show a 95% accuracy rate. Zhang et al. [25] proposed an algorithm that uses only visible spectral bands to determine the shadow size of buildings. They tested the method in complex urban areas in Toronto and Beijing and found the building detection rate to be over 95%. Zhou and Myint [26] proposed a shadow pattern classification system (ShadowClass) to summarize various shadow shapes of buildings into pattern categories. Wang et al. [27] calculated the building heights in the Pudong district of Shanghai, China, with the formula derived from the relationship between the sun, satellites, and buildings and obtained an average absolute error of 4.45 m. Qi et al. [28] calculated building heights from Google Earth images. They obtained 0.98 m RMSE by using the angles in Google Earth images of 21 different buildings. Garzelli [29] estimated building heights using the full range of multi-angle spatial and spectral information provided by the WorldView-2 system. They estimated building heights with an error of about 3 m using the nadir image of building tops and images obtained from four different angles. Li et al. [30] used a deep

neural network called RoofNet to detect building heights using street scene data. In the study, 92.8% of the buildings were estimated with 4 m RMSE and below.

However, the main disadvantage of calculating a building's height from the shadow is that the time of image capture and the angle of the sun are constantly changing. Therefore, a separate optimization may be required for each image or each building. In addition to shadow-based methods, there are also single or multi-image building estimation models. Kim et al. [31] proposed a semi-automatic method that extracts 3D building information from a single satellite image. In the approach, they calculated building heights using building shadows and estimated building heights with an RMSE of 1.66 m. Sun et al. [32] determined building heights from a single SAR image and obtained an RMSE of 3.51 m. Li et al. [33] developed a VVH indicator that combines dual polarization information (i.e., VV and VH) from Sentinel-1 data and applied it in seven different cities in the USA. As a result of the study, they obtained 1.5 m RMSE between Sentinel-1 data and ICESat data. Chen et al. [34] developed an algorithm using Gaofen-7 imagery to determine the height of more than 700,000 buildings. The RMSE between the algorithm results and LiDAR reference data ranged from 4.06 m to 8.06 m. Huang et al. [35] calculated building heights for the whole of China using ALOS World 3D data. They estimated building heights with an RMSE of 4.98 m when compared to ground truth data. There are also studies that combine different data to estimate building heights. Wu et al. [11] used radar and optical data to estimate the height of 2020 buildings in China. In the study, an RMSE of 6.1 m and a correlation of 0.77 were obtained. Zhang et al. [36] estimated building height by calculating building roofs from GF-7 stereo image and ground elevations from Digital Surface Model (DSM) and obtained 2.31 m MAE between aerial LiDAR. Bshouty and Dalyot [37] calculated building heights by combining OpenStreetMap vectors with photos from smartphones.

Although ICESat-2 is designed to monitor changes in sea ice [38], it is also used in areas such as terrain modeling [23,39,40] and estimating urban building heights [12,41,42]. Zhao et al. [43] combined Google Earth imagery with ICESat-2 photons to estimate the height of approximately 16,000 buildings with an MAE of 4.08 m. Yang et al. [44] estimated the height of 430 buildings, with a margin of error of 13.2%, by combining Quickbird images and ICESat GLAS data. Wu et al. [11] used ICESat-2 ATL03 photons and high-resolution remote sensing imagery to estimate building heights at the urban scale. The results of their proposed method show that they can estimate building heights with an MAE of 4.7 m. Lao et al. [42] classified ATL03 photons with the random sample consensus (RANSAC) algorithm to determine building heights from ICESat-2 data and obtained RMSEs ranging from 0.35 m to 0.45 m with reference heights. Cai et al. [12] determined the heights of 17,399 buildings in NYC using ICESat-2 LiDAR altimetry, building footprints and landcover raster data. In the study, building heights were estimated with 8.1 m RMSE and 3 m MAE.

GEDI is a LiDAR full waveform system mounted on the International Space Station (ISS). GEDI collects data between 51.6° north and south in the WGS84 coordinate system [45]. GEDI data provide important information about the land surface and have been used in many studies to determine canopy and forest heights [46,47]. However, interestingly, GEDI data have been used to determine building heights in a limited number of studies [48].

ICESat-2 and GEDI LiDAR data provide higher vertical precision than other satellite altimetry systems. In addition, since both satellites follow different patterns and visit the same region at different times, it helps to expand the represented coverage. ICESat-2 altimetry data can be obtained in .h5 or .csv formats. GEDI data can only be obtained in a .h5 format. In addition, the fact that the data contain multi-layered parameters facilitates pre-processing operations.

In the literature, studies using ICESat-2 and GEDI data usually require more additional data. For example, Zhao et al. [43] used Google Earth data for building heights. Similarly, Wu et al. utilized high-resolution remote sensing imagery in addition to ICESat-2 data. Cai et al. [12] determined building heights using global DEM, building footprints and landcover data. Unfortunately, it is not always possible to access various data sources. The

motivation behind this study is to determine building heights and create base data for applications such as energy demand, energy consumption, urban planning, and population planning and management in large-scale areas, using the freely available ICESat-2 and GEDI LiDAR altimeters. This paper presents a method to determine building heights using only airborne LiDAR data and building footprints. ICESat-2 and GEDI data were used together to calculate further building heights. Using building footprints in NYC and LA and ICESat-2 and GEDI data, photons falling on buildings and photons falling outside buildings were separated from each other. A distance algorithm was then developed to determine the nearest ground elevation for each building. Two contributions were made to this study. First, it was determined which methodology should be followed when determining photons falling on top of buildings in urban areas using ICESat-2 and GEDI photons. Second, it was determined how to obtain ground elevation values of buildings with minimal data. In the rest of the paper, the study areas are presented in Section 2, and the data and methodology used in this study are presented. Then, an analysis and discussion of the results are given in Section 3. Finally, the conclusions are presented in Section 4.

2. Materials and Methods

2.1. Study Area

NYC is one of the largest cities on the east coast of the United States, known as a global metropolis [49]. With a population of 8,258,000 as of 1 July 2023, NYC is the most populous city in the United States [50]. The city is a global center of influence in many sectors, such as finance, media, fashion, and technology, and has proven its economic strength with a gross regional product of USD 1514 billion in 2021 [12].

NYC covers an area of 783.8 km² and consists of five boroughs: Manhattan, Brooklyn, Queens, Bronx, and Staten Island. Each of these boroughs has its own distinct cultural and architectural character. For example, Manhattan is famous for its world-renowned high-rise buildings and dense business district, while Brooklyn is noted for its historic residences and artistic communities. By 2024, there were 1,083,057 buildings in the city [51]. These buildings are of different heights and shapes.

LA is one of the largest and most important cities in the world, located on the west coast of the United States. With over 3,898,747 inhabitants in 2024, it is the second most populous city in the United States [52]. In 2022, it had the second largest economy in the US, with a gross regional product of USD 1227 billion [53].

LA covers an area of 1302.76 km² and consists of several districts: Downtown, Hollywood, Westside, San Fernando Valley, and South Los Angeles. Each district has its own character, architecture and urban landscape. From the famous movie studios and iconic walkways of Hollywood to the expansive beaches of Santa Monica and the historic buildings of Pasadena, the city's diversity is reflected in its neighborhoods. By 2024, there are 1,122,422 buildings in LA, covering nearly 70% of the total area [54].

Most of the buildings used in this study are residential. However, there are also business centers and official institutions among these buildings.

Figure 1 shows the location of the study area covered by multiple beams of ICESat-2 and GEDI ground tracks.

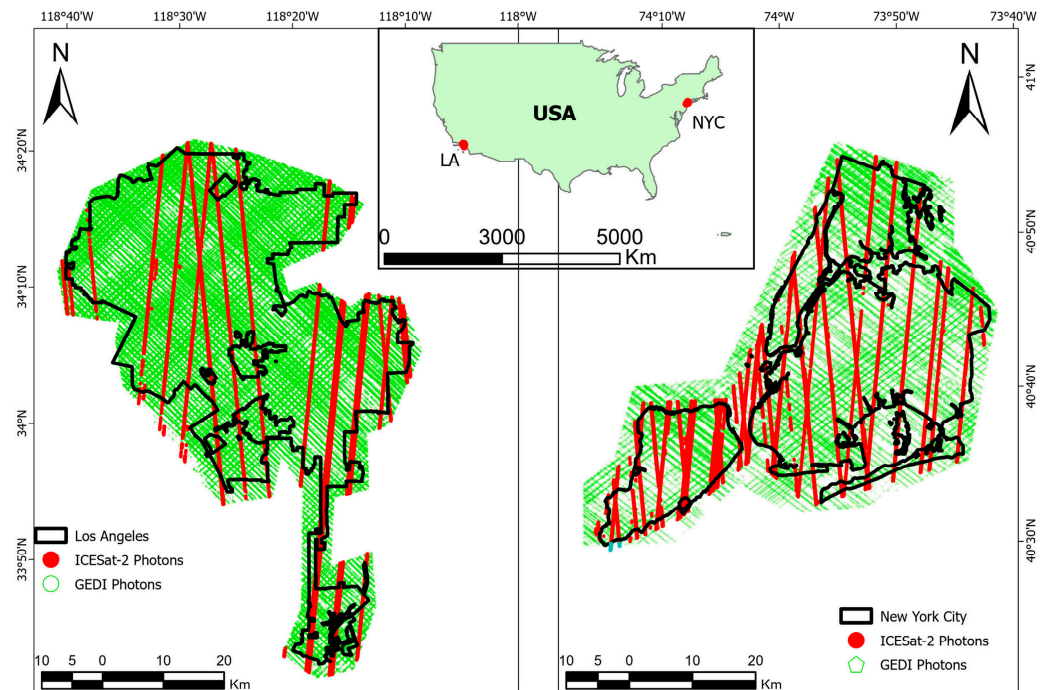


Figure 1. Study area map. The red line represents the path of ICESat-2. The green grid represents the path of GEDI.

2.2. Data

2.2.1. ICESat-2/ATLAS

ICESat was the first satellite with a laser altimeter and has a temporal resolution of 91 days [55]. The Geoscience Laser Altimeter System (GLAS, Brisbane, Australia) instrument was operational on the satellite from 2003 until 2009. During its 7 years of operation, it collected data only in February/March, May/June, and October/November [56]. The ICESat-2 (Ice, Cloud, and Land Elevation Satellite 2) is a satellite mission launched by NASA (National Aeronautics and Space Administration) in September 2018 [57], equipped with the Advanced Terrain Laser Altimeter System (ATLAS, Nacka Municipality, Sweden) as its primary instrument. The ATLAS instrument uses laser pulses to determine the ellipsoidal heights of glaciers, seas, inland waters, and land topography with unprecedented precision. The data products of ICESat-2/ATLAS are divided into layers containing a total of 21 standard data products ranging from ATL00 to ATL21 [57,58]. The ATL03 package provides many parameters for each photon, including longitude, latitude, ellipsoidal height, time, confidence labels, and height adjustment information. In this study, the ATL03/V006 product was used to estimate the heights of buildings in NYC and LA.

2.2.2. GEDI

GEDI is operated by NASA and produces high-resolution LiDAR observations of the Earth's vertical structure [45]. Deployed in 2018 to the International Space Station, GEDI collects elevation data from the Earth with a footprint accuracy of approximately 10 m between 51.6° South and 51.6° North latitudes. In this study, GEDIV002 L2A products were used in addition to ICESat-2 data to calculate building heights in NYC and LA.

2.2.3. Building Footprints NYC

NYC's building footprints consist of geometric data indicating the boundaries and locations of all buildings in the city. It is a project led by the New York City Department of Information Technology and Telecommunications (DoITT, New York, NY, USA), the Department of Buildings (DOB), and the Department of City Planning (DCP) to produce NYC's architectural footprints. The dataset contains various information such as building

construction date, building condition, and building height. The data are freely accessible at <https://data.cityofnewyork.us/Housing-Development/Building-Footprints/nqwf-w8eh> (accessed on 24 June 2024). The data are available in different formats such as shapefile, GeoJSON, and csv depending on the intended use. In this study, footprint data were obtained in shapefile format and used in ArcGIS Pro environment to associate with altimeter data.

2.2.4. Building Footprints LA

Building footprints in LA consist of geometric data specifying the boundaries and locations of all buildings in the city. The data are managed and updated by the Los Angeles Department of City Planning (DCP) and is used for urban planning, emergency management, urban analysis and many other applications. The dataset contains polygon georeferenced geometry of each building and contains additional information such as building age, building height, type of use, etc. The data are freely accessible at https://geohub.lacity.org/datasets/813fcefde1f64b209103107b26a8909f_0/explore (accessed on 24 June 2024). The data is available in different formats such as shapefile, kml, GeoJSON, csv depending on the intended use. In this study, footprint data were provided in shapefile format and associated with altimeter data.

2.3. Methodology

The methodology of this study involves the estimation of building height using ICESat-2 ATL03 and GEDI Level2A data and building footprints. The methodology includes 4 main steps: Pre-noise removal for ICESat-2 and GEDI products, separation of ground and building photons, determination of building height by identifying the appropriate photons for building and ground, and accuracy assessment of height estimates. The methodology of this study is shown in Figure 2.

2.3.1. Preliminary Denoising of ICESat-2

The ATL03 V006 dataset was used in this study. ATL03 data were downloaded in .h5 (HDF) format. The ATL03 package provides two different confidence parameters for each photon: photon quality (quality_ph) and photon signal confidence (signal_conf_ph) [44]. The signal_conf_ph provides 7 different confidence levels (−2, −1, 0, 1, 2, 3, 4) for 5 different surfaces (land, ocean, sea ice, land ice, and inland water). Here, −2 and −1 represent photons not related to these surfaces, while 0 and 1 represent noise. In addition, 2, 3 and 4 represent low, medium and high confidence signal, respectively. The “quality_ph” takes 4 different values (0, 1, 2, and 3) representing nominal, possible afterpulse, possible impulse response effect, and possible tep, respectively.

In this study, photons labeled quality_ph = 0 and signal_conf_ph = 4 reflected from the land surface were preferred. In this way, the photons with the highest confidence for building heights were used (Figure 2, upper right corner).

2.3.2. Preliminary Denoising of GEDI

The GEDI L2A dataset downloaded in HDF format provides two different confidence parameters for each photon: quality_flag and sensitivity. The quality_flag value represents whether the photon meets the criteria based on energy, sensitivity, amplitude, real-time surface tracking quality and difference to a DEM. Sensitivity provides a sensitivity threshold for users to model the ground surface well. This is usually considered to be 0.9 for land.

In this study, quality_flag = 1 and sensitivity > 0.9 were first used to remove noise from the GEDI data (Figure 2, upper right corner).

2.3.3. Preprocessing of Building Footprints

To improve accuracy in comparing estimated building heights to measured building heights, buildings constructed after the operational dates of ICESat-2 and GEDI satellites (2018) were not taken into account. Additionally, buildings with unknown heights and

values of “0” were also removed from the dataset. The buildings we found in LA were less than one meter high and were not considered to be truly representative of a building, so these data were removed from the dataset (Figure 2, middle left).

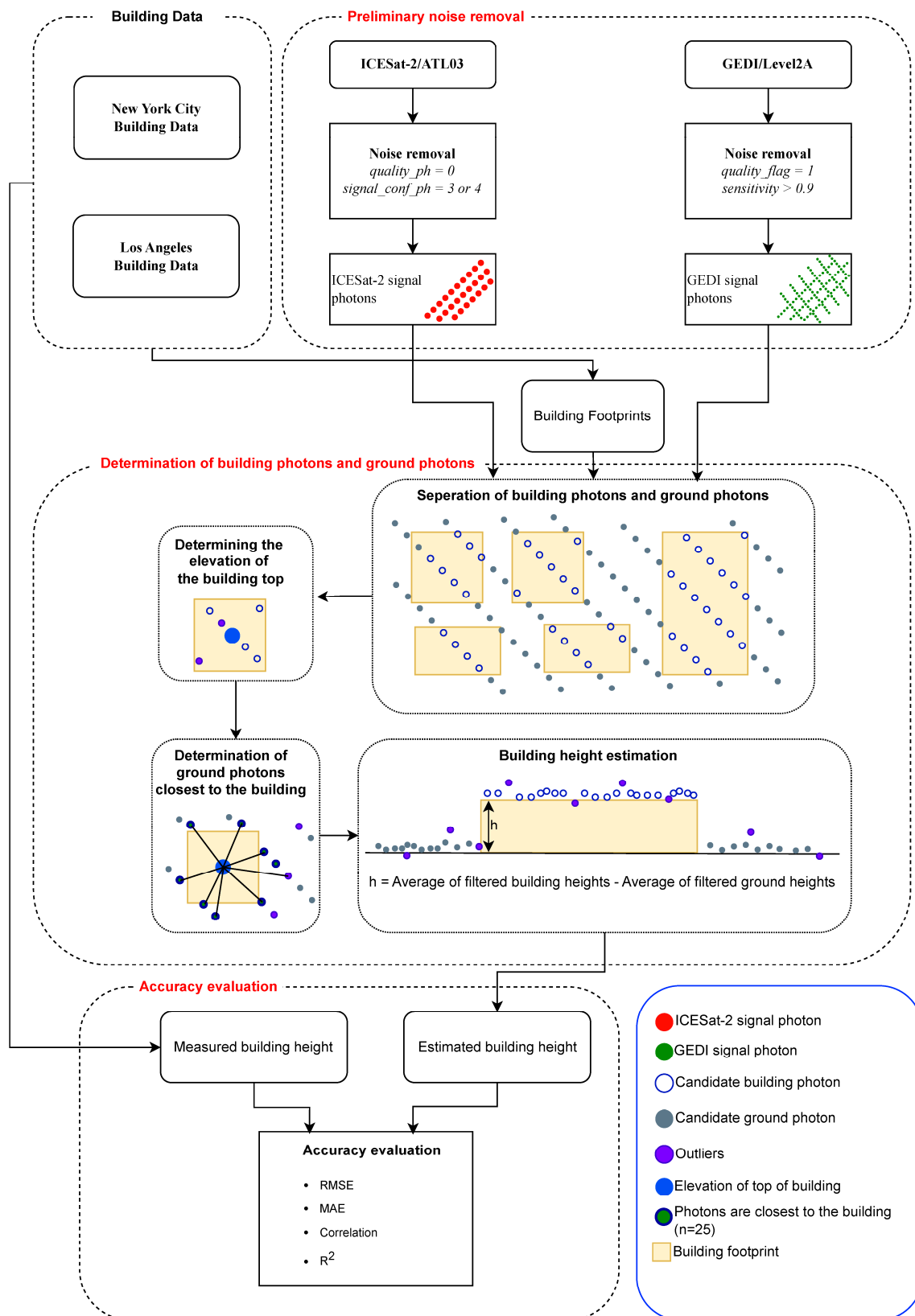


Figure 2. The flow chart of the study.

2.3.4. Extracting Building Photons and Ground Photons

To automatically determine the heights of a large number of buildings in NYC and LA, ICESat-2 and GEDI photons had to be separated into a building photon and a ground photon. To separate the photons as accurately and precisely as possible and to determine the heights of buildings more precisely, we performed the following series of operations:

- [1] First, the downloaded shapefile building footprints and laser altimeter data were merged in ArcGIS Pro, and photons falling on the building boundary were labeled as building photon candidates, while others were labeled as ground photon candidates;
- [2] Photons falling within each building were then evaluated and outliers were removed using the interquartile range (IQR) approach [59]. The interquartile range (IQR) approach, which identifies data beyond 1.5 times the difference between the first and third quartiles [59] (Figure 2, middle right), was implemented by applying Equations (1)–(3).

$$IQR = Q_3 - Q_1 \quad (1)$$

$$Outliers < Q_1 - 1.5 * IQR \quad (2)$$

$$Outliers > Q_3 + 1.5 * IQR \quad (3)$$

Q_1 and Q_3 represent the first and third quartiles, respectively.

Building heights were determined by averaging the extreme values of building photons. Here, the arithmetic average of all photons was taken without considering the factors that would cause height changes on the roofs of the buildings (solar panels, antennas, etc.).

- [3] By averaging the latitude and longitude values of the photons within the building boundary, the average location of the photons falling on each building was determined as the building top point. The N number of points closest to the building top were determined as candidate ground photons, and the outliers in these data were removed according to the IQR method. Then, the average of these photons was taken to obtain the ground elevation. Here, N was taken as 10, 25, 50, 100, 250, and 500, respectively, but examining 10 photons caused too much bias, especially in wooded areas. The 50–500 values also did not give the desired ground height as they represented the heights of points quite far from the building. Therefore, the ground height for all buildings was determined based on the 25 closest points (non-filtered count) (Figure 2, middle left);
- [4] Building heights were determined by subtracting the average elevation of the top of each building from the ground elevation of that building (Figure 2).

2.3.5. Accuracy Evaluation

The RMSE is defined as the square of the average square difference between the estimated building heights and the measured building height. RMSE is calculated by Equation (4). The RMSE ensures that the error units are in the same units as the data and is more susceptible to major errors. Low RMSE values indicate that the model performs better. The MAE is the average of the absolute differences between the estimated building heights and the measured building height. The MAE gives the average size of errors and does not take into account the aspects of the errors. The MAE presents how much deviation there is in the model estimate in a simple and understandable way. Compared to RMSE, it reflects the impact of major errors less. MAE is calculated by Equation (5). R^2 is a statistic showing the correlation between estimated and measured building heights. A higher R^2 value indicates a better match between the estimate and the measurement data. R^2 is calculated by Equation (6). The Pearson correlation coefficient measures the linear relationship between two variables and gives values between -1 and 1 . As the R value approaches $+1$, it represents a strong and positive relationship between the estimated values and the observed values. This relationship is strong but negative as the R value

approaches -1 . $R = 0$ means there is no linear relationship between the two variables. R is calculated by Equation (7).

$$RMSE = \sqrt{\frac{1}{n} \sum (H_{estimated} - H_{measured})^2} \quad (4)$$

$$MAE = \frac{\sum |H_{estimated} - H_{measured}|}{n} \quad (5)$$

$$R^2 = 1 - \frac{\sum (H_{estimated} - H_{measured})^2}{\sum (H_{estimated} - \bar{H}_{measured})^2} \quad (6)$$

$$R = \frac{\sum_{i=1}^n (H_{measured,i} - \bar{H}_{measured})(H_{estimated,i} - \bar{H}_{estimated})}{\sqrt{\sum_{i=1}^n (H_{measured,i} - \bar{H}_{measured})^2 \sum_{i=1}^n (H_{estimated,i} - \bar{H}_{estimated})^2}} \quad (7)$$

$H_{estimated}$: estimated building heights based on ICESat-2 and GEDI data;

$H_{measured}$: observed building height (reference data);

n : number of variables.

3. Results and Discussion

In this study, ICESat-2 ATL03 and GEDI Level2A data were used to estimate the heights of 84,045 buildings in LA and 50,654 buildings in NYC. The results were compared with reference building heights in LA and NYC (Figure 3). Figure 3 shows that 41,451 buildings in NYC were determined with ICESat-2, yielding a building height of 8.06 m RMSE and 6.20 m MAE. With GEDI data, the height of 10,698 buildings were estimated, with RMSE and MAE values of 11.31 m and 8.71 m, respectively. With the combined evaluation of ICESat-2 and GEDI data, the height of 50,654 buildings was estimated with RMSE and MAE values of 8.28 m and 6.24 m, respectively. In LA, 69,094 buildings were determined with ICESat-2 and 15,959 buildings were determined with GEDI. Using ICESat-2 and GEDI data together, the height of 84,045 buildings was estimated with 6.42 m RMSE and 4.66 m MAE.

RMSE values show how much the estimated building heights deviate from the measured heights on a squared average. The average RMSE of 50,654 buildings in NYC is 8.28 m (Figure 3). Figure 4 shows the analysis of different building height estimates for different RMSE bounds. When the distribution of buildings according to their RMSE amounts is analyzed, ~4% of the buildings were below 0.5 m RMSE. In addition, 46% of the buildings were determined with 5 m RMSE. In LA, a total of 84,045 buildings were determined with 6.42 m RMSE. ~3% of buildings were estimated below 0.5 m RMSE. In addition, 67% of buildings were estimated with <5 m RMSE. A 5 m RMSE value corresponds to ~1.5 building floor height. Although this accuracy is not sufficient to determine the height of buildings, it is a significant advance for estimating energy consumption in cities, countries, and the world.

NYC and LA have a wide range of building heights. However, the predominant building height limits in both cities are between 3 m and 10 m. In NYC, 70% of buildings have an estimated height between 3 m and 10 m and in LA, 87% of buildings have an estimated height between 3 m and 10 m (Figure 5). The variation in land use and the presence of dense tree cover, especially in much of LA, is an obstacle in determining building heights. To overcome this problem, an extra threshold was added for the detection of ground photons, with a maximum threshold of 1 m between the minimum and maximum value of ground photons, so that RMSE values for buildings <10 m were below 3 m.

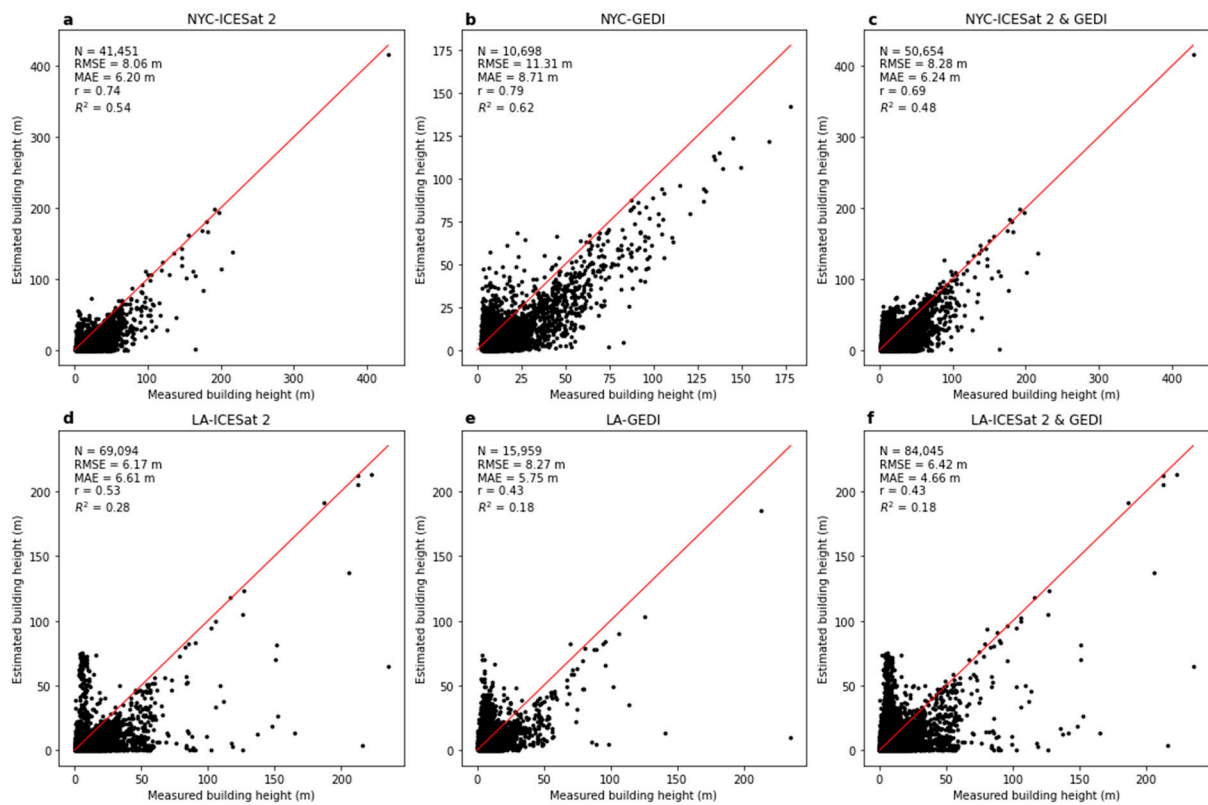


Figure 3. Accuracy of building height estimation: (a–c) show the accuracy of ICESat-2, GEDI and combined data for NYC, respectively; (d–f) show the accuracy of ICESat-2, GEDI, and combined data for LA, respectively.

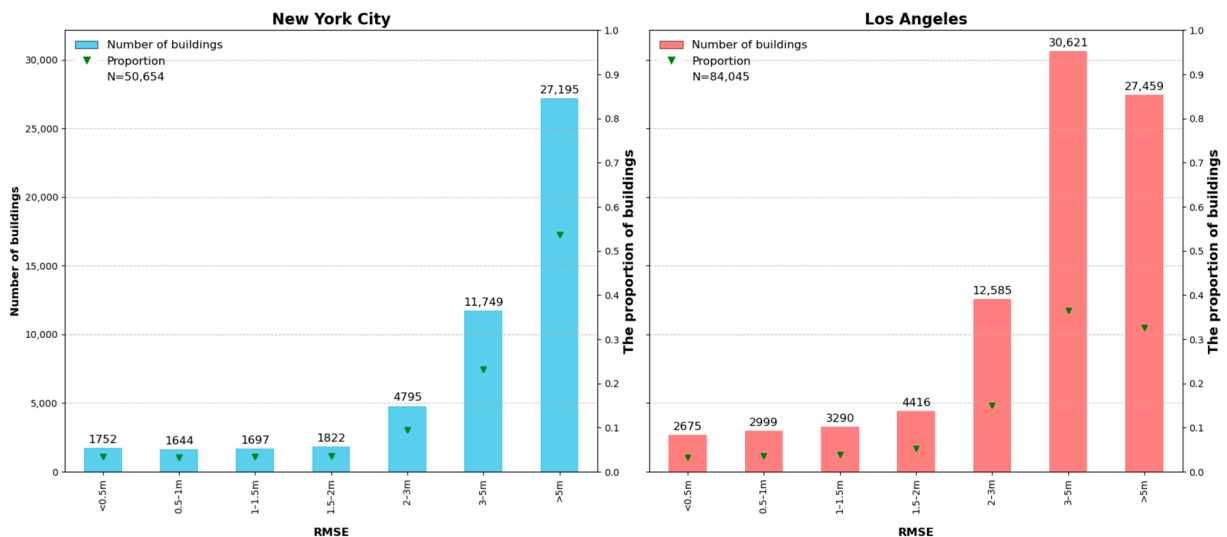


Figure 4. Number of buildings estimated with different RMSE values for NYC (left) and LA (right).

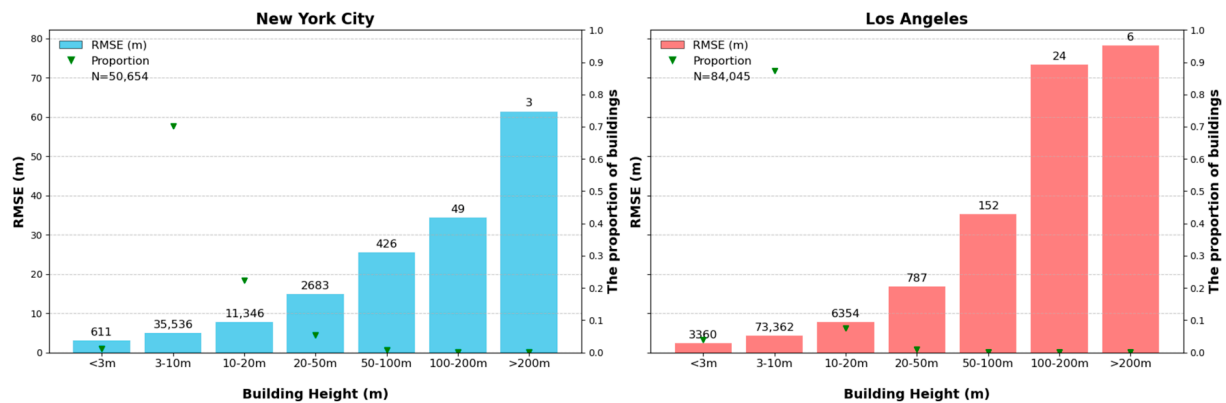


Figure 5. RMSE values of buildings of different heights and their proportions to the total number of buildings.

The variability of the building area directly affects the number of photons falling at the building boundary. In buildings with large footprints, the number of photons representing the building can be more than 500, whereas in buildings with small footprints there may be only a few photons. Therefore, it is important to see how the number of photons affects the result when estimating building heights. Figure 6 compares the RMSE values of buildings represented by different numbers of photons. Figure 6 shows that for both NYC and LA, the RMSEs of buildings represented by 5–10 photons are better than the others. Similarly, RMSE values are quite high for buildings with more than 100 photons. The most important reason for this is that many photons represent a single building in buildings with large areas. As the number of photons increases, the difference between the minimum and maximum height values of the photons representing the building increases. However, there may be many objects (solar panels, antennas, etc.) on the top of the building that will affect the building height.

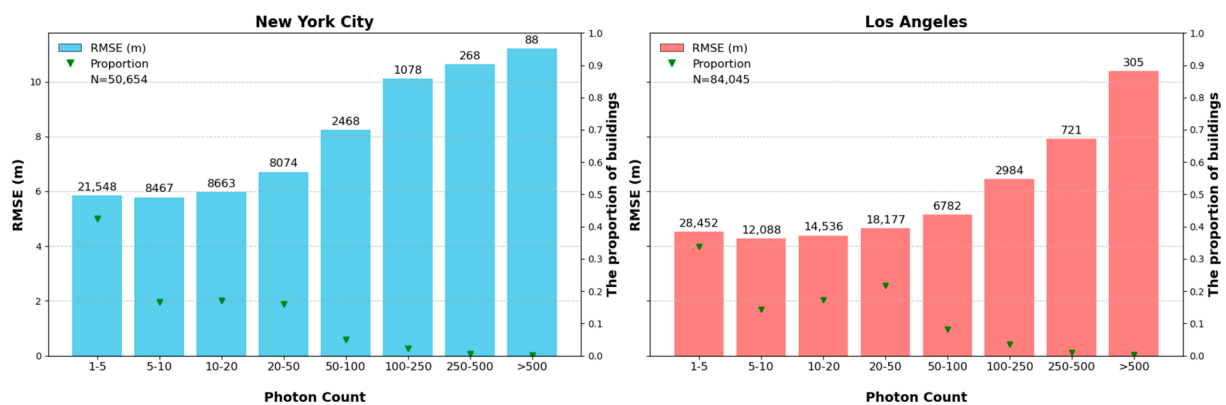


Figure 6. Photon counts and RMSE values used to estimate building height.

To summarize, the automatic estimation of building heights is difficult for many reasons. The top of the building is not flat, there are objects of different heights on the building roofs, and there are reflective surfaces on the building roofs, making it difficult to determine the height of the building tops. Similarly, there are various difficulties in detecting ground photons. It is difficult to remove from the dataset signals that are reflected from various objects on the ground such as vehicles, garbage containers, trees, etc. and do not represent the ground. In addition, the presence of reflective surfaces leads to the inaccurate determination of ground elevations.

Nevertheless, this study presents a methodology for estimating building heights using ICESat-2 and GEDI laser altimeter data provided free of charge by NASA, and building footprints provided free of charge by the USA. Although similar studies have

been conducted in the literature, there has not yet been a study that estimates building heights with this scope and different datasets. Zhao et al. [43] estimated the heights of ~16,000 buildings with an MAE of 4.08 m by combining Google Earth images and ICESat-2 photons. In this study, the heights of 84,045 buildings in LA were determined with a MAE of 4.66 m. Cai et al. [12] determined the heights of 17,399 buildings in NYC, which is also the study area in this study, with an RMSE of 8.1 m using only ICESat-2 data. In this study, they determined the height of 41,451 buildings in NYC with an RMSE of 8.06 m using only ICESat-2 data. In other words, with the same dataset, more buildings with the same RMSE were identified. Also, unlike Cai et al. [12], this study did not use NYC land use raster data and aimed to determine the maximum building height with the least possible dataset.

The major limitation of this study is the presence of signal reflecting objects on building roofs and around buildings. In addition, this study is currently only available in regions where building footprints are available, and their accuracy is closely related to the spatial accuracy of the building footprints. Despite these challenges, this study demonstrates a successful methodology for estimating the height of many buildings with the minimum possible data. Given these limitations, future work will focus on the following topics.

- [1] Spatially accurate determination of building footprints and determination of building height on a global scale by obtaining building footprints worldwide;
- [2] Improving the accuracy of photon representation by using high-resolution multispectral satellite imagery to detect reflective surfaces that affect the quality of building photons and ground photons;
- [3] More extensive machine learning and deep learning algorithms are planned to be used to separate building photons and ground photons.

4. Conclusions

In this study, an innovative approach is proposed to determine building heights by integrating building footprints and ICESat-2 and GEDI data provided free of charge by NASA. A total of 134,699 buildings, 50,654 in NYC and 84,045 in LA, were estimated. The RMSE values for NYC and LA are 8.28 m and 6.42 m, respectively. Furthermore, 46% of the buildings in NYC and 67% of the buildings in LA were estimated with RMSE less than 5 m. This study shows that ICESat-2 data are more accurate than GEDI data in determining building heights. However, thanks to the grid pattern of GEDI data, the height of many buildings can be estimated with a small number of photons. In this respect, the combined use of ICESat-2 and GEDI data provides an advantage in estimating the height of more buildings. The proposed method is highly accurate and applicable for buildings with various architectural features and various environmental factors. This makes the proposed method usable in various geographical areas and for buildings of various complexity. Since the study is based on the availability of building footprints and satellite altimetry data, it can be applied worldwide if sufficient data are available. However, in regions where seasonal effects are severe (buildings partially or completely covered with snow), it can be difficult to estimate building height. To overcome this obstacle, it is necessary to focus on measurements outside the winter months of the region.

Due to the presence of signal-obstructing surfaces and frequently changing topography in urban areas, the quality of photon data obtained from ICESat-2 and GEDI is likely to be lower in these regions. Consequently, future research should investigate the integration of various data sources (such as SAR, radar altimeter, DEM, optical imagery) to overcome this limitation.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The author declares no conflicts of interest.

References

- Urban, M.C.; Alberti, M.; De Meester, L.; Zhou, Y.; Verrelli, B.C.; Szulkin, M.; Schmidt, C.; Savage, A.M.; Roberts, P.; Rivkin, L.R.; et al. Interactions between climate change and urbanization will shape the future of biodiversity. *Nat. Clim. Chang.* **2024**, *14*, 436–447. [CrossRef]
- Leong, D.P.; Teo, K.K.; Rangarajan, S.; Lopez-Jaramillo, P.; Avezum, A., Jr.; Orlandini, A. World Population Prospects 2019. Department of Economic and Social Affairs Population Dynamics. New York (NY): United Nations; 2019 (<https://population.un.org/wpp/Download/>, accessed 20 September 2020). The decade of healthy ageing. Geneva: World Health Organization. *World* **2018**, *73*, 362k2469.
- Moreno-Monroy, A.I.; Schiavina, M.; Veneri, P. Metropolitan areas in the world. Delineation and population trends. *J. Urban Econ.* **2021**, *125*, 103242. [CrossRef]
- Aldhshan, S.R.; Abdul Maulud, K.N.; Wan Mohd Jaafar, W.S.; Karim, O.A.; Pradhan, B. Energy consumption and spatial assessment of renewable energy penetration and building energy efficiency in Malaysia: A review. *Sustainability* **2021**, *13*, 9244. [CrossRef]
- Chuah, J.W. Analysis and Optimization of Building Energy Consumption. Ph.D. Thesis, Princeton University, Princeton, NJ, USA, 2013; pp. 1–184.
- Han, J.; Meng, X.; Zhou, X.; Yi, B.; Liu, M.; Xiang, W.-N. A long-term analysis of urbanization process, landscape change, and carbon sources and sinks: A case study in China's Yangtze River Delta region. *J. Clean. Prod.* **2017**, *141*, 1040–1050. [CrossRef]
- Wang, Y.; Sun, G.; Wu, Y.; Rosenberg, M.W. Urban 3D building morphology and energy consumption: Empirical evidence from 53 cities in China. *Sci. Rep.* **2024**, *14*, 12887. [CrossRef] [PubMed]
- Leng, H.; Chen, X.; Ma, Y.; Wong, N.H.; Ming, T. Urban morphology and building heating energy consumption: Evidence from Harbin, a severe cold region city. *Energy Build.* **2020**, *224*, 110143. [CrossRef]
- Li, Z.; Ma, J.; Jiang, F. Exploring the effects of 2D/3D building factors on urban energy consumption using explainable machine learning. *J. Build. Eng.* **2024**, *97*, 110827. [CrossRef]
- Gong, P.; Li, X.; Wang, J.; Bai, Y.; Chen, B.; Hu, T.; Liu, X.; Xu, B.; Yang, J.; Zhang, W.; et al. Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. *Remote Sens. Environ.* **2020**, *236*, 111510. [CrossRef]
- 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]
- Cai, P.; Guo, J.; Li, R.; Xiao, Z.; Fu, H.; Guo, T.; Zhang, X.; Li, Y.; Song, X. Automated Building Height Estimation Using Ice, Cloud, and Land Elevation Satellite 2 Light Detection and Ranging Data and Building Footprints. *Remote Sens.* **2024**, *16*, 263. [CrossRef]
- IEA Buildings. Paris. 2022. Available online: <https://www.iea.org/reports/buildings> (accessed on 23 October 2024).
- Olasolo-Alonso, P.; López-Ochoa, L.M.; Las-Heras-Casas, J.; López-González, L.M. Energy performance of buildings directive implementation in southern European countries: A review. *Energy Build.* **2023**, *281*, 112751. [CrossRef]
- Zheng, G.; Feng, Z.; Jiang, M.; Tan, L.; Wang, Z. Predicting the Energy Consumption of Commercial Buildings Based on Deep Forest Model and Its Interpretability. *Buildings* **2023**, *13*, 2162. [CrossRef]
- De Cristo, E.; Evangelisti, L.; Guattari, C.; De Lieto Vollaro, R. An Experimental Direct Model for the Sky Temperature Evaluation in the Mediterranean Area: A Preliminary Investigation. *Energies* **2024**, *17*, 2228. [CrossRef]
- Resch, E.; Bohne, R.A.; Kvamsdal, T.; Lohne, J. Impact of urban density and building height on energy use in cities. *Energy Procedia* **2016**, *96*, 800–814. [CrossRef]
- Schug, F.; Frantz, D.; van der Linden, S.; Hostert, P. Gridded population mapping for Germany based on building density, height and type from Earth Observation data using census disaggregation and bottom-up estimates. *PLoS ONE* **2021**, *16*, e0249044. [CrossRef]
- Pan, T.C.; Goh, K.S.; Megawati, K. Empirical relationships between natural vibration period and height of buildings in Singapore. *Earthq. Eng. Struct. Dyn.* **2014**, *43*, 449–465. [CrossRef]
- Kadhim, N.; Mourshed, M. A shadow-overlapping algorithm for estimating building heights from VHR satellite images. *IEEE Geosci. Remote Sens. Lett.* **2017**, *15*, 8–12. [CrossRef]
- Goetz, M. Towards generating highly detailed 3D CityGML models from OpenStreetMap. *Int. J. Geogr. Inf. Sci.* **2013**, *27*, 845–865. [CrossRef]
- Biljecki, F.; Ledoux, H.; Stoter, J. Generating 3D city models without elevation data. *Comput. Environ. Urban Syst.* **2017**, *64*, 1–18. [CrossRef]
- Li, B.; Xie, H.; Tong, X.; Liu, S.; Xu, Q.; Sun, Y. Extracting accurate terrain in vegetated areas from ICESat-2 data. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *117*, 103200. [CrossRef]
- Lee, T.; Kim, T. Automatic building height extraction by volumetric shadow analysis of monoscopic imagery. *Int. J. Remote Sens.* **2013**, *34*, 5834–5850. [CrossRef]
- Zhang, H.; Xu, C.; Fan, Z.; Li, W.; Sun, K.; Li, D. Detection and Classification of Buildings by Height from Single Urban High-Resolution Remote Sensing Images. *Appl. Sci.* **2023**, *13*, 10729. [CrossRef]
- 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.* **2022**, *16*, 180–190. [CrossRef]

27. Wang, X.; Yu, X.; Ling, F. Building heights estimation using ZY3 data—A case study of Shanghai, China. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 1749–1752.
28. Qi, F.; Zhai, J.Z.; Dang, G. Building height estimation using Google Earth. *Energy Build.* **2016**, *118*, 123–132. [\[CrossRef\]](#)
29. Garzelli, A. Building height retrieval from WorldView-2 multi-angular images. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1163–1166.
30. Li, B.; Chen, X.; Lin, Z. Building Height Restoration Method of Remote Sensing Images Based on Faster RCNN. In Proceedings of the 2022 IEEE 34th International Conference on Tools with Artificial Intelligence (ICTAI), Macao, China, 31 October–2 November 2022; pp. 955–959.
31. Kim, T.; Javzandulam, T.; Lee, T.Y. Semiautomatic reconstruction of building height and footprints from single satellite images. In Proceedings of the 2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 23–28 July 2007; pp. 4737–4740.
32. Sun, Y.; Hua, Y.; Mou, L.; Zhu, X.X. Large-scale building height estimation from single VHR SAR image using fully convolutional network and GIS building footprints. In Proceedings of the 2019 Joint Urban Remote Sensing Event (JURSE), Vannes, France, 22–24 May 2019; pp. 1–4.
33. 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. [\[CrossRef\]](#)
34. Chen, P.; Huang, H.; Liu, J.; Wang, J.; Liu, C.; Zhang, N.; Su, M.; Zhang, D. Leveraging Chinese GaoFen-7 imagery for high-resolution building height estimation in multiple cities. *Remote Sens. Environ.* **2023**, *298*, 113802. [\[CrossRef\]](#)
35. Huang, H.; Chen, P.; Xu, X.; Liu, C.; Wang, J.; Liu, C.; Clinton, N.; Gong, P. Estimating building height in China from ALOS AW3D30. *ISPRS J. Photogramm. Remote Sens.* **2022**, *185*, 146–157. [\[CrossRef\]](#)
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. Bshouty, E.; Dalyot, S. Calculating OpenStreetMap Building Heights from Single User-Generated Photographs. In Proceedings of the IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 855–858.
38. Petty, A.A.; Kurtz, N.T.; Kwok, R.; Markus, T.; Neumann, T.A. Winter Arctic sea ice thickness from ICESat-2 freeboards. *J. Geophys. Res. Ocean.* **2020**, *125*, e2019JC015764. [\[CrossRef\]](#)
39. Osama, N.; Yang, B.; Ma, Y.; Freeshah, M. A digital terrain modeling method in urban areas by the ICESat-2 (Generating precise terrain surface profiles from photon-counting technology). *Photogramm. Eng. Remote Sens.* **2021**, *87*, 237–248. [\[CrossRef\]](#)
40. Narin, O.G.; Abdikan, S.; Gullu, M.; Lindenbergh, R.; Balik Sanli, F.; Yilmaz, I. Improving global digital elevation models using space-borne GEDI and ICESat-2 LiDAR altimetry data. *Int. J. Digit. Earth* **2024**, *17*, 2316113. [\[CrossRef\]](#)
41. 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. [\[CrossRef\]](#)
42. 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. [\[CrossRef\]](#)
43. 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. [\[CrossRef\]](#)
44. 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. [\[CrossRef\]](#)
45. Dubayah, R.; Blair, J.B.; Goetz, S.; Fatoyinbo, L.; Hansen, M.; Healey, S.; Hofton, M.; Hurtt, G.; Kellner, J.; Luthcke, S.; et al. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth’s forests and topography. *Sci. Remote Sens.* **2020**, *1*, 100002. [\[CrossRef\]](#)
46. Fayad, I.; Baghdadi, N.; Alcarde Alvares, C.; Stape, J.L.; Bailly, J.S.; Scolforo, H.F.; Cegatta, I.R.; Zribi, M.; Le Maire, G. Terrain slope effect on forest height and wood volume estimation from GEDI data. *Remote Sens.* **2021**, *13*, 2136. [\[CrossRef\]](#)
47. Morin, D.; Planells, M.; Baghdadi, N.; Bouvet, A.; Fayad, I.; Le Toan, T.; Mermoz, S.; Villard, L. Improving heterogeneous forest height maps by integrating GEDI-based forest height information in a multi-sensor mapping process. *Remote Sens.* **2022**, *14*, 2079. [\[CrossRef\]](#)
48. 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. [\[CrossRef\]](#)
49. Civerolo, K.; Hogrefe, C.; Lynn, B.; Rosenthal, J.; Ku, J.Y.; Solecki, W.; Cox, J.; Small, C.; Rosenzweig, C.; Goldberg, R.; et al. Estimating the effects of increased urbanization on surface meteorology and ozone concentrations in the New York City metropolitan region. *Atmos. Environ.* **2007**, *41*, 1803–1818. [\[CrossRef\]](#)
50. NYC Planning. New York City Department of City Planning. Available online: <https://www.nyc.gov/site/planning/planning-level/nyc-population/nyc-population-current-estimates.page> (accessed on 24 June 2024).
51. NYC Open Data. New York City Building Footprints. Available online: <https://data.cityofnewyork.us/Housing-Development/Building-Footprints/nqwf-w8eh> (accessed on 24 June 2024).
52. USCB. United States Census Bureau. Available online: https://data.census.gov/profile/Los_Angeles_city_California?g=160XX00U50644000 (accessed on 24 June 2024).

53. FRED. Federal Reserve Bank of St. Louis. Available online: <https://fred.stlouisfed.org/series/NGMP31080> (accessed on 24 June 2024).
54. Los Angeles City Building Footprints. Available online: <https://geohub.lacity.org/datasets/813fcefde1f64b209103107b26a8909f/explore> (accessed on 24 June 2024).
55. Schutz, B.E.; Zwally, H.J.; Shuman, C.A.; Hancock, D.; DiMarzio, J.P. Overview of the ICESat mission. *Geophys. Res. Lett.* **2005**, *32*. [[CrossRef](#)]
56. Feng, Y.; Yang, L.; Zhan, P.; Luo, S.; Chen, T.; Liu, K.; Song, C. Synthesis of the ICESat/ICESat-2 and CryoSat-2 observations to reconstruct time series of lake level. *Int. J. Digit. Earth* **2023**, *16*, 183–209. [[CrossRef](#)]
57. Neumann, T.A.; Martino, A.J.; Markus, T.; Bae, S.; Bock, M.R.; Brenner, A.C.; Brunt, K.M.; Cavanaugh, J.; Fernandes, S.T.; Hancock, D.W.; et al. The Ice, Cloud, and Land Elevation Satellite-2 Mission: A global geolocated photon product derived from the advanced topographic laser altimeter system. *Remote Sens. Environ.* **2019**, *233*, 111325. [[CrossRef](#)] [[PubMed](#)]
58. Gao, M.; Xing, S.; Zhang, G.; Zhang, X.; Li, P. Assessment of ICESat-2's horizontal accuracy using an iterative matching method based on high-accuracy terrains. *Remote Sens.* **2023**, *15*, 2236. [[CrossRef](#)]
59. Yuan, C.; Gong, P.; Bai, Y. Performance assessment of ICESat-2 laser altimeter data for water-level measurement over lakes and reservoirs in China. *Remote Sens.* **2020**, *12*, 770. [[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.