

## Article

# Evaluation of NASA's GEDI Lidar Observations for Estimating Biomass in Temperate and Tropical Forests

Mei Sun <sup>1</sup>, Lei Cui <sup>2,\*</sup>, Jongmin Park <sup>3</sup>, Mariano García <sup>4</sup> , Yuyu Zhou <sup>5</sup> , Carlos Alberto Silva <sup>6</sup> , Long He <sup>2</sup>, Hu Zhang <sup>2</sup>  and Kaiguang Zhao <sup>3</sup>

- <sup>1</sup> Beijing Engineering Research Center for Global Land Remote Sensing Products, Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
- <sup>2</sup> College of Urban and Environmental Sciences, Tianjin Normal University, Tianjin 300387, China
- <sup>3</sup> School of Environment and Natural Resources, The Ohio State University, Wooster, OH 43210, USA
- <sup>4</sup> Universidad de Alcalá, Department of Geology, Geography and Environment, Environmental Remote Sensing Research Group, Colegios 2, 28801 Alcalá de Henares, Spain
- <sup>5</sup> Department of Geological and Atmospheric Sciences, Iowa State University, Ames, IA 50011, USA
- <sup>6</sup> School of Forest, Fisheries, and Geomatics Sciences, University of Florida, Gainesville, FL 32611, USA
- \* Correspondence: cuil@tjnu.edu.cn

**Abstract:** Accurate estimation of forest aboveground biomass (AGB) is vital for informing ecosystem and carbon management. The Global Ecosystem Dynamics Investigation (GEDI) instrument—a new-generation spaceborne lidar system from NASA—provides the first global coverage of high-resolution 3D altimetry data aimed specifically for mapping Earth's forests, but its performance is yet to be tested for large parts of the world. Here, our goal is to evaluate the accuracies of GEDI in measuring terrain, forest vertical structures, and AGB in reference to independent airborne lidar data over temperate and tropical forests in North America. We compared GEDI-derived elevations and canopy heights (e.g., relative height percentiles such as RH50 and RH100) with those from the Shuttle Radar Topography Mission (SRTM) or from two airborne lidar systems: the Laser Vegetation Imaging Sensor (LVIS) and Goddard's Lidar, Hyperspectral and Thermal system (G-LiHT). We also estimated GEDI's geolocation errors by matching GEDI waveforms and G-LiHT pseudo-waveforms. We assessed the predictive power of GEDI metrics in estimating AGB using Random Forests regression. Results showed that GEDI-derived ground elevations correlated strongly those from LVIS, G-LiHT, and SRTM ( $R^2 > 0.91$ ), but with nonnegligible RMSEs of 5.7 m (G-LiHT), 3.1 m (LVIS), and 10.9 m (SRTM). GEDI canopy heights had poorer correlation with LVIS (e.g.,  $R^2 = 0.44$  for RH100) than with G-LiHT (e.g.,  $R^2 = 0.60$  for RH100). The estimated horizontal geolocation errors of GEDI footprints averaged 6.5 meters, comparable to the nominal accuracy of 9 m. Correction for the locational errors improved the correlation of GEDI vs G-LiHT canopy heights significantly, on average by 53% (e.g.,  $R^2$  from 0.57 to 0.82 for RH50). GEDI canopy metrics were useful for predicting AGB ( $R^2 = 0.82$  and RMSE = 19.1 Mg/Ha), with the maximum canopy height RH100 being the most useful predictor. Our results highlight the importance of accommodating or correcting for GEDI geolocation errors for estimating forest characteristics and provide empirical evidence on the utility of GEDI for monitoring global biomass dynamics from space.

**Keywords:** GEDI; airborne lidar; forest structure; ground elevation; aboveground biomass; carbon



**Citation:** Sun, M.; Cui, L.; Park, J.; García, M.; Zhou, Y.; Silva, C.A.; He, L.; Zhang, H.; Zhao, K. Evaluation of NASA's GEDI Lidar Observations for Estimating Biomass in Temperate and Tropical Forests. *Forests* **2022**, *13*, 1686. <https://doi.org/10.3390/f13101686>

Academic Editor: Nikolay Strigul

Received: 8 September 2022

Accepted: 8 October 2022

Published: 13 October 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Carbon neutrality has been put on the agenda of the United Nations as an important goal, and how to achieve this goal is presently an urgent problem for the major economic powers [1]. Forest carbon sinks are among the main means to achieve carbon neutrality [2,3]—forests capture carbon dioxide from the atmosphere and transform it into biomass through photosynthesis. To some extent, the aboveground biomass (AGB) can largely characterize the carbon sequestration capacity of forests [4,5]. Therefore, accurate estimation

of forests' AGB is of great significance for conducting carbon neutrality research. Existing studies have shown that lidar is one of the most effective ways to estimate forests' AGB by obtaining forest vertical structure and ground elevation information [6–8].

Lidar data from airborne platforms have been widely used to estimate forest structure parameters at stand or regional scales, such as Goddard's Lidar, Hyperspectral and Thermal (G-LiHT) and the Land, Vegetation, and Ice Sensor (LVIS). The G-LiHT project was led by the National Aeronautics and Space Administration (NASA) to support their scientific research on satellite observation verification and calibration [9], such as validating the Ice Cloud and Land Elevation Satellite (ICESat-2) satellite-borne lidar observations over forest ecosystems [10]. G-LiHT data have also been used to map forest vertical structure, AGB, and carbon storage [11–13]. The LVIS is another widely used airborne lidar with the advantage of providing low-cost observational data for NASA-related researchers and scientific missions [14]. For example, LVIS lidar data have been successfully applied to tree height and AGB estimation [15–17] and satellite observation verification [18].

Spaceborne lidar systems have been proven to have the potential to extend lidar observations on unprecedented scales, from regional to continental and global extents. The Geoscience Laser Altimeter System (GLAS) aboard the ICESat was the first spaceborne lidar system and was successfully applied to global canopy height mapping [19,20] and national-scale biomass mapping [21,22]. However, due to equipment failure, ICESat GLAS ceased operation in 2009. ICESat-2/ATLAS, the successor to ICESat GLAS, was launched in September 2018 equipped with a newly designed photon-counting lidar sensor (ATLAS). Although ICESat-2 observations have been applied to map forest vertical structure parameters [23,24], studies have shown that the ATLAS sensor is extremely sensitive to noise, which in turn leads to an increase in uncertainty [25].

Although ICESat and ICESat-2 provide means for measuring forest structure parameters on a global scale, they were not originally designed for forest observation; thus, they often suffer from problems such as relatively low spatial resolution and susceptibility to noise interference. The Global Ecosystem Dynamics Investigation (GEDI) onboard the international space station (ISS) is a new spaceborne lidar system designed for forest observation. The inventor claims that GEDI can provide high-quality measurements of forest vertical structure with a spatial resolution of ~25 m at footprint level, especially in temperate and tropical forests [26]. Researchers have started using available GEDI data for forest studies, including canopy height mapping [27], AGB estimation [28–30], fuel load mapping [31], and vertical leaf area index profile mapping [32].

GEDI provides an opportunity to improve the accuracy of multiple forest structure parameters globally, such as AGB. However, using GEDI data to derive accurate forest AGB requires a comprehensive understanding of GEDI's measurement performance, such as percentile height and ground elevation. Percentile height—which is the canopy vertical position extracted from the integrated received waveform energy—is a description of the forest vertical structure; studies have documented that it is sensitive to variations in biomass [33]. Ground elevation extracted from the lidar received waveform is another important variable that can determine the accuracy of forest biomass [34]. Furthermore, the geolocation error of GEDI should be figured out. As described in the official GEDI documentation, the geolocation information released by the GEDI product has an error of about 9 m [35]; this uncertainty factor will increase the uncertainty of the high-level products, such as biomass, particularly over heterogeneous canopies.

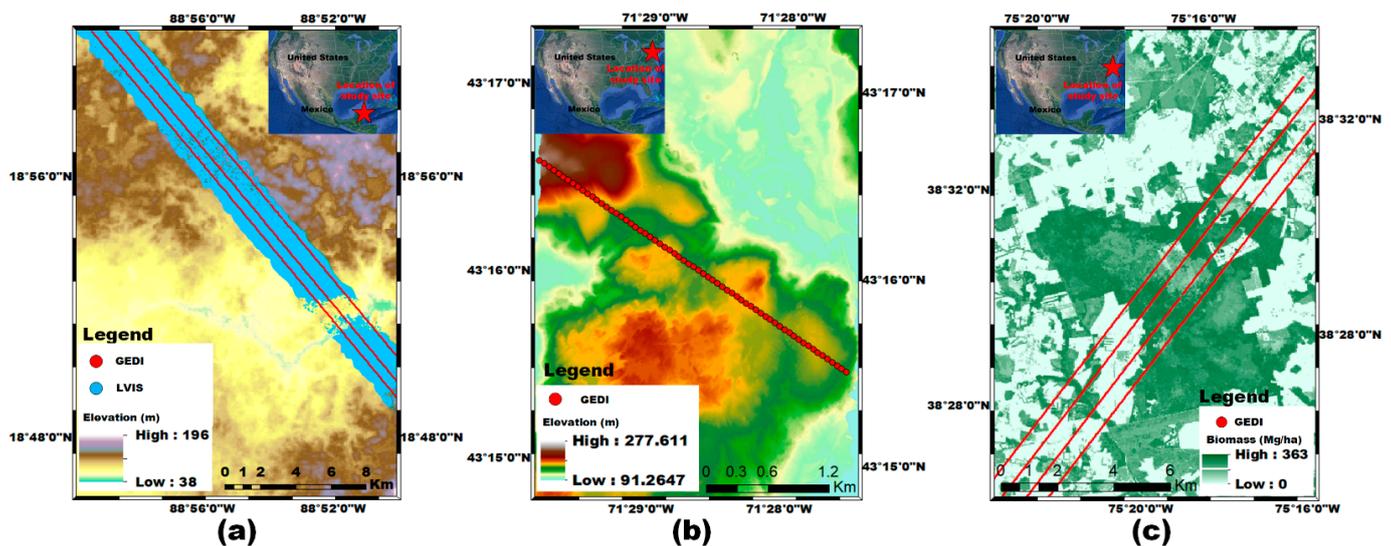
As the new generation of spaceborne lidar, GEDI was specially designed for observing forest dynamics; therefore, it should be a good choice for biomass estimation. The overall goal of this study was to evaluate the information content of GEDI for forest AGB estimation by using G-LiHT discrete lidar data, LVIS full-waveform lidar data, and SRTM data in low and mid-latitudes of North America. Our study areas covered the forest structure characteristics of tropical and temperate regions. Specifically, we (i) evaluated the ground elevation from GEDI with that derived from G-LiHT, LVIS, and SRTM data; (ii) evaluated GEDI typical percentile height metrics (25th, 50th, 75th, and 100th percentile of relative

heights: RH25, RH50, RH75, and RH100) based on airborne lidar observations; (iii) assessed the geolocation error of the GEDI footprint centroid by comparing the profiles of energy measured by GEDI and G-LiHT; and (iv) used the Random Forests (RF) machine learning method to evaluate the ability of GEDI to estimate forest AGB determined by forest vertical structure and ground elevation.

## 2. Materials

### 2.1. Study Area

Our study area included three sites in the United States (US) and Mexico. Site a was located in Quintana Roo, Mexico (Figure 1a) where the terrain is relatively gentle, with an altitude of 38–196 m; the forest type is dominated by tropical forests, including evergreen and deciduous forests; and the canopy height ranges between 10 and 35 m. Site b was located in New Hampshire, US (Figure 1b) where the altitude ranges from 90 to 280 m; the land cover is homogeneous, mainly coniferous forests with heights varying from 15 to 35 m. Site c was located in New Jersey, US (Figure 1c) where the terrain is relatively flat with an average elevation of about 200 m; forest resources are quite rich; and the forest type is mostly deciduous. Due to the rich forest resources, site c has become a widely used research site for biomass estimation [36].



**Figure 1.** GEDI footprints at three sites. (a) Site a was located in Quintana Roo, Mexico; the base map of this site comprises the elevation information from SRTM data. (b) Site b was located in New Jersey, US; the base map of this site comprises the elevation information provided by airborne lidar observations. (c) Site c was located in New Hampshire, US; the base map of this site is a biomass map derived from airborne lidar observations.

### 2.2. GEDI Data

GEDI was successfully launched at the end of 2018 and mounted on the ISS, and it began acquiring and processing science data in April 2019 [26]. Based on the collected ~25 m spatial resolution GEDI full-waveform lidar data, several products have been produced and released, including footprints and gridded data. These products were divided into four different levels depending on the degree of data processing: level 1: raw waveform data and geographic location information; level 2: footprint-level canopy height and percentile heights; level 3: grid forest structure parameters; and level 4: biomass products. The GEDI level 1 data product (i.e., L1) was further developed into two separate products: level 1A (L1A) and level 1B (L1B). They all include the raw waveform and geolocation information; the difference between them is that the L1A data product contains fundamental instrument engineering and housekeeping data. L1B was used in this study (Table 1).

Furthermore, the L1B product includes the quality flag parameter to assist the user in performing quality filtering—we removed GEDI shots flagged as unsuitable for measuring the canopy structure, i.e., the “stale\_return\_flag” was not equal to 1, and the “degrade” was less than 1. We used GEDI data obtained on days 112, 165, and 191 in 2019, and the local particle IDs for these data were as follows:

- (1) GEDI01\_B\_2019112201147\_O02034\_T01337\_02\_003\_01
- (2) GEDI01\_B\_2019165230622\_O02859\_T04183\_02\_003\_01
- (3) GEDI01\_B\_2019191095142\_O03254\_T00065\_02\_003\_01

**Table 1.** Key information extracted from GEDI product used in this study.

Products	Attributes	Record Name in GEDI File
L1B	Unique shot identifier	shot_number
	Latitude	latitude_bin0
	Longitude	longitude_bin0
	Number of sample intervals of received waveform	rx_sample_count
	Starting address of received waveform	rx_sample_start_index
	Record file of received waveform	rxwaveform
	Height of start of received waveform, relative to WGS-84 ellipsoid (ellipsoidal height)	elevation_bin0
	Quality level—indicates that a “stale” cue point from the coarse search algorithm is being used	stale_return_flag
	Quality level—greater than zero if shot occurs during a degrade period, zero otherwise	degrade
	Starting address of transmitted waveform	tx_sample_start_index
	Number of sample intervals of transmitted waveform	tx_sample_count
Starting address of transmitted waveform	txwaveform	

### 2.3. LVIS Data

We used airborne LVIS full-waveform lidar data as the reference to evaluate the performance of GEDI observations in estimating percentile height metrics and ground elevation. LVIS is an imaging laser altimeter system that uses a 1064 nm wavelength laser and three detectors to detect the ground surface topography and 3D structure [14]. In May–June 2019, NASA used the LVIS equipment to conduct an observation mission in the southeastern United States and Central America; the purpose was to provide calibration and validation for GEDI. We used the LVIS data collected in Central America from the above mission to conduct our research; the nominal footprint diameter was about 25 m, and the acquisition date was 14 June 2019.

### 2.4. G-LiHT Data

G-LiHT data were used to evaluate the GEDI performance in estimating percentile height metrics, ground elevation, and GEDI geolocation error. G-LiHT collects the scanning lidar data with the Rigel VQ-480 airborne laser scanning instrument. At a standard working height of 335 m, the G-LiHT ranging accuracy can reach 25 mm and the ground resolution is 0.1 m [9]. Here, we used the G-LiHT 3D-derived point cloud data collected in New Hampshire, USA on 14 June 2017.

### 2.5. SRTM Data

Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global elevation data were used as other reference data to analyze the accuracy of GEDI in detecting ground elevation information. The SRTM 1 Arc-Second Global elevation data are the processing result of C-band InSAR observations, with a spatial resolution of 30 m [37]. Regarding vegetation-free areas, the radar phase center should correspond to the ground surface. In vegetated landscapes, such as forests, the phase center is usually located within the canopy [38]. Since the elevation from SRTM data is referenced to the Earth Gravitational Model (EGM 96)

geoid, it is the orthometric height. To make sure SRTM elevation had the same reference as GEDI elevation, we used the MATLAB 2021a (Math Works) *egm96* geoid function to convert it to ellipsoidal height.

### 2.6. *GlobeLand30 Land Cover Product*

We used the 2020 version of *GlobeLand30* land cover product to identify and exclude non-forest-covered GEDI footprints. *GlobeLand30* is a source of high-spatial-resolution global land cover data developed by the National Geomatics Center of China (NGCC) with a spatial resolution of 30 m; it is derived from Landsat, Chinese HJ-1, and GF-1 images [39]. The evaluation of *GlobeLand30 V2020* led by the Aerospace Information Innovation Institute of the Chinese Academy of Sciences showed that the overall accuracy of the 2020 version reached 85.72%, and the kappa coefficient was 0.82 [40].

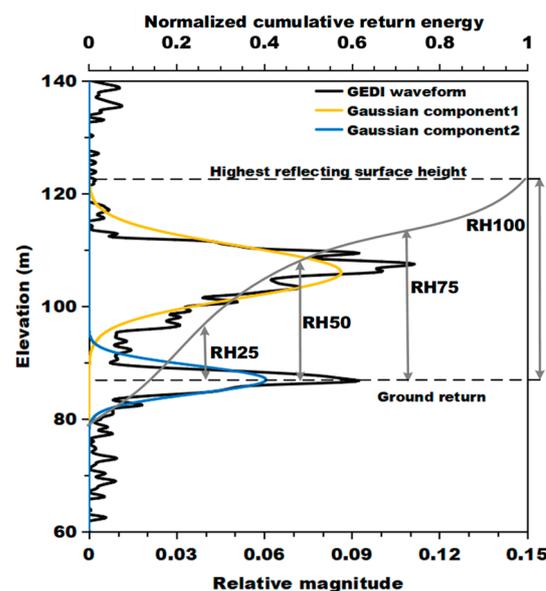
### 2.7. *Carbon Monitoring System*

To assess the ability of GEDI on estimating biomass, we introduced spatially continuous biomass data. This biomass data were provided by the NASA Carbon Monitoring System (CMS) program, estimated by using leaf-off lidar data and gridded by a spatial resolution of 30 m [41]. In terms of the accuracy of this biomass data, the ground validation results showed that the root mean square error (RMSE) was 89.3 Mg/ha, and the validation based on the US Forest Inventory data showed that this product could achieve an accuracy of RMSE = 58.2 Mg/ha [41]. The time range of this biomass data corresponds to the period from 2004 to 2014, and we selected the products of 2014 to conduct our research.

## 3. Methods

### 3.1. *GEDI and LVIS Data Processing*

The ground position is one of the key factors to accurately estimate biomass from full-waveform lidar data. Here, the ground position was identified from full-waveform lidar data based on a Gaussian decomposition method. The processing steps mainly included Gaussian filtering, Gaussian decomposition, and the identification of the ground position (Figure 2). Based on previous research conclusions [42], we took the peak position of the last Gaussian decomposition result as the ground position.



**Figure 2.** Process of estimating terrain elevation and percentile height metrics from raw lidar received waveform.

Identifying noise thresholds is key for waveform lidar data analyses in order to determine the effective waveform range. It is usually calculated separately for the start and end of the waveform because generally the noise level is lower at the start of the signal than at the end [43]. The equation for calculating noise threshold ( $nt$ ) is as follows:

$$nt = Mean + n \times Std \quad (1)$$

where  $Mean$  and  $Std$  are the mean and standard deviation of the waveform noise; and  $n$  is an empirical value. For the start of the noise threshold,  $Mean$  and  $Std$  were calculated based on the first 50 bins of the waveform; and for the end they were calculated based on the last 50 bins;  $n$  was set equal to 4 based on previous research [44].

Based on the waveform noise thresholds, the effective waveform was located. We then used the effective waveform to calculate the 25th, 50th, 75th, and 100th percentile heights (RH25, RH50, RH75, and RH100), as shown in Figure 2. First, the total waveform energy was calculated by summing the waveform energy from the beginning to the end of the signal. Afterward, the RH25, RH50, RH75, and RH100 were calculated by comparing the corresponding accumulated percentile energy with total energy. All these percentile heights refer to the ground position.

### 3.2. G-LiHT Lidar Data Processing

The accuracy of measured forest vertical structure is a critical factor that determines the accuracy of biomass estimation. To evaluate the performance of GEDI on the measurement of forest vertical structure, we introduced G-LiHT airborne discrete point cloud data; however, the data type was not consistent between the two. To achieve comparability between G-LiHT point cloud data and GEDI full-waveform data, we took a strategy to convert the G-LiHT point cloud data into full-waveform data. Details of our strategy are as follows.

Measurements of full-waveform lidar depend on the characteristics of the instrument, such as the intensity of laser emission energy, the shape of the emitted laser pulse, and the resolution of the receiver. The GEDI lidar receives the emitted Gaussian pulse to achieve target observation, with a time resolution of 1 ns. Therefore, the energy distribution of the GEDI waveform in the horizontal and vertical directions is satisfied by Gaussian distribution, as shown in Figure 3. In the horizontal distribution, the energy distribution on the ground footprint can be approximately expressed as:

$$W_h(x, y) = \exp\left\{-\left[(x - x_0)^2 + (y - y_0)^2\right]/2\sigma^2\right\} \quad (2)$$

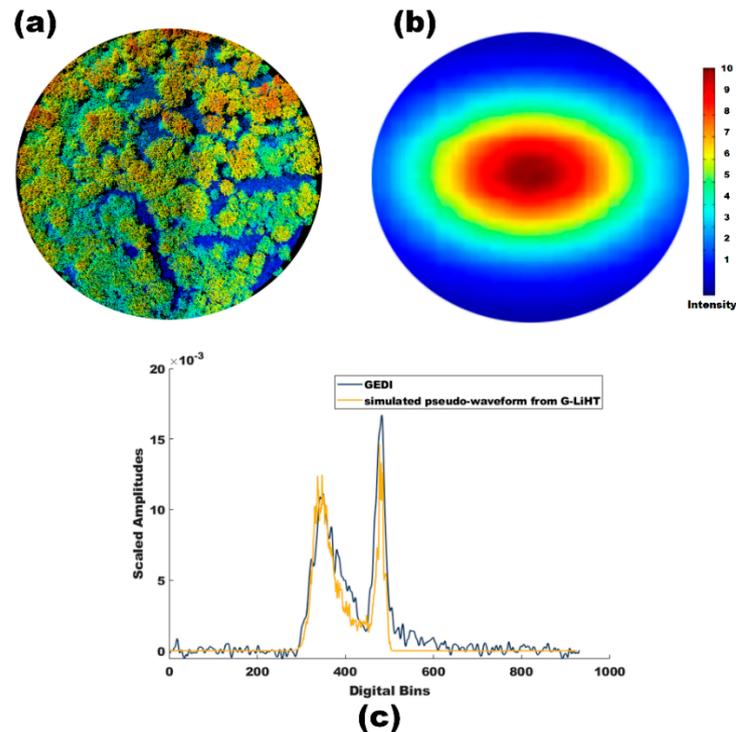
where  $(x_0, y_0)$  is the centroid point of the footprint, and  $\sigma$  is the width of the footprint ( $\sigma$  is related to footprint radius by  $r = 2\sigma$ ). In terms of the vertical distribution, the energy distribution is usually determined by the upward reflected energy intensity of the basic observed element; here, we used the reflection intensity information of the point cloud data to represent this item.

A waveform can be simulated from point cloud data by convoluting the horizontal and vertical energy distribution weighting function [45]. In the simulation process, we used an effective field of view (FOV) 126 m in diameter, which was determined by the orbital height and the FOV of the GEDI (diameter =  $2 \times H \times \tan(\text{FOV}/2)$ ). Overall, given any airborne lidar laser echoes  $(x_i, y_i, z_i)$ , the simulated waveform can be expressed as:

$$WV(z) = \left[ \sum_{i \in U} I_i \times W_h(x_i, y_i) \right] \times Int(z) \quad (3)$$

$$U = \left\{ i : \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \leq 126, |z_i - z| \leq \frac{\Delta h}{2} \right\} \quad (4)$$

where  $\Delta h$  is the waveform height interval—it was set to 0.15 m, consistent with GEDI observations;  $U$  is the collection of airborne lidar points that can be seen by the GEDI telescope; and  $i$  is a weighting parameter that represents the attributes of the  $i$ th point, which is determined by partial hits. All points can be weighted equally ( $i = 1$ ) when ignoring partial hits. Previous studies have documented that partial hits can be ignored when the footprint radius of point cloud data is small enough [34,46]. The footprint radius of G-LiHT airborne lidar data is around 0.1 m; thus, we could reasonably assume that partial reflections were almost non-existent in our application scenarios.



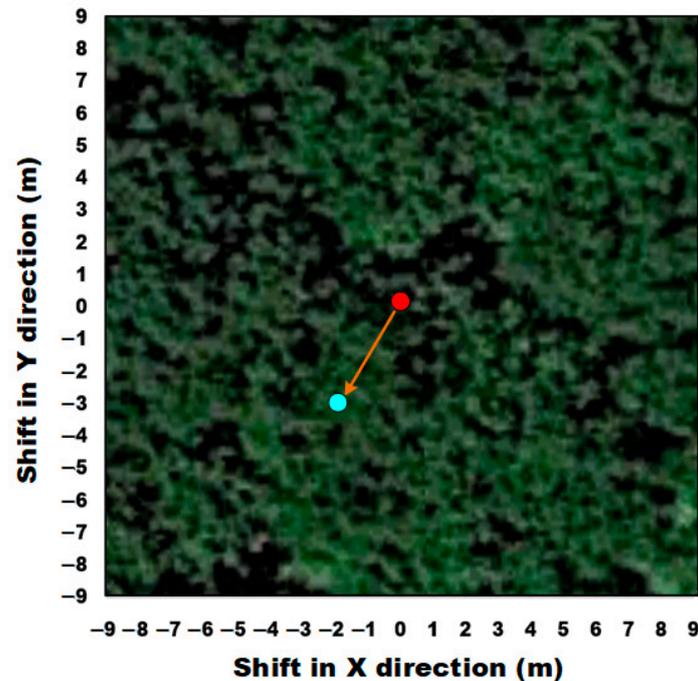
**Figure 3.** Example of simulating the waveform from G-LiHT airborne discrete lidar data. (a) Distribution of airborne lidar points within a GEDI footprint. (b) Distribution of the Gaussian pulse energy, which was used as the weighting term for simulating the waveform. (c) Simulated waveform based on distribution of the Gaussian pulse energy, as compared to the observed GEDI waveform.

Following the above steps, we obtained a simulated waveform to represent the vertical profile of the airborne lidar points frequency within the GEDI footprint. Although the simulated waveform cannot completely represent the real GEDI lidar receiving waveform, it can effectively describe the vertical distribution indicated by reflecting surfaces in the forest canopy on a large footprint scale. The true waveform represents the spatial distribution of the laser beam intensity along the entire beam path [45]. Therefore, our simulated waveform was expected to be extremely similar to the true GEDI waveform. The simulated waveform was used to conduct the evaluation with the GEDI-observed waveform.

### 3.3. Assessing the Uncertainty of GEDI Geolocation Information

The GEDI science team states that the geolocation data provided by the GEDI product are subject to random errors, with an error of about  $\pm 9$  m [35]. This uncertainty factor will affect the effectiveness of the GEDI observation in biomass estimation to a certain extent. In this study, we adopted a strategy to quantitatively evaluate GEDI geolocation errors to address this uncertainty by referring to previous research [34]. Our strategy assumed that the geolocation information from the airborne lidar was correct since the georeferencing errors of the airborne lidar can be negligible compared to those of GEDI. The strategy mainly included two steps (Figure 4): step one was to shift the center of the footprint along

the x-axis and y-axis in a  $(-9, 9)$ -m window with an interval of 1 m; step two was to take the correlation between the GEDI waveform and the simulated waveform as an evaluation criterion to locate the potentially correct geolocation of the GEDI footprint centroid.



**Figure 4.** Example of locating the potentially correct geolocation of the GEDI footprint centroid by shift-searching procedure. Red point is nominal location obtained from GEDI product; cyan point is the potentially correct geolocation of GEDI obtained through shift-searching procedure; simulated waveform of this point has the best correlation with GEDI waveform.

### 3.4. Estimating AGB Using Random Forests

We used the Random Forests machine learning approach to model AGB and assess the ability of GEDI-acquired forest vertical structure and ground elevation metrics for estimating AGB across our study sites. We selected the Random Forests machine learning method as many studies have shown this model's good performance in biomass prediction [47,48]. In the training process of the Random Forests, we used CMS AGB data as the objective variable, which were extracted based on the pixel's value at the center position. Four typical percentile heights (RH25, RH50, RH75, and RH100) from GEDI observation were used as input for biomass estimation, all calculated referring to the ground position. Furthermore, we used the 10-fold cross-validation strategy for Random Forests training, in which we randomly divided the collected data into ten equal-sized subsamples. Of the ten subsamples, one was retained as the validation data, and the remaining subsamples were used as training data. We also used the Gini index to assess the importance of each percentile height of GEDI to the predicted biomass. The Gini index, which is widely used to analyze the sensitivity of the input variables in the Random Forests model to the output [49], is an indicator of node purity; the smaller the Gini index, the higher the purity of the node. Higher node purity usually means the corresponding variable is sensitive to the output.

### 3.5. Accuracy Assessment

The basis for assessing the accuracy of GEDI observations in forest vertical structure, ground elevation, and biomass consisted of three statistical items: mean bias, the coefficient of determination ( $R^2$ ), and the root mean square error (RMSE), which are shown as follows:

$$R^2 = \frac{\sum_{i=1}^n (V_i - \bar{V})^2 - \sum_{i=1}^n (V_i - \bar{V}_i)^2}{\sum_{i=1}^n (V_i - \bar{V})^2} \quad (5)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (V_i - \bar{V}_i)^2}{n}} \quad (6)$$

$$\text{Bias} = \frac{\sum_{i=1}^n (V_i - \bar{V}_i)}{n} \quad (7)$$

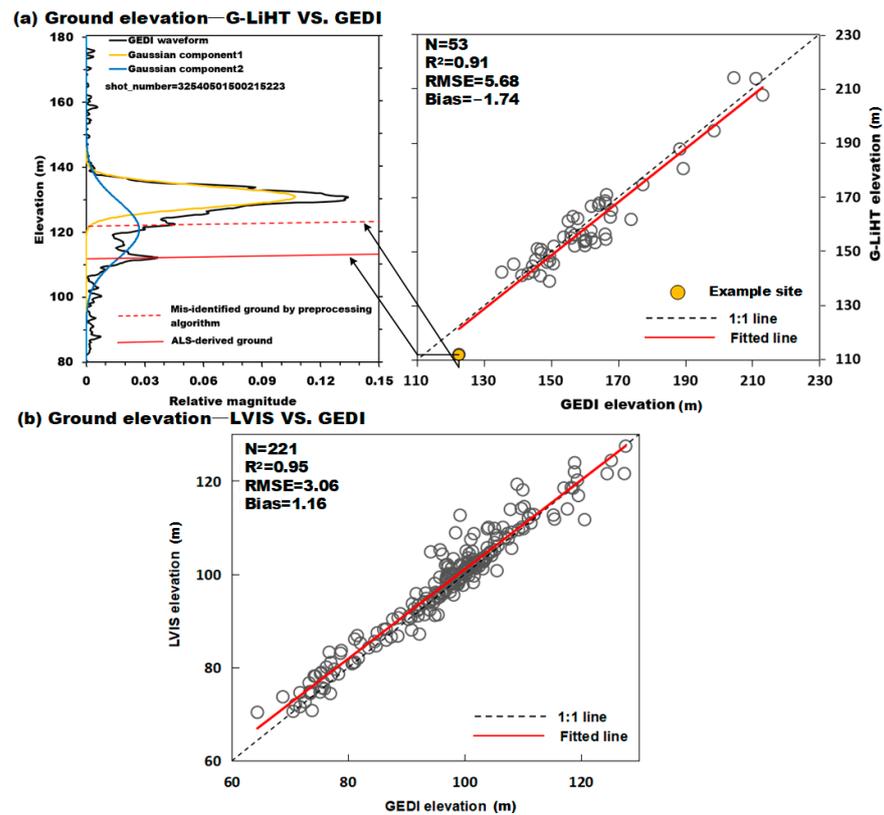
where  $V_i$  is the result extracted from the GEDI observation,  $\bar{V}_i$  is the reference value,  $\bar{V}$  is the average result of reference values, and  $n$  is the amount of data involved in the comparison.

## 4. Results

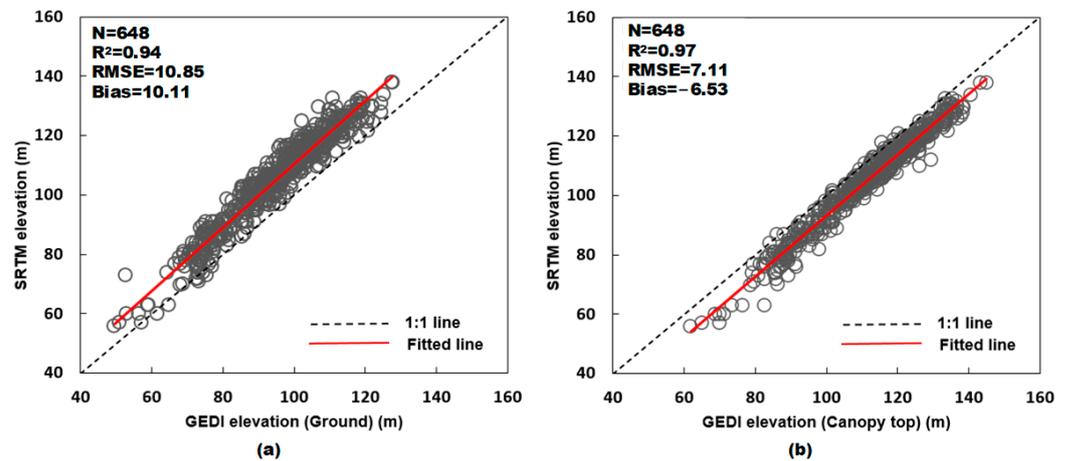
### 4.1. Ground Elevation Assessment

We used airborne lidar (LVIS and G-LiHT) data and SRTM data to evaluate the ability of GEDI to measure the ground elevation across our study sites (Figure 5). The ground elevation information provided by GEDI and G-LiHT was strongly correlated, with  $R^2 = 0.91$ , RMSE = 5.68 m, and bias =  $-1.74$  m (Figure 5a). The GEDI- and LVIS-lidar-derived ground information also yielded a strong correlation, with  $R^2 = 0.95$ , RMSE = 3.06 m, and bias = 1.16 m (Figure 5b). Overall, the consistency of the comparison results shows the stable performance of the GEDI lidar system. However, it can be noted that the comparison data are distributed on both sides of the 1:1 line (Figure 5a,b), while there are some obviously inconsistent points. These inconsistent points may mostly be caused by inaccurate recognition of the ground position in the data preprocessing step (Figure 5a).

Comparison of the GEDI ground elevation with SRTM data also showed a strong correlation, with  $R^2 = 0.94$  (Figure 6a). However, there was an overall bias of 10.11 m, indicating that the elevation from SRTM data was higher than that from GEDI. We also used the elevation of the canopy top position from the GEDI product and compared it with the SRTM elevation; there was an overall bias of  $-6.53$  m (Figure 6b). Combining the two comparison results (Figure 6), we can determine that the elevation provided by SRTM data is located at approximately 39% of the distance from the canopy top to the ground, on average.



**Figure 5.** Comparison scatter plots of the ground elevation measured by GEDI and airborne lidar. Comparison data are distributed on both sides of the 1:1 line. (a) Comparison scatter plot between GEDI and G-LiHT; the yellow point is an example to show the reason for these inconsistent comparison points. (b) Comparison scatter plot between GEDI and LVIS.



**Figure 6.** Comparison scatter plots of elevation provided by GEDI and SRTM: (a) ground elevation between SRTM and GEDI, and (b) ground elevation and GEDI canopy top elevation.

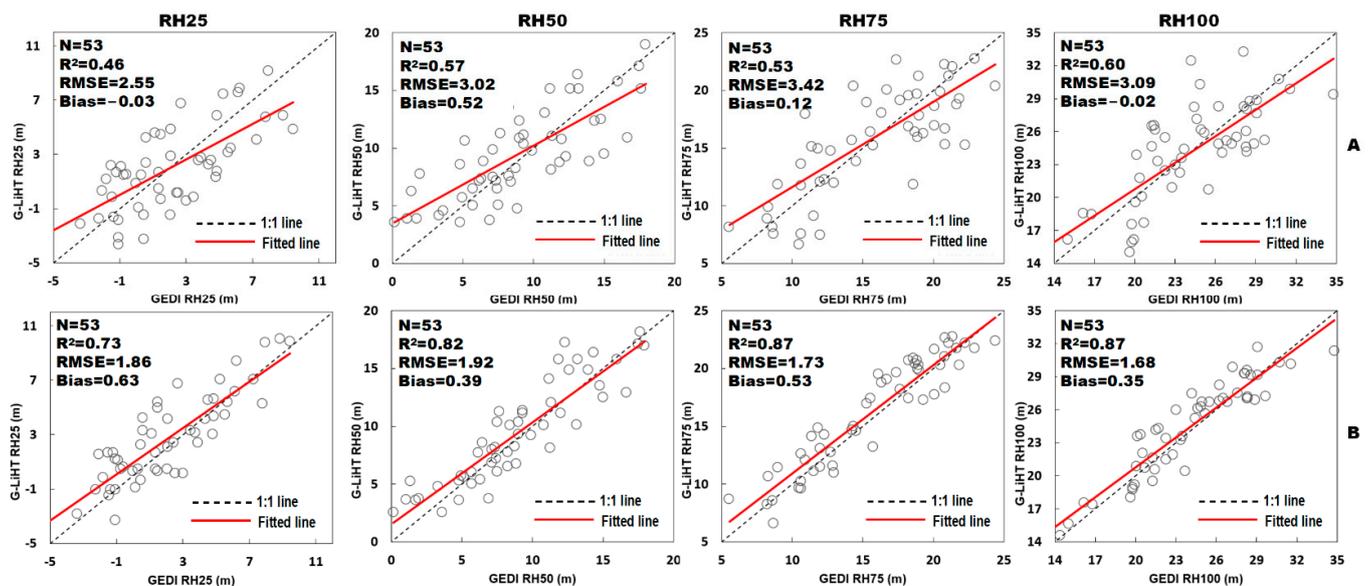
#### 4.2. Percentile Height Assessment

Percentile heights of the full-waveform lidar data indicate the density of forest canopy elements at a certain height in the forest area. The combination of different percentile heights can describe the vertical distribution of canopy elements, such as RH25, RH50, RH75, and RH100; these kinds of canopy vertical height metrics are key factors in determining the forest biomass. In this section, we used the RH25, RH50, RH75, and RH100 percentile heights obtained from G-LiHT and LVIS airborne lidar data as the reference to

assess the values obtained by GEDI. The evaluation result can provide a guideline for using GEDI observation to estimate forest biomass.

#### 4.2.1. Comparisons of Percentile Heights between GEDI and G-LiHT

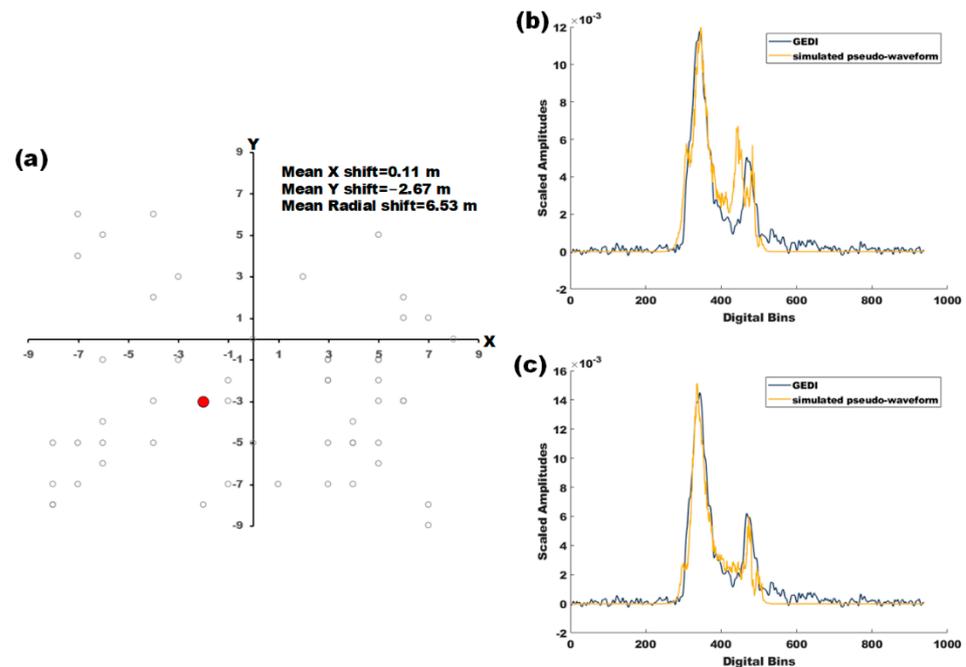
A moderate correlation of percentile heights was found between GEDI and G-LiHT (Figure 7). Percentile height RH100 has the relatively best correlation, with  $R^2 = 0.60$ , RMSE = 3.09 m, and bias =  $-0.02$  m. Percentile height RH25 has the relatively worst correlation, with  $R^2 = 0.46$ , RMSE = 2.55 m, and bias =  $-0.03$  m. The scatter plots in Figure 7 show that percentile heights from these two data sources are not consistent. For instance, data points are loosely distributed on both sides of the 1:1 line. It can be observed how the spread of the points around the trend line increases as the height decreases. This phenomenon resulted from the occlusion of mid- and understory vegetation by upper vegetation, preventing their being appropriately captured by the lidar sensors, particularly for the discrete-return sensor from which the pseudo-waveforms were simulated. Furthermore, RH25 yielded negative values in some cases, mostly due to the lidar work principles. The GEDI lidar emitted a Gaussian waveform, so there will be a weakening process after the emitted energy touches the ground. In other words, waveform energy still exists below the ground position (i.e., the last Gaussian peak). We calculated the percentile height metrics by starting the integration from the end position of the energy, so 25% of the integrated energy is usually lower than the ground position in the case of shrubland or young regenerating forests, and there will be a negative result.



**Figure 7.** Comparisons of four typical percentile heights (RH25, RH50, RH75, and RH100) provided by GEDI and G-LiHT lidar observations: graphs in row (A) use data without the geolocation error revised; graphs in row (B) use data that have undergone geolocation revision.

The GEDI product instructions documented random errors of about 9 m in the geolocation information provided by the products. Here, we used a shifting strategy (as described in Section 3.3) to identify potential geolocation errors. The potential geolocation error of GEDI was estimated to be 6.53 m, and the specific information on the shifting process is shown in Figure 8a. Furthermore, we provide an example to show how the waveform changed when undergoing the shifting process (Figure 8b,c)—which is obvious considering that the geolocation error can significantly improve the correlation between GEDI and airborne lidar observation. We recalculated the percentile heights of the simulated airborne lidar waveform with explicit consideration of the geolocation error, which in turn were used to compare with GEDI. The comparison results showed an increased correlation;

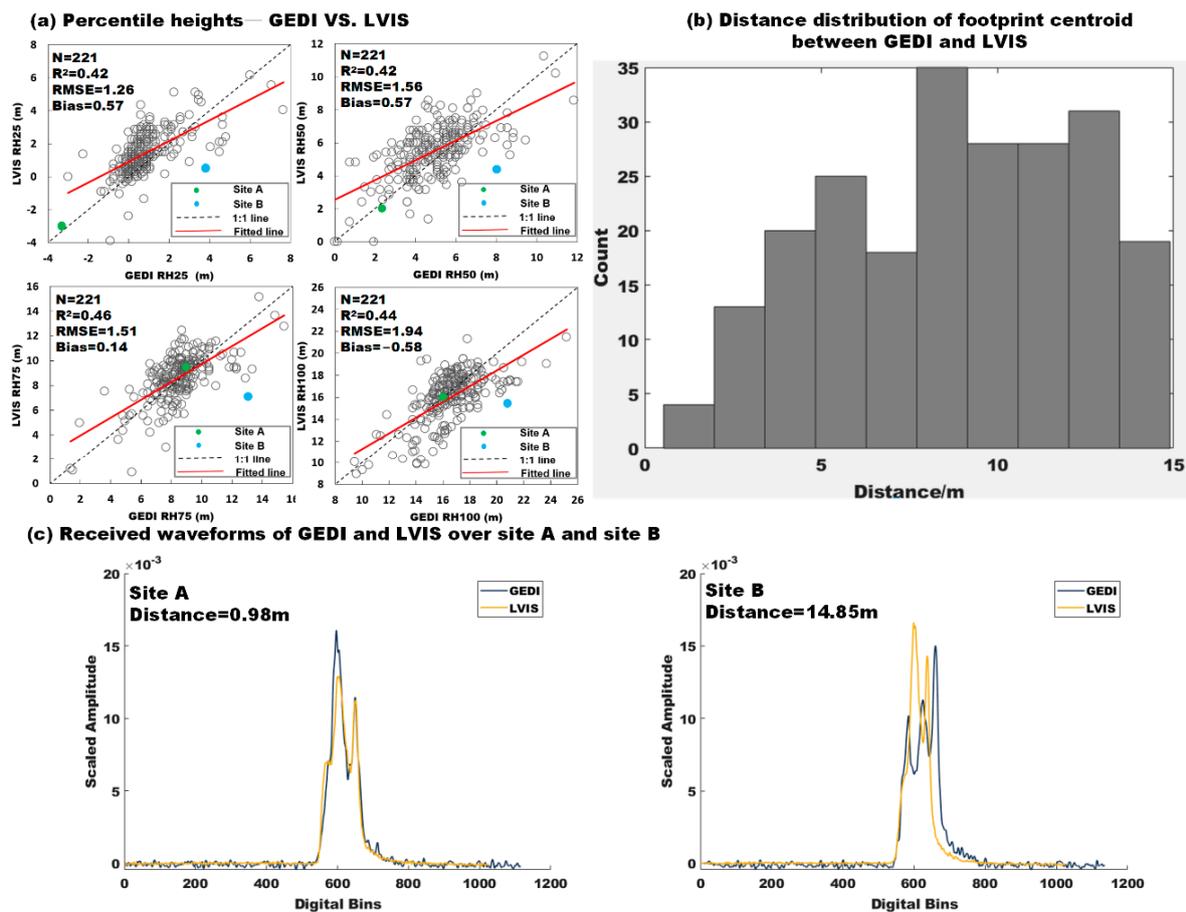
for example, the correlation of RH50 increased from 0.57 to 0.82 (Figure 7). The RMSE of four different percentile heights yielded less than 2 m, indicating a strong consistency between GEDI and G-LiHT observation. The above results illustrate the impact of geolocation errors on GEDI observations, which may increase the uncertainty in applying GEDI, such as modeling forest biomass. In addition, it should be noted that these estimated potential geolocation errors may not represent true errors due to some specific factors, such as the similarity in land cover that may cause a failure in the evaluation result based on our strategy.



**Figure 8.** Diagram of shift-searching process to estimate the geolocation error of the GEDI: (a) is the potential geolocation error of the used GEDI data in site b; (b) is the comparison of the GEDI received waveform and simulated waveform derived from G-LiHT data; here, we used the footprint centroid of GEDI without the geolocation error revised; and (c) is also the comparison of the waveform between GEDI and G-LiHT; it should be noted that here the used footprint centroid of GEDI has the geolocation error revised, the red point shown in (a) is the offset of the footprint centroid.

#### 4.2.2. Comparisons of Percentile Heights between GEDI and LVIS

LVIS full-waveform lidar data were also used as a reference to measure the ability of GEDI to obtain forest vertical structure. Figure 9a presents comparison results of percentile heights from LVIS and GEDI, showing a moderate correlation between the two (RH25:  $R^2 = 0.42$ ; RH50:  $R^2 = 0.57$ ; RH75:  $R^2 = 0.46$ ; RH100:  $R^2 = 0.44$ ). In theory, they should have a better relationship than the results show for the following reasons: firstly, the footprint size of the collected LVIS is almost the same as that of GEDI; secondly, the lidar sensors of LVIS and GEDI work similarly [46]; thirdly, we applied the same data preprocessing algorithm to them. However, the comparison result is not as good as expected, mostly because of the geolocation mismatch of the observed footprints. At our comparison site, almost no footprints between GEDI and LVIS completely overlapped. Figure 9b presents the distance between the footprint centroids of the LVIS and GEDI data. The maximum distance between the two centroids is about 15 m, which is close to the radius of GEDI. In other words, almost half the objects observed by LVIS and GEDI do not overlap.



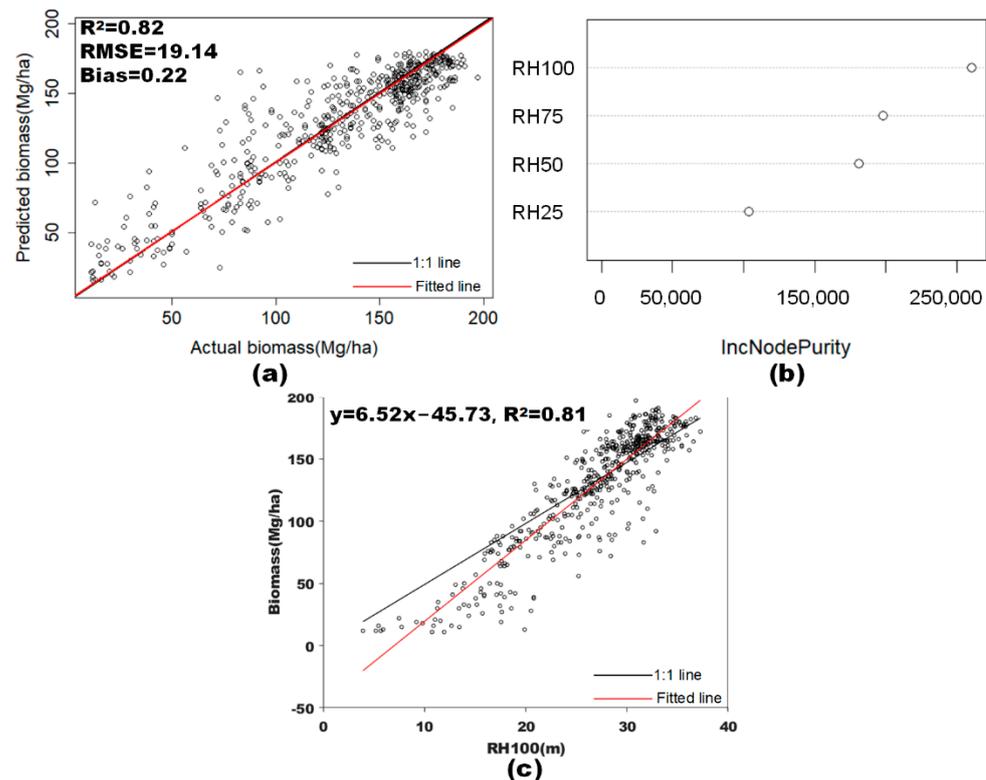
**Figure 9.** (a) Comparison of four typical percentile heights (RH25, RH50, RH75, and RH100) derived from GEDI and LVIS lidar data; (b) distance distribution of footprint centroids between GEDI and LVIS; (c) set of comparison cases used to explain the effect of inconsistency of observations on comparison result.

We selected two typical sites to further explain the effect of the inconsistency of observations on our comparison results (Figure 9c). As the LVIS and GEDI observations showed an almost coincident footprint centroid at site a, the received lidar waveforms were the same for both (Figure 9c). Therefore, percentile height metrics derived from site a showed good agreement (Figure 9a). For site b, however, the footprint centroids between LVIS and GEDI were not coincident, with a 14.85 m distance between them (Figure 9c); the received lidar waveform and percentile height metrics from site b showed a relatively significant bias (Figure 9a). The above evidence can explain the impact of geolocation differences between GEDI and LVIS on our validation analysis to some extent.

Although the significance of using LVIS data to conduct the evaluation analysis for GEDI has been reduced to a certain extent due to the inconsistent geolocation between LVIS and GEDI observation, our comparison can also reflect the reliability of GEDI observation. There are two reasons that can support our view. First, although the correlation between LVIS and GEDI is not very strong, the RMSE and bias are very small: the maximum RMSE and bias are 1.94 m and  $-0.58$  m, respectively; considering that the land cover in our study is homogeneous, this result can illustrate the consistency between GEDI and LVIS observation. Second, our example shows that sites with consistent geolocation between GEDI and LVIS show good agreement on forest structure observation.

#### 4.3. Aboveground Biomass Estimation

We used the Random Forest method to explore the role and performance of GEDI observations on biomass estimates; the ground elevation and percentile heights verification analysis in Sections 4.1 and 4.2 further motivated this analysis. Figure 10 illustrates the biomass estimated based on four typical percentile heights (i.e., RH25, RH50, RH75, and RH100); the result showed a good agreement between predicted and reference biomass, with  $R^2 = 0.82$ ,  $RMSE = 19.14$  Mg/ha, and  $bias = 0.22$  Mg/ha.



**Figure 10.** Performance of GEDI observations on biomass estimation: (a) comparison scatterplot of predicted and reference biomass; (b) role of percentile heights in biomass prediction; (c) regression results for biomass estimated based on RH100.

The advantage of Random Forests is that it can evaluate the importance of input variables to the output result. Figure 10b shows the importance of each percentile height calculated from the node purity of the Random Forests in predicting biomass. Four typical percentile heights all played a crucial role in biomass estimates, of which RH100 was the variable most sensitive to the variation in biomass. Table 2 shows the ability of each percentile height to estimate AGB, which was calculated based on the linear regression method. The  $t$ -test was used to test if the regression equations for the biomass estimates were statistically significant—the  $p$ -value of 0.00 was less than the default significance level of 0.05, indicating a significant linear regression relationship between the response biomass and the predictor variables. Results showed that the RH100 was able to explain 81% of the variance associated with the reference biomass; the RH25 predicted the relatively worst biomass, which was able to explain 35% of the variance in biomass.

It can be noted that biomass estimates in the low–medium part are not as good as those in the high part (Figure 10a). The main reason for the inaccurate prediction of low and medium biomass is due to the difference in observation times between airborne lidar and GEDI observations. More specifically, the reference biomass data we used was from the airborne lidar observation in 2014, while the GEDI lidar data was obtained in 2019. Biomass in the range of a low-to-medium value usually corresponds to young and

middle-aged forests, and biomass with a high value corresponds to mature forests. Young and middle-aged forests tend to grow vigorously, while mature forests grow slowly [50]. Therefore, there would be a significant increase in tree height for young and middle-aged forests from 2014 to 2019; however, the tree height variety in mature forests might be minimal. This tree growth pattern is consistent with our evaluation results for biomass. Furthermore, regression results based on RH100 for biomass showed that the biomass was underestimated when the tree height was less than 20 m (Figure 10c). This result can further explain the rationality of our conjecture. Although the biomass maps we used are inconsistent with the GEDI observations in terms of time, they are coincident and matched in space and scale. Furthermore, using this kind of biomass map as the “true value” can avoid representative problems with ground-based reference data, i.e., whether the ground-observed result can represent the footprint of GEDI.

**Table 2.** The regression coefficients and root mean square error:  $AGB = a * GEDI\text{-derived RH} + b$ , where a is the slope and b is the intercept of the linear model.

GEDI-Derived RH	a	b	R <sup>2</sup>	t-Test
RH25	4.42	97.03	0.35	0.00
RH50	4.81	59.95	0.56	0.00
RH75	5.24	29.77	0.67	0.00
RH100	6.52	−45.73	0.81	0.00

## 5. Discussion

### 5.1. Limitations of the Current Lidar Waveform Processing Method for Ground Detection

In the step of full-waveform lidar data processing, ground detection is one of the most challenging parts, especially for the observations over a complex canopy structure or steep terrain areas. A complex canopy structure leads to the received waveform containing multi-peaks, and the steep slope leads to the waveform near the ground being extended; these two factors will make higher peaks above the ground position in the waveform easily mistaken for the ground. The above situation was the main reason for the error in the ground elevation extracted from the full-waveform lidar data, such as GEDI, which also appeared in our study. However, there were almost no suitable automatic detection algorithms for a complex waveform that could accurately identify the ground position. Although some researchers have used a manual extraction strategy to identify the ground position from the waveform data to reduce the possible errors of the automatic detection algorithm [34], this strategy was not suitable to apply across the large spatial domain. Furthermore, the manual extraction strategy is susceptible to interference from subjective factors in actual operation. Overall, efforts are still required to extract accurate ground position information from full-waveform lidar data.

### 5.2. Difference in Ground Elevation between SRTM and GEDI

In the comparison between SRTM and GEDI ground elevation, there is an inconsistent result in that elevation information provided by SRTM data is located within the canopy from approximately 39% of the distance from the canopy top to the ground on average. This difference may result from the different characteristics of lidar and InSAR. SRTM elevation is derived from the C-band InSAR observation, and thus the detected so-called ground position is the scattering center. In forested areas, the scattering center is usually located within the canopy, not the ground, due to the limited penetration ability of InSAR [51–53]. However, the lidar has a better penetration ability, which can detect the ground position in forest areas [54,55]. A previous study documented that there was a 40% overall bias for comparison between SRTM data and lidar-detected ground elevation in forest areas [56], which is close to the experimental results of this study. Overall, GEDI has better performance in measuring ground elevation than InSAR equipment, especially in dense forest areas. Accordingly, GEDI is expected to be an effective means to revise the error of the

ground elevation in dense forest areas from the InSAR observation, which can further help improve the accuracy of forest biomass estimation.

### *5.3. Role of Lidar Emitted Energy for Canopy Detection*

Some inconsistent points appeared in the comparison between LVIS and GEDI in terms of percentile height metrics. Although the geolocation error of GEDI is one reason for this inconsistency, as described in the Section 4, the difference in the emitted power of the lidar equipment also needs to be mentioned. The emitted power of LVIS and GEDI is not the same in terms of not only GEDI's coverage lasers but also its full-power lasers [57]. The difference in the emitted power of lidar will lead to a different capability to reach the bottom of the canopy, and this will have a certain effect on the obtained percentile height metrics. How the emitted power of lidar affects the derived height metrics, which in turn affect biomass estimation, is a topic for investigation in the next step of research.

### *5.4. Selection of Percentile Heights in Biomass Prediction*

Previous studies have documented that percentile height RH50, also known as the height of mean energy (HOME) of the waveform, was the most sensitive percentile height variable to the variation in biomass [58]. Our experimental results differ from previous studies—RH100 is the most sensitive percentile height to the biomass variation. This is primarily because our used CMS biomass map was derived from leaf-off lidar data. In the leaf-off state, the lidar mainly obtains information from thicker branches, which means the main trunk is the main object detected by the lidar. For this reason, the biomass estimated based on this kind of point cloud data should be closely related to the tree height, which explains why RH100 is the most sensitive variable to the variation in biomass in this study. From the above analysis and evidence, we can reasonably consider that the RH100 height should be more efficient than the RH50 height in leaf-off forest biomass prediction. However, it should be noted that the above conjecture may still require additional conditions to hold, such as the distribution of canopy heights in the sample plot. Therefore, clear conclusions about the performance of RH100 for leaf-off forest biomass prediction need to be further explored in subsequent studies.

## **6. Conclusions**

This study investigated the ability of GEDI-derived forest structure and ground elevation observations for estimating AGB in temperate and tropical forest ecosystems across low-to-middle latitudes of North America. An evaluation was conducted by comparing against the NASA LVIS-, G-LiHT-, and SRTM-derived products. In terms of ground elevation, the elevation obtained from GEDI showed good consistency with both LVIS and G-LiHT lidar observation. However, the SRTM elevation provided by GEDI observation was located at approximately 39% of the distance from the canopy top to the ground, on average, which occurs in dense forest conditions. For forest vertical structure, the typical percentile heights extracted from GEDI were highly correlated with airborne lidar observation, especially considering the geolocation error of GEDI. Using the evaluation results of ground elevation and forest vertical structures as guidance, we estimated forest AGB based on GEDI observations. Results showed that GEDI observations could effectively capture biomass variations, for example, RH100 percentile height extracted from GEDI could explain more than 80% of leaf-off forests' biomass variations.

GEDI, the latest generation of Earth observation lidar equipment mounted on the ISS, can provide vertical observations for forests between 50° N and 50° S on a global scale. Compared with previous full-waveform spaceborne lidar equipment, such as ICESat GLAS, GEDI has higher spatial resolution, broader spatial coverage, and denser sampling in its Earth observation capabilities—all this improved performance makes GEDI a good choice for forest biomass estimation. A better understanding of GEDI characteristics is crucial in applying GEDI observations, such as forest biomass estimation. In this study, we provide

an example to evaluate the information content of GEDI observation, and the results can serve to instruct the efficient use of GEDI observation to estimate forest biomass.

**Author Contributions:** Conceptualization, M.S. and L.C.; data curation, M.S., H.Z. and L.H.; formal analysis, M.S. and L.C.; funding acquisition, H.Z.; software, M.S. and L.C.; writing—original draft, M.S.; and writing—review and editing, L.C., K.Z., M.G., Y.Z., J.P. and C.A.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (Nos. 41971306, 42090013, and 41971288).

**Data Availability Statement:** All satellite remote sensing data used in this study are openly and freely available. GEDI data were available at <https://lpdaac.usgs.gov> (accessed on 16 March 2022). LVIS data were available at <https://nsidc.org> (accessed on 16 March 2022). G-LiHT data were available at <https://gliht.gsfc.nasa.gov/index.php?section=34> (accessed on 16 March 2022). SRTM data were available at <http://srtm.csi.cgiar.org/srtmdata/> (accessed on 16 March 2022). CMS biomass data were available at [https://carbon.nasa.gov/cgi-bin/available\\_archived\\_products.pl](https://carbon.nasa.gov/cgi-bin/available_archived_products.pl) (accessed on 16 March 2022). GlobeLand30 landcover data were available at [http://www.globallandcover.com/defaults\\_en.html?type=data&src=/Scripts/map/defaults/En/browse\\_en.html&head=browse](http://www.globallandcover.com/defaults_en.html?type=data&src=/Scripts/map/defaults/En/browse_en.html&head=browse) (accessed on 16 March 2022).

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

## References

- Chen, J.M. Carbon Neutrality: Toward a Sustainable Future. *Innovation* **2021**, *2*, 100127. [[CrossRef](#)] [[PubMed](#)]
- Nunes, L.J.R.; Meireles, C.I.R.; Pinto Gomes, C.J.; Almeida Ribeiro, N.M.C. Forest Contribution to Climate Change Mitigation: Management Oriented to Carbon Capture and Storage. *Climate* **2020**, *8*, 21. [[CrossRef](#)]
- Loaiza Usuga, J.C.; Rodriguez Toro, J.A.; Ramirez Alzate, M.V.; Lema Tapias, A.D.J. Estimation of Biomass and Carbon Stocks in Plants, Soil and Forest Floor in Different Tropical Forests. *Forest. Ecol. Manag.* **2010**, *260*, 1906–1913. [[CrossRef](#)]
- Qi, W.; Saarela, S.; Armston, J.; Stahl, G.; Dubayah, R. Forest Biomass Estimation Over Three Distinct Forest Types Using TanDEM-X InSAR Data and Simulated GEDI Lidar Data. *Remote Sens. Environ.* **2019**, *232*, 111283. [[CrossRef](#)]
- Stark, S.C.; Leitold, V.; Wu, J.L.; Hunter, M.O.; de Castilho, C.V.; Costa, F.R.C.; McMahon, S.M.; Parker, G.G.; Shimabukuro, M.T.; Lefsky, M.A.; et al. Amazon Forest Carbon Dynamics Predicted by Profiles of Canopy Leaf Area and Light Environment. *Ecol. Lett.* **2012**, *15*, 1406–1414. [[CrossRef](#)] [[PubMed](#)]
- Brede, B.; Calders, K.; Lau, A.; Raunonen, P.; Bartholomeus, H.M.; Herold, M.; Kooistra, L. Non-Destructive Tree Volume Estimation through Quantitative Structure Modelling: Comparing UAV Laser Scanning with Terrestrial LIDAR. *Remote Sens. Environ.* **2019**, *233*, 111355. [[CrossRef](#)]
- Zhao, K.; Suarez, J.C.; Garcia, M.; Hu, T.; Wang, C.; Londo, A. Utility of Multitemporal Lidar for Forest and Carbon Monitoring: Tree Growth, Biomass Dynamics, and Carbon Flux. *Remote Sens. Environ.* **2018**, *204*, 883–897. [[CrossRef](#)]
- Dubayah, R.O.; Sheldon, S.L.; Clark, D.B.; Hofton, M.A.; Blair, J.B.; Hurtt, G.C.; Chazdon, R.L. Estimation of Tropical Forest Height and Biomass Dynamics Using Lidar Remote Sensing at La Selva, Costa Rica. *J. Geophys. Res.-Biogeosci.* **2010**, *115*, G00E09. [[CrossRef](#)]
- Cook, B.D.; Corp, L.A.; Nelson, R.F.; Middleton, E.M.; Morton, D.C.; McCorkel, J.T.; Masek, J.G.; Ranson, K.J.; Vuong, L.; Montesano, P.M. NASA Goddard’s LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. *Remote Sens.* **2013**, *5*, 4045–4066. [[CrossRef](#)]
- Xing, Y.; Huang, J.; Gruen, A.; Qin, L. Assessing the Performance of ICESat-2/ATLAS Multi-Channel Photon Data for Estimating Ground Topography in Forested Terrain. *Remote Sens.* **2020**, *12*, 2084. [[CrossRef](#)]
- Corp, L.A.; Cook, B.D.; McCorkel, J.; Middleton, E.M. Data products of NASA Goddard’s LiDAR, Hyperspectral, and Thermal Airborne Imager (G-LiHT). *Next-Gener. Spectrosc. Technol. VIII* **2015**, *9482*, 333–344.
- Marrs, J.; Ni-Meister, W. Machine Learning Techniques for Tree Species Classification Using Co-Registered LiDAR and Hyperspectral Data. *Remote Sens.* **2019**, *11*, 819. [[CrossRef](#)]
- Nelson, R.; Margolis, H.; Montesano, P.; Sun, G.; Cook, B.; Corp, L.; Andersen, H.; DeJong, B.; Paz Pellat, F.; Fickel, T.; et al. Lidar-Based Estimates of Aboveground Biomass in the Continental US and Mexico Using Ground, Airborne, and Satellite Observations. *Remote Sens. Environ.* **2017**, *188*, 127–140. [[CrossRef](#)]
- Blair, J.B.; Rabine, D.L.; Hofton, M.A. The Laser Vegetation Imaging Sensor: A Medium-Altitude, Digitisation-Only, Airborne Laser Altimeter for Mapping Vegetation and Topography. *ISPRS J. Photogramm.* **1999**, *54*, 115–122. [[CrossRef](#)]
- Cui, L.; Jiao, Z.; Dong, Y.; Sun, M.; Zhang, X.; Yin, S.; Ding, A.; Chang, Y.; Guo, J.; Xie, R. Estimating Forest Canopy Height Using MODIS BRDF Data Emphasizing Typical-Angle Reflectances. *Remote Sens.* **2019**, *11*, 2239. [[CrossRef](#)]

16. Silva, C.A.; Saatchi, S.; Garcia, M.; Labriere, N.; Klauberg, C.; Ferraz, A.; Meyer, V.; Jeffery, K.J.; Abernethy, K.; White, L.; et al. Comparison of Small- and Large-Footprint Lidar Characterization of Tropical Forest Aboveground Structure and Biomass: A Case Study From Central Gabon. *IEEE J.-Stars* **2018**, *11*, 3512–3526. [[CrossRef](#)]
17. Huang, W.; Sun, G.; Dubayah, R.; Cook, B.; Montesano, P.; Ni, W.; Zhang, Z. Mapping Biomass Change After Forest Disturbance: Applying LiDAR Footprint-Derived Models at Key Map Scales. *Remote Sens. Environ.* **2013**, *134*, 319–332. [[CrossRef](#)]
18. Sun, G.; Ranson, K.J.; Kimes, D.S.; Blair, J.B.; Kovacs, K. Forest Vertical Structure From GLAS: An Evaluation Using LVIS and SRTM Data. *Remote Sens. Environ.* **2008**, *112*, 107–117. [[CrossRef](#)]
19. Wang, Y.; Li, G.; Ding, J.; Guo, Z.; Tang, S.; Wang, C.; Huang, Q.; Liu, R.; Chen, J.M. A Combined GLAS and MODIS Estimation of the Global Distribution of Mean Forest Canopy Height. *Remote Sens. Environ.* **2016**, *174*, 24–43. [[CrossRef](#)]
20. Simard, M.; Pinto, N.; Fisher, J.B.; Baccini, A. Mapping Forest Canopy Height Globally with Spaceborne Lidar. *J. Geophys. Res.-Biogeosci.* **2011**, *116*, G04021. [[CrossRef](#)]
21. Huang, H.; Liu, C.; Wang, X.; Zhou, X.; Gong, P. Integration of Multi-Resource Remotely Sensed Data and Allometric Models for Forest Aboveground Biomass Estimation in China. *Remote Sens. Environ.* **2019**, *221*, 225–234. [[CrossRef](#)]
22. Lefsky, M.A.; Harding, D.J.; Keller, M.; Cohen, W.B.; Carabajal, C.C.; Espirito-Santo, F.D.; Hunter, M.O.; de Oliveira, R. Estimates of Forest Canopy Height and Aboveground Biomass Using ICESat. *Geophys. Res. Lett.* **2005**, *32*, L22S02. [[CrossRef](#)]
23. Li, W.; Niu, Z.; Shang, R.; Qin, Y.; Wang, L.; Chen, H. High-Resolution Mapping of Forest Canopy Height Using Machine Learning by Coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 Data. *Int. J. Appl. Earth. Obs.* **2020**, *92*, 102163. [[CrossRef](#)]
24. Narine, L.L.; Popescu, S.; Neuenschwander, A.; Zhou, T.; Srinivasan, S.; Harbeck, K. Estimating Aboveground Biomass and Forest Canopy Cover with Simulated ICESat-2 Data. *Remote Sens. Environ.* **2019**, *224*, 1–11. [[CrossRef](#)]
25. Popescu, S.C.; Zhou, T.; Nelson, R.; Neuenschwande, A.; Sheridan, R.; Narine, L.; Walsh, K.M. Photon Counting LiDAR: An Adaptive Ground and Canopy Height Retrieval Algorithm for ICESat-2 Data. *Remote Sens. Environ.* **2018**, *208*, 154–170. [[CrossRef](#)]
26. 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](#)]
27. Potapov, P.; Li, X.; Hernandez-Serna, A.; Tyukavina, A.; Hansen, M.C.; Kommareddy, A.; Pickens, A.; Turubanova, S.; Tang, H.; Silva, C.E.; et al. Mapping Global Forest Canopy Height through Integration of GEDI and Landsat Data. *Remote Sens. Environ.* **2021**, *253*, 112165. [[CrossRef](#)]
28. Puliti, S.; Breidenbach, J.; Schumacher, J.; Hauglin, M.; Klingenberg, T.F.; Astrup, R. Above-Ground Biomass Change Estimation Using National Forest Inventory Data with Sentinel-2 and Landsat. *Remote Sens. Environ.* **2021**, *265*, 112644. [[CrossRef](#)]
29. Silva, C.A.; Duncanson, L.; Hancock, S.; Neuenschwander, A.; Thomas, N.; Hofton, M.; Fatoyinbo, L.; Simard, M.; Marshak, C.Z.; Armston, J.; et al. Fusing Simulated GEDI, ICESat-2 and NISAR Data for Regional Aboveground Biomass Mapping. *Remote Sens. Environ.* **2021**, *253*, 112234. [[CrossRef](#)]
30. Duncanson, L.; Neuenschwander, A.; Hancock, S.; Thomas, N.; Fatoyinbo, T.; Simard, M.; Silva, C.A.; Armston, J.; Luthcke, S.B.; Hofton, M.; et al. Biomass Estimation From Simulated GEDI, ICESat-2 and NISAR Across Environmental Gradients in Sonoma County, California. *Remote Sens. Environ.* **2020**, *242*, 111779. [[CrossRef](#)]
31. Leite, R.V.; Silva, C.A.; Broadbent, E.N.; Do Amaral, C.H.; Liesenberg, V.; Alves De Almeida, D.R.; Mohan, M.; Godinho, S.; Cardil, A.; Hamamura, C.; et al. Large Scale Multi-Layer Fuel Load Characterization in Tropical Savanna Using GEDI Spaceborne Lidar Data. *Remote Sens. Environ.* **2022**, *268*, 112764. [[CrossRef](#)]
32. Rishmawi, K.; Huang, C.; Zhan, X. Monitoring Key Forest Structure Attributes across the Conterminous United States by Integrating GEDI LiDAR Measurements and VIIRS Data. *Remote Sens.* **2021**, *13*, 442. [[CrossRef](#)]
33. Goetz, S.; Dubayah, R. Advances in Remote Sensing Technology and Implications for Measuring and Monitoring Forest Carbon Stocks and Change. *Carbon. Manag.* **2011**, *2*, 231–244. [[CrossRef](#)]
34. Popescu, S.C.; Zhao, K.; Neuenschwander, A.; Lin, C. Satellite Lidar Vs. Small Footprint Airborne Lidar: Comparing the Accuracy of Aboveground Biomass Estimates and Forest Structure Metrics at Footprint Level. *Remote Sens. Environ.* **2011**, *115*, 2786–2797. [[CrossRef](#)]
35. GEDI Project Science Office. GEDI Instrument Overview. 2022. Available online: <https://gedi.umd.edu/instrument/instrument-overview/> (accessed on 1 February 2022).
36. Huang, W.; Dolan, K.A.; Swatantran, A.; Johnson, K.D.; Tang, H.; O'Neil-Dunne, J.; Dubayah, R.; Hurtt, G. High-Resolution Mapping of Aboveground Biomass for Forest Carbon Monitoring System in the Tri-State Region of Maryland, Pennsylvania and Delaware, USA. *Environ. Res. Lett.* **2019**, *14*, 95002. [[CrossRef](#)]
37. Cowan, D.; Cooper, G. The Shuttle Radar Topography Mission—a New Source of Near-Global Digital Elevation Data. *Explor. Geophys.* **2005**, *36*, 334–340. [[CrossRef](#)]
38. Sarabandi, K.; Lin, Y.C. Simulation of Interferometric SAR Response for Characterizing the Scattering Phase Center Statistics of Forest Canopies. *IEEE Trans. Geosci. Remote* **2000**, *38*, 115–125. [[CrossRef](#)]
39. Chen, J.; Cao, X.; Peng, S.; Ren, H. Analysis and Applications of GlobeLand30: A Review. *Isprs. Int. J. Geo-Inf.* **2017**, *6*, 230. [[CrossRef](#)]
40. Jun, C.; Ban, Y.; Li, S. Open Access to Earth Land-Cover Map. *Nature* **2014**, *514*, 434. [[CrossRef](#)]

41. Dubayah, R.O.A.S. *LiDAR Derived Biomass, Canopy Height and Cover for Tri-State (MD, PA, DE) Region, V2*; ORNL DAAC: Oak Ridge, TN, USA, 2018.
42. Yang, X.; Wang, C.; Pan, F.; Nie, S.; Xi, X.; Luo, S. Retrieving Leaf Area Index in Discontinuous Forest Using ICESat/GLAS Full-Waveform Data Based On Gap Fraction Model. *Isprs. J. Photogramm.* **2019**, *148*, 54–62. [[CrossRef](#)]
43. Hofton, M.; Blair, J.B.; Story, S.; Yi, D. Algorithm Theoretical Basis Document (ATBD) for GEDI Transmit and Receive Waveform Processing for L1 and L2 Products. 2019. Available online: [https://lpdaac.usgs.gov/documents/581/GEDI\\_WF\\_ATBD\\_v1.0.pdf](https://lpdaac.usgs.gov/documents/581/GEDI_WF_ATBD_v1.0.pdf) (accessed on 1 February 2022).
44. Cui, L.; Jiao, Z.; Zhao, K.; Sun, M.; Dong, Y.; Yin, S.; Zhang, X.; Guo, J.; Xie, R.; Zhu, Z.; et al. Retrieving Forest Canopy Elements Clumping Index Using ICESat GLAS Lidar Data. *Remote Sens.* **2021**, *13*, 948. [[CrossRef](#)]
45. Blair, J.B.; Hofton, M.A. Modeling Laser Altimeter Return Waveforms Over Complex Vegetation Using High-Resolution Elevation Data. *Geophys. Res. Lett.* **1999**, *26*, 2509–2512. [[CrossRef](#)]
46. Hancock, S.; Armston, J.; Hofton, M.; Sun, X.; Tang, H.; Duncanson, L.I.; Kellner, J.R.; Dubayah, R. The GEDI Simulator: A Large-Footprint Waveform Lidar Simulator for Calibration and Validation of Spaceborne Missions. *Earth Space Sci.* **2019**, *6*, 294–310. [[CrossRef](#)] [[PubMed](#)]
47. Karlson, M.; Ostwald, M.; Reese, H.; Sanou, J.; Tankoano, B.; Mattsson, E. Mapping Tree Canopy Cover and Aboveground Biomass in Sudano-Sahelian Woodlands Using Landsat 8 and Random Forest. *Remote Sens.* **2015**, *7*, 10017–10041. [[CrossRef](#)]
48. Gleason, C.J.; Im, J. Forest Biomass Estimation From Airborne LiDAR Data Using Machine Learning Approaches. *Remote Sens. Environ.* **2012**, *125*, 80–91. [[CrossRef](#)]
49. Pandit, S.; Tsuyuki, S.; Dube, T. Estimating Above-Ground Biomass in Sub-Tropical Buffer Zone Community Forests, Nepal, Using Sentinel 2 Data. *Remote Sens.* **2018**, *10*, 601. [[CrossRef](#)]
50. Meko, D.; Cook, E.R.; Stahle, D.W.; Stockton, C.W.; Hughes, M.K. Spatial Patterns of Tree-Growth Anomalies in the United States and Southeastern Canada. *J. Clim.* **1993**, *6*, 1773–1786. [[CrossRef](#)]
51. Zhao, X.; Su, Y.; Hu, T.; Chen, L.; Gao, S.; Wang, R.; Jin, S.; Guo, Q. A Global Corrected SRTM DEM Product for Vegetated Areas. *Remote Sens. Lett.* **2018**, *9*, 393–402. [[CrossRef](#)]
52. Su, Y.; Guo, Q. A Practical Method for SRTM DEM Correction Over Vegetated Mountain Areas. *Isprs. J. Photogramm.* **2014**, *87*, 216–228. [[CrossRef](#)]
53. Hofton, M.; Dubayah, R.; Blair, J.B.; Rabine, D. Validation of SRTM Elevations Over Vegetated and Non-Vegetated Terrain Using Medium Footprint Lidar. *Photogramm. Eng. Rem. Sens.* **2006**, *72*, 279–285. [[CrossRef](#)]
54. Su, Y.; Guo, Q.; Ma, Q.; Li, W. SRTM DEM Correction in Vegetated Mountain Areas through the Integration of Spaceborne LiDAR, Airborne LiDAR, and Optical Imagery. *Remote Sens.* **2015**, *7*, 11202–11225. [[CrossRef](#)]
55. Wang, C.; Glenn, N.F. Integrating LiDAR Intensity and Elevation Data for Terrain Characterization in a Forested Area. *IEEE Geosci. Remote Sens.* **2009**, *6*, 463–466. [[CrossRef](#)]
56. Carabajal, C.C.; Harding, D.J. SRTM C-band and ICESat Laser Altimetry Elevation Comparisons as a Function of Tree Cover and Relief. *Photogramm. Eng. Rem. Sens.* **2006**, *72*, 287–298. [[CrossRef](#)]
57. Fayad, I.; Baghdadi, N.; Lahssini, K. An Assessment of the GEDI Lasers' Capabilities in Detecting Canopy Tops and their Penetration in a Densely Vegetated, Tropical Area. *Remote Sens.* **2022**, *13*, 2969. [[CrossRef](#)]
58. Drake, J.B.; Dubayah, R.O.; Clark, D.B.; Knox, R.G.; Blair, J.B.; Hofton, M.A.; Chazdon, R.L.; Weishampel, J.F.; Prince, S.D. Estimation of Tropical Forest Structural Characteristics Using Large-Footprint Lidar. *Remote Sens. Environ.* **2002**, *79*, 305–319. [[CrossRef](#)]