This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2024.3522688

An On-Orbit Calibration Model for Full-Waveform Laser Altimeters by Matching Single-Modal and Multi-Modal Waveforms

Qi Xu, Huan Xie, *Senior Member, IEEE*, Peiqi Huang, Min Ji, Hao Wu, Changda Liu, Shijie Liu, *Member, IEEE*, Xiaohua Tong, *Fellow, IEEE*, Yanmin Jin, Yuan Sun, Binbin Li, Kuifeng Luan, Genghua Huang, Hao Zhang, Gang Chen

***Abstract*—Accurate pointing and ranging calibration are essential for laser altimeters to obtain the geolocation accuracy required for further scientific application. Waveform matching is an economical and effective on-orbit calibration technique that involves simulating waveforms based on the vertical structure within potential correct footprints acquired from high-resolution digital surface/elevation model and comparing them with the recorded waveforms to correct the footprint geolocation. We propose an on-orbit calibration model using waveform matching by integrating multi-modal waveforms and single-modal waveforms. The model fully leverages the characteristics of multi-modal waveforms that can be accurately located through waveform matching, and the single-modal waveforms that can precisely measure elevation, achieving an enhancement in performance over classical model. To validate the effectiveness of the improved calibration model, we conducted experiments using Global Ecosystem Dynamics Investigation (GEDI R02) to correct the potential positioning bias in GEDI measurements. The eight beams were adjusted between −2 m to 8 m in the north and east directions, with corresponding adjustments in laser pointing from 2.5 arcsec to 4 arcsec in both roll and pitch. Range corrections varied from −2 cm to −15 cm. The results demonstrate that the geolocation accuracy is markedly improved for all eight beams, with the elevation bias reduced by approximately 10 cm and the planar bias restricted to within 3 m (coverage beams) and 1 m (full-power beams). The model can be used to effectively calibrate full-waveform laser altimeters, and also provides new insights into the verification of the system accuracy.**

***Index Terms*— calibration and validation; waveform matching; full-waveform laser altimeters; GEDI**

I. INTRODUCTION

SPACEBORNE laser altimetry technology is rapidly becoming a new method for observing and exploring the Earth, which is capable of quickly and accurately acquiring three-dimensional (3D) surface information at a global scale. The laser altimeter accurately records the emission and echo times of laser pulses to precisely obtain ranging information [1, 2]. When combined with the satellite's position information (from the Global Navigation Satellite System (GNSS)) and the laser pointing angle (determined by the satellite attitude and the mounting information of the laser altimeter), the laser altimeter can calculate the coordinates of the detection target [3]. In the beginning, laser altimeters were utilized for the investigation of celestial bodies such as planets and asteroids located in the distant reaches of space [4]. Notable examples of such systems include the Clementine Laser Image Detection and Ranging (LiDAR) system [5], the Mars Orbiter Laser Altimeter [6], the Mercury Laser Altimeter [7], the CE-series laser altimeters [8], the SELENE Laser Altimeter [9], the Lunar Orbiter Laser Altimeter [10], and the BepiColombo Laser Altimeter [11]. The data obtained by these systems provide an important basis for researchers to establish the 3D topography of observed celestial bodies. Subsequently, the application of laser altimetry has gradually been extended to Earth observation. Full-waveform systems such as the Ice, Cloud, and land Elevation Satellite Geoscience Laser Altimeter System (ICESat/GLAS) [12], the Global Ecosystem Dynamics Investigation (GEDI) instrument [13], GaoFen 7 (GF-7) [14], and photon-counting systems such as the Ice, Cloud and land Elevation Satellite-2 Advanced Topographic Laser Altimeter System (ICESat-2/ATLAS) [15] provide a large amount of

This research was supported by the National Natural Science Foundation of China under Grant 42325106 and 42221002, the Shanghai Academic Research Leader Program under Grant 23XD1404100, the Shanghai Science and Technology Innovation Action Plan Program under Grant 22511102900, and the Fundamental Research Funds for the Central Universities of China. *(Corresponding author: Huan Xie)*

Q. Xu, H. Xie, P. Huang, M. Ji, H. Wu, C. Liu, S. Liu, X. Tong, Y. Jin, Y. Sun and B. Li are with the College of Surveying and Geo-Informatics, Tongji University, and the Shanghai Key Laboratory of Space Mapping and Remote Sensing for Planetary Exploration, Shanghai, China (e-mail: smxj_xq@tongji.edu.cn; huanxie@tongji.edu.cn; 2131962@tongji.edu.cn; 1851023@tongji.edu.cn; 2210208@tongji.edu.cn; cdliu@tongji.edu.cn; liusjtj@tongji.edu.cn; xhtong@tongji.edu.cn; jinyanmin@tongji.edu.cn; sunyuan@tongji.edu.cn; libinbin@tongji.edu.cn).

K. Luan is with the College of Marine Sciences and the Shanghai Engineering Research Center of Estuarine and Oceanographic Mapping, Shanghai Ocean University (e-mail: kfluan@shou.edu.cn).

G. Huang is with Shanghai Institute of Technical Physics, Chinese Academy of Science, Shanghai, China; (e-mail: genghuah@mail.sitp.ac.cn)

H. Zhang and G. Chen are with Xi'an Research Institute of Surveying and mapping, Xi'an, China (e-mail: isfetz@126.com; splliter_ch@163.com)

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2024.3522688

high-accuracy, long time-series terrain information. These data play an increasingly important role in high-accuracy 3D terrain mapping [16], vegetation height detection and biomass monitoring [17, 18], ice sheet mass balance [19] and change monitoring [20], physical oceanography [21-23], and other research fields [15].

The ability of a spaceborne laser altimetry system to acquire highly precise geolocation information of surface returns hinges on accurate instrument mounting information for the satellite platform, along with high-accuracy measurements of ranging, satellite positioning, attitude, and laser pointing. The process of launch, which is characterized by rocket thrust [24], as well as satellite platform jitter [25] during the satellite's orbital operation, can introduce errors in the laser altimeter's ranging and pointing capabilities. Moreover, it has been observed that ancillary elements, including atmospheric delay and variations in the space environment, frequently contribute to substantial deviations between the theoretically computed geolocation and the actual geolocation of the footprint [24]. For instance, with China's GF-7 satellite, the maximum pointing angle correction for beam 2 exceeds 0.1°[26], which is significantly beyond the design requirements of the altimetry system. Therefore, on-orbit calibration of laser altimetry systems is of vital importance. The accuracy of calibrating the pointing and ranging deviations introduced by the various factors during the satellite's orbital operation directly impacts the on-orbit performance of the laser altimetry system.

For full-waveform spaceborne laser altimeters, several on-orbit calibration methods are commonly used, including the ocean sweep maneuver [27], field calibration tests based on detectors [24, 28], terrain profile matching between altimeter measurements and other digital surface model (DSM)/digital elevation model (DEM) products [29], and waveform matching [30]. The ocean sweep maneuver involves attitude maneuvers of the satellite platform when transiting over calm ocean surfaces. This method analyzes the residuals (the difference between the modeled range and the recorded range obtained by the altimetry system) to correct pointing and ranging errors. This on-orbit correction method requires the satellite platform to have stable attitude control capability and precise attitude measurement capability. Field calibration tests involve capturing the laser footprint using detectors, and correcting the pointing and ranging errors based on the residuals between the captured and recorded positions. The positions of the laser footprints are predicted based on historical satellite position and attitude information. Near the predicted position, the detectors are deployed within a certain range. By analyzing the position of the detectors illuminated during the satellite's transit, accurate laser footprint positions can be obtained. An error equation is then established based on the laser positioning model to solve the pointing and ranging errors [31, 32]. However, this is always a complex task, which typically necessitates support in terms of personnel, equipment, and funding [24]. It has been reported that thousands of detectors were used for the field calibration test of the GF-7 laser altimeter, and the process lasted for more than one month. The third method, terrain profile matching, involves matching the entire track terrain profile recorded by the laser altimeter with existing DEM profiles to correct the pointing and ranging errors. However, compromises exist between the solution's accuracy and the scope of the DEM. For instance, a DEM with a high resolution and accuracy but limited to a small area could actually be less effective for calibration purposes than a DEM encompassing a broader area, albeit with a lower resolution and accuracy [33]. Generally speaking, the calibration accuracy of the profile matching method is limited, and it is difficult to obtain high-accuracy calibration parameters solely by this method. However, this method can quickly remove large-scale pointing biases. Therefore, the profile matching method is often used to initialize laser pointing, provide initial footprint positions for detector layout in the field calibration tests, and establish initial footprints and search ranges for the waveform matching method. The fourth method employs waveform matching to correct the position of the footprints and solve the calibration parameters. By comparing the similarity in shape between the simulated waveforms and recorded waveforms at potential positions, the corrected positions of the footprints can be obtained. The calibration parameters for pointing and ranging are then solved by analyzing the relationship between the positions of the footprints before and after correction. It is well known that the waveform can deliver a precise depiction of the vertical distribution of the surfaces intercepted by the laser footprint. To facilitate this, high-resolution DSM data, replete with intricate vertical distribution details, can be employed to simulate waveforms within a specified vicinity of the recorded footprint's position. Through a meticulous comparison of the recorded waveforms with simulated waveforms, it is possible to ascertain the footprint's position with enhanced accuracy, thereby refining the calibration process [26, 30, 34].

The waveform matching technique obviates the need for the deployment of additional detectors and lengthy experimental durations extending up to a month. Consequently, in comparison to field calibration tests, this method presents a cost-effective calibration solution, particularly for laser altimetry satellites lacking maneuver capabilities for the ocean sweep maneuver. In the realm of successfully launched laser altimetry systems, the waveform matching method has been extensively employed in calibration and validation experiments. Introduced by Blair and Hofton [34], waveform matching involves the generation of high-resolution elevation datasets, incorporating the energy distribution within a footprint, to create an elevation distribution function. This function, when convolved with the emitted wave, can yield simulated waveforms that can be compared with airborne LiDAR data (such as Land, Vegetation, and Ice Sensor (LVIS) data) to adjust footprint position inaccuracies. This method has been essential in assessing the planimetric accuracy of ICESat/GLAS. Several scholars have attempted to solve laser altimeter pointing and ranging calibration parameters using the classical model [31, 32] of least-squares adjustment (hereinafter referred to as the classical calibration model). A common approach involves selecting multi-modal footprints, seeking potential footprint locations through waveform feature matching, establishing an equation relating the geolocation error to the pointing and ranging errors, and further solving for the calibration parameters. Further advancements were made by Liu et al. [26], who significantly enhanced the method's computational efficiency through a pyramid approach that reduces the search step size. This approach can precisely locate the footprints with the multi-modal waveform, aiding in the derivation of accurate calibration parameters, as demonstrated with GF-7, where the elevation measurement accuracy (EMA) was improved to 0.35 ± 0.50 m. Similarly, Li et al. [30] refined the correlation function in waveform matching, considering multiple waveform parameters to more accurately determine the footprint positions, thereby improving ICESat's altimetric accuracy from 2 m to 0.5 m.

Although the waveform matching methods offer a cost-effective approach for calibrating laser altimeters, several challenges remain. Notably, the classical calibration model, which employs least-squares adjustment, only incorporates multi-modal waveforms in the calculations, overlooking single-modal waveforms due to their lack of distinctive waveform features. This omission gives rise to two primary concerns: firstly, the prevalence of multi-modal waveforms is limited, making up less than 20% of waveforms, according to our statistics, a figure that decreases further in barren landscapes such as deserts. For low-frequency full-waveform LiDAR altimeters, relying solely on 20% of multi-modal waveforms for calibration is unlikely to achieve optimal calibration results; secondly, despite the fact that multi-modal waveforms aid in pinpointing the precise locations of footprints through waveform matching, accurately extracting elevation information from these footprints is complex due to the intricate elevation structures they present. This also leads to the calibration results being unable to achieve desirable levels of accuracy in elevation measurements. As a result, we have noted that the classical calibration model maintains certain elevation biases following calibration correction. For instance, Li et al. [35] identified an elevation bias of ~0.3 m using the classical calibration model. Another considerable challenge lies in accurately determining the footprint locations through individual waveform matching. The accuracy of waveform matching is greatly affected by terrain features that are similar in nature. For example, forested areas with uniform tree heights exhibit bi-modal elevation structures, reflecting the characteristics of tree canopies and the ground surface, which results in highly similar simulation waveforms for multiple locations. Identifying the precise footprint location among these potential candidates is therefore challenging. Liu et al. [26], Li et al. [35], and others have adopted the point of maximum correlation as the accurate footprint location. Nonetheless, when comparing the calibration outcomes derived from waveform matching methods to those obtained through field calibration tests, a planar error of ~11 m was observed, indicating that improvements can still be made to the waveform matching strategies.

In this paper, based on the discussion of the afore-mentioned issues, we have made contributions in two aspects:

1) We introduce an improved on-orbit calibration model for full-waveform laser altimeters. This advanced model synthesizes the benefits of utilizing footprints with multi-modal waveforms for planar position determination and leverages the robustness of single-modal waveforms for elevation extraction. Our approach surpasses the constraints of former model, which relied solely on multi-modal waveform footprints for calibration via waveform matching. For low-frequency laser altimeters, increasing the utilization rate of data is crucial for improving calibration results. Consequently, this innovative strategy substantially augments the calibration outcomes, with a notable enhancement in ranging calibration precision.

2) We have refined the waveform matching procedure by developing the *Global Average Correlation Coefficient Matrix*, which incorporates considerations for random errors and corrects for potential systematic biases. Following this enhancement, we conduct a constrained window search to ascertain the optimal matching position, thereby mitigating the inaccuracies that may occur from direct localization via maximum correlation matching, particularly in regions with similar terrain where such mismatches are more prevalent. This methodological advancement addresses the critical challenges associated with mismatches in waveform data analysis.

II. METHODS

In this section, we describe the methodologies based on the discussion presented earlier. In Section II-A, we introduce the recorded waveforms of full-waveform altimeters and analysis the characteristics of the single- and multi- modal waveforms. In Section II-B, we briefly introduce the precision positioning model and the classical calibration model [32], while thoroughly deriving the improved on-orbit calibration model proposed in this paper, which integrates the characteristics of both single-modal and multi-modal waveforms. In Section II-C, we focus on a discussion of the waveform matching optimization process, detailing the optimized waveform matching method proposed in this paper.

A. Recognition and Characteristics of Recorded Waveforms of Full-Waveform Altimeters

The recorded waveform is the digitized sampling of echoes generated by the interaction between the emitted waveform and terrestrial surfaces. Generally, Blair et al.[34] classified the received waveforms into two categories based on their shape. The first type consists of waveforms with a simple shape, closely resembling the emitted waveform, which we refer to as *"single-modal waveform"* (**Fig. 1**-(a)). The second type involves waveforms with a complex shape, often containing multiple modes, with each mode representing a distinct surface type within the footprint. These are referred to as *"multi-modal waveform"* (**Fig. 1**-(b), (c)).

Both types of waveforms have their own distinct characteristics. For multi-modal waveforms, the footprint contains a more diverse elevation distribution, making the

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2024.3522688

waveform generated within the footprint unique compared to waveforms from surrounding areas. This uniqueness is a key factor in improving positioning accuracy through waveform matching, which is critical for calibration purposes. In contrast, single-modal waveforms typically represent footprints with more uniform surfaces, such as flat and barren ground, ocean surfaces, or ice sheets. This simplicity allows researchers to easily and accurately resolve the range information for these footprints.

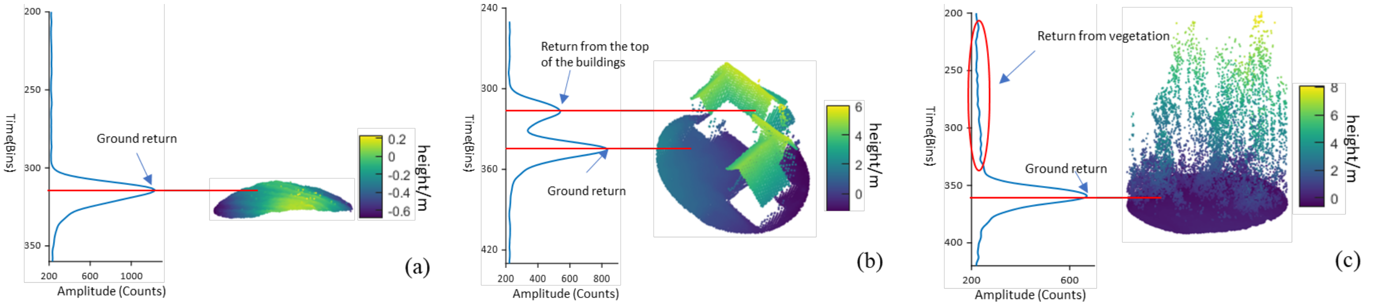


Fig. 1. The elevation distribution within the footprint corresponds to the received waveform. For each subplot, the left side displays the waveform received by the full-waveform laser altimeter, while the right side shows the corresponding point cloud data within the footprint. The figure illustrates three different surface types: bare flat ground, buildings, and vegetation. (a) represents a typical single-modal waveform, while (b) and (c) are typical examples of multi-modal waveforms.

B. Precise Footprint Positioning Model and Calibration Model

The precise footprint positioning model calculates the exact locations of the footprints by utilizing laser ranging data, supplemented with auxiliary information such as laser pointing, satellite attitude, and position data. Section II-B-1) introduces this content. However, the ranging and auxiliary information used in the model often contains errors, leading to deviations in the calculated positions of the footprints. These deviations consist of systematic and random errors. The systematic errors mainly include pointing and ranging biases. Pointing bias arises due to forces acting on the satellite during launch and on-orbit operations, resulting in discrepancies between the actual on-orbit and laboratory-measured pointing. Ranging bias is primarily due to misalignment of the system emission timing. The random errors are mainly caused by platform jitter, leading to random fluctuations in the geolocation of the footprints. While random errors are stochastic and unpredictable, systematic errors can be resolved through calibration models. In Section II-B-2), we discuss the positioning model under the influence of errors, detailing how ranging errors and pointing errors affect the positioning results. In addition, we introduce the classical calibration model, which solves the systematic errors in ranging and pointing by using the obtained corrected locations of the footprints. In Section II-B-3), we explain how footprints with single-modal waveforms are incorporated into the calibration model to enhance the precision of the calibration results.

1) Precise Footprint Positioning Model

Numerous scholars [29, 30, 36] have conducted extensive research on footprint positioning models for laser altimetry systems. The essence of these models lies in establishing the relationships between the Optical Bench Frame (OBF), Earth-Centered Inertial (ECI), and Earth-Centered Fixed (ECF) coordinate systems. This process involves the calculation of footprint coordinates by combining high-accuracy ranging data with spacecraft position, attitude, and other data. The formula for precise footprint position determination is as described in Equation (1) and **Fig. 2**:

$$\vec{X}_{ECI}^{BP} = \vec{X}_{ECI}^{IRP} + R_{Body\to ECI}(\vec{x}_{off} + \rho\cdot R_{OBF\to Body}\cdot \hat{p}) \qquad (1)$$

where $\vec{X}_{ECI}^{BP}$ represents the position of the bounce point (the footprint) in the ECI system. $\vec{X}_{ECI}^{IRP}$ denotes the position of the instrument reference point in the ECI system, which is typically the location of the GNSS receiver. $R_{Body\to ECI}$ is the rotation matrix from the satellite body to the ECI coordinate system, measured by the satellite's star tracker. $\vec{x}_{off}$ indicates the offset vector of the laser firing point relative to the GNSS receiver within the body coordinate system. ρ is the measured value of ranging. $\hat{p}$ signifies the laser emission direction, which is typically $[0 \quad 0 \quad 1]^T$. $R_{OBF\to Body}$ is the rotation matrix from the laser coordinate system to the body coordinate system.

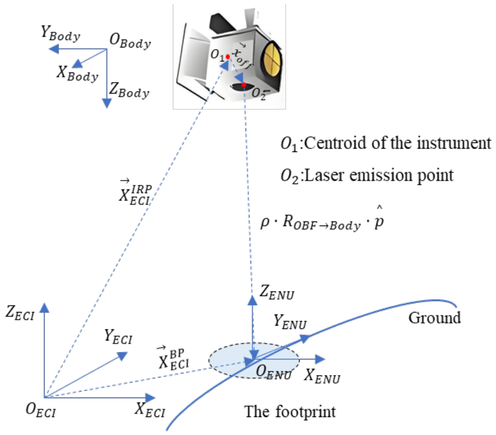


Fig. 2. The geometry model for precise footprint position determination of laser altimeters.

2) Error-Affected Geometric Positioning Model and Classical Calibration Model

According to the description of Equation (1), we consider the possible systematic and random error components in the calculation of $\vec{X}_{ECI}^{BP}$. The systematic error component originates solely from the systematic errors in ranging and pointing of laser[37]. The difference in ranging calibration values is denoted as ρ', and the discrepancy in pointing is represented by the rotation matrix R_m. The random error component mainly comprises instrument positioning error, measurement deviations of the instrument attitude by star tracker, the random error in laser ranging, and the jitter in laser pointing. These random errors are collectively represented as $\vec{X}_e$. The positional error of the instrument is generally less than 0.1 m. The attitude measurement error of the star tracker is approximately 3 arcsec, which can result in a planar error of about 5 m to 10 m on the ground. The random error in laser ranging is approximately 0.1 m, which equivalently affects the elevation error by 0.1 m. Laser pointing jitter is approximately 0.5 arcsec, which can lead to a planar error of about 1m to 3m. Taking these error conditions into account, the error-affected geometric positioning model can be represented as shown in Equation (2) [38, 39]:

$$\vec{X}_{ECI}^{BP} = \vec{X}_{ECI}^{IRP} + R_{Body\to ECI}[\vec{x}_{off} + (\rho + \rho') \cdot R_{Laser\to Body} \cdot R_m \cdot \hat{p}] + \vec{X}_e \qquad (2)$$

R_m is the rotation matrix corresponding to the correction amount for pointing errors, considering that the mounting angles biases are minimal, which can be expressed as shown in Equation (3), and we provide the detailed derivation in Appendix-A.

$$R_m = \begin{bmatrix} 1 & -\kappa & \varphi \\ \kappa & 1 & -\omega \\ -\varphi & \omega & 1 \end{bmatrix} \qquad (3)$$

where φ, ω and κ represent the mounting bias angles of the pointing direction in the pitch, roll and yaw directions, respectively. Considering that the laser emission direction always satisfies $\hat{p} = [0 \quad 0 \quad 1]^T$, the yaw direction deviation angle does not affect the result. Therefore, Equation (3) can be further simplified to Equation (4).

$$R_m = \begin{bmatrix} 1 & 0 & \varphi \\ 0 & 1 & -\omega \\ -\varphi & \omega & 1 \end{bmatrix} \qquad (4)$$

The classical calibration model for solve the pointing and ranging errors was established by Yi et al. [31] and Xie et al. [32] using least-squares adjustment. The error equation is as shown in Equation (5):

$$V_{classic}(\varphi, \omega, \rho') = \vec{X}_{ECI}(\varphi, \omega, \rho') - \vec{X'}_{ECI} \qquad (5)$$

where $\vec{X}_{ECI}(\varphi, \omega, \rho')$ represents the calculated position of the footprint under parameters φ, ω, and ρ'. $\vec{X'}_{ECI}$ denotes the reference position of the footprint, as determined through waveform matching. $V_{classic}(\varphi, \omega, \rho')$ represents the residual under parameters φ, ω, and ρ'. It is important to note that only footprints with multi-modal waveforms are selected for inclusion in these calculations, as footprints with single-modal waveforms lack a distinct elevation distribution, rendering them unsuitable for obtaining accurate matching positions through the waveform matching method. Following the principle of the least-squares method, the objective equation for solving Equation (5) is formulated as shown in Equation (6):

$$\sum_{i=1}^{n} V_{classic}^T V_{classic} = min \qquad (6)$$

where n represents the number of high-quality multi-modal footprints. The solution to this equation, $\vec{k}^*$, is the set of parameters obtained through the calibration process.

3) The Improved On-Orbit Calibration Model

The classical calibration model typically utilizes only the footprints with multi-modal waveforms, while the single-modal waveforms are often excluded, due to their incompatibility with waveform matching. However, it is noteworthy that elevation calculations from single-modal waveforms tend to be more accurate than those from multi-modal waveforms. To fully leverage the characteristics of both single- and multi-modal waveforms, we developed an improved model. This model allows footprints with multi-modal waveforms to control the calculation of the pointing correction, while footprints with single-modal waveforms govern the determination of the ranging correction. The flowchart of the improved model is shown in **Fig.3**.

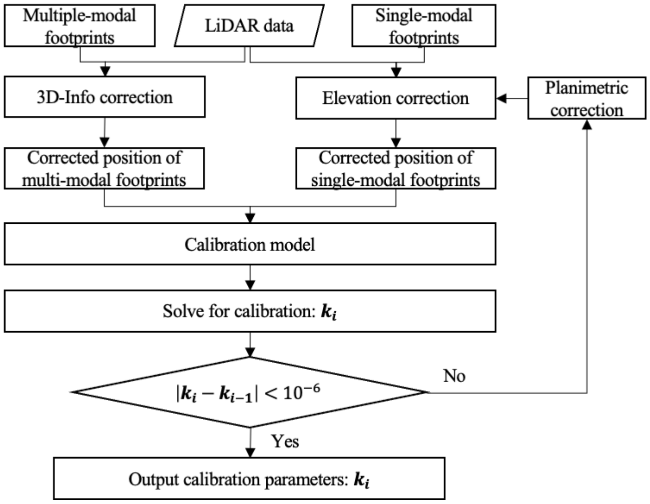


Fig. 3. Flowchart of the improved on-orbit calibration model. $\boldsymbol{k_i}$ represents the solution value of the calibration parameter after the i-th iteration.

Firstly, it is important to note that, in the ECI coordinate system, it is not possible to distinguish between the planimetric direction and the elevation direction. Therefore, it is necessary to convert the ECI coordinates into the east-north-up (ENU) coordinate system. This transformation allows for a clear distinction between the planimetric direction and the elevation direction, facilitating more accurate adjustment and calibration specific to each direction.

$$\vec{X}_{ENU}^{BP} = R_{ECF\to ENU} \cdot R_{ECI\to ECF} \cdot (\vec{X}_{ECI}^{BP} - \vec{X}_{ECI}^{Center}) \qquad (7)$$

where, $\vec{X}_{ECI}^{Center}$ represent the position of the origin of the ENU coordinate system, which is generally substituted with the footprint position calculated using pre-calibration values. It is noteworthy that for each footprint, an ENU coordinate system is established. Consequently, $R_{ECF\to ENU}$ varies with each footprint to eliminate the effects of earth's curvature.

Subsequently, we can derive a new error equation, as described in Equation (8):

$$V_{improved}(\varphi,\omega,\rho') = \vec{X}_{ENU}(\varphi,\omega,\rho') - \vec{X'}_{ENU} \quad (8)$$

where, $V_{improved}(\varphi,\omega,\rho')$ represents the residual under parameters φ, ω, and ρ' of the improved model. $\vec{X}_{ENU}(\varphi,\omega,\rho')$ represents the calculated position of the footprint under parameters φ, ω, and ρ' in the ENU coordinate system. $\vec{X'}_{ENU}$ denotes the reference position of the footprint, as determined through waveform matching or other method.

Contrary to the classical calibration model, the footprints with single-modal waveforms and those with multi-modal waveforms are both included in the calculation. The acquisition of $\vec{X'}_{ENU}$ differs between footprints with multi- and single-modal waveforms. For footprints with multi-modal waveforms, the acquisition is achieved using the waveform matching method, as introduced in Section II-C. In the case of footprints with single-modal waveforms, due to the lack of matching information, precise location data are unattainable. Therefore, in the initial iteration phase, we do not change its planar position information; we only modify its elevation information based on high- resolution elevation data. Therefore, $\vec{X}_{ENU} = (0,0,\Delta h)$ is used as the initial value, where the elevation Δh is extracted using high-resolution elevation data at the (0,0) position in the ENU coordinate system.

According to the characteristics of the waveforms, multi-modal waveforms, after matching, can provide more accurate planar positions, while single-modal waveforms are more precise in elevation measurements. Therefore, we attempt to construct a new calibration model to ensure that the calculated $\vec{k}$ better conforms to the planar position information of the multi-modal waveforms and the elevation information of the single-modal waveforms. Since we need the result $\vec{k}$ to align more closely with the planar correction results of the multi-modal waveforms and the elevation correction results of the single-modal waveforms, the model may produce significant residual errors in the elevation of multi-modal waveforms and the planar corrections of single-modal waveforms. To address this issue, we introduce the $C\vec{t}$ to represent the coordinate correction amount in the uncontrolled direction. Consequently, the error equation after linearization can be expressed as shown in Equation (9), and the objective equation is as shown in Equation (10):

$$V_{improved} = B_{improved}\vec{k} + C\vec{t} - \hat{L} \quad (9)$$

$$\sum_{i=1}^{n} V_{improved}^{T} V_{improved} = min \quad (10)$$

where $\vec{k}$ is $[\varphi \quad \omega \quad \rho']^T$, which represents the pointing and ranging corrected value (calibration parameters) to be determined. $B_{improved}$ is $\frac{\partial V_{improved}}{\partial \vec{k}}$. $\vec{t}$ represents the corrected value of the footprint position in the ENU coordinate system, and is a 3×1 vector. $\hat{L}$ is the constant term after linearization, represented as a 3×1 vector. n represents the number of footprints. C represents the control coefficient of $\vec{t}$. The optimal solution $\vec{k}^*$ for this model can still be determined using the least-squares criterion. Since the multi-modal and single-modal footprints do not control the elevation and planar directions, respectively, it is necessary to consider the 3D coordinate correction in the uncontrolled direction for both types of footprints. Therefore, the definition of C differs for single- and multi- modal footprints.

$$C_{multi} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a \end{bmatrix} \quad (11)$$

$$C_{single} = \begin{bmatrix} b & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (12)$$

where, a, b, c are equal to $\frac{\partial V_{improved\_U}}{\partial \vec{X}_{ENU\_U}}$, $\frac{\partial V_{improved\_E}}{\partial \vec{X}_{ENU\_E}}$, $\frac{\partial V_{improved\_N}}{\partial \vec{X}_{ENU\_N}}$. $V_{improved\_E}$, $V_{improved\_N}$, $V_{improved\_U}$ represent the residuals in the east, north, and up directions, respectively. $\vec{X}_{ENU\_E}$, $\vec{X}_{ENU\_N}$, $\vec{X}_{ENU\_U}$ represent the coordinates of the footprint in the east, north, and up directions in the ENU coordinate system. From Equation (8), it can be deduced that $V_{improved}$ and $\vec{X}_{ENU}$ exhibit a linear relationship. Therefore, $C_{multi} = diag(0,0,1)$, $C_{single} = diag(1,1,0)$.

Based on the principle of least squares, the normal equation corresponding to Equation (10) can be formulated as shown in Equation (13):

$$\begin{bmatrix} B_{improved}^T B_{improved} & B_{improved}^T C \\ C^T B_{improved} & C^T C \end{bmatrix} \begin{bmatrix} \vec{k} \\ \vec{t} \end{bmatrix} = \begin{bmatrix} B_{improved}^T \hat{L} \\ C^T \hat{L} \end{bmatrix} \quad (13)$$

Due to the large number of footprint points involved in the calculation, the number of unknowns in the above equation is too large to directly solve using matrix inversion. Given that the number of unknowns for the 3D coordinate correction $\vec{t}$ is significantly greater than the system error parameters to be solved, we use the elimination method to remove these unknown parameters, thereby obtaining the solution for the correction of the parameters $\vec{k}$. The adjustment normal equation as shown in Equation (14).

$$[N_{11} - N_{12}N_{22}^{-1}N_{12}^T]\vec{k} = L_1 - N_{12}N_{22}^{-1}L_2 \quad (14)$$

where, N_{11} represents $B_{improved}^T B_{improved}$, N_{12} represents $B_{improved}^T C$, N_{22} represents $C^T C$, L_1 represents $B_{improved}^T \hat{L}$, L_2 represents $C^T \hat{L}$. So we could solve $\vec{k}$ as delineated in Equations (15)–(17):

$$\vec{k}^{*} = W_b^{-1} M_b \tag{15}$$

$$W_b = B_{improved}^{T} B_{improved} - B_{improved}^{T} C (C^T C)^{-1} C^T B_{improved} \tag{16}$$

$$M_b = B_{improved}^{T} \hat{L} - B_{improved}^{T} C (C^T C)^{-1} C^T \hat{L} \tag{17}$$

C. The Optimized Matching Process for Multi-Modal Waveforms

Obtaining an accurate position of the actual footprint using waveform matching is a crucial step in the calibration process. The accuracy of the matching position directly influences the accuracy of the calibration model results described in the previous section. A reasonable assumption must be stated: within the same track data, the geolocation errors of the footprints exhibit a consistent trend, which is determined by the systematic errors in the laser system's ranging and pointing measurements. This assumption has been made in the Cal/Val of satellite laser systems by previous research[34, 37, 40, 41]. Unlike their approaches, the algorithm proposed in this paper also considers the random errors of the laser footprints, meaning that the footprint geolocation errors fluctuate randomly around a relative position. Hence, we initially define a search area around the initial recorded footprint location by the laser altimeter. The size of this region is determined based on the planimetric accuracy of the altimeter system. It is imperative that this region covers the actual laser footprint location, as this is a necessary condition for the success of the algorithm. Waveform simulation is performed based on various waveform parameters, including the footprint size, full-waveform sampling frequency, emitted waveform shape, and spatio-temporal distribution. Within the search area, a fixed step length approach is used, applying this process to all the potential footprint locations. Consequently, a set of simulated waveforms is generated in this region. The optimized matching process is performed to obtain an accurate actual footprint position. In the following sections, the detailed matching process is introduced.

1) Waveform Simulation and the Correlation Coefficient Between Simulated and Recorded Waveforms

Waveform simulation is a key technique in waveform matching. Accurately simulating the waveform at potential footprint locations is essential for precise determination of laser footprint positions. The waveform simulation technique was initially proposed by Blair et al. [34] and subsequently refined by Hancock et al. [42]. In this study, we utilized the simulation technique developed by Hancock et al. [42], modeling the transmitted wave energy from both the spatial and temporal domains using a Gaussian function. High-resolution elevation data are used for the waveform simulation. For each point in the footprint, convolution is performed with the transmitted waveform. The convolution waveforms generated by all the points are then accumulated along the ranging axis to produce the simulated waveform within the footprint. The Pearson correlation coefficient, $Corr_{P_r,Ps}$, between the simulated and recorded waveforms is then calculated. The calculation formula for $Corr_{P_r,Ps}$ is provided in Equation (18).

$$Corr_{P_r,Ps} = \frac{n\sum P_r(i)Ps(i) - \sum P_r(i)\sum Ps(i)}{\sqrt{n\sum P_r(i)^2 - (\sum P_r(i))^2}\sqrt{n\sum P_S(i)^2 - (\sum P_S(i))^2}} \tag{18}$$

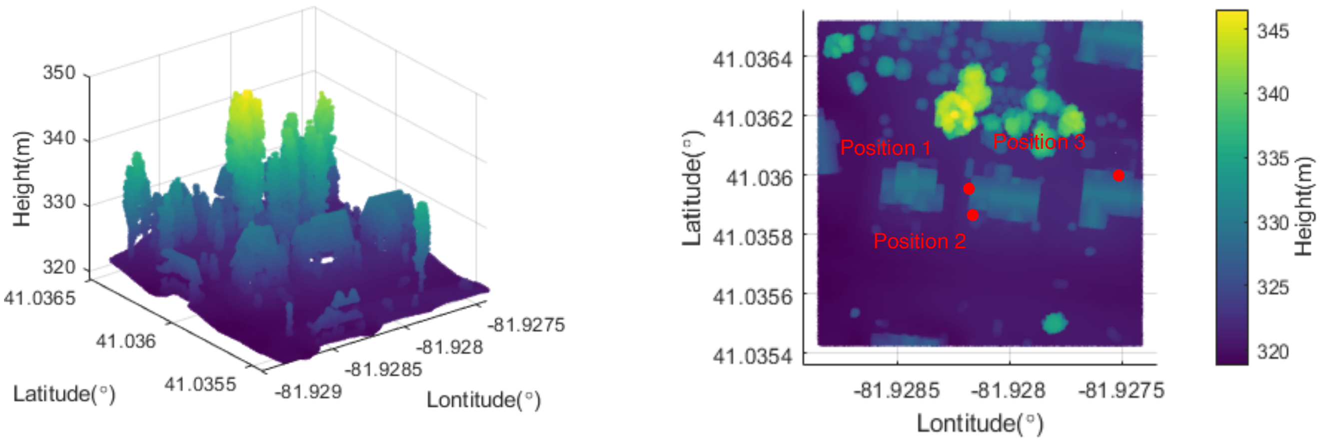
where, n represents the samples of the waveform, $P_r(i)$, $Ps(i)$ represent the i-th sample of the received waveform and simulated waveform, respectively.

To eliminate the uncertainty of energy loss during laser pulse transmission in the atmosphere and detection by the instrument, both the recorded and the simulated waveforms are normalized.

2) Optimized Waveform Matching Method Constrained by the Global Average Correlation Coefficient Matrix

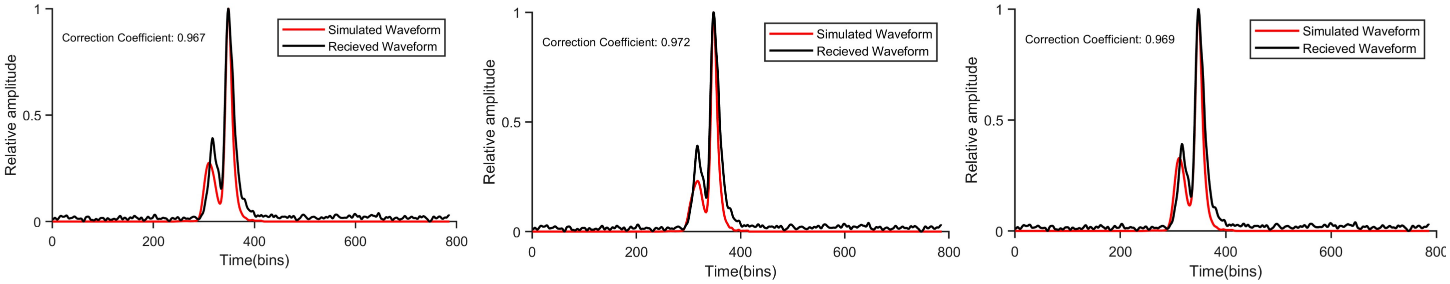
In the previous approaches [26, 30], the point with the highest correlation within the search area was identified as the optimal matching point for individual waveforms. However, experimental analysis has shown that this can be influenced by similar ground features. For example, the simulated waveform for any area at the edge of a house can be very similar to the recorded waveform (such as Position 1 and Position 2 in **Fig. 4**), making it difficult to accurately obtain the actual footprint position through matching, leading to potential mismatches. If there are multiple houses of a similar height in the area (a common scenario in the residential areas of cities, such as Position 2 and Position 3 in **Fig. 4**), this can result in a significant discrepancy between the matching position and the actual laser footprint location, severely impacting the accuracy of the ultimately derived calibration parameters.

Xu et al. [41] proposed a matching approach that utilizes the correlation of the entire track's footprints as a constraint to search for the optimal matching point. Inspired by this idea, we developed an optimal matching method constrained by the global average correlation coefficient matrix M_a'. We used M_a' to reset the initial planimetric address and then performed a local search within the precisely defined search range around this position to determine the optimal matching location (**Fig. 5**). The specific steps are as follows:



(a) Laser point cloud side view

(b) Laser point cloud top view



(c) Simulated waveform of Position 1

(d) Simulated waveform of Position 2

(e) Simulated waveform of Position 3

Fig. 4. An example of waveform similarity using a GEDI waveform at the edge of a house. Positions 1, 2, and 3 are potential geolocations of the GEDI footprint. (a)–(b) Side view and top view of the airborne point cloud near the footprint. (c)–(e) Comparisons of the simulated waveforms for Positions 1, 2, and 3 with the GEDI waveform, respectively.

a) Calculate correlation coefficient matrix M. Centering on the initial footprint position, define a search area with side length l and perform a search with a fixed step size b. At each search position, conduct waveform simulation, compare the simulated waveform with the received waveform, and calculate the Pearson correlation coefficient. The correlation coefficients for all positions within the search area form a correlation coefficient matrix M.

b) Generate the corrected correlation coefficient matrix M'. In Section II-B-2), we analyzed that footprint geolocation errors consist of both systematic and random errors. The system errors of pointing and ranging causes a consistent offset of geolocation errors in the footprints, while random errors cause the footprints to randomly vary around the this offset. To ensure that the systematic error offset consistently achieves the highest waveform correlation and to reduce the impact of random errors on finding the systematic errors offset position, M needs to perform maximum value filtering to generate a corrected correlation coefficient matrix M'. The filter window size is denoted as $w \times w$, where w represents the magnitude of the laser footprint's random error.

c) Determine the planar deviation caused by systematic errors over the entire track. For all the corrected matrices M' of multi-modal waveforms, calculate the mean of these matrices to obtain the global correlation average matrix. Identify the position (x^o, y^o) with the highest correlation coefficient in this matrix, which represents the obtained systematic error offset.

d) Conduct a local search to determine the optimal matching position for each multimodal footprint. Around the offset (x^o, y^o) of each multi-modal footprint, define a search range with a radius of w. Identify the position (x_i^*, y_i^*) within this range that has the highest correlation coefficient M for i-th waveform, and use this as the optimal planar correction for this footprint.

For elevation determination, two approaches are considered. The first approach is treating the waveform as a whole; for instance, in the GEDI system, the entire signal waveform undergoes Gaussian fitting, and the centroid of the fitted curve is used to calculate the corresponding elevation value. The second approach is decomposing the waveform. This involves decomposing the waveform and extracting the elevation of the mode with the lowest height as the position of the footprint. Typically, the lowest mode corresponds to the ground signal, and ground elevation tends to be more stable. In the experiments conducted in this study, we contrasted both methods, evaluating their respective merits and drawbacks.

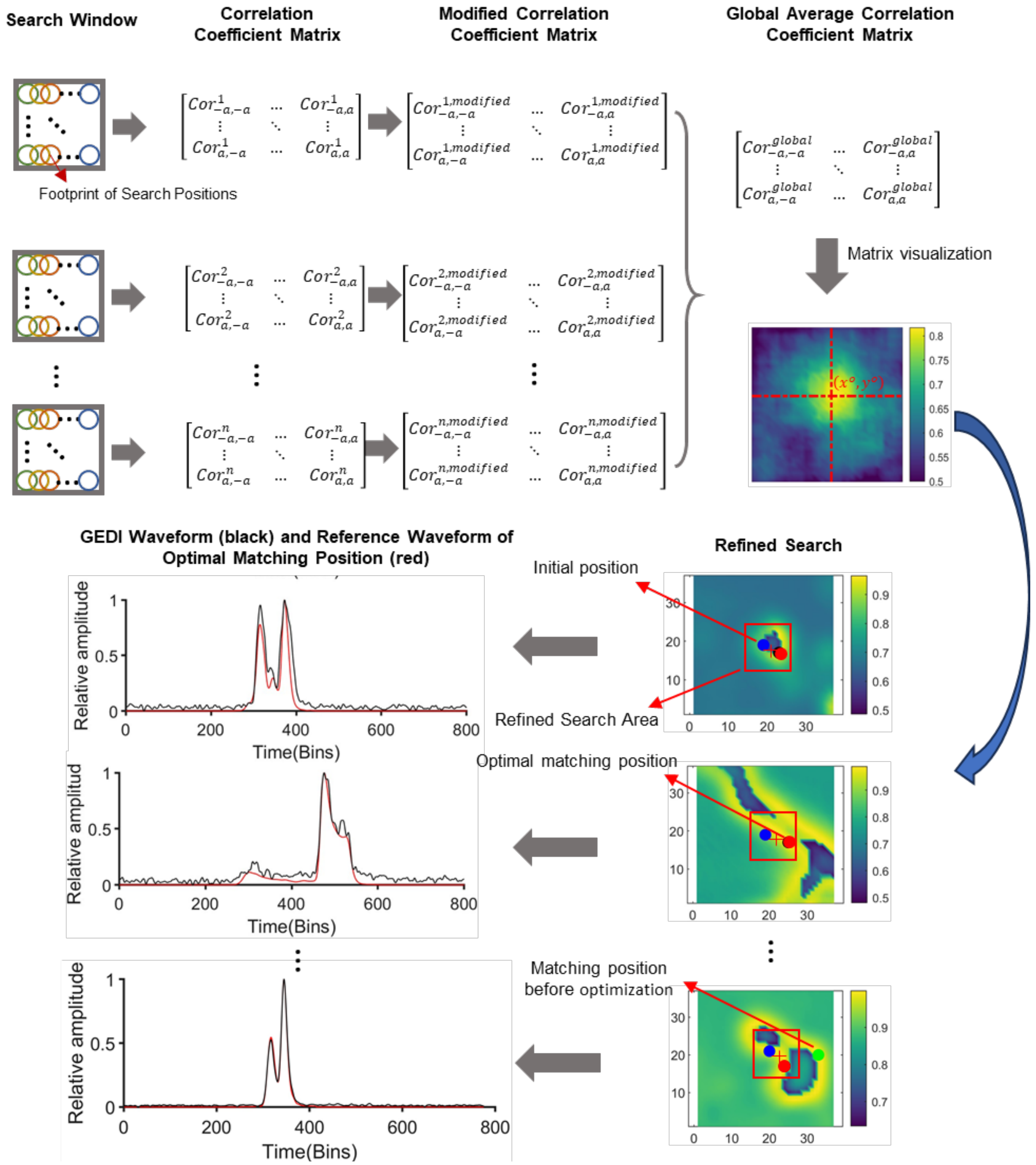


Fig. 5. Explanation of the optimized waveform matching method. In the "Refined Search" column, the blue dots represent the “Initial Position”, the green dots indicate the “matching position before optimization” (the position where the correlation between the simulated waveform and the received waveform is highest), and the red dots represent the “optimal matching position”. The figure provides three matching examples. In the first two examples, the points with the highest correlation coefficient coincide with the optimal matching position. However, in the third example, the point with the highest correlation lies outside the search area, so the optimal matching position is corrected to the red dot's position.

III. DATA AND STUDY AREAS

To validate the effectiveness of the improved on-orbit calibration model proposed in this paper, we selected various laser altimeter datasets for calibration. The available public datasets include ICESat and GEDI datasets. ICESat, operational from 2003 to 2009, was not chosen due to the limited availability of high-accuracy elevation data corresponding to its observation period. In contrast, GEDI, which was launched in 2018 and is still operational, offers a higher sampling frequency of 242 Hz, compared to ICESat’s 40 Hz, providing a more abundant data volume for the waveform matching and calibration tasks. Moreover, GEDI's planimetric positioning accuracy of ~10 m (R02)[40], while consistent with its scientific objective of high-resolution ecosystem structure measurements, limits its performance in terrain mapping. So,

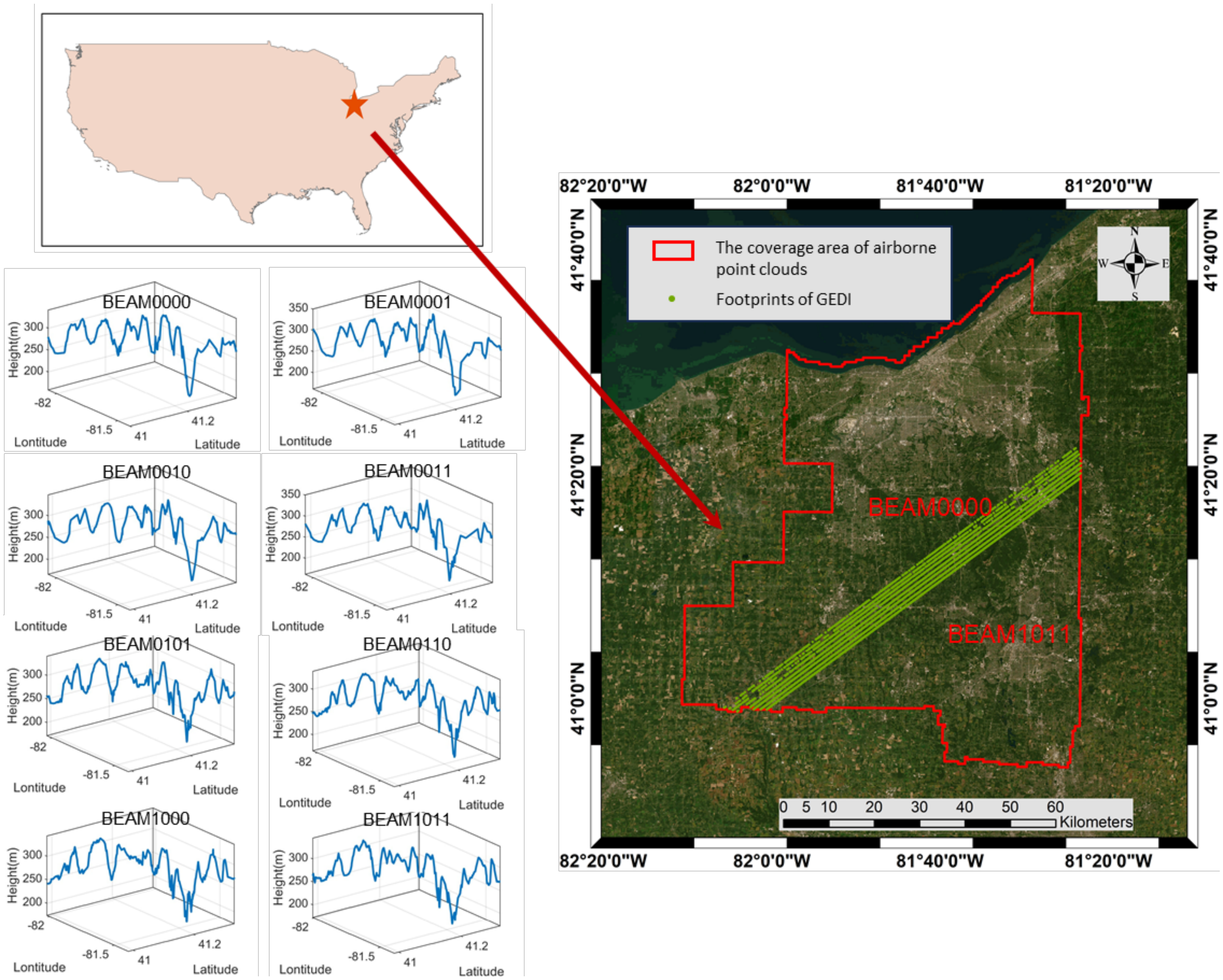
This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2024.3522688

the second release of GEDI data (R02) are selected as the research data. Through calibrating the existing R02 GEDI datasets, our goal is to enhance their planimetric and altimetric accuracy, thereby amplifying their utility in terrain measurement. In addition, we selected 3D Elevation Program (3DEP) data for the waveform simulation, matching the simulated waveforms with recorded ones to acquire accurate footprint location information. The 3DEP data feature a sufficiently large coverage area, high resolution, and a high accuracy, which is necessary for this task. We first introduce the study area in Section III-A, followed by descriptions of the GEDI data and 3DEP data in Sections III-B and III-C, respectively.

A. Study Area

For the selection of the study area, we followed these criteria: 1) Both the 3DEP and GEDI datasets needed to possess observation data of high quality. Given that suboptimal observation conditions with GEDI data can lead to additional observation errors, selecting data from periods with better observation conditions was crucial. 2) The observation years for both datasets needed to be closely aligned. This approach aims to reduce the impact of both natural factors, such as earthquakes, hurricanes, and vegetation growth, and anthropogenic factors, including urban development and deforestation, which can induce changes in terrain and land-cover types, thereby influencing the waveform matching outcomes. 3) The observation seasons of both datasets needed to coincide. As 3DEP data are predominantly collected in the leaf-off season, it was imperative that the GEDI data chosen for comparison were also from the leaf-off season. This ensured the consistency between the simulated and recorded canopy modes, thus improving the simulation accuracy.

Within the constraints of the afore-mentioned criteria, Ohio, USA, was selected as the study area. Within this region, the “2019-2020 USGS Lidar: Ohio Statewide-Phase 1” dataset from 3DEP and the “2022032180417_O17786_02” dataset from GEDI were chosen as the research data. The time interval between the two datasets is two years, with the collection seasons aligning. The selected GEDI dataset exhibits an effective strong beam data ratio exceeding 99% (TABLE I) (determined using the quality_flag tag, which is discussed in Section III-B). A schematic of the study area is shown in **Fig 6**.



This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2024.3522688

Fig. 6. The study site. The red outline represents the coverage area of the airborne point clouds. The eight images on the left show the profiles of the eight beams of GEDI. The green dots represent the GEDI footprints used in this study. The uppermost track represents BEAM 0000, and the lowermost track represents BEAM 1011. The ground tracks of the remaining six beams are arranged sequentially between them.

TABLE I
DATA SITUATION FOR EACH GEDI BEAM, INCLUDING THE TOTAL NUMBER OF FOOTPRINTS, NUMBER OF USABLE FOOTPRINTS, AND PROPORTION OF USABLE FOOTPRINTS

Beam ID	0000	0001	0010	0011	0101	0110	1000	1011
Number of footprints in the ROI	1286	1266	1260	1221	1188	1170	1151	1119
Number of footprints meeting quality standards	328	417	501	608	1177	1167	1135	1116
Proportion	25.5%	32.9%	39.8%	49.8%	99.1%	99.7%	98.6%	99.7%

B. GEDI Data

The GEDI instrument is a unique full-waveform multibeam laser altimeter mounted on the International Space Station (ISS), which is anticipated to yield over 10 billion waveform measurements over its two-year nominal mission. These measurements will primarily target the vertical structure of temperate and tropical forests [2, 13]. As the sole space-based mission dedicated to mapping the vertical layout of vegetation, GEDI's key scientific output is a detailed 1 × 1 km² grid mapping above-ground biomass density (AGBD) across latitudes ranging from 51.6°N to 51.6°S [16, 43]. GEDI comprised of three identify 1064 nm lasers, which generate a total of eight ground beams [2, 36]. The laser referred to as "Coverage" is split into two transects and, through beam dithering units (BDUs), produce four ground tracks (0000, 0001, 0010, and 0011). The other two lasers, known as "Full Power" lasers, only undergo dithering to generate four additional ground tracks (0101, 0110, 1000, and 1011) [2, 36]. The pulse energy of the coverage laser is 5 mJ, while the full power pulse has an energy of 10 mJ [44, 45]. The GEDI instrument captures complete waveforms at each footprint, spaced 60 m along track and approximately 600 m across its orbital tracks, with each footprint spanning roughly 25 m in diameter [16] (**Fig. 7.**).

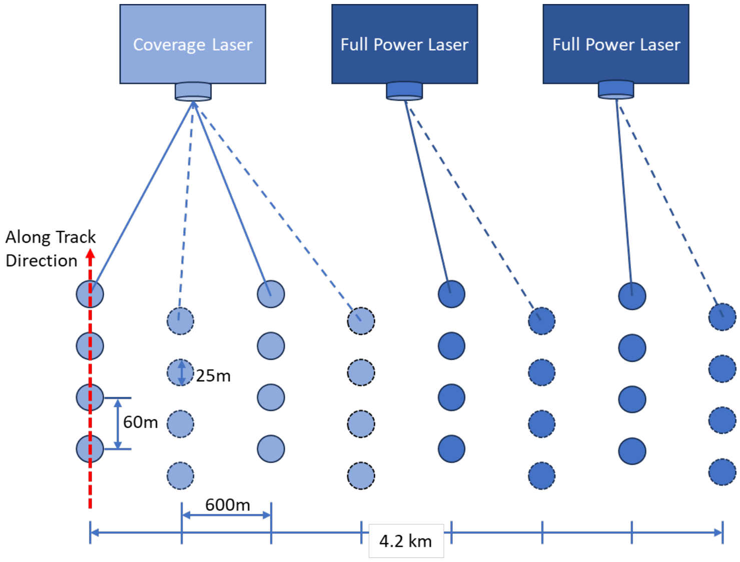


Fig. 7. Schematic diagram of GEDI working mode *(redrawn from Hofton et. al [2])*. On the far left is the "Coverage Laser," while the two on the right are the "Full Power Laser". The Coverage Laser generates four beams, with IDs from left to right labeled as 0000, 0001, 0010, and 0011. The two full power lasers generate additional four beams, with IDs from left to right labeled as 0101, 0110, 1000, and 1011. Beams 0001, 0011, 0110, and 1011 are generated through dithering and are represented by dashed lines in the footprint on the diagram.

The L1B and L2A products were used for the experiments. The L1B product offers geolocated, corrected, and smoothed waveform data, including ancillary, geolocation, and geophysical correction details for each laser pulse across all eight beams. The L2A product, derived from the L1B waveforms, provides elevation and relative height metrics. From the L1B dataset, we extracted both the transmitted and received waveforms, along with the satellite positional data. The L2A dataset yielded footprint geolocations and quality flags. We aligned these datasets using their respective *shot_number*. The details of the specific datasets utilized are provided in TABLE II.

C. 3DEP Data

The 3D Elevation Program (3DEP), managed by the National Geospatial Program of the United States Geological Survey (USGS), is designed to meet the growing demand for high-quality topographic data and various other 3D representations of the nation's natural and constructed features. This collaborative initiative encompasses the collection of airborne LiDAR data across the United States, characterized by high resolution and accuracy. The program achieves a planimetric resolution finer than 1 m, equating to more than two LiDAR points per square meter. The resolution and accuracy can vary among the different contractual projects, but generally, in open areas, the planimetric accuracy is expected to be better than 15 cm, and the vertical accuracy better than 10 cm. Each project report includes a detailed accuracy assessment as part of its open metadata, with over 70% of the contiguous United States already covered with LiDAR data that meet these specifications.

The selected dataset "2019-2020 USGS Lidar: Ohio Statewide-Phase 1" covers specific regions of Ohio gathered from

November 2019 to May 2020. The data were collected by the Leica Terrain Mapper airborne LiDAR mapping system. This system operates at 300–5500 m above ground level (AGL), with a 20°–40° scan angle. It can scan 250 lines per second and achieve a scan frequency of up to 125 Hz. The nominal point spacing is 0.35 m, which means eight points per square meter. The accuracy of the dataset is 13 cm (1 σ) in the horizontal direction and 5 cm (1 σ) in the elevation direction. It is important to note that the dataset uses the NAD83/Ohio North (EPSG:3734) system for the horizontal coordinates and the North American Vertical Datum of 1988 (NAVD88) for the elevation, differing from the WGS84 coordinate system utilized by GEDI. To align with GEDI's system, we converted the coordinate system of the airborne LiDAR point cloud to WGS84, employing the Python library "pyproj" together with NASA's VDatum tool.

TABLE II
L1B AND L2A PRODUCT DATASETS

Product	Group	Dataset	Description
L1B & L2A	/BEAMxxxx/	shot_number	Unique shot ID. Used to align the L1B and L2A products.
L1B	/BEAMxxxx/	rxwaveform, rx_sample_start_index, rx_sample_count, txwaveform, tx_sample_count, tx_sample_start_index	Uses xx_sample_count and xx_sample_start_index to identify the location of each waveform, and obtains the corrected receive (RX) or transmit (TX) waveform of each shot.
	/BEAMxxxx/geolocation/	delta_time, altitude_instrument, latitude_instrument, longitude_instrument	Transmit time of the shot and the geolocation information of the instrument at the laser transmit time.
L2A	/BEAMxxxx/	lat_lowestmode, lon_lowestmode, elev_lowestmode, latitude_1gfit, longitude_1gfit, elevation_1gfit	The longitude, latitude, and elevation information of the lowest mode or a single Gaussian fit to the waveform relative to the reference ellipsoid
	/BEAMxxxx/	quality_flag, degrade_flag	quality_flag is used to select validation waveforms, and degrade_flag indicates the degradation state of the pointing and/or positioning information. The two datasets are used to select validation data.

IV. RESULTS AND ANALYSIS

We employed the proposed calibration method with the selected dataset to calibrate the potential systematic errors in the GEDI system. The experimental design is elucidated in Section IV-A, where a concise overview of the parameters selected in the experiment, along with the experimental procedure, is provided. Subsequently, in Sections IV-B and IV-C, the calibration procedure and the outcomes of the validation are delineated, respectively.

A. Experimental Design

To validate the calibration results, we initially divided the selected GEDI dataset into control and check sets. To ensure the reliability of the algorithm, this division was random. For the full-power beams and coverage beams, we randomly selected 300 and 200 footprints, respectively, as the control set, with the remaining footprints designated as the check set. Given the varying number of beams, the number of footprints in the check set also varied. For the coverage beams, this number fluctuated between 100 and 400, whereas, for the full-power beams, it ranged between 800 and 900. For each beam's control set, we utilized the dataset of waveform counts from the GEDI L1B product, classifying the waveforms into single-modal and multi-modal categories. For the multi-modal waveforms, the correlation was calculated between the simulated and recorded waveforms to ascertain the optimal waveform matching position (x_i^*, y_i^*) for each multi-modal waveform, employing the optimized waveform matching method previously described. The side length l of search area and fixed step size b is set to 70 m and 2 m, respectively. The filter window size w is set to 20 m, which is twice the planar positioning accuracy. Subsequently, the footprint position information for both the single- and multi-modal waveforms was input into the calibration model, as described in Section II-B-1), utilizing the process outlined in **Fig. 3** for calculating the calibration parameters for each beam.

B. Results of the Calibration Process

We compiled the initial optimized matching point (x^o, y^o) for all eight beams (TABLE III). This provided a comprehensive assessment of the geolocation accuracy across the different beams, aiding in the fine-tuning of the system's calibration.

TABLE III
THE INITIAL OPTIMIZED MATCHING POINT (x^o, y^o) FOR EACH BEAM, AND THE MAXIMUM CORRELATION COEFFICIENT IN M_a'

BEAM ID	x^o/m	y^o/m	Max correlation coefficient in M_a'
0000	0	8	0.9352
0001	4	6	0.9469
0010	0	6	0.9466
0011	4	0	0.9142
0101	2	4	0.9880
0110	−2	8	0.9844
1000	0	2	0.9781
1011	2	4	0.9764
Average	1.25	4.25	0.9587

Analysis of the data in TABLE III reveals that seven out of the eight beams exhibit a slight deviation in both the positive X-axis (eastward) and Y-axis (northward) directions, with the sole exception being beam 0110, whose optimal matching position is at (−2 m, 8 m). The uniformity in the beams' directional bias demonstrates the reliability of the results and also suggests a systematic deviation in the pointing of GEDI's beams, leading to planar offsets. The average deviation for the eight beams in the X-axis direction (1.25 m) is marginally smaller than that in the Y-axis direction (2.25 m), which aligns with the planimetric positioning accuracy for the second release of GEDI products, as reported by the GEDI team [40].

Based on the global deviations of each beam, we further analyzed the distribution of the optimal planar correction (x_i^*, y_i^*) of each multi-modal footprint around each beam's systematic error offset (x^o, y^o). The density figures of the footprint distribution (**Fig. 8**) clearly show a pattern where the density is high at the center and gradually decreases toward the edges. When comparing the full-power and coverage beams, it is apparent that the number of multi-modal waveforms in the coverage beams is significantly less, amounting to only one-sixth of that in the strong beams. This discrepancy does not proportionally match the 2:3 ratio of the control set points. In the case of the coverage beams, the discrepancy is largely due to the beam energy, leading to the fewer multi-modal waveforms achieving the quality standards (*quality_flag* equals 1). As previously discussed, the planar deviation of footprints is composed of both systematic and random errors. The systematic error arises from the inaccuracies in pointing angles, while the random error is induced by factors such as attitude measurement errors, satellite platform vibrations, and laser pointing variations. In the simulation and matching process of the waveforms, random errors can also be introduced. The distribution of the optimal matches reflects the random error distribution across the various beams. From the distribution of these optimal matches, it is evident that the planar deviations caused by these factors are not completely random but tend to follow a pattern akin to a Gaussian distribution.

For the single-modal waveforms, we simulated the waveform and computed the elevation of each footprint using waveform fitting. This is a meticulous process. it aligns with GEDI's method of calculating elevation, thereby minimizing the potential bias in the elevation computation. The calibration model's solution is iterative, with each computation resulting in a positional shift of the footprints of the single-modal waveforms until the shift stabilizes. During this process, we continuously updated the elevation of the footprints using the afore-mentioned method.

Nevertheless, while the elevation accuracy of the single-modal waveforms is generally higher than that of the multi-modal waveforms, certain special circumstances (such as steep slopes) can lead to a suboptimal elevation accuracy. Therefore, we also performed a selection of the single-modal waveforms. For each single-modal waveform's corresponding LiDAR points, we selected the points labeled as ground (a classification label of 2 in las/laz files) and calculated the elevation difference within the footprint. If the difference exceeded 3 m, the footprint was excluded from the current iteration.

Using the selected footprints with both single- and multi-modal waveforms, the calibration process was conducted for each of the eight beams individually. We then corrected the positions of the footprints with the calibrated parameters. An elevation comparison was conducted between the corrected positions from GEDI and the extracted positions from the airborne LiDAR points. The results obtained in the control set are listed in TABLE IV. The post-calibration residuals in the control set are noticeably reduced. For the full-power beams, the original altimetry bias of approximately 10 cm is reduced to about 1 cm after calibration, with the mean absolute error (MAE) dropping from around 20 cm to 12 cm, and the root-mean-square error (RMSE) from around 23 cm to 16 cm. For the coverage beams, the original altimetry bias of approximately 15 cm is reduced to less than 10 cm, with the MAE and RMSE dropping from around 25 cm and 30 cm to approximately 15 cm and 20 cm, respectively. These numerical indicators suggest that the residuals from the calibration solution are very small.

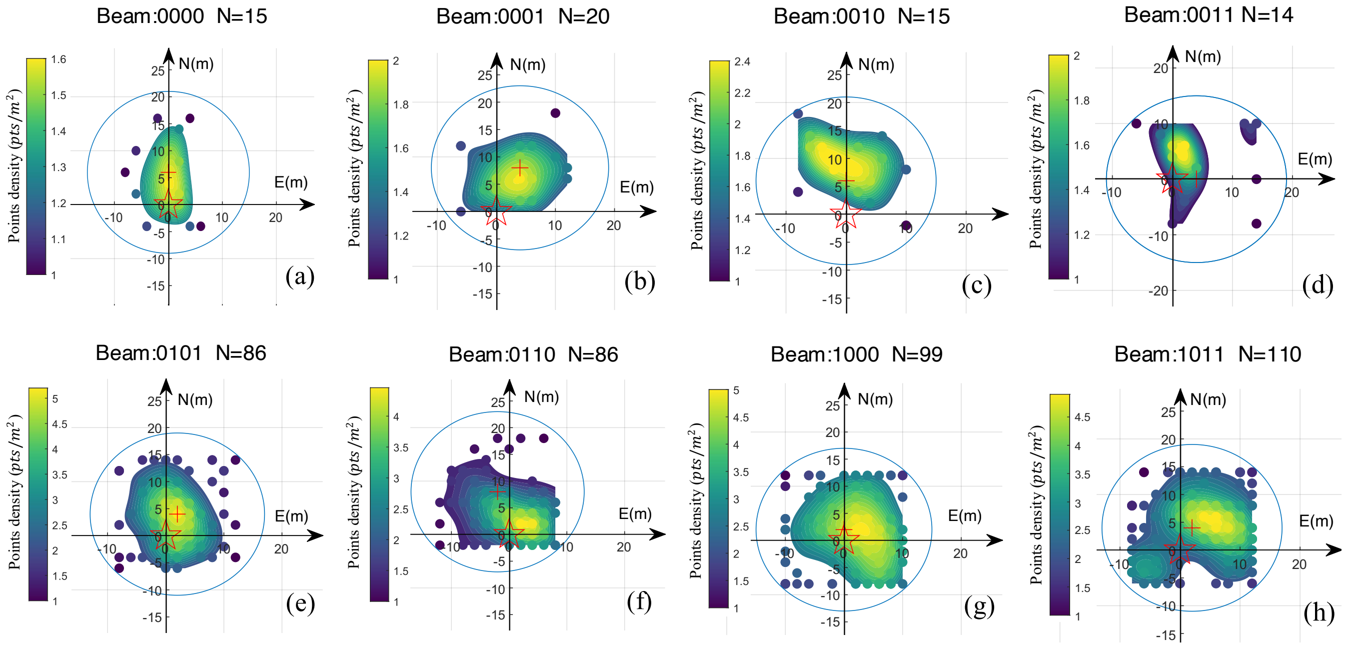


Fig. 8. Points density figures of the optimal matching positions for footprints with multi-modal waveforms. (a)-(h) represent the points density figures within each beam, and *N* denotes the number of multi-modal waveforms selected. The red pentagram represents the initial position, the red cross marks the position of the cumulative maximum value, and the blue circles indicate the search range surrounding the cumulative maximum position.

TABLE IV
ELEVATION DIFFERENCE BETWEEN GEDI AND 3DEP BEFORE AND AFTER CALIBRATION IN THE CONTROL SET

Status	Index	0000	0001	0010	0011	0101	0110	1000	1011
Before calibration	Bias	0.174 m	0.186 m	0.187 m	0.094 m	0.053 m	0.122 m	0.134 m	0.151 m
	MAE	0.246 m	0.284 m	0.256 m	0.236 m	0.135 m	0.178 m	0.206 m	0.199 m
	RMSE	0.322 m	0.377 m	0.342 m	0.312 m	0.180 m	0.230 m	0.274 m	0.247 m
After calibration	Bias	0.095 m	0.057 m	0.020 m	0.011 m	0.037 m	0.002 m	0.005 m	0.004 m
	MAE	0.168 m	0.160 m	0.117 m	0.149 m	0.126 m	0.115 m	0.145 m	0.114 m
	RMSE	0.225 m	0.203 m	0.165 m	0.221 m	0.168 m	0.162 m	0.228 m	0.145 m

C. Validation Results for the Check Set

It was crucial to evaluate the model's generalizability using the check set. We conducted a detailed analysis of the calibration results in terms of the planimetric and altimetric accuracy. Firstly, we corrected the positions of the footprints in the check set using the calculated calibration parameters. From the perspective of planimetric accuracy, we selected the multi-modal waveforms from each beam's check set and searched for the optimal waveform match, comparing the bias and MAE in the X- and Y-directions, as shown in TABLE V.

Directly assessing the planimetric waveform search results, it is evident that the bias and MAE are significantly decreased post-calibration. In particular, the bias in the X-direction is notably reduced, with the four coverage beams showing a decrease to within 3 m and an MAE reduction of over 1 m. Similarly, the full-power beams exhibit a reduction in the X-direction bias to within 1 m and a decrease in the MAE ranging from 0.2 m to 0.9 m. Compared to the X-direction, the Y-direction bias is generally lower. After calibration, the bias in the Y-direction for all the beams is within 2 m, with the MAE levels in the Y-direction comparable to those in the X-direction for the different beams.

We also selected the single-modal waveforms and recalculated the positions using the computed corrections for pointing and ranging. The differences in elevation were then compared in terms of bias, MAE, and RMSE. The results (**Fig. 9** and TABLE VI) from the check set largely align with those from the control set, showing varying degrees of reduction in bias, MAE, and RMSE across all eight beams. Notably, beams 0010, 0110, 1000, and 1011 exhibit a bias of approximately 1 cm in the check set. The other beams, 0001, 0010, and 0101, show bias values of less than 10 cm, except for beam 0000, which has a post-correction bias of greater than 10 cm. This discrepancy is likely due to the limited number of valid points in both the control and check sets for this beam, potentially hindering the effective correction of the systematic biases.

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2024.3522688

Despite its lesser performance compared to the other seven beams, the 0.14 m elevation bias of beam 0000 still signifies a high accuracy. This not only highlights the advantages of spaceborne laser altimetry in elevation measurement, but also demonstrates the efficacy of the improved calibration method proposed in this paper.

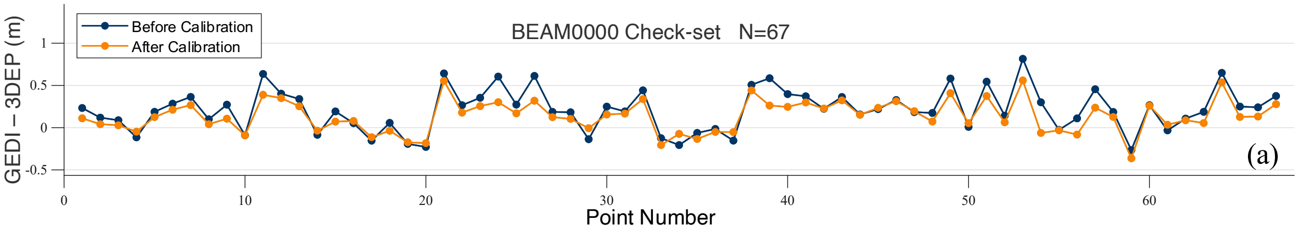
TABLE V
PLANIMETRIC RESULTS OF THE CALIBRATION FOR THE CHECK SET

Beam ID	Number of multi-modal footprints	Status	Bias of X	Bias of Y	MAE of X	MAE of Y
0000	12	BC	3.142 m	2.475 m	5.801 m	4.734 m
		AC	−1.667 m	2.000 m	5.000 m	4.000 m
0001	13	BC	7.981 m	1.920 m	8.841 m	3.918 m
		AC	2.769 m	−1.077 m	6.154m	3.754 m
0010	22	BC	8.214 m	1.088 m	8.698 m	5.962 m
		AC	−0.455 m	0.545 m	5.000 m	5.818 m
0011	32	BC	5.025 m	2.792 m	6.594 m	5.522 m
		AC	1.875 m	−1.500 m	5.250 m	4.750 m
0101	226	BC	3.387 m	1.854 m	5.662 m	4.975 m
		AC	−0.549 m	0.478 m	4.920 m	4.743 m
0110	232	BC	3.596 m	0.078 m	5.831 m	4.728 m
		AC	−0.759 m	0.086 m	4.948 m	4.724 m
1000	260	BC	1.910 m	2.251 m	5.590 m	5.729 m
		AC	0.531 m	1.392 m	5.346 m	5.392 m
1011	289	BC	2.954 m	3.022 m	6.150 m	5.859 m
		AC	−0.706 m	0.069 m	5.744 m	5.439 m

**BC* means before calibration, *AC* means after calibration

TABLE VI
ELEVATION DIFFERENCE BETWEEN GEDI AND 3DEP BEFORE AND AFTER CALIBRATION IN THE CHECK SET

Status	Index	0000	0001	0010	0011	0101	0110	1000	1011
Before calibration	Bias	0.212 m	0.156 m	0.157 m	0.131 m	0.089 m	0.123 m	0.128 m	0.143 m
	MAE	0.268 m	0.268 m	0.255 m	0.236 m	0.149 m	0.168 m	0.183 m	0.195 m
	RMSE	0.324 m	0.357 m	0.325 m	0.309 m	0.186 m	0.225 m	0.233 m	0.252 m
After calibration	Bias	0.136 m	0.053 m	-0.010 m	0.044 m	0.071 m	0.006 m	0.010 m	-0.016 m
	MAE	0.188 m	0.141 m	0.145 m	0.153 m	0.132 m	0.124 m	0.130 m	0.122 m
	RMSE	0.232 m	0.204 m	0.227 m	0.209 m	0.163 m	0.176 m	0.177 m	0.169 m



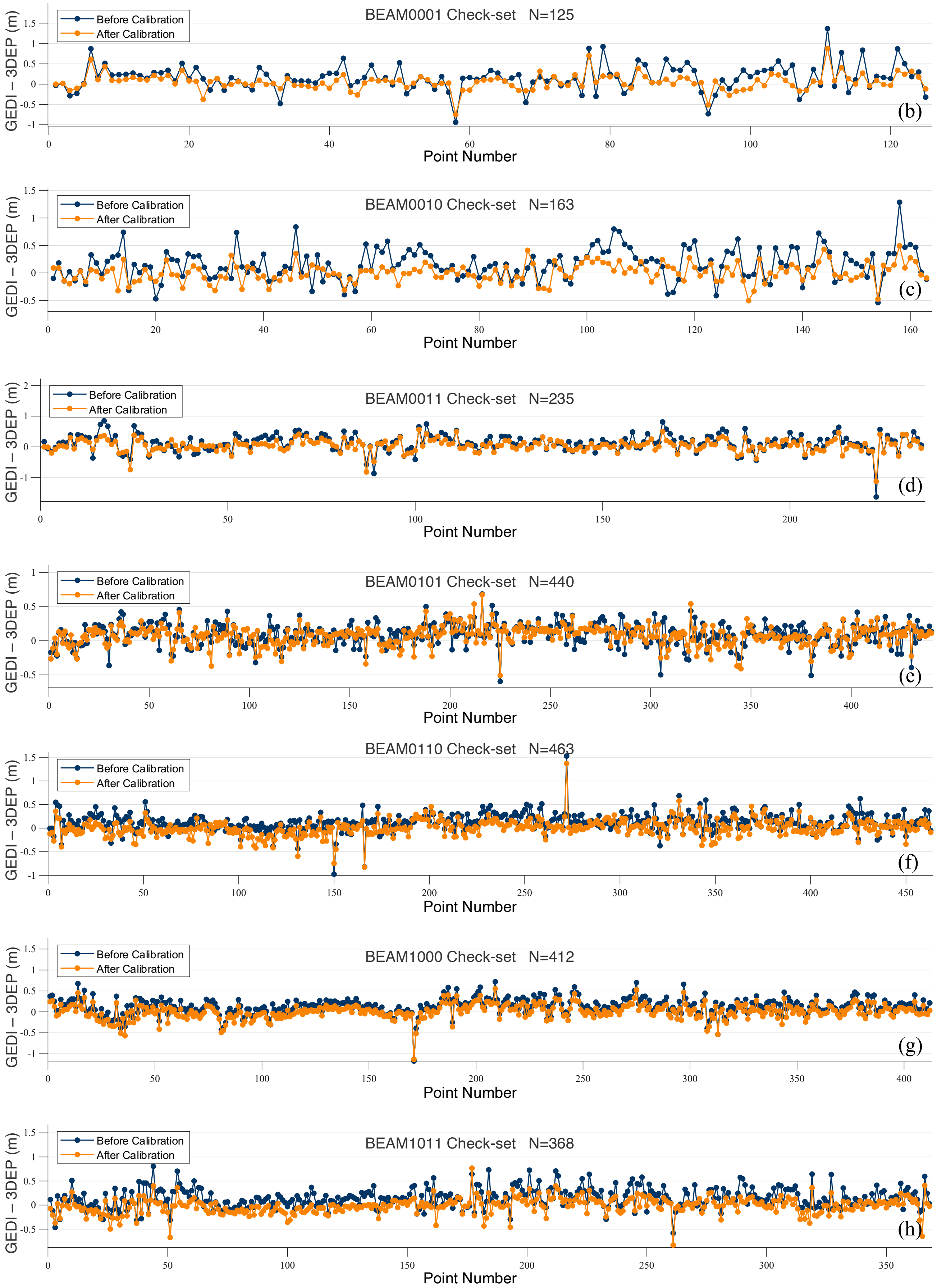

BEAM0001 Check-set N=125
Before Calibration
After Calibration
GEDI – 3DEP (m)
1.5
1
0.5
0
-0.5
-1
0
20
40
60
80
100
120
Point Number
(b)
BEAM0010 Check-set N=163
Before Calibration
After Calibration
GEDI – 3DEP (m)
1.5
1
0.5
0
-0.5
0
20
40
60
80
100
120
140
160
Point Number
(c)
BEAM0011 Check-set N=235
Before Calibration
After Calibration
GEDI – 3DEP (m)
2
1
0
-1
0
50
100
150
200
Point Number
(d)
BEAM0101 Check-set N=440
Before Calibration
After Calibration
GEDI – 3DEP (m)
1
0.5
0
-0.5
0
50
100
150
200
250
300
350
400
Point Number
(e)
BEAM0110 Check-set N=463
Before Calibration
After Calibration
GEDI – 3DEP (m)
1.5
1
0.5
0
-0.5
-1
0
50
100
150
200
250
300
350
400
450
Point Number
(f)
BEAM1000 Check-set N=412
Before Calibration
After Calibration
GEDI – 3DEP (m)
1.5
1
0.5
0
-0.5
-1
0
50
100
150
200
250
300
350
400
Point Number
(g)
BEAM1011 Check-set N=368
Before Calibration
After Calibration
GEDI – 3DEP (m)
1.5
1
0.5
0
-0.5
0
50
100
150
200
250
300
350
Point Number
(h)

Fig. 9. Elevation difference between GEDI and 3DEP before and after calibration in the check set. (a)-(h) represent the elevation difference figures within each beam, and *N* denotes the number of check-points in check-set. The blue line represents the difference before calibration, and the orange line represents the difference after calibration.

TABLE VII

COMPARISON OF THE CALIBRATION RESULTS OF THE TWO MODELS

Beam ID	Corrections of the classical calibration model			Corrections of the improved calibration model			The differential outcomes between the two models		
	Roll (arcsec)	Pitch (arcsec)	Range (m)	Roll (arcsec)	Pitch (arcsec)	Range (m)	Roll (arcsec)	Pitch (arcsec)	Range (m)
0000	1.780	−1.568	−0.504	1.803	−1.568	−0.058	−0.023	0.000	−0.446
0001	3.755	−0.806	−0.288	3.772	−0.805	−0.127	−0.017	−0.001	−0.161
0010	3.658	−2.219	−0.462	3.677	−2.217	−0.038	−0.019	−0.002	−0.424
0011	2.460	0.856	−0.871	2.485	0.861	−0.052	−0.025	−0.005	−0.819
0101	1.981	−0.548	−0.210	1.983	−0.548	-0.115	−0.002	0.000	−0.095
0110	1.757	−1.241	−0.020	1.758	−1.241	0.025	−0.001	0.000	−0.045
1000	0.793	−0.041	−0.265	0.799	−0.039	0.060	−0.006	−0.002	−0.325
1011	2.305	0.169	−0.524	2.315	0.172	−0.024	−0.010	−0.003	−0.500

We also observed an intriguing phenomenon. Although the correction of elevation is primarily related to the correction of the ranging measurement, we noticed that the post-correction elevation residuals do not simply shift downwards uniformly; instead, they exhibit smaller fluctuations and reduced MAE. We found that the uncorrected GEDI products tend to slightly overestimate the ground elevation (by approximately 10 to 20 cm), with the elevation difference between GEDI and 3DEP exceeding 1 m for some footprints. After applying the calibration parameters for correction, we noted that the overestimated elevations are adjusted, primarily due to the corrections in the ranging values. However, we also observed corrections in underestimations of elevation (notably in beam 0011, although similar phenomena were observed in the other beams as well), which we believe is more related to the correction of the footprint's planar position. The adjustment of the planar location leads to more accurate elevation measurements.

It is important to note that in this dataset, we found that even before calibration, the accuracy of the GEDI product had already achieved a promising result. As shown in TABLE VI, the RMSE for the coverage beams is less than 0.4 m, and for the full power beams, the RMSE for elevation is less than 0.3 m. However, this does not imply that GEDI products consistently achieve such accuracy. First, the dataset we selected is limited, consisting of only a single track of data. Second, we filtered this track based on observation conditions: it was collected under clear weather, and the instrument was in good operating condition. Therefore, we can only analyze the effectiveness of our algorithm by examining the changes in RMSE before and after calibration.

V. DISCUSSION

A. Comparison Between the Classical Calibration Model and the Improved Calibration Model

To better highlight the advantages of the improved calibration model, we compared the results of the classical calibration model with those of the improved calibration model. From the results (Table VI), it can be seen that the pointing calibration outcomes of the two models are similar, which is understandable since both models use the same waveform matching process. However, there is a significant difference in the ranging correction, especially evident in the coverage beams, with beam 0011 showing a ranging correction of up to 0.8 m, which is the largest discrepancy among the eight beams.

To compare the calibration effects of the two models, we applied the correction results of the two models to correct the positions of the footprints and compared the performance of the footprint elevation measurement with the beams. We conducted a histogram analysis of the height differences between GEDI and 3DEP for both the single- and multi- modal waveforms, in both the full-power and coverage beams (**Fig. 10**). We also calculated the mean and median deviations for each group (**Fig. 10**). Compared to the distribution of the elevation bias in the multi-modal waveforms, the distribution in the single-modal waveforms is more concentrated, consistent across both coverage beams and full-power beams. However, the multi-modal waveforms show more deviation outliers, which are especially pronounced in the coverage beams.

We conducted an in-depth comparison of the calibration results obtained by the two models. The difference in the elevation calculation between the two models is about 0.5 m for the coverage beams and 0.2 m for the full-power beams. Using the improved model, the mean and median elevation bias for the single-modal waveforms are both within 3 cm. For the multi-modal waveforms, the mean and median bias show slight differences, with the median deviations remaining within 5 cm for both the coverage and full-power beams, and the mean deviations being −0.42 m and −0.25 m, respectively. When using the classical calibration model for calculation, the performance in terms of the mean deviation for the multi-modal waveforms is around 1 cm, but the other indicators are not ideal.

We also analyzed the different performances of the two models with the full-power and coverage beams in detail. The

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2024.3522688

classical calibration model derives the elevation correction from the lowest mode elevation measurements of the multi-modal waveforms, generally corresponding to the ground mode within the footprint. This model eliminates the mean elevation bias in the multi-modal waveforms but is affected by the numerous elevation errors in these footprints. In contrast, the improved model uses the single-modal waveforms with less than 3 m of elevation change within the footprint. The ranging accuracy of the single-modal waveforms is higher than that of the multi-modal waveforms, due to the stronger energy concentration and reduced terrain complexity in footprints with single-modal waveforms, leading to the good performance of the improved model.

In terms of model stability, the improved model clearly has great advantages. The classical calibration model shows an inconsistent performance in the full-power and coverage beams, with slightly smaller deviations in the full-power beams (around 20 cm) but larger deviations in the coverage beams (around 50 cm). We believe that this may be related to the signal-to-noise ratio and waveform recognition. In the coverage beams, noise can more easily affect the start and end positions of the signal wave. We selected two typical waveforms of coverage beams, one where GEDI overestimates the ground elevation and one where it underestimates the ground elevation (**Fig. 11**). By simulating the waveforms and using airborne LiDAR points corresponding to the footprints, we could deduce the correct position of the ground mode. **Fig. 11-a)** displays a situation where, under the influence of noise waves, the waveform decomposition algorithm mistakenly identifies the original noise wave as a signal mode and erroneously assigns its corresponding elevation as the ground elevation, resulting in an underestimated ground elevation. **Fig. 11-b)** presents a situation where the ground has a slope; affected by noise, the algorithm's recognized lowest mode position significantly deviates from the simulated lowest mode position, leading to an overestimated ground elevation.

Previous research[46, 47] has shown that the ranging error, $d\rho$, of the spaceborne laser altimeter is positively correlated with the slope of the laser footprint, S, and negatively correlated with the echo energy, E. **Fig.10** indicates that the multi-modal group in the coverage beam has a higher gross error rate (proportion of points with deviations exceeding 3 m) than the other three groups (multi-modal and single-modal in full-power beam, and single-modal group in coverage beam), theoretically due to significantly lower echo energy in this group (multi-modal footprint experiences more energy loss in the transmission link, while the full-power beam has twice the energy of the coverage beam). **Fig. 11-b)** further analyzes that the altitudinal deviation for footprints with a significant slope increases, as sloped terrain increases the spread of the echo, which presents additional challenges in accurately extracting ground elevations on sloped surfaces.

For the performance of different models' corrected results under varying slopes, we also conducted statistical analysis (**Fig. 12**). From the graph, it is evident that regardless of whether it is under coverage beams or full power beams, as the slope increases, errors gradually enlarge in the improved model, classical model, and GEDI results. full power beams generally exhibit smaller errors compared to coverage beams, reflecting the positive correlation between greater emission energy and higher altimetric precision, especially on steeper slopes, where full power beams show greater advantages. The improved model shows near-zero deviation within a 0° to 5° slope range, while the GEDI result and classical model both tend to overestimate ground elevation. As the slope increases, all three groups gradually show a tendency to underestimate ground slope, with errors increasing with slope. The results from the improved model are very close to those of GEDI, but an interesting phenomenon occurs in the 5° to 10° and 10° to 15° ranges, where GEDI's deviation is smaller. However, as the slope further increases, the deviation in the improved model results slightly outperforms GEDI. We believe this phenomenon occurs because the improved model includes corrections for the planar position. The results of the classical model and improved model resemble two parallel lines, with their differences stemming from corrections in range measurement values. In the three slope groups exceeding 10°, the results of the classical model are superior to those of both GEDI and the improved model.

Of course, we must also consider another scenario where the errors in matching could lead us to mistakenly believe that the ground elevation has been overestimated or underestimated. The matching errors are primarily due to a low signal-to-noise ratio, which allows noise waveforms to affect the judgment of signal modal, resulting in higher correlations at incorrect matching positions. Another potential cause could be the instrumental differences between airborne LiDAR data and spaceborne laser altimeters in recording reflectance values, leading to incomplete waveform matches. When mismatches occur, the resulting altimetry deviations at sloped locations are significantly higher than in flat areas, which is related to the nearly 90° zenith angle of the laser altimeter[39].

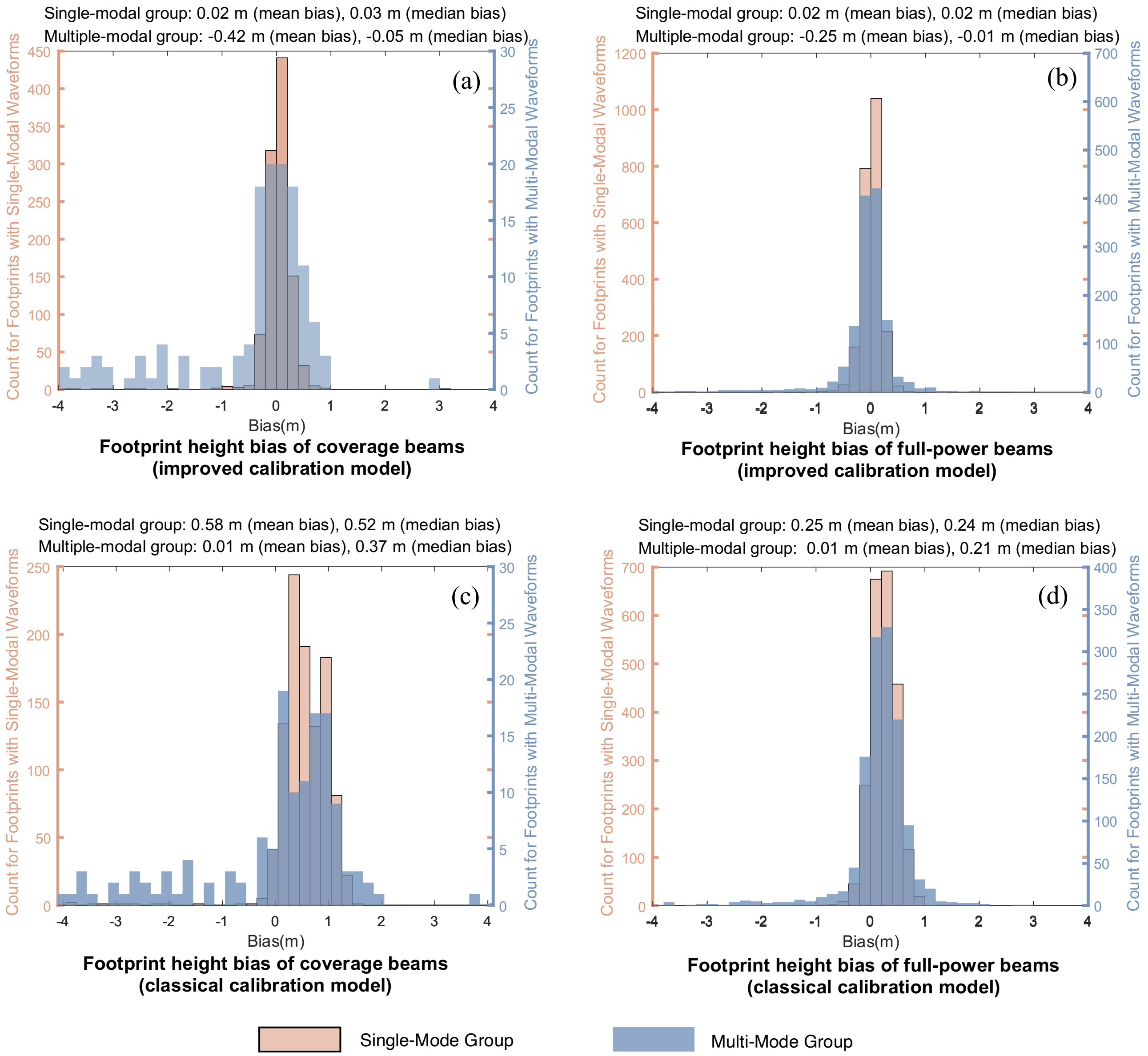


Fig. 10. Statistical histograms of the elevation measurement performance of the two models across coverage beams and full-power beams. The upper row shows the outcome of the improved calibration model in both types of beams, and the lower row depicts the outcome of the classical calibration model. The pink and blue colorations denote the single-modal and multi-modal groups, respectively.

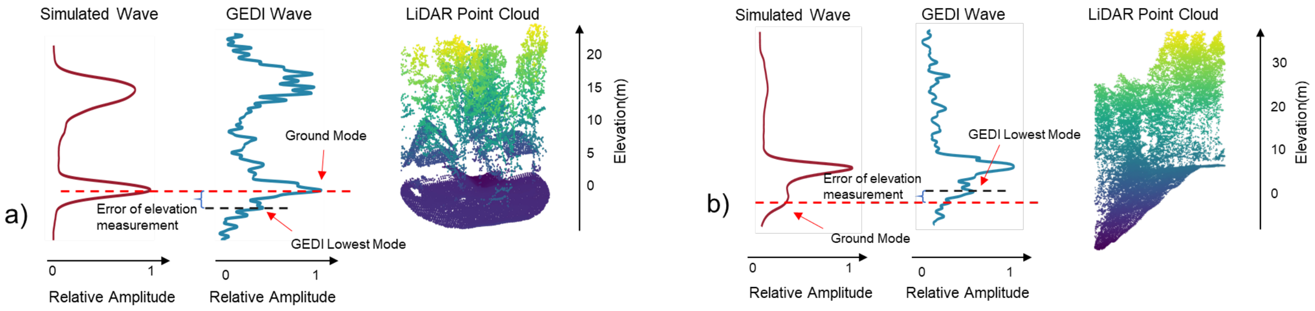


Fig. 11. Two representative waveforms and their corresponding ground features within footprints. a) exemplifies cases where GEDI underestimates terrain elevation, while b) exemplifies cases where GEDI overestimates terrain elevation. Each figure contains three subplots: the first column of subplots represents the waveforms simulated using LiDAR point clouds, the second column represents the waveform data recorded by GEDI, and the third column depicts the conditions of the LiDAR point cloud

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2024.3522688

data.

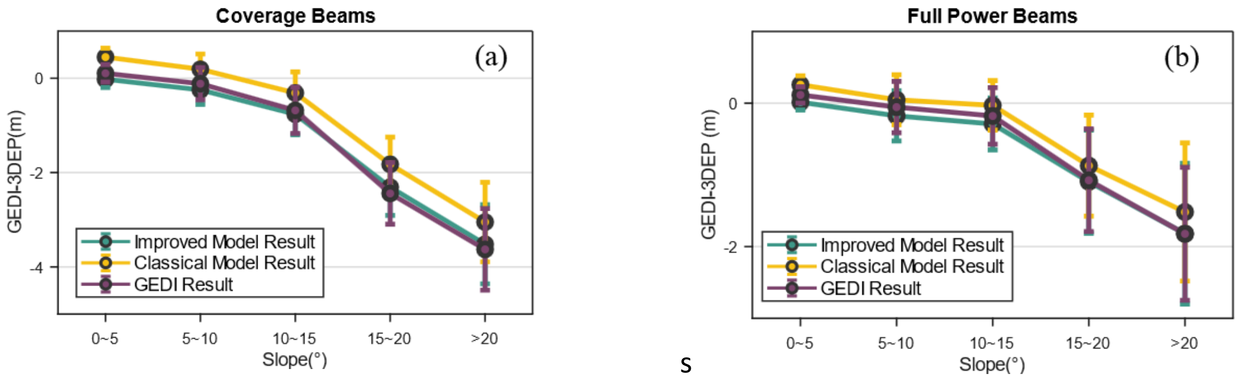
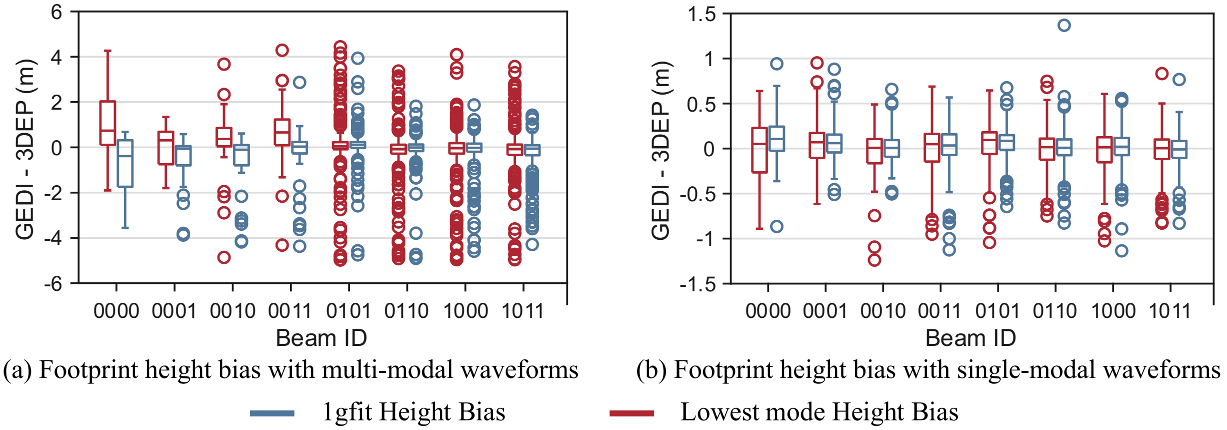


Fig. 12. Errorbar in different slope of the improved model result, classical model result, and GEDI result. a) represents the errors of three groups under different slopes in coverage beams. b) represents the errors of three groups under different slopes in full power beams.

B. Comparison of the Lowest Mode Elevation and "1gfit" Elevation in GEDI

The accuracy of the elevation of the single-modal waveforms significantly impacts the calibration results for ranging, as discussed in the previous section. In GEDI L2A products, two types of elevation information are provided for a footprint: the lowest mode elevation and the "1gfit" elevation. Generally speaking, the lowest mode elevation tends to represent the ground elevation of the footprint, while the "1gfit" elevation corresponding to the center of a single gaussian fit to the waveform, reflecting the overall elevation of the ground features within the footprint. Both are expressions of the footprint elevation, and this section focuses on discussing the accuracy of these two types of elevation.

We compared the two types of elevations obtained from the eight beams with the elevation values obtained from simulating waveforms with the airborne LiDAR point cloud. **Fig. 13-a)** represents the elevation differences at the multi-modal waveforms between the two methods, while **Fig. 13-b)** represents the elevation differences at the single-modal waveforms. The figures show that the lowest mode elevation is more stable than the 1gfit elevation because, in all the comparison groups, the lowest mode elevation box is shorter than the 1gfit elevation box in terms of the box and whiskers. However, there are differences in performance within specific groups. The most significant difference is observed in the coverage beams' multi-modal group, where the GEDI 1gfit obtained footprint elevations are clearly overestimated, while the lowest mode elevations show a slight underestimation. As analyzed earlier, this group's data have a low signal-to-noise ratio, and the elevation calculation results are significantly affected by noise. Nonetheless, the lowest mode elevation results are more reliable than the 1gfit elevation. In the full-power beams' multi-modal groups and all the single-modal groups, the lowest elevation not only has a more concentrated elevation difference distribution, but also maintains nearly the same median elevation difference as the 1gfit elevation.



(a) Footprint height bias with multi-modal waveforms

(b) Footprint height bias with single-modal waveforms

Fig. 13. Performance of the 1gfit elevation versus the lowest mode elevation in footprints with multi-modal and single-modal waveforms. Figure a) depicts their performance in multi-modal waveforms, while figure b) illustrates their performance in single-modal waveforms.

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2024.3522688

C. The Discussion of The Parameters Setting of The Optimized Matching Process

In regard to the numerous parameters set for our proposed method, a detailed discussion is desired to enable more laser altimeters to utilize this method for calibration calculations.

There are several parameters associated with the matching process of multi-modal wave peaks, including the search area side length l, step size b, and the filter window w. The side length of the search area is related to the initial positioning accuracy, typically set at 2 to 3 times the precision of initial positioning. However, during the initial phases of a laser altimeter operation, ground deviations can reach several hundred meters or even exceed 1 km (Beam 2 of GF-7)[26]. Under such circumstances, conducting a broad-range waveform matching search is highly impractical. The reasons include, firstly, that an excessively large search range can lead to significantly prolonged computation times, and secondly, that a larger search area increases the probability of encountering similar terrains within the region, which makes pinpointing the optimal matching position exceedingly challenging. Typically, for laser altimetry systems not performing ocean-maneuvers, a "two-step" strategy[35] is often employed. This involves initially reducing the system's bias within a certain range using terrain profile matching, where the pointing angle deviation is generally less than 20 arcsec. This is followed by precise waveform matching to determine the final calibration parameters.

Regarding the side length for searching, it needs to be larger than the resolution of the high-resolution elevation product used (typically high-precision airborne point clouds or large-scale DEM/DSM products), with a usual step size setting of 1m to 2m. The size of the filter window w is designed based on the size of the footprint's random error, which generally originates from the satellite platform's attitude error and the laser emission jitter. These factors are related to the overall design of the laser altimetry system and are typically set within the range of 5m to 20m. This approach allows for effective calibration while minimizing potential errors due to similar terrain features within the search area.

VI. CONCLUSION

In this paper, we have introduced a improved on-orbit calibration model for full-waveform laser altimeters, which integrates the characteristics of footprints with single-modal and multi-modal waveforms. Utilizing recorded waveforms and high-resolution elevation data such as airborne LiDAR data, the model can estimate and correct the systematic errors in the ranging and pointing of laser altimeters, potentially improving geolocation accuracy. Detailed experiments were conducted using on-orbit GEDI data, yielding the following conclusions:

1. The proposed model can effectively correct the systematic deviations in the geolocation of laser altimeters. By dividing the selected GEDI data into control and check sets, the systematic bias in the pointing and ranging for the eight GEDI beams was calculated and subsequently validated. The validation results indicated that the planimetric bias can be constrained within 3 m for the coverage beams and within 1 m for the full-power beams. The elevation bias across all eight beams showed a significant improvement, with all but one beam achieving a post-calibration bias within 10 cm. The RMSE of the elevation bias for all the coverage beams was corrected to within 0.25 m, and for all the full-power beams to within 0.2 m.

2. We examined the accuracy of the elevation obtained from multi-modal and single-modal waveforms and the impact of different elevation types using the classical calibration model and the improved model. A comparison of the validation results of the two calibration methods revealed that using the improved model for calibration performs well in both the coverage beams and full-power beams. However, when using the classical calibration model, the results are satisfactory in the full-power beams but poorer in the coverage beams, which is mainly due to the influence of outliers.

3. The method proposed in this paper has significant practical value, which can be demonstrated in the following two aspects: First, in the period of the initial launch of spaceborne laser altimeter system, the model presented in this paper can effectively calibrate the laser's pointing and ranging, thereby improving the overall accuracy of the spaceborne laser altimeters. Second, when laser products are used for in-depth applications, it is possible to extract data from the full track data to correct potential horizontal and elevation errors, thus enhancing the measurement accuracy of the entire track.

However, the method proposed in this paper still relies on waveform matching for horizontal position correction, which imposes certain requirements on the data collection time interval between high-resolution elevation data and laser altimetry data. This limitation restricts the method's applicability in some cases. Further research is required to improve the matching method, with the goal of enhancing the model's broader applicability.

APPENDIX

A. The Derivation Process of the Rotation Matrix R_m

We define φ, ω and κ as the misalignment angles of the laser pointing in the pitch (Y-axis), roll (X-axis), and yaw (Z-axis) directions, respectively. Additionally, we define the rotation sequence of the axes as ZYX (first rotating about the Z-axis, then the Y-axis, and finally the X-axis).

The rotation matrix for a rotation about the X-axis can be expressed as Equation (a-1).

$$R_x(\omega)=\begin{bmatrix}1 & 0 & 0\\ 0 & cos(\omega) & -sin(\omega)\\ 0 & sin(\omega) & cos(\omega)\end{bmatrix} \tag{a-1}$$

The rotation matrix for a rotation about the Y-axis can be expressed as Equation (a-2).

$$R_y(\varphi)=\begin{bmatrix}cos(\varphi) & 0 & sin(\varphi)\\ 0 & 1 & 0\\ -sin(\varphi) & 0 & cos(\varphi)\end{bmatrix} \tag{a-2}$$

The rotation matrix for a rotation about the Z-axis can be expressed as Equation (a-3).

$$R_z(\kappa) = \begin{bmatrix} cos(\kappa) & -sin(\kappa) & 0 \\ sin(\kappa) & cos(\kappa) & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{(a-3)}$$

Considering that φ, ω and κ are very small angles (typically small than 20 arcseconds), we can approximate $cos(\omega)$, $cos(\varphi)$, and $cos(\kappa)$ as 1. $sin(\omega)$, $sin(\varphi)$, and $sin(\kappa)$ can be approximated by φ, ω, and κ, respectively. Therefore, Equations (a-1), (a-2), and (a-3) can be simplified to (a-4), (a-5), and (a-6).

$$R_x(\omega) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\omega \\ 0 & \omega & 1 \end{bmatrix} \qquad \text{(a-4)}$$

$$R_y(\varphi) = \begin{bmatrix} 1 & 0 & \varphi \\ 0 & 1 & 0 \\ -\varphi & 0 & 1 \end{bmatrix} \qquad \text{(a-5)}$$

$$R_z(\kappa) = \begin{bmatrix} 1 & -\kappa & 0 \\ \kappa & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{(a-6)}$$

For the defined ZYX rotation sequence, the rotation matrix R_m can be expressed as Equation (a-7).

$$R_m = R_x(\omega)R_y(\varphi)R_z(\kappa) \qquad \text{(a-7)}$$

The result is in Equation (a-8).

$$R_m = \begin{bmatrix} 1 & -\kappa & \varphi \\ \kappa & 1 & -\omega \\ -\varphi & \omega & 1 \end{bmatrix} \qquad \text{(a-8)}$$

REFERENCES

[1] T. A. Neumann *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.,* vol. 233, p. 111325, 2019.

[2] *Algorithm Theoretical Basis Document (ATBD) for GEDI Transmit and Receive Waveform Processing for L1 and L2 Products*.

[3] B. Schutz, "Laser footprint location (geolocation) and surface profiles," *Geoscience laser altimeter system (GLAS): algorithm theoretical basis document version,* vol. 3, pp. 11-19, 2002.

[4] H. Zhou, Y. Chen, Y. Ma, S. Li, H. Juha, and L. Pei, "The effect of terrestrial surface slope and roughness on laser footprint geolocation error for spaceborne laser altimeter," *Photogramm. Eng. Remote Sens.,* vol. 84, no. 10, pp. 647-656, 2018.

[5] K. K. Williams and M. T. Zuber, "Measurement and analysis of lunar basin depths from Clementine altimetry," *Icarus,* vol. 131, no. 1, pp. 107-122, 1998.

[6] M. T. Zuber *et al.*, "The Mars Observer laser altimeter investigation," *J. Geophys. Res. E,* vol. 97, no. E5, pp. 7781-7797, 1992.

[7] R. E. Gold *et al.*, "The MESSENGER mission to Mercury: scientific payload," *Planet. Space Sci,* vol. 49, no. 14-15, pp. 1467-1479, 2001.

[8] C. Li *et al.*, "Laser altimetry data of Chang'E-1 and the global lunar DEM model," *Sci. China Earth Sci.,* vol. 53, pp. 1582-1593, 2010.

[9] H. Araki *et al.*, "Lunar global shape and polar topography derived from Kaguya-LALT laser altimetry," *Science,* vol. 323, no. 5916, pp. 897-900, 2009.

[10] D. E. Smith *et al.*, "Initial observations from the lunar orbiter laser altimeter (LOLA)," *Geophys. Res. Lett.,* vol. 37, no. 18, 2010.

[11] K. Gunderson, N. Thomas, and M. Rohner, "A laser altimeter performance model and its application to BELA," *IEEE Trans. Geosci. Remote Sens.,* vol. 44, no. 11, pp. 3308-3319, 2006.

[12] J. B. Abshire *et al.*, "Geoscience laser altimeter system (GLAS) on the ICESat mission: on - orbit measurement performance," *Geophys. Res. Lett.,* vol. 32, no. 21, 2005.

[13] R. Dubayah *et al.*, "The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography," *Sci. Remote Sense.,* vol. 1, p. 100002, 2020.

[14] X. Tang *et al.*, "Overview of the GF-7 laser altimeter system mission," *Earth Space Sci.,* vol. 7, no. 1, p. e2019EA000777, 2020.

[15] T. Markus *et al.*, "The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation," *Remote Sens. Environ.,* vol. 190, pp. 260-273, 2017, doi: 10.1016/j.rse.2016.12.029.

[16] A. Liu, X. Cheng, and Z. Chen, "Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals," *Remote Sens. Environ.,* vol. 264, p. 112571, 2021.

[17] P. Potapov *et al.*, "Mapping global forest canopy height through integration of GEDI and Landsat data," *Remote Sens. Environ.,* vol. 253, p. 112165, 2021.

[18] L. Duncanson *et al.*, "Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California," *Remote Sens. Environ.,* vol. 242, p. 111779, 2020.

[19] L. S. Sørensen *et al.*, "Mass balance of the Greenland ice sheet (2003–2008) from ICESat data–the impact of interpolation, sampling and firn density," *Cryosphere,* vol. 5, no. 1, pp. 173-186, 2011.

[20] H. Ewert, A. Groh, and R. Dietrich, "Volume and mass changes of the Greenland ice sheet inferred from ICESat and GRACE," *J. Geodyn.,* vol. 59, pp. 111-123, 2012.

[21] M. Bagnardi, N. T. Kurtz, A. A. Petty, and R. Kwok, "Sea Surface Height Anomalies of the Arctic Ocean From ICESat - 2: A First Examination and Comparisons With CryoSat-2," *Geophys. Res. Lett.,* vol. 48, no. 14, p. e2021GL093155, 2021.

[22] Y. Yu, D. T. Sandwell, S. T. Gille, and A. B. Villas Bôas, "Assessment of ICESat-2 for the recovery of ocean topography," *Geophys. J. Int.,* vol. 226, no. 1, pp. 456-467, 2021.

[23] R. Kwok *et al.*, "Surface height and sea ice freeboard of the Arctic Ocean from ICESat-2: Characteristics and early results," *J. Geophys. Res. Oceans,* vol. 124, no. 10, pp. 6942-6959, 2019.

[24] J. Xie *et al.*, "ZY3-02 laser altimeter footprint geolocation prediction," *Sensors,* vol. 17, no. 10, p. 2165, 2017.

[25] X. Tong *et al.*, "Framework of jitter detection and compensation for high resolution satellites," *Remote Sensing,* vol. 6, no. 5, pp. 3944-3964, 2014.

[26] R. Liu and J. Xie, "Calibration of the laser pointing bias of the GaoFen-7 satellite based on simulation waveform matching," *Opt. Express,* vol. 29, no. 14, pp. 21844-21858, 2021.

[27] S. B. Luthcke, D. D. Rowlands, D. J. Harding, J. L. Bufton, C. C. Carabajal, and T. A. Williams, "ICESat Laser Altimeter Pointing, Ranging and Timing Calibration From Integrated Residual Analysis: A Summary of Early Mission Results," in *Flight Mechanics Symposium 2003*, 2003, no. NASA/CP-2003-212246: National Aeronautics and Space Administration.

[28] J. Xie *et al.*, "A Geometric Calibration Method Without a Field Site of the GF-7 Satellite Laser Relying on a Surface Mathematical Model," *IEEE Trans. Geosci. Remote Sens.,* vol. 61, pp. 1-14, 2022.

[29] X. Tang *et al.*, "GF-7 dual-beam laser altimeter on-orbit geometric calibration and test verification," *Sinica Acta Geod. Cartogr. Sin.,* vol. 50, no. 3, p. 384, 2021.

[30] S. Li, C. Li, G. Zhang, and Y. Wang, "Geometric calibration of satellite laser altimeters based on waveform matching," *Photogramm. Rec.,* vol. 36, no. 174, pp. 104-123, 2021.

[31] H. Yi, S. Li, Y. Ma, K. Huang, H. Zhou, and G.-Y. Shi, "On-orbit calibration of satellite laser altimeters based on footprint detection," *Acta. Phys. Sin.,* vol. 66, no. 13, 2017.

[32] J. Xie *et al.*, "In - orbit geometric calibration and experimental verification of the ZY3 - 02 laser altimeter," *Photogramm. Rec.,* vol. 33, no. 163, pp. 341-362, 2018.

[33] B. E. Schutz, "GLAS altimeter post-launch calibration/validation Plan," *Version,* vol. 1, pp. 58-61, 2001.

[34] J. B. Blair and M. A. Hofton, "Modeling laser altimeter return waveforms over complex vegetation using high-resolution elevation data," *Geophys. Res. Lett.,* vol. 26, no. 16, pp. 2509-2512, 1999.

[35] G. Li *et al.*, "The method of GF-7 satellite laser altimeter on-orbit geometric calibration without field site," *Sinica Acta Geod. Cartogr. Sin.,* pp. 401-412, 2022.

[36] S. B. Luthcke, T. Rebold, T. Thomas, and T. Pennington, "Algorithm Theoretical Basis Document (ATBD) for GEDI Waveform Geolocation for L1 and L2 Products," ed: NASA, 2019.

[37] S. Filin, "Calibration of spaceborne laser Altimeters-an algorithm and the site selection problem," *IEEE Trans. Geosci. Remote Sens.,* vol. 44, no. 6, pp. 1484-1492, 2006.

[38] S. Filin, *Calibration of Airborne and Spaceborne Laser Altimeters Using Natural Surfaces*. Ohio State University, 2001.

[39] S. Luthcke, D. D. Rowlands, J. J. McCarthy, D. E. Pavlis, and E. Stoneking, "Spaceborne laser-altimeter-pointing bias calibration from range residual analysis," *J. Spacecr. Rockets,* vol. 37, no. 3, pp. 374-384, 2000.

[40] H. Tang *et al.*, "Evaluating and mitigating the impact of systematic geolocation error on canopy height measurement performance of GEDI," *Remote Sens. Environ.,* vol. 291, p. 113571, 2023.

[41] Y. Xu, S. Ding, P. Chen, H. Tang, H. Ren, and H. Huang, "Horizontal Geolocation Error Evaluation and Correction on Full-Waveform LiDAR Footprints via Waveform Matching," *Remote Sensing,* vol. 15, no. 3, p. 776, 2023.

[42] S. Hancock, J. Armston, M. Hofton, X. Sun, and R. Dubayah, "The GEDI Simulator: A Large-Footprint Waveform Lidar Simulator for Calibration and Validation of Spaceborne Missions," *Earth Space Sci.,* vol. 6, no. 2, 2019.

[43] P. L. Patterson *et al.*, "Statistical properties of hybrid estimators proposed for GEDI—NASA's global ecosystem dynamics investigation," *Environ. Res. Lett.,* vol. 14, no. 6, p. 065007, 2019.

[44] I. Fayad, N. Baghdadi, and K. Lahssini, "An assessment of the GEDI lasers' capabilities in detecting canopy tops and their penetration in a densely vegetated, tropical area," *Remote Sensing,* vol. 14, no. 13, p. 2969, 2022.

[45] S. Hancock, C. McGrath, C. Lowe, I. Davenport, and I. Woodhouse, "Requirements for a global lidar system: spaceborne lidar with wall-to-wall coverage," *R. Soc. Open Sci.,* vol. 8, no. 12, p. 211166, 2021.

[46] B. Li *et al.*, "High-accuracy laser altimetry global elevation control point dataset for satellite topographic mapping," *IEEE Trans. Geosci. Remote Sens.,* vol. 60, pp. 1-16, 2022.

[47] C. S. Gardner, "Ranging performance of satellite laser altimeters," 1992.