Forest Canopy Height Estimation Using Satellite-Borne LiDAR and Optical Remote Sensing Data

Yanan Yan*

School of Cangzhou No. 1 Secondary School, Cangzhou, Hebei, China

*Corresponding author: ff9829@outlook.com

Abstract. Forests are the main carbon sinks in terrestrial ecosystems, and accurately estimating forest canopy height is the basis for calculating terrestrial ecosystem carbon storage. Conducting large-scale forest canopy height research is crucial for monitoring changes in forest carbon stocks and assessing forest carbon sequestration capacity. This paper takes the forest in Maoershan, Heilongjiang Province, as the research object, using data from the Global Ecosystem Dynamics Investigation (GEDI) LiDAR satellite to perform forest canopy height inversion for Maoershan. The results show that the root mean square error (RMSE) value for terrain inversion accuracy in the Maoershan area is 4.49 m, and the mean absolute error (MAE) value is 3.33 m. Inversions based on GEDI data of forested terrain show that strong beams are superior to coverage beams, the daytime effect is better in humid areas, while the nighttime effect is better in arid areas. Furthermore, GEDI data are significantly more precise in determining understory terrain than in estimating canopy height. Increasing slope obviously causes a deterioration in the accuracy of GEDI data.

Keywords: Forest canopy height; Satellite-borne LiDAR; Optical remote sensing.

1. Introduction

The global issue of climate change has become a major challenge facing humanity [1], with greenhouse gases generated and emitted by human production activities leading to an increasingly severe greenhouse effect [2]. To minimise environmental impact, nations worldwide have successively set targets such as ‘carbon reduction’ and ‘carbon neutrality’. China has also established clear climate change objectives, including achieving ‘carbon peak’ by 2030 and ‘carbon neutrality’ by 2060 [3]. Forests, as vital components of terrestrial ecosystems, perform multiple ecological functions including carbon sequestration, water conservation, oxygen release, and influencing precipitation patterns [4]. They play an irreplaceable role in achieving global carbon balance and mitigating global warming [5].

Forest canopy height is the most crucial structural parameter for forests, serving to monitor both biomass and carbon stocks. Consequently, the ability to accurately and rapidly estimate and map forest canopy height is essential [6]. Traditional data acquisition methods involve field surveys, yet this approach consumes substantial resources, being costly, time-consuming and labour-intensive, rendering it unsuitable for large-scale monitoring [7]. With the advancement of remote sensing technology, optical remote sensing has become the most widely applied technical method in Earth observation and spatial information fields [8]. Utilising optical remote sensing to acquire extensive horizontal structural information of forest canopies has effectively reduced resource consumption, establishing itself as the primary source for obtaining forest canopy height data [9]. However, optical remote sensing possesses limited penetration capabilities in densely forested areas, typically yielding only surface characteristics of the forest canopy and struggling to provide detailed information on vertical forest structure [10]. Light Detection and Ranging (LiDAR) represents a modern optical remote sensing technology that employs active laser detection to acquire high-precision spatial data on forest structure and understory topography. By emitting laser energy and receiving the returned signal, it rapidly and directly captures three-dimensional spatial information of ground features [11]. Spaceborne lidar technology addresses the limitations of optical remote sensing techniques [12], providing robust technical support for extracting vertical structural characteristics of forest canopies. [13].

The world's first satellite-borne lidar system, the Ice, Cloud, and Land Elevation Satellite (ICESat-1), equipped with the Geoscience Laser Altimeter System (GLAS), has been successfully employed in scientific research on forest canopy height [14]. Due to the relatively large spot size of GLAS, it is unable to provide high-precision sampling for forest height products. Furthermore, the GLAS system ceased operations in 2009 [15], leaving no new relevant data products currently available. In November 2018, the United States deployed the full-waveform LiDAR sensor GEDI (Global Ecosystem Dynamics Investigation) aboard the International Space Station. Utilising GEDI data, a global forest canopy height model was successfully established. Research indicates that this model can yield more precise forest canopy data parameters [16]. Therefore, this paper investigates forest canopy height in the Maershan region of Heilongjiang Province based on GEDI data.

2. Data and Methods

2.1. Study Area

The study area is situated in Maoershan Town, Shangzhi City, Heilongjiang Province, on the western slope of the Zhangguangcai Ridge. Its geographical coordinates are 45°20′ to 45°25′ north latitude and 127°30′ to 127°34′ east longitude. The region experiences a temperate continental monsoon climate, characterised by long, cold, dry winters and short, warm, rainy summers. The terrain predominantly consists of low mountains and hills formed from Jurassic acidic volcanic rocks. Elevation slopes from south to north, with an average altitude of 300–400 metres. The prominent peak of Mount Mao'ershan rises abruptly, resembling a crown in shape. Rivers within the area serve as secondary tributaries of the Songhua River (e.g., the Ash River), and the region includes the Xiquanyan Reservoir (with a capacity of 500 million cubic metres). Regarding forest resources, the Maoshan Forest Farm covers a total area of 27,858 hectares, with a forest coverage rate of 83.29% (forestry land: 26,067 hectares) and a total timber volume of approximately 1.88 million cubic metres.

2.2. GEDI Spaceborne Lidar Data

The Global Ecosystem Dynamics Initiative (GEDI) is a spaceborne lidar system launched aboard the International Space Station (ISS) by NASA on 5 December 2018. This system carries the world's first multi-beam linear-array altimeter, primarily designed to acquire vertical structural information of tropical and temperate forests (as shown in Figure 1) [17]. GEDI comprises three identical near-infrared neodymium-doped yttrium aluminium garnet (Nd:YAG) lasers. These lasers emit short pulses of 1064 nm wavelength (14 ns duration) at a pulse repetition frequency of 242 Hz and an energy of 10 mJ towards the Earth's surface. The emitted laser beam has a divergence angle of 56 mrad. Two lasers operate at full power, while one laser is split into two beams, collectively generating four beams. Beam deflection units (BDUs) alter the laser beam direction by 1.5 mrad, enabling the laser spot to traverse 600 metres rapidly across the ground surface. This ultimately produces eight ground tracks (Table 1) (Figure 1).

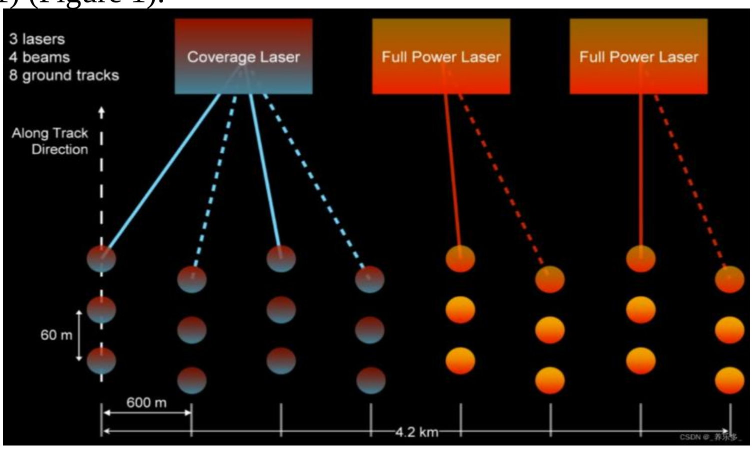

Fig. 1 Schematic diagram of ground sampling for GEDI laser [17]

Table 1. Main parameters description of GEDI

Parameter Name	Parameter value	Parameter Name	Parameter value
Operating altitude	400 km	Laser wavelength	1064 nm
Coverage	51.6°S~51.6°N	Crossing distance	60 m
Transmission frequency	242 Hz	Track distance	600 m
Spot diameter	25 m	Pulse width	14 ns
Number of tracks	8	Pulse intensity	10 mJ
Parameter Name	Parameter value	Parameter Name	Parameter value
Operating altitude	400 km	Laser wavelength	1064 nm

GEDI waveform data acquired from eight orbits undergoes processing by the NASA team. Based on data storage types, GEDI product files can be categorised into four levels, with the contents of each level detailed in Table 2.

Table 2. GEDI product document information

Product	Data products	Product
L1A	Raw waveform data	L1A
L1B	Waveform data following geolocation	L1B
L2A	Spot size canopy height index	L2A
L2B	Spot size biophysical parameter	L2B
L3	Grid-based Level 2 products	L3
L4A	Above-ground biomass products	L4A
L4B	Grid-based aboveground biomass density	L4B

2.3. Parameter Extraction and Data Preprocessing

Due to interference from both internal and external factors such as satellite position, attitude, and atmospheric conditions, GEDI L2A data exhibit certain inaccuracies in estimating ground elevation and forest canopy height. To enhance data quality, feature parameter extraction and screening must be performed on downloaded GEDI data to exclude invalid GEDI patches, thereby ensuring the acquisition of high-quality patch data. Specific parameters are detailed in Table 3.

Table 3. Parameter List for Extracting GEDI L2A Data

Parameter Name	Parameter Meaning
Lat lowestmode	The central latitude of the lowest mode
Lon lowestmode	The central longitude of the lowest mode
Sensitivity	Sensitivity
Elev lowestmode	Ground elevation value, relative to the elevation of the lowest mode centre of the reference ellipsoid
Mean sea surface	Average sea level height above the wobble Ellipsoid
Surface flag	Surface marking
Quality flag	Quality Mark
Degrade flag	Degradation indicator
Landsat Treecover	Canopy closure of all vegetation exceeding 5 metres in height
Modis treecover	Percentage of tree cover
Solar elevation	Solar altitude angle, the angular distance between the horizontal plane and the Sun in the sky
Solar azimuth	Azimuth of the Sun, measured from north, with east being positive.
Rh (90-100)	Relative forest height values at 1% intervals

To enhance the accuracies of acquired satellite-borne LiDAR data, data screening is required in addition to parameter extraction. In this study, the specific screening criteria for GEDI data are as follows:

(1) Sensitivity ≥ 0.9 indicates superior spot quality, featuring favourable signal-to-noise ratio and quality, enabling better penetration through forest canopy height to reach the ground surface. Therefore, spots with sensitivity < 0.9 are excluded.

(2) Surface_flag = 1 signifies that the lowest elevation mode lies within 300m of the DEM, demonstrating good spatial matching. Consequently, spots with surface_flag = 1 are retained.

(3) quality_flag=1 indicates the waveform meets standards based on energy, sensitivity, amplitude, and real-time surface tracking quality, qualifying as a valid waveform. Spots with quality_flag=1 are thus retained.

(4) degrade_flag ≠ 1. If this value is 1, it indicates degraded pointing and/or positioning information, potentially affected by interference or positional factors. Spots with degrade_flag=1 are therefore discarded. (5) rx_assess_flag ≠ 1. If this value is 1, it indicates potential errors within the waveform. Spots with rx_assess_flag=1 are consequently discarded.

(6) selected_mode_flag ≠ 0 and selected_mode_flag ≠ 1: This indicates deletion of both spots failing the mode reselection criterion (when selected_mode_flag = 0) and spots lacking a compliant mode (when selected_mode_flag = 1). Consequently, spots with selected_mode_flag = 0 and selected_mode_flag = 1 are removed.

(7) (elev_lowestmode - mean_sea_surface) - digital_elevation_model | < 50. If the absolute difference between the (ground elevation - mean sea level) obtained from the satellite-borne LiDAR data and TanDEM-X exceeds 50m, it is considered erroneous data affected by cloud cover. Consequently, the corresponding spot is deleted.

2.4. Accuracy Validation

To evaluate the accuracy of GEDI-derived forest canopy height, this study employs four statistical metrics: the Coefficient of Determination (R^2), Mean Error (Bias), Standard Deviation (σ), and Root Mean Square Error (RMSE). R^2, Bias, σ, and RMSE were employed to validate the accuracy of ground elevation extracted from GEDI data. R^2, Bias, and RMSE were further utilised to assess the precision of canopy height estimation from GEDI data, alongside evaluating model and mapping accuracy. R^2 indicates the linear relationship between two datasets, ranging from 0 to 1. Values closer to 1 signify higher correlation between the datasets and greater model accuracy. RMSE, meanwhile, yields values greater than zero.

Where xi denotes the ground elevation/vegetation canopy reference value for the ith observation, yi represents the ground elevation/vegetation canopy estimate derived from GEDI data for the ith observation, $\overline{x}$ signifies the mean of the ground elevation/vegetation canopy estimates, and n indicates the total number of samples.

3. Results

3.1. Accuracy validation

The topographic elevation RMSE for 7,731 GEDI sample data points in China's Maoershan forest region was 4.49 m, with an MAE of 3.33 m. In the grouped experiments, the results yielded were: Terrain elevation measured by the daytime strong beam: MAE of 2.86 m, RMSE of 3.90 m; Terrain elevation measured by the nighttime strong beam: MAE of 4.66 m, RMSE of 5.96 m; Terrain elevation measured by the daytime coverage beam: MAE of 2.85 m, RMSE of 3.81 m; while the nighttime covered beam yielded a MAE of 5.38m and an RMSE of 6.72m (as shown in Figure 2). The experiments conducted in the forest region of Maoershan China, indicate that the daytime strong beam and covered beam deliver nearly identical performance, both significantly outperforming nighttime measurements in terrain height accuracy. Within the nighttime groupings, the strong beam demonstrated markedly superior measurement efficacy compared to the covered beam.

3.2. Error Analysis

Results for the Maoershan region of China: For the 0-5° grouping, terrain inversion accuracy achieved an MAE of 0.97m and an RMSE of 1.74m; For the 5-10° grouping, MAE was 1.75m and RMSE was 2.59m; For the 10-15° grouping, MAE was 2.82m and RMSE was 3.78m; The 15–20° group yielded an MAE of 3.64 m and an RMSE of 4.66 m; the 20–30° group showed an MAE of 4.68

m and an RMSE of 5.74 m; the ≥30° group produced an MAE of 5.80 m and an RMSE of 6.95 m (as per Table 4). As slope increases, RMSE exhibits a linear upward trend, indicating that slope significantly impacts the elevation measurement accuracy of GEDI data. On gentle terrain, GEDI provides relatively precise elevation results; however, measurement errors become more pronounced with steeper slopes. For high-precision measurements, it is advisable to avoid areas with steep gradients or employ scientific methods for terrain correction prior to data utilization (Table 4).

Table 4. The accuracy of GEDI-derived terrain elevation under different slopes in the Maoshan area

slope /(°)	MAE/m	R^2	RMSE/m
0-5	0.97	1	1.74
≥5-10	1.75	1	2.59
≥10-15	2.82	1	3.78
≥15-20	3.64	1	4.66
≥20-30	4.68	1	5.74
＞30	5.8	1	6.95

The significance of slope gradient, aspect, laser beam intensity (strong/weak), and solar elevation during data acquisition (daytime/night-time) on ground elevation extraction errors in GEDI data is shown in Figure 2. It is evident that slope exerts the most significant influence on ground elevation errors, indicating that accurate elevation extraction proves particularly challenging in areas with pronounced inclines. Sun elevation angle during data acquisition (daytime/night-time) ranks second in impact, likely due to external noise interference such as solar noise affecting waveform signals. Slope aspect and laser beam intensity also exert a certain degree of influence on ground elevation errors in GEDI data extraction.

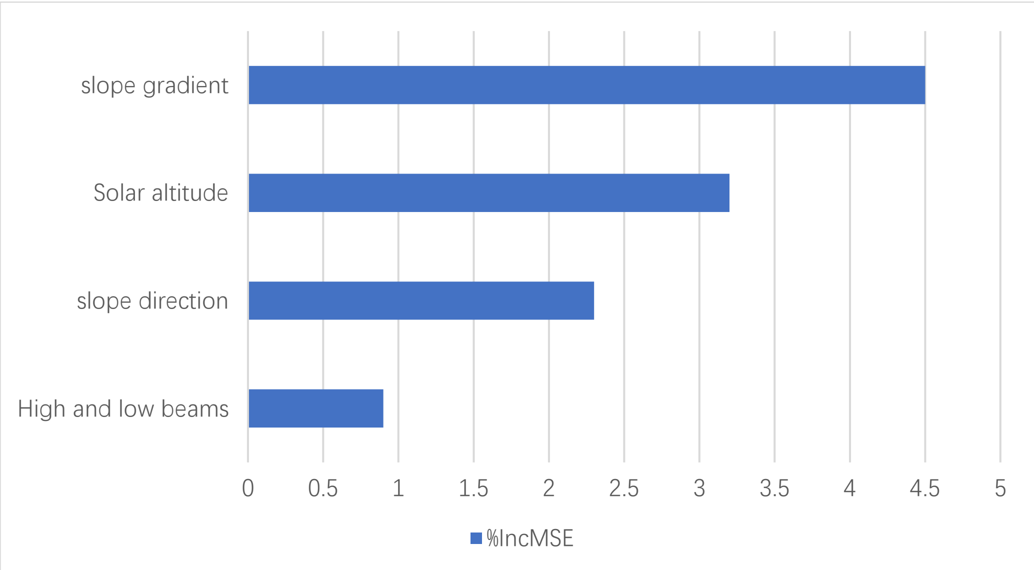

Fig. 2 The influence importance of each factor on ground elevation retrieval [17]

4. Conclusion

In the experimental study, the vegetation cover in areas exposed to strong sunlight during the day was 91.6%, while 90.35% of the vegetation cover was shaded at night, considering other research findings, the cause is likely due to the relatively sparse vegetation in the areas traversed by the daytime coverage beam, not inconsistent with the findings in coniferous forest regions. Thus, the daytime coverage beam accuracy is slightly higher than that of the strong beam, where night-time high-intensity beams are more accurate than coverage beams. The background noise of the GEDI detector exceeds solar noise levels, thus solar background noise does not constitute the primary factor for performance discrepancies between daytime and night-time observations. Given that the Maoershan site experiences a temperate humid climate, variations in cloud cover, temperature differentials, precipitation, and other factors attributable to humid versus arid climates are considered potential sources of error. Consequently, the performance of GEDI data during daytime and night-time is not fixed but must be determined based on specific local climatic conditions. Performance is generally superior during daytime in humid regions, while arid regions exhibit superior night-time performance.

Slope gradient and vegetation cover are both critical factors influencing the performance of GEDI data. In areas with slopes below 20°and vegetation cover below 60%, the terrain inversion accuracy is exceptionally high. However, as slope increases and vegetation cover rises, the ability of GEDI data to accurately invert the terrain beneath the canopy diminishes. This occurs because in steep terrain, the full waveform data becomes confused due to the mixing of ground echoes with vegetation echoes, leading to reduced accuracy with increasing slope. In areas of high vegetation cover, GEDI laser energy suffers attenuation when penetrating the canopy, and multi-layered canopies further degrade accuracy, resulting in diminished inversion precision with increasing vegetation coverage.

1) GEDI data inversion of understory topography yields superior results with the high-beam configuration over the coverage beam. Performance is optimal during daytime in humid regions and at night in arid areas.

2) The GEDI laser may overestimate canopy height; applying error correction prior to use enhances the accuracy of canopy height measurements.

3) GEDI elevation precision exhibits a declining trend with increasing ground slope. Data performance is excellent in relatively gentle terrain, whereas steep areas suffer reduced accuracy due to the co-spatial overlap of ground and vegetation echoes.

References

[1] Pan Y., Birdsey R. A., Fang J. A large and persistent carbon sink in the world’s forests. Science, 2011, 333(6045): 988-993.

[2] Zhang Q. Review of research on global climate change and its influencing factors. Climatic and Environmental Research, 2018, 23(5): 545-559.

[3] Song G. On social action subjects towards the goal of peaking carbon dioxide emissions and achieving carbon neutrality and measures in China. Urban and Environmental Studies, 2021, (4): 47-60.

[4] Zhou T. Frontier scientific issues in climate change supporting the “Future Earth” program. Advances in Earth Science, 2015, 30(9): 993-1003.

[5] Zhou Y. Carbon storage and balance of major forest ecosystems in China. Acta Ecologica Sinica, 2006, 26(7): 2071-2081.

[6] Wang C., Luo S. Z., Xi X. H., et al. Influence of voxel size on forest canopy height estimates using full-waveform airborne LiDAR data. Remote Sensing, 2019, 11(22): 2676.

[7] Yang T. Mapping Forest canopy height in China using satellite-borne GLAS LiDAR and MODIS optical data. Journal of Remote Sensing, 2017, 21(1): 1-12.

[8] Gasparri N. I., Parmuchi M. G., Bono J., et al. Assessing multi-temporal Landsat 7 ETM+ images for estimating above-ground biomass in subtropical dry forests of Argentina. Journal of Arid Environments, 2010, 74(10): 1231-1238.

[9] Zhang B. An overview of optical remote sensing information technology and its applications. Remote Sensing Information, 2014, 29(6): 1-7.

[10] Lu D., Chen Q., Wang G., et al. Aboveground Forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates. Remote Sensing of Environment, 2012, 124: 286-300.

[11] Zhao F. Review of airborne LiDAR applications in forest resource surveys. Forest Inventory and Planning, 2013, 38(1): 1-6.

[12] Wang Y., Li G., Ding J., et al. A combined GLAS and MODIS estimation of the global distribution of mean forest canopy height. Remote Sensing of Environment, 2016, 174: 24-43.

[13] Pang Y. Progress and trends in satellite-borne LiDAR for forest detection. Journal of Remote Sensing, 2019, 23(4): 855-864.

[14] Wen H. Principles and applications of ICESat/GLAS laser altimetry. Journal of Geodesy and Geoinformation Science, 2012, 41(3): 35-42.

[15] Enßle F., Heinzel J., Koch B. Accuracy of vegetation height and terrain elevation derived from ICESat/GLAS in forested areas. Remote Sensing of Environment, 2014, 151: 32-44.

[16] Yuan Z. Accuracy evaluation and influencing factors of GEDI-derived ground elevation and forest canopy height. Journal of Remote Sensing, 2021, 25(7): 1401-1415.

[17] Xie D. The GEDI space-based laser altimetry system of the United States and its applications. Laser & Optoelectronics Progress, 2020, 57(10): 100004.