GISCIENCE & REMOTE SENSING
2025, VOL. 62, NO. 1, 2507985
https://doi.org/10.1080/15481603.2025.2507985

Progressive noise photons removal from ICESAT-2 data based on the characteristics of different types of noise

Zhenyang Hui[a,b,c], Li Zhang[a,b,c], Shuanggen Jin[d,e], Wenbo Chen[a,b,c], Penggen Cheng[a,b,c] and Yao Yevenyo Ziggah[f]

[a]National Key Laboratory of Uranium Resources Exploration-Mining and Nuclear Remote Sensing, East China University of Technology, Nanchang, China; [b]School of Surveying and Geoinformation Engineering, East China University of Technology, Nanchang, China; [c]Jiangxi Key Laboratory of Watershed Ecological Process and Information, East China University of Technology, Nanchang, China; [d]School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo, China; [e]Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai, China; [f]Faculty of Mineral Resources Technology, University of Mines and Technology, Tarkwa, Ghana

ABSTRACT

Removing noise photons from Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) data is crucial for various applications of the photon-counting LiDAR system. Existing methods for noise photon removal often struggle with parameter tuning, lack robustness, and may compromise accuracy across different datasets. To address these issues, this study proposes an innovative progressive noise removal method. Unlike conventional approaches that treat all noise photons uniformly, our method first categorizes noise photons into isolated, low-density clustered, and outer clustered types based on their unique spatial distribution characteristics. Each type is then targeted with specific denoising techniques, resulting in higher denoising efficiency and better signal photon preservation. Specifically, isolated noise photons are automatically identified using a multi-thresholding strategy based on the maximum between-clustering variance algorithm without requiring parameter tuning. Low-density clustered noise photons are removed using the ellipse-based photon counting method, where the Douglas-Peucker algorithm is utilized to align the ellipse's major axis with the locally calculated terrain slope. Outer clustered noise photons are also automatically detected through a box plots analysis technique based on local elevation distributions. The efficacy of the proposed method was evaluated using diverse datasets containing strong and weak signals, as well as various land covers. Experimental results demonstrate that the proposed method outperformed five traditional denoising methods in terms of both denoising effectiveness and signal photon fidelity. Furthermore, testing on datasets with diverse land covers showcased the robustness of the proposed method.

ARTICLE HISTORY
Received 13 January 2025
Accepted 13 May 2025

KEYWORDS
Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2); noise photons removal; maximum between-clustering variance; Douglas-peucker; box plots analysis

1. Introduction

The Spaceborne LiDAR, as a key tool for innovative earth observation, offers unique benefits in gathering 3D surface information over large areas and inverting forest heights due to its high orbit and wide observational range (Li et al. 2021; Xu et al. 2023). On 15 September 2018, NASA launched the Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2), equipped with the Advanced Topographic Laser Altimeter System (ATLAS). ATLAS utilizes micro-pulse, multi-beam photon-counting LiDAR technology for the first time (Abdalati et al. 2010). The ATLAS system is known for its low energy consumption, high detection sensitivity, and high repetition frequency, which reduces the need for laser power and increases sampling frequency (Cao et al. 2020a; Neuenschwander and Pitts 2019). Consequently, ATLAS data is widely used in tasks such as elevation control points extraction (Lian et al. 2022), monitoring glacier elevation changes (He et al. 2024), inversing shallow water bathymetry (Ye et al. 2024), extracting forest canopy heights (Mansouri, Jafari, and Dehkordi 2024; Wang et al. 2024), and estimating biomass carbon storage (Neuenschwander et al. 2024; Varvia et al. 2024).

However, photon-counting LiDAR emits weak signals that are easily interfered with by atmospheric scattering, solar radiation, and instrument artifacts during target detection (Neuenschwander et al. 2020). This interference results in a significant amount of randomly distributed background noise in the recorded point cloud, which hampers the accurate extraction of signal photons (Jiao et al. 2021). Therefore, effective removal of noise photons has become a crucial step in processing photon-counting LiDAR data (Huang et al. 2019; Kui et al. 2023; Zhu et al. 2020).

Various methods for noise photons removal have been proposed recently. These methods can be categorized into three categories (Liu et al. 2023; Pan et al. 2024;

CONTACT Penggen Cheng pgcheng@ecut.edu.cn

Qin et al. 2024): 2D image processing-based methods, local statistical parameter-based methods, and density spatial clustering-based methods.

In methods based on 2D image processing, the profiled photons are converted into a 2D image and then processed using image processing techniques to eliminate noise photons (Chen and Pang 2015; Magruder and Kevin 2012; Pan et al. 2024). Magruder and Kevin (2012) organized the profile photons into 2D raster images and utilized image processing technologies, specifically the Canny operator detection boundary algorithm, to identify and remove noise photons effectively. Chen and Pang (2015) on the other hand, employed the classic active contours method and applied the Chan-Vese segmentation model to detect potential signal photons. However, there is a need to enhance the stability and broad applicability of this approach across extensive areas and diverse ground surfaces. While this kind of method can reduce noise photons to some extent, the rasterization process results in the loss of geometric information from the photons, decreasing algorithm accuracy (Jiao et al. 2021; Xia et al. 2014; Zhu et al. 2020).

The second category relies on local statistical parameters, which are widely utilized. This involves calculating various parameters such as distance, elevation, point density, eigenvectors, etc., for each photon locally. Subsequently, by utilizing distribution characteristics like histograms to establish global thresholds based on these parameters, noise and signal classification becomes feasible (R. Liu et al. 2024; Nie et al. 2018; Xia et al. 2014; F. Xie et al. 2017; Zhu et al. 2018; Cao et al. 2020a). Xia et al. (2014) introduced a denoising algorithm based on local distance statistics and applied least squares fitting to determine local curve parameters, achieving satisfied overall accuracy. Xie et al. (2017) tackled terrain influences by defining adjustable filtering kernels that effectively eliminated near-ground noise photons in steep terrain areas. Zhu et al. (2018) devised an enhanced noise photons filtering algorithm based on local statistics with adaptive threshold determination capabilities. Nie et al. (2018) used an automatic approach for noise photon removal, successfully reducing edge effects and addressing inconsistent noise photon density challenges; however, algorithm accuracy may decrease in areas with very dense or sparse vegetation cover. Cao et al. (2020b) implemented and evaluated the Differential Regression and Gaussian Adaptive Nearest Neighbor (DRAGANN) algorithm in ATL08, achieving denoising accuracy above 88%, though slightly lower accuracy was observed in vegetation scenes due to additional empirical parameters required.

In the third category of density spatial clustering-based methods, denoising is achieved by characterizing the dense spatial distribution of signal photons and the relatively sparse distribution of noise photons (Zhang and Kerekes 2015; He et al. 2023; Wang, Pan, and Glennie 2016, Chen et al. 2019; Zhang et al. 2024; Zhu et al. 2021). In consideration of the horizontal aggregation of signal photons, Zhang et al. (2014) enhanced the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm by transitioning from a circular search shape to an elliptical one. Through experimentation, B. W. Chen, Pang, Li, Lu, et al. (2019) and B. W. Chen, Pang, Li, North, et al. (2019) demonstrated that horizontal elliptical search shapes outperform circular and vertical ones. X. Wang, Pan, and Glennie (2016) introduced a denoising algorithm based on Bayesian decision theory, assuming a uniform photon distribution in space. Ma et al. (2019) proposed a novel adaptive signal photon detection method that matches geographic coordinates with a land cover database to obtain surface type information. The method employs an improved DBSCAN algorithm and JONSWAP wave spectrum algorithm, combined with adaptive thresholds, to detect signal photons. This approach enables adaptive adjustment based on different land cover types, enhancing the accuracy and efficiency of signal photon detection. Z. Zhang et al. (2020) introduced a novel land/snow classification technique utilizing the background noise from photon-counting LiDAR. This method involves analyzing the noise photon distribution patterns of various land cover types by developing a noise model that considers factors such as solar incidence angle, terrain slope, and surface reflectance properties. Experimental results from the Karakoram Plateau demonstrate that this approach can achieve a classification accuracy exceeding 93% without the need for optical images. Zhang et al. (2021) established for the first time a correlation model of signal-to-noise ratio (SNR) between strong and weak beams of ICESat-2 in mountainous areas. By deriving the geometric relationship of photon propagation paths, this paper revealed the auxiliary mechanism of strong beam data for weak beam signal extraction.

To address parameter sensitivity, Zhu et al. (2021) revised the search area shape in the Ordering Points to Identify the Clustering Structure (OPTICS) algorithm to an ellipse for photon denoising. They also implemented adaptive detection of signal and noise photons in photon data using a distance threshold determined by the Otsu method, effectively extracting signal photons in complex terrain while reducing input parameter sensitivity. However, this algorithm is computationally intensive and time-consuming. Furthermore, G. H. He et al. (2023) adapted a density clustering algorithm with an adaptive mountain slope, achieving better adaptation in forested areas with complex topography. Zhang et al. (2021) proposed a parameter-free noise removal algorithm based on quadtree isolation for photon-counting LiDAR data, achieving adaptive signal-noise separation through spatial partitioning and tree depth analysis. The algorithm demonstrates improved efficiency and accuracy compared to traditional methods such as DBSCAN, especially in complex terrains and low signal-to-noise scenarios. Building on the success of the quadtree isolation method, Zhang, Xing, Xu, Li, et al. (2023) and Zhang, Xing, Xu, Zhang, et al. (2023) further refined the approach into the Pre-Pruning Quadtree Isolation (PQI) method for extracting bathymetric photons. For instance, Zhang, Xing, Xu, Li, et al. (2023) introduced a PQI method with dynamic threshold adjustment for bathymetric photon extraction from ICESat-2 data. By integrating spatial pruning and depth-adaptive thresholds derived from photon elevation histograms, PQI effectively suppresses noise and enhances photon extraction accuracy, surpassing traditional methods with an F1-score of 92.71%. In a similar vein, Zhang, Xing, Xu, Zhang, et al. (2023) applied the PQI method to automatically extract nearshore bathymetric photons from ICESat-2 data. Through a pruning step to prevent over-segmentation of noise photons and leveraging Otsu's method for dynamic threshold adjustment, PQI achieves high extraction accuracy (F1-score: 93.96%) and versatility across different underwater terrains and data acquisition times.

Tian and Shan (2023) presented an innovative gravity-based photon density model and directional region growing algorithm for detecting signal photons from ICESat-2 data. Their approach enhances accuracy and robustness in challenging terrains when compared to the ATL03 and ATL08 algorithms. Liu et al. (2024) introduced a denoising technique that relies on adaptive parameter density clustering. This method utilized numerical simulations to adjust the key parameters (such as neighborhood radius Eps and minimum number of points MinPts) to address the issue of denoising spaceborne photon-counting laser altimeter point clouds with varying noise densities.

Moreover, the growing utilization of machine learning and deep learning has introduced new approaches to photons processing (Agca et al. 2024; Chen et al. 2020; Kong and Pang 2024; Lu et al. 2021). Chen et al. (2020) proposed a machine learning-based method for detecting potential signal photons from photon-counting LiDAR data. Lu et al. (2021) developed a denoising and classification algorithm based on convolutional neural networks. In terms of deep learning methods, Lin and Knudby (2023) introduced the PointNet++ deep neural network model into ICESat-2 laser altimetry data for bathymetric photon extraction for the first time. Compared with supervised classification methods based on Random Forest or SVM, PointNet++ demonstrated an improvement in F1-score on the test set, with the false detection rate reduced. Z. Leng et al. (2023) implemented the Long Short-Term Memory (LSTM) deep learning model for the reconstruction of ICESat-2 bathymetric signals. By integrating active-passive data fusion, this approach enhances the accuracy and reliability of signal reconstruction. Qin et al. (2024) integrated the GoogLeNet model with the Convolutional Block Attention Module (CBAM) for signal photons extraction. The GoogLeNet model expands the network width through Inception modules, enhancing its ability to capture features. CBAM strengthens the network's ability to learn key features by utilizing both channel and spatial attention mechanisms, thereby improving the accuracy and robustness of photon extraction. Liu et al. (2024) proposed an end-to-end deep neural network based on VOJA-Net, which utilizes semantic information from a multilevel decoder to learn multilevel feature representations to improve denoising performance. Testing on the ICESat-PC dataset demonstrated that VOJA-Net outperformed traditional methods such as DBSCAN and advanced deep learning models like PointNet++ and Point Transformer, achieving remarkable F1 scores of 93.34% and a mean Intersection-over-Union (mIoU) of 73.40%.

Nevertheless, these algorithms heavily rely on training samples, and the variability in training samples due to diverse terrains and surface coverages can negatively impact denoising performance.

While numerous photon point denoising methods have been proposed, two main challenges remain unresolved. Firstly, most denoising methods require complex parameters tuning, with these parameters significantly influencing the denoising results. For example, the classical DBSCAN-based denoising method requires defining the neighboring distance and the minimum number of points, both of which have a substantial impact on the denoising outcomes (Leng et al. 2022; Liu et al. 2024). Similarly, the Local Outlier Factor (LOF) denoising method also necessitates parameter adjustment, with the k nearest neighbors directly affecting its denoising performance (Chen et al. 2019). Secondly, existing denoising methods typically apply a uniform strategy to all noisy points without considering their different characteristics. In this study, noisy points are categorized into isolated, low-density clustered, and outer clustered noise photons based on their distinct attributes.

Isolated noise photons are typically identified by larger spacing distances between photons. While low-density clustered noise photons may have spacing distances similar to signal points, their neighboring photons usually have fewer neighbors compared to signal photons. The outer clustered noise photons have a similar number of neighbors as signal photons, but they show sudden increases in elevation compared to their neighboring signal photons. Managing these various types of noise photons presents a significant challenge. It is evident from these descriptions that each type of noise photon has distinct characteristics. Employing a single denoising approach with uniform thresholds for all types of noisy points may lead to suboptimal denoising results.

To tackle these challenges, this study introduces a progressive method for removing noise photons in ICESAT-2 data based on different noise characteristics. The study categorizes the noise photons into three groups and gradually removes them based on their unique characteristics, primarily by introducing an automatic multi-thresholding technique, an ellipse-based photon counting method, and a box plots analysis approach. The effectiveness of the proposed method was assessed using diverse datasets with diverse features.

2. Methodology

In this paper, the noise photons are classified into three categories: isolated noise photons, low-density clustered noise photons, and outer clustered noise photons. Each type possesses distinct characteristics. To effectively filter out the noise photons, this study detects and eliminates them gradually based on their unique attributes. The flowchart of the proposed method is illustrated in Figure 1. Initially, isolated noise photons are identified by comparing their point spacing distance with that of signal photons. Thresholds for detecting isolated noise photons are determined using an automatic multi-thresholding strategy. Subsequently, the Douglas-Peucker algorithm is utilized to merge data segments with similar terrain slopes to adaptively calculate terrain slopes in local regions, followed by the utilization of an ellipse-based photon counting method along the primary terrain slope direction to remove low-density clustered noise photons. While the terrain slope-based ellipse clustering method effectively removes a portion of noise photons, some clustered noise photons still remain. These noise photons have a higher point density but are spatially distant from the signal photons, termed as outer clustered noise photons. To address this issue, these outer clustered noise photons are eliminated based on their local elevation distributions using a proposed box plots analysis technique. In summary, this paper consists of three main steps: I. Isolated noise photons removal based on a multi-thresholding strategy, II. Adaptive calculation of terrain slopes and removal of low-density clustered noise photons, and III. Outer clustered noise photons removal based on the box plots analysis.

2.1. Removal of isolated noise photons based on a multi-thresholding strategy

In Figure 2a, the raw photons are segmented into multiple sections using filtering windows with a window size set to 50 m (Lian et al. 2022). From Figure 2a, it is evident that the point spacing distances (the distance between a point and the nearest neighboring point) of signal photons are considerably smaller than those of noise photons. Within each window section, the average neighboring distance for each photon can be computed, and its distribution is depicted in Figure 2b. Note that the average neighboring distance is defined as the mean distance between a photon and its N closest neighboring points. In this study, we have set N as a constant value of 55. The subsequent Discussion section will further explore the impact of this parameter setting.

The distribution of average neighboring distances appears as a combination of two Gaussian distributions. The left Gaussian distribution corresponds to signal photons, while the right one represents noise photons. This distinction arises from the fact that the point spacing distances of signal photons are significantly smaller than those of noise photons. To eliminate these isolated noise photons, a threshold for each window section must be determined as shown in Figure 2b.

To establish the threshold within each filtering window, the maximum between-clustering variance algorithm is employed, which is a commonly used adaptive threshold segmentation technique in computer vision and image processing fields (Li et al. 2014; Xie et al. 2023). This method aims to convert a grayscale image into a binary image by computing an optimal threshold that maximizes the variance between two classes of pixels – foreground pixels and background pixels. In this study, the two classes that require segmentation are signal and noise photons. Thus, the between-class variance g for this case is defined as Eq (1).

$$g = w_0(u_0 - u)^2 + w_1(u_1 - u)^2 \tag{1}$$

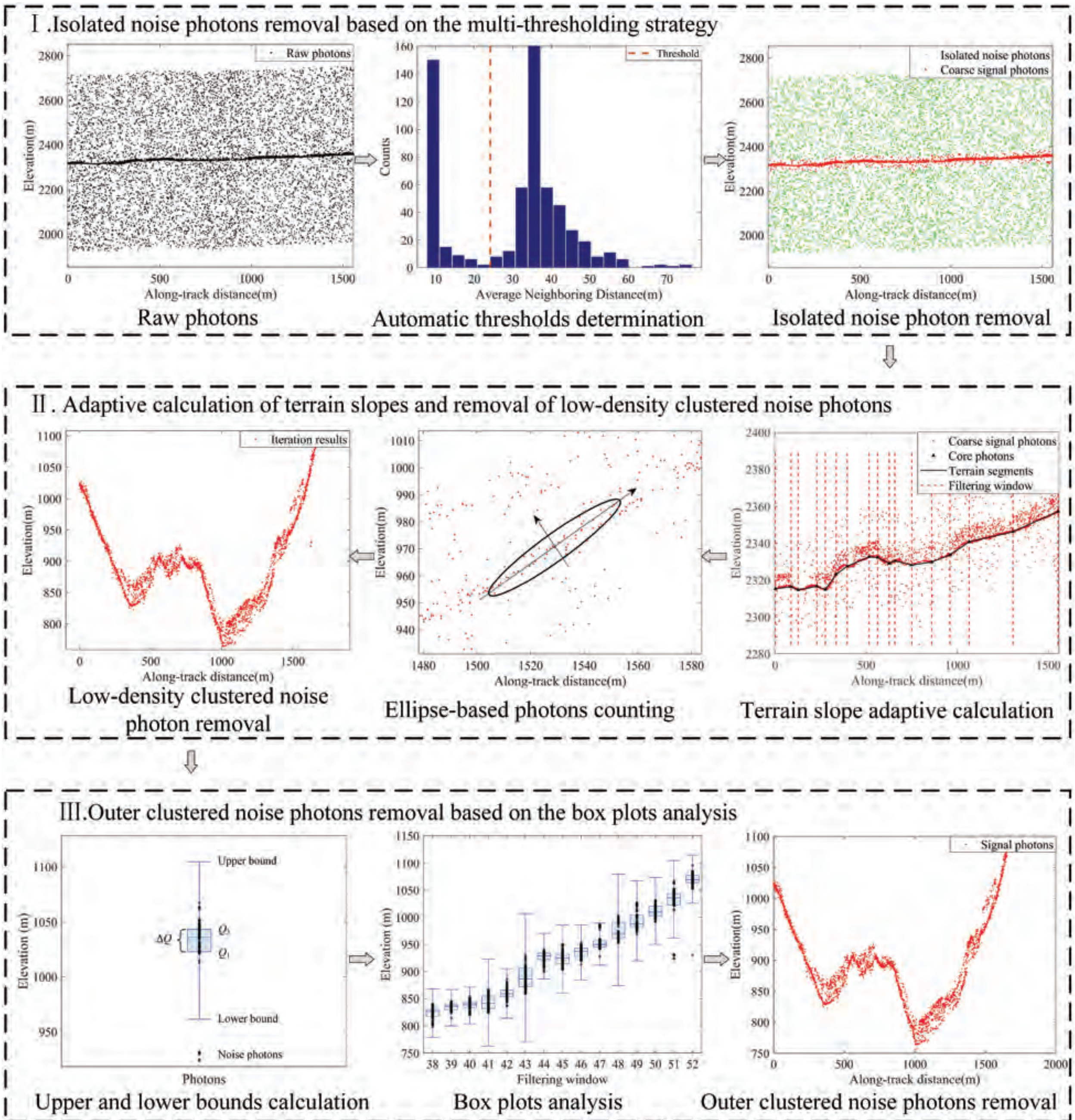


Figure 1. The flowchart of the proposed method.

where w_0 and w_1 the proportions of signal and noise photons, respectively, and u_0 and u_1 are the mean average neighboring distance values of signal and noise photons. u is the total mean average neighboring distance value of all the raw photons within the filtering window. Through the application of the maximum between-clustering variance method, the threshold for distinguishing signal and isolated photons can be determined, and those photons with an average neighboring distance exceeding the calculated threshold are identified as isolated noise photons and subsequently removed.

2.2. Adaptive calculation of terrain slopes and removal of low-density clustered noise photons

After determining the threshold within each filtering window, most of the isolated noise photons can be successfully removed, as demonstrated in Figure 3a. Nonetheless, a few noise photons remain mixed with the signal photons because their average point spacing distances are smaller than the self-adaptive calculated threshold. Consequently, although these noise photons are also clustered, their neighboring points generally less than

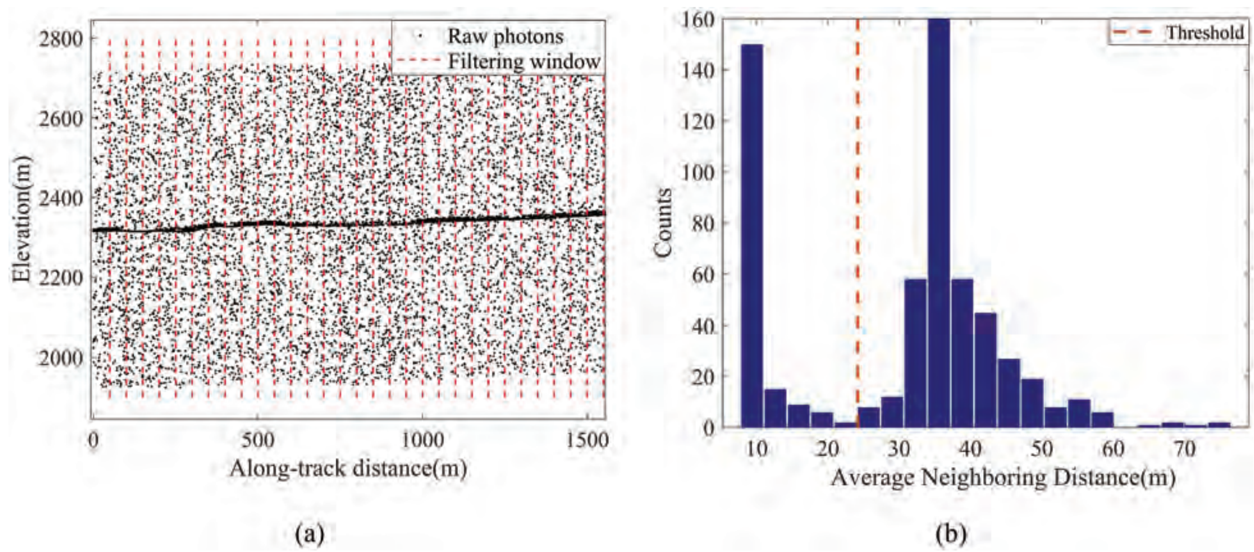


Figure 2. Isolated noise photons removal. (a) Row photons with filtering windows. (b) The average neighboring distance distribution within each filtering window.

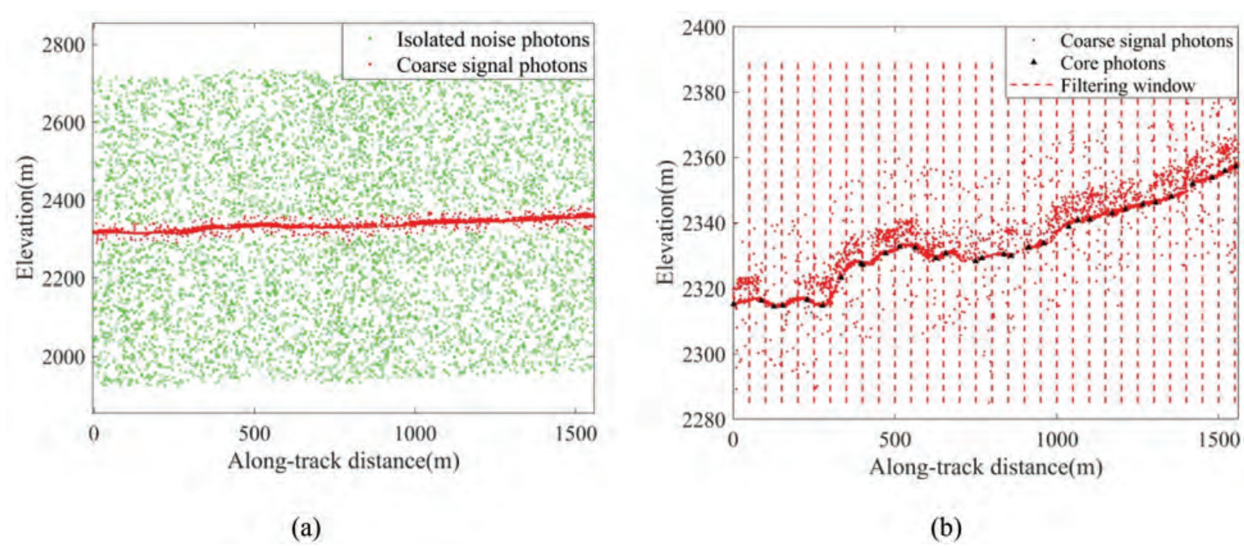


Figure 3. Coarse signal photons and core photons. (a) Coarse signal photons after isolated noise photons removal. (b) Core photons identification based on elevation frequency histogram.

those of the signal photons, indicating that they are low-density clustered. This distinction is evident from Figure 3b. To address these low-density clustered noise photons, this study proposes an ellipse-based photon counting method. The rationale behind this method is that the neighboring photons of signal photons typically outnumber those of noise photons.

To implement the ellipse-based photon counting method, terrain slope calculation is essential. This calculation allows for the automatic determination of the major axis direction of the ellipse, facilitating the preservation of signal photons across varying terrain slopes. Initially, the elevation frequency histogram for photons within each filtering window is constructed to select the core photon with the highest elevation frequency, depicted as a black triangle in Figure 3b. Note that the filtering window mentioned here is the same with the filtering window shown in Figure 2a. Within each filtering window, the elevations of photons are counted, and the point with the highest elevation frequency is selected as the core point.

Given that core photons with highest elevation frequencies are typically situated on terrain features, they effectively reflect terrain fluctuations. Therefore, the core photons within each two adjacent filtering windows can be used to calculate local terrain slopes. To accurately identify the core photons, specifying the resolution of the frequency histogram is crucial. Generally, a smaller resolution is recommended for gentle terrain, whereas a larger resolution is more suitable for steep terrain. In this study, we have set the resolution to 1 m for gentle terrain and 15 m for steep terrain.

The proposed ellipse-based photon counting method aligns the direction of the ellipse's major axis with the local terrain slope; hence, merging adjacent terrain segments with similar slopes is prioritized.

This study employs the Douglas-Peucker algorithm to merge segments with similar terrain slopes into distinct sections (Vučetić, Petrović, and Strunje 2007). The Douglas-Peucker algorithm simplifies a curve represented by a series of points by recursively dividing a line segment defined by the first and last points and identifying the farthest point from this line. If this distance exceeds a specified tolerance, it is added to the simplified curve, and this process is iterated on the resulting line segments, as depicted in Figure 4a.

As this study aims to combine segments with similar terrain slopes, it requires that the elevation change between two consecutive core photons within neighboring terrain segments be minimal. This requirement is akin to the Douglas-Peucker algorithm, which retains feature points with larger distance residuals while eliminating those with smaller residuals. Here, a tolerance of 1.5 m is set for the Douglas-Peucker algorithm to eliminate the close point, as it is employed in merging terrain segments with comparable slope characteristics. If the elevation difference between two successive core photons surpasses 1.5 m, it signifies differences in slope attributes between the corresponding terrain segments.

After implementing the Douglas-Peucker algorithm, it is observed that the terrains are segmented into distinct sections, as evidenced by the altered filtering window sizes shown in Figure 4b. Within each section, the terrain slope exhibits similarity, implying that the major axis direction of the ellipse for all photons within the section can be set as a constant value, equivalent to the terrain slope of that specific terrain section.

In this study, the ellipse is defined as $E(a, b, \varphi)$. a and b denote the major and minor axes of an ellipse, respectively, and φ represents the major axis direction of the ellipse, corresponding to the calculated terrain slope. In this paper, an empirical ratio of a to b is established at 6:1 according to the recommendations of Chen, Pang, Li, Lu, et al. (2019) and Zhu et al. (2021). The minor axis of the ellipse (b) can range from 2 m to 6 m. Generally, when the terrain slope is gentle, a smaller value for b can be used, whereas for steeper terrain slopes, a larger value for b is more appropriate. This parameterization does not significantly impact the results, as will be further discussed in the Discussion section.

Typically, signal photons exhibit proximity to one another compared to noise photons. Additionally, the spatial density distribution of signal photons in horizontal and vertical directions is uneven. The abundance of signal photons in the horizontal direction along the terrain exceeds that in the vertical direction. By aligning an ellipse with its major axis direction matching the terrain slope, a significant proportion of signal photons can be encompassed within the ellipse.

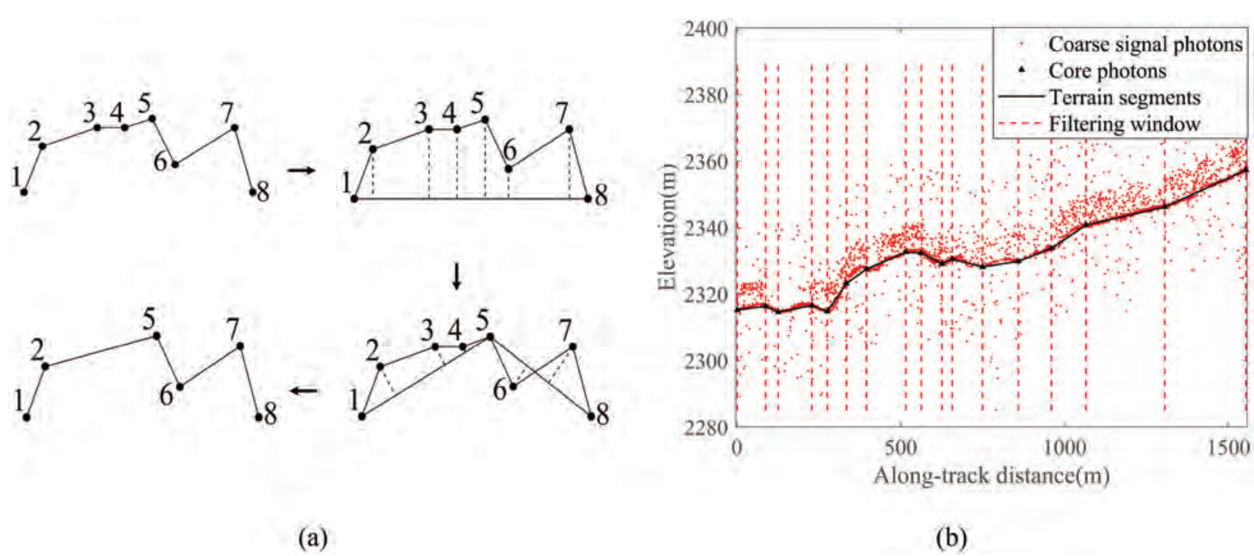


Figure 4. Terrain slope adaptive calculation by applying the Douglas-Peucker algorithm. (a) The principle of the Douglas-Peucker algorithm. (b) Terrain slope calculation using the Douglas-Peucker algorithm.

By establishing a minimum number of photons within the ellipse, it becomes possible to identify low-density clustered noise photons. The minimum number of photons within an ellipse is denoted as *minpts*, which can be automatically calculated using a method proposed by (Liu et al. 2024). Specifically, if a photon has fewer neighboring points than *minpts*, it is classified as a noisy photon; otherwise, it is considered a signal photon and its neighbors are added to a point set. Subsequently, each photon in the set is examined; if a photon has more neighbors than *minpts*, its neighboring photons are also incorporated into the set. This iterative process continues until all photons have been visited. The sequential steps of this iterative ellipse-based connected growth method are outlined below:

Step 1: Randomly select an unvisited point p_k from point set $\{p_i\}$ and add a corresponding visited flag, that is $VisitedFlag(p_k) = 1$.

Step 2: Calculate the number of points (N_{p_k}) within the ellipse of p_k, of which direction of the major axis is equal to the terrain slope.

Step 3: If $N_{p_k} > minpts$, p_k is labeled as a signal point and the neighboring points of p_k are added into a point set $\{p_{neighbors}\}$.

Step 4: Traverse the points in $\{p_{neighbors}\}$, and repeat Step 2 and Step 3 until all the points in $\{p_{neighbors}\}$ have been visited.

Step 5: If all the points in $\{p_i\}$ are visited, stop; otherwise go to Step 1.

2.3. Removal of outer clustered noise photons using the box plots analysis technique

After using the ellipse-based method for counting photons, most clustered noise photons can be effectively eliminated. However, some clustered noise photons with higher point densities remain undetected, as seen in Figure 5a. These unreleased clustered noise photons are typically located far from the signal photons. Therefore, this study categorizes these noise photons as outer clustered noise photons and proceeds to remove them based on their elevation distributions. That is the results of removing low-density clustered noise photons are divided into sections using a filtering window. The elevation distributions of photons within each filtering window are depicted using box plots in Figure 5b. Within each filtering window, the first quartile (Q_1), third quartile (Q_3) and interquartile range ($\Delta Q = Q_3 - Q_1$) of photon elevations can be calculated. Since outer clustered noise photons are usually distant from signal photons, upper and lower bounds for identifying this type of error are defined as $(Q_3 + 3\Delta Q)$ and $(Q_1 - 3\Delta Q)$, respectively. Photons falling outside these bounds are classified as outer clustered noise photons and removed. By analyzing box plots of photons within each filtering window, as shown in Figure 5c, outer clustered noise photons are successfully eliminated, resulting in the final removal of noise photons shown in Figure 5d.

3. Experimental results and analysis

3.1. Study sites

This study utilizes experimental data from eight groups located at two distinct geographical sites in the United States for testing, as illustrated in Figure 6a. Study Site 1 (depicted in Figure 6b) is situated in Yellowstone National Park (YELL) in Wyoming (44°50′N ~ 45°0′N, 110°20′W ~ 111°43′W), USA. It experiences an annual average temperature of 3.4°C, an annual average precipitation of approximately 493 mm, and an elevation range of 1847–2244 m. The primary vegetation consists of grassland, shrubland, and evergreen forest, with an average canopy height of 14 m.

Study Site 2 (Figure 6c) is positioned in Great Smoky Mountains National Park (GRSM) in southeastern Tennessee(35°30′N ~ 35°50′N, 83°10′W ~ 83°45′W), USA. It has a subtropical humid climate, with an annual average precipitation of 1375 mm, an annual average temperature of 13.1°C, and an elevation range of 426–1978 m. The predominant vegetation types include deciduous forest and evergreen forest, with an average canopy height of 30 m.

The transit area chosen within Study Site 1 is primarily characterized by evergreen forest dominated by pine trees, featuring gentle terrain and minimal topographic relief. Conversely, the transit area within Study Site 2 comprises mainly coniferous forest dominated by fir and hemlock, showcasing rugged terrain and significant topographic variability.

Both study sites have collected data during both daytime and nighttime periods, with specifics outlined in Table 1. The contrasting terrains and surface cover types between the two sites provide valuable insights into the denoising impact on ICESat-2 data across various temporal, geographical, and topographical settings.

3.2. Data sources

3.2.1. Icesat-2 data

ICESat-2 employs the Advanced Topographic Laser Altimeter System (ATLAS), which integrates micro-pulse multi-beam photon-counting LiDAR technology. It releases six beams of laser pulses with repetition frequency at a rate of 10 kHz, organized into three pairs along its orbital trajectory. Each pair consists of a strong

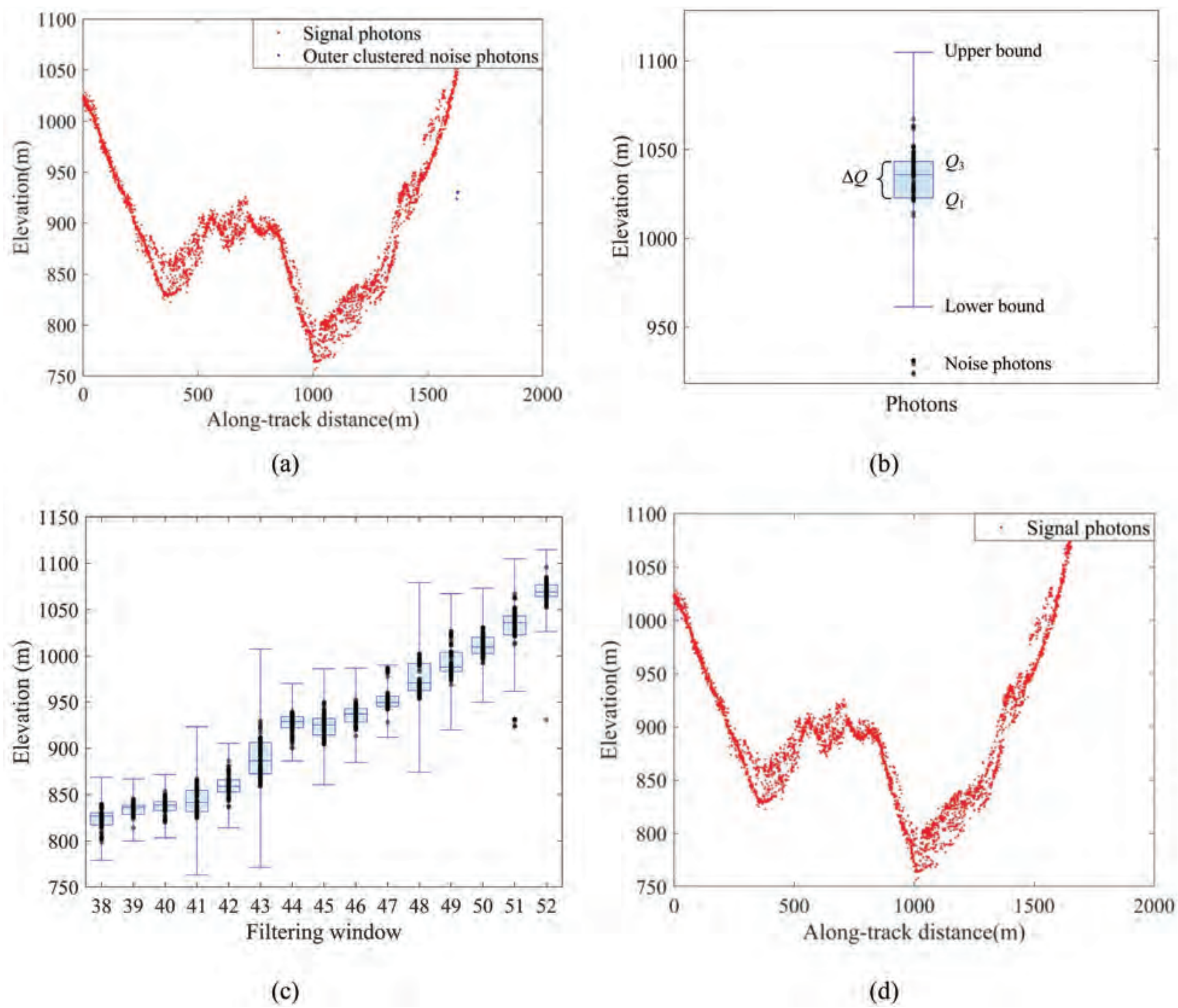


Figure 5. Outer clustered noise photons removal. (a) The denoising result after low-density clustered noise photons removal. (b) The upper and lower bounds for outer clustered noise photons removal. (c) Box plots analysis for the photons within each filtering window. (d) The final denoising result after outer clustered noise photons removal.

signal (160μJ) and a weak signal (40μJ), maintaining an energy ratio of 4:1 between them. The cross-track distance between pairs is approximately 3.3 km, while the spacing within each pair is around 90 m. This configuration results in overlapping footprints on the Earth's surface with a diameter of about 17 m and an along-track spacing of roughly 0.7 m (Zhu et al. 2021). ICESat-2 provides 22 standard data products designated as ATL00-ATL21, categorized into Level 0, Level 1, Level 2, and Level 3 groups. This research primarily utilizes the Level 2 product ATL03 and the Level 3 product ATL08. ATL03 comprises global geolocated photon data containing positional information for each photon event such as time, latitude, longitude, and elevation. ATL08 offers terrestrial vegetation height data by classifying the geolocated photons from ATL03 into noise, ground, canopy, and top-of-canopy categories, particularly in vegetated areas (Neuenschwander et al. 2024). The National Snow & Ice Data Center (NSIDC) provides various ICESat-2/ATLAS data products free of charge to the global community. This study leverages the latest release, Version 6, of ICESat-2 data stored in HDF5 file format. The relevant data for the study sites is accessible for download at no cost from the following link: https://nsidc.org/data/atl03.

3.2.2. *Airborne LiDAR data*

The National Ecological Observatory Network (NEON), established by the National Science Foundation in the United States, aims to collect high-quality, standardized data on climate change and land-use change from 81 locations nationwide (47 terrestrial and 34 aquatic). This data serves to investigate critical ecological and environmental matters, predict environmental change

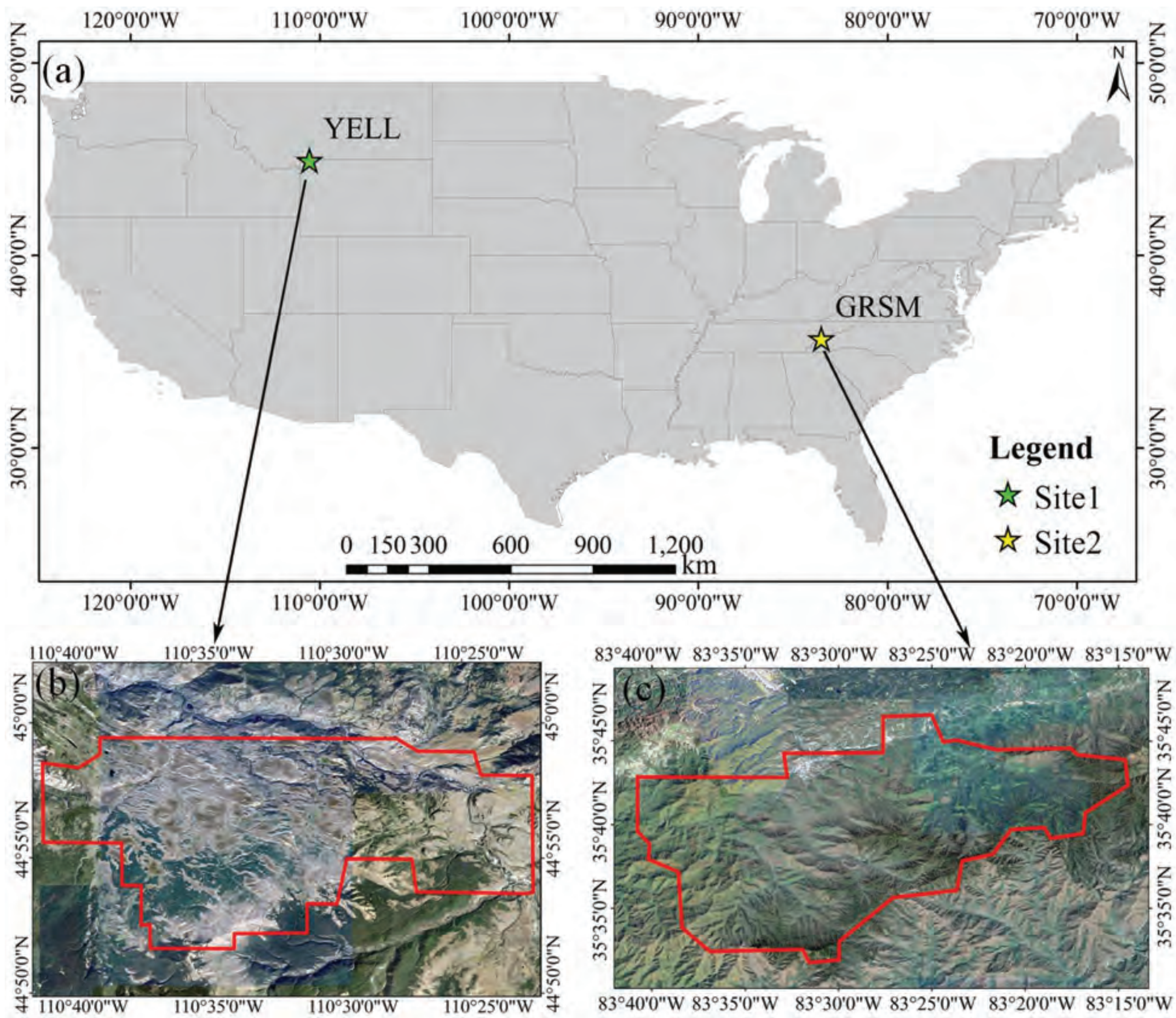


Figure 6. Geographical locations and terrain features of the study sites. (a) Geographical locations of the study sites. (b) Terrain feature for study site 1 (YELL). (c) Terrain feature for study site 2 (GRSM).

Table 1. Detailed information of ICESat-2 strong beam data for the study sites.

Study sites	Datasets	Time	Data type	Terrain slope
YELL	data1	Daytime	gt3l/strong	gentle
	data2	Daytime	gt1r/strong	gentle
	data3	Nighttime	gt1l/strong	gentle
	data4	Nighttime	gt1l/strong	gentle
GRSM	data5	Daytime	gt1r/strong	steep
	data6	Daytime	gt3r/strong	steep
	data7	Nighttime	gt1r/strong	steep
	data8	Nighttime	gt1l/strong	steep

trajectories, and propose pertinent strategies. In this investigation, NEON's airborne LiDAR data products – encompassing the Digital Terrain Model (DTM) and Digital Surface Model (DSM) – are selected as benchmark values to validate the accuracy of the proposed algorithm. These datasets have a resolution of 1 m and are provided in TIFF format. The airborne data for Study Site 1 and Study Site 2 were acquired in June 2022 and September 2022 respectively, closely aligning with the ATLAS data collection dates to minimize discrepancies resulting from inconsistent data acquisition times. To obtain the reference signal photons, this paper utilizes multi-value extraction on DTM and DSM points in ArcGIS based on the latitude and longitude coordinates of ICESat-2 photons, the corresponding DTM and DSM values for each photon are determined. This is followed by contour line generation for DTM and DSM through interpolation. Finally, based on the spatial distribution characteristics of the photon point clouds, points falling between the DSM and DTM contour lines in the airborne data are classified as signal photons, while those outside this range are categorized as noise photons.

The relevant data can be accessed from the following site: https://www.neonscience.org/data-collection/LiDAR. Figure 7 demonstrates the alignment of ICESat-2 photons with DTMs and DSMs. In the figure, the green lines depict the ICESat-2 beam data as depicted in Figure 7a,b. As previously stated, the DTMs and DSMs were created using NEON's airborne LiDAR. The alignment of ICESat-2 photons with DTMs and DSMs is determined by their specific latitude and longitude coordinates. Figure 7c–f displays the overlapping results of ICESat-2 photons with DTMs and DSMs.

3.3. Accuracy metrics

This paper employs four accuracy metrics, namely Recall (R), Precision (P), F1 Score (F_1) and Accuracy (Acc) to quantitatively assess the denoising performance of the proposed method. These metrics are defined by Eqs (2)-(5).

$$R = \frac{TP}{TP + FN} \tag{2}$$

$$P = \frac{TP}{TP + FP} \tag{3}$$

$$F_1 = \frac{2P * R}{P + R} \tag{4}$$

$$Acc = \frac{TP + TN}{TP + TN + FN + FP} \tag{5}$$

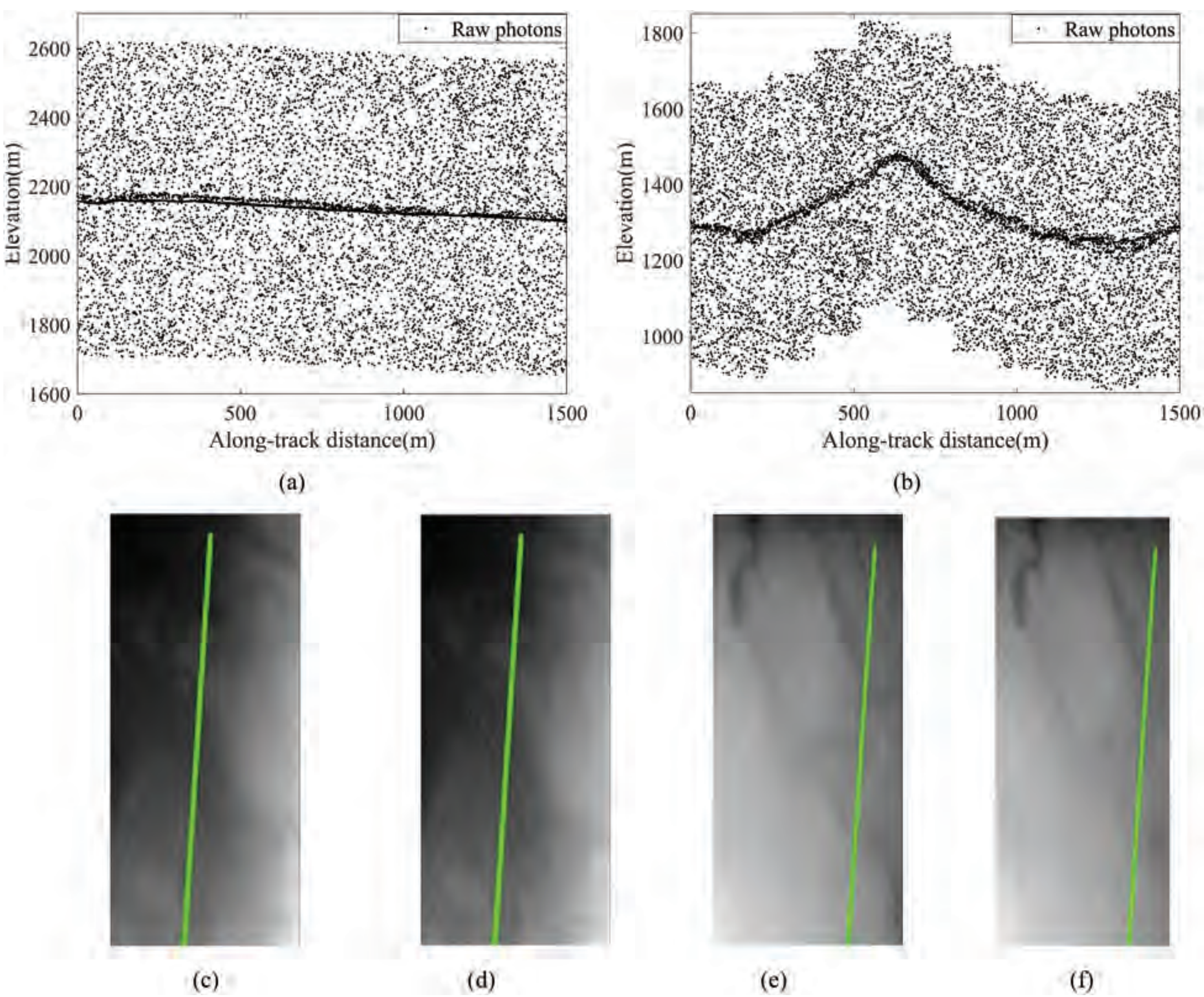


Figure 7. Photons aligned with DTMs and DSMs. (a) Raw photons of data 1. (b) Raw photons of data 8. (c) Overlapped photons with DTM provided by NEON for data 1. (d) Overlapped photons with DSM provided by NEON for data 1. (e) Overlapped photons with DTM provided by NEON for data 8. (f) Overlapped photons provided by NEON with DSM for data 8.

where *TP* represents the number of correctly detected signal photons, *TN* denotes the correct identification of noise photons, *FP* indicates the misclassification of noise photons as signal photons, and *FN* refers to the misclassification of signal photons as noise photons. The research acquires reference outcomes and performs accuracy assessments using the DTM and DSM generated from NEON airborne LiDAR data. Specifically, the DTM and DSM corresponding to the test region are extracted based on the latitude and longitude of the signal photons. Contour lines of the DTM and DSM are then created through interpolation and extrapolation. Photons falling within these boundaries (with the DTM representing the ground boundary and the DSM representing the canopy top boundary) are designated as reference signal photons, while those outside these boundaries are categorized as noise photons (Huang et al. 2022). Figure 8 illustrates the alignment between the DTM and DSM boundaries with the photons. By comparing the signal and noise photon reference outcomes derived from the DTM and DSM with those obtained through the proposed method, a confusion matrix is established to compute *TP*, *TN*, *FP*, and *FN*.

3.4. Experimental comparison and analysis

Five traditional denoising techniques have been chosen for comparison: the local distance statistics-based approach (LDS), the classical DBSCAN-based denoising method (DBSCAN), the Differential, Regressive, and Gaussian Adaptive Nearest Neighbor filtering method (DRAGANN), the Modified DBSCAN method and the Modified Ordering Points to Identify the Clustering Structure (OPTICS) method.

The LDS method identifies noise photons based on local distance statistical histograms. Initially, this method computes distances between each pair of photons. Subsequently, it calculates the sum of neighboring distances for *n* nearest points for each photon to form a frequency histogram. Typically, signal photons exhibit lower sum values while noisy ones display larger values due to signal photons being closely clustered compared to noisy ones. By setting a threshold at the mean value plus *t* times the standard deviation based on this frequency histogram, noisy photon points can be distinguished.

The DBSCAN-based method identifies outliers by primarily utilizing density-reachability concepts to group closely positioned points and detects outliers as points situated in low-density regions. The fundamental concept behind DBSCAN revolves around two crucial parameters: *eps* (epsilon) and *minpts* (minimum points). *eps* determines the radius of the neighborhood surrounding a point, while *minpts* specifies the minimum number of points necessary within this neighborhood to classify an area as dense. Points with a limited number of neighboring points below the *minpts* threshold are marked as outliers.

DRAGANN is a classical denoising method employed in generating ATL08 products and serves as an official denoising approach for processing ICESat-2 photon data. The core principle behind DRAGANN lies in disparities in local point density between noisy and signal photons. A bimodal distribution is observed in local density distribution histograms due to noise photons being sparsely distributed and signal photons densely distributed, with noise on one end and signal on another. Gaussian curves are utilized to fit histograms associated with noise and signal, respectively. The density at the intersection point of these curves serves as a threshold; classifying photons with local densities below this threshold as noise photons which are subsequently removed.

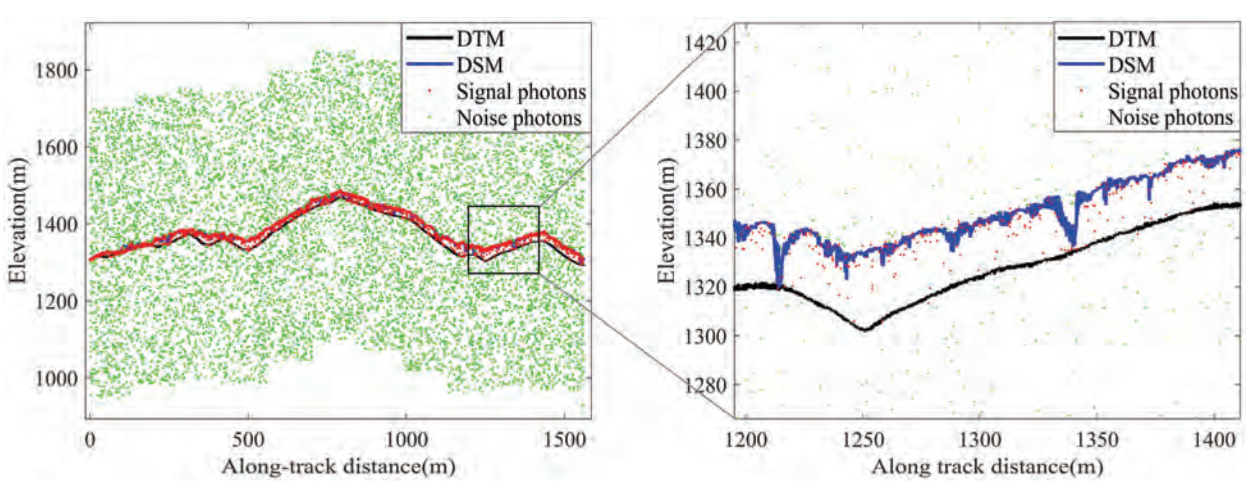


Figure 8. Alignment between the DTM and DSM boundaries and the photons.

As mentioned in our previous review work, the elliptic neighborhood may be more suitable for denoising the ICESat-2 data. Thus, several researchers have tried to modify the traditional DBSCAN method to make it more suitable for processing photon data. Zhang and Kerekes (2015) and Zhang et al. (2025) have modified the traditional circular search area to an elliptical shape, catering to the characteristic of higher point density in the horizontal direction of photon-counting laser altimeter point clouds. As a result, this study conducted additional testing on the modified DBSCAN approach.

Zhu et al. (2021) modified the circular search area in the OPTICS algorithm to an elliptical shape to accommodate the distribution characteristics of photon data, which features a higher photon density in the horizontal direction than in the vertical. By employing the improved OPTICS algorithm, a distance ordering of all photons was generated, and distance thresholds were automatically set using the Otsu method, effectively distinguishing signal photons from noise photons.

To quantitatively evaluate the performance of the proposed method, this study further calculated accuracy indicators based on Eqs (2)-(5) for eight selected datasets and compared the four accuracy indicators (*P*, *R*, F_1, and *Acc*) with those of five other methods. The comparison results of accuracy indicators are presented in Table 2.

Regarding precision (*P*), the proposed method achieved the highest mean precision value of 0.978, outperforming the other three methods. Notably, the DRAGANN method had the lowest mean precision value of 0.775, significantly lower than that of the proposed method as shown in Figure 9. Precision measures the proportion of true positive predictions among all positive predictions, indicating that the proposed method can more accurately identify signal photons.

In terms of recall (*R*), DBSCAN, DRAGANN, and the proposed method achieved similar mean recall values ranging from 0.925 to 0.930. On the other hand, Modified DBSCAN performed the worst with a mean recall value of 0.880. This highlights the strong ability of the proposed method in detecting more signal photons.

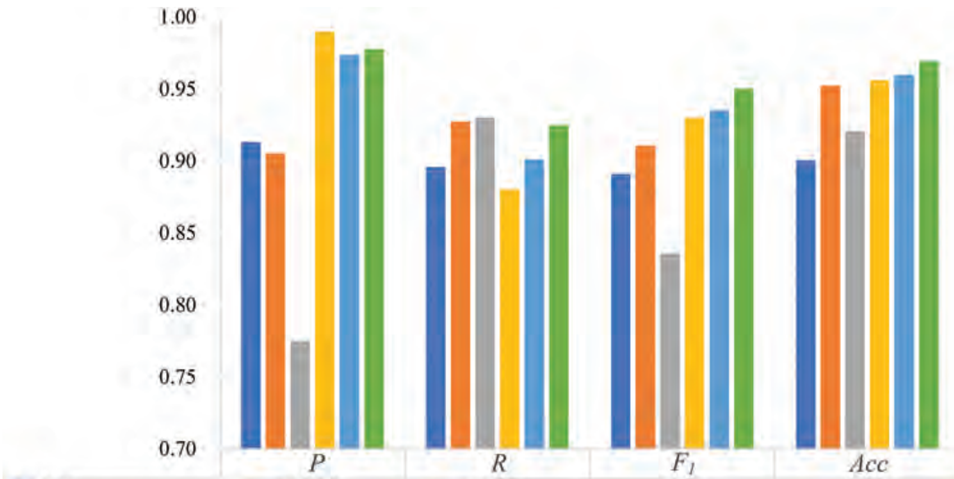
The F_1 score is a statistical measure used to assess binary classification model performance, being the harmonic mean of precision and recall. A higher F_1 score indicates better overall performance. The proposed method demonstrated the best performance with the highest mean F_1 score of 0.950, signifying a balance between precision and recall.

Accuracy (*Acc*) provides a direct measure of overall model performance, with higher accuracy indicating fewer prediction errors. The proposed method achieved a mean accuracy of 0.969, surpassing the values obtained by the other five methods. This suggests that the proposed method delivers superior signal and noise photon classification results by effectively detecting and removing noise photons.

As illustrated in Table 1, the eight datasets can be categorized into two main groups: those collected during the day and those collected at night, as well as datasets in gentle and steep terrains. Therefore, this

Table 2. Accuracy indicators comparison for the tested datasets.

		data1	data2	data3	data4	data5	data6	data7	data8	mean
LDS	*P*	0.908	0.950	0.997	0.500	0.999	0.979	0.998	0.974	0.913
	R	0.983	0.845	0.914	1.000	0.834	0.853	0.843	0.894	0.896
	F_1	0.944	0.895	0.954	0.667	0.909	0.912	0.914	0.933	0.891
	Acc	0.969	0.945	0.960	0.500	0.940	0.954	0.961	0.974	0.900
DBSCAN	*P*	0.873	0.991	0.979	0.998	0.945	0.884	0.909	0.662	0.905
	R	0.982	0.837	0.935	0.925	0.891	0.903	0.965	0.980	0.927
	F_1	0.923	0.907	0.957	0.960	0.917	0.894	0.936	0.789	0.910
	Acc	0.958	0.950	0.963	0.938	0.949	0.948	0.974	0.936	0.952
DRAGANN	*P*	0.941	0.988	0.906	0.921	0.522	0.504	0.725	0.693	0.775
	R	0.992	0.893	0.953	0.957	0.882	0.887	0.948	0.927	0.930
	F_1	0.966	0.938	0.929	0.938	0.656	0.643	0.822	0.793	0.836
	Acc	0.981	0.968	0.942	0.909	0.835	0.863	0.934	0.934	0.921
Modified DBSCAN	*P*	0.993	0.999	0.997	1.000	1.000	0.998	0.974	0.957	0.990
	R	0.972	0.862	0.930	0.927	0.845	0.723	0.897	0.886	0.880
	F1	0.982	0.925	0.962	0.962	0.916	0.838	0.934	0.920	0.930
	A	0.989	0.960	0.967	0.941	0.945	0.906	0.971	0.970	0.956
Modified OPTICS	*P*	0.976	0.997	0.995	1.000	0.996	0.916	0.955	0.951	0.973
	R	0.980	0.867	0.928	0.918	0.857	0.848	0.932	0.880	0.901
	F1	0.978	0.927	0.960	0.957	0.921	0.881	0.944	0.914	0.935
	A	0.987	0.961	0.966	0.933	0.949	0.939	0.976	0.967	0.960
The proposed method	*P*	0.986	0.998	0.991	0.997	0.990	0.968	0.938	0.954	0.978
	R	0.982	0.872	0.940	0.938	0.886	0.886	0.961	0.933	0.925
	F_1	0.984	0.931	0.965	0.967	0.935	0.925	0.950	0.943	0.950
	Acc	0.991	0.963	0.970	0.949	0.959	0.962	0.979	0.979	0.969



	P	R	F_1	Acc
LDS	0.913	0.896	0.891	0.900
DBSCAN	0.905	0.927	0.910	0.952
DRAGANN	0.775	0.930	0.836	0.921
Modified DBSCAN	0.990	0.880	0.930	0.956
Modified OPTICS	0.973	0.901	0.935	0.960
The proposed method	0.978	0.925	0.950	0.969

Figure 9. Comparison of mean accuracy indicators among these six methods.

study further examines the performance of the proposed method across different dataset characteristics. Figure 10a,b present comparisons of mean F_1 score and accuracy among the six methods for datasets collected during the day and at night, respectively. It is evident that the DBSCAN, Modified DBSCAN and Modified OPTICS methods exhibit satisfied performance for both daytime and nighttime datasets, while the DRAGANN method performs poorly for both types. Additionally, it is apparent that, in terms of both F_1 score and accuracy, the proposed method surpasses the other five methods.

In addition to comparing accuracy indicators, this paper also analyzed the time efficiency of the different methods. The comparison results are presented in Table 3. It can be observed from Table 3 that DBSCAN, DRAGANN and Modified DBSCAN demonstrated similar computation times. The average computation time for the eight datasets is approximately 2 s. The LDS method exhibited a longest computation time due to the necessity of calculating local distance statistics, which is typically time-consuming. The proposed method had the longer average computation time at 4.67 seconds. This is because it involves three steps to remove noise based on spatial distribution characteristics, leading to a longer process for achieving the final denoising result. It should be noted that the extended computation time may restrict the practical application of the proposed method in real-time scenarios.

4. Illustration on the performance for the datasets collected under different conditions

As tabulated in Table 1, the datasets under evaluation consist of four distinct types: ICESat-2 photons gathered during both daytime and nighttime, as well as ICESat-2 photons located in gentle and steep terrains. Consequently, this study selected four different datasets (data1, data3, data5, and data8) with diverse characteristics for thorough comparison.

From the illustration in Figure 11a, it is evident that the dataset is situated in a flat terrain with a gentle slope. Referring to Table 1, it can be noted that this dataset was collected during daytime. Figure 11b shows the referenced denoising result, where red points denote signal photons and green points represent noise photons. Figure 11c showcases the denoising results obtained using the LDS method. A comparison with the reference result shown in Figure 11c reveals that the LDS method encounters challenges in managing noise photons near the terrain and tends to misclassify canopy photons as noise photons, as indicated by the labeled ellipses in Figure 11c. This difficulty arises from the fact that the LDS primarily identifies noise photons based on local distance statistical histograms. While signal photons with smaller local distances are effectively preserved, those with larger sum of values are prone to erroneous classification. Moving on to Figure 11d, this depicts the denoising outcomes achieved through the DBSCAN

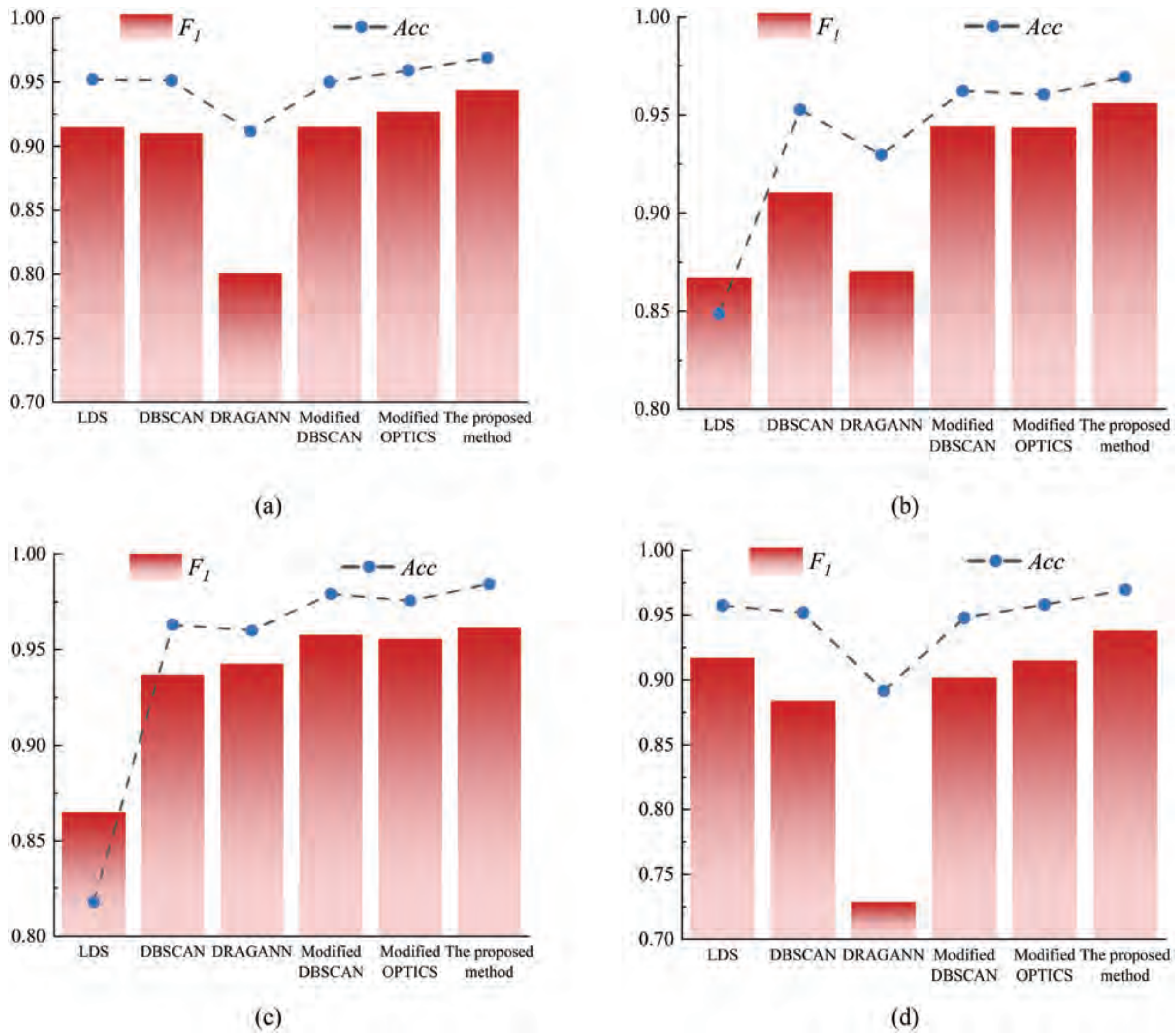


Figure 10. Comparisons of mean F_1 score and accuracy among the six methods for datasets of different characteristics. (a) and (b) present comparisons of mean F_1 score and accuracy among the six methods for datasets collected during the day and at night, respectively. (c) and (d) depict comparisons of mean F_1 score and accuracy among different methods for datasets in gentle and steep terrains, respectively.

Table 3. Comparison of time efficiency among different methods (unit: second).

	data1	data2	data3	data4	data5	data6	data7	data8	mean
LDS	8.81	4.33	8.67	17.81	3.14	7.72	7.23	10.03	8.47
DBSCAN	2.87	1.1	1.06	7.31	0.88	1.96	1.82	2.31	2.41
DRAGANN	1.54	1.4	1.71	3.64	1.42	1.84	1.65	2.73	1.99
Modified DBSCAN	1.91	0.98	1.39	5.11	0.76	1.43	1.44	1.74	1.85
Modified OPTICS	1.65	1.02	1.55	20.33	0.97	1.46	1.26	1.35	3.70
The proposed method	3.00	1.78	3.38	21.95	1.69	1.96	1.74	1.89	4.67

method. The efficacy of this approach is heavily influenced by two key parameters (*eps* and *minpts*). Inappropriate parameter settings can yield incorrect photon denoising outcomes, as depicted in Figure 11d. When compared to the LDS and DBSCAN methods, it is evident that the DRAGANN method delivers superior denoising results, albeit some signal photons close to the terrain are still misclassified as noise photons, as shown in Figure 11e. Figure 11f,g presents the results obtained using the Modified DBSCAN and Modified OPTICS methods. It is clear that a significant limitation of these two methods is their tendency to mistakenly

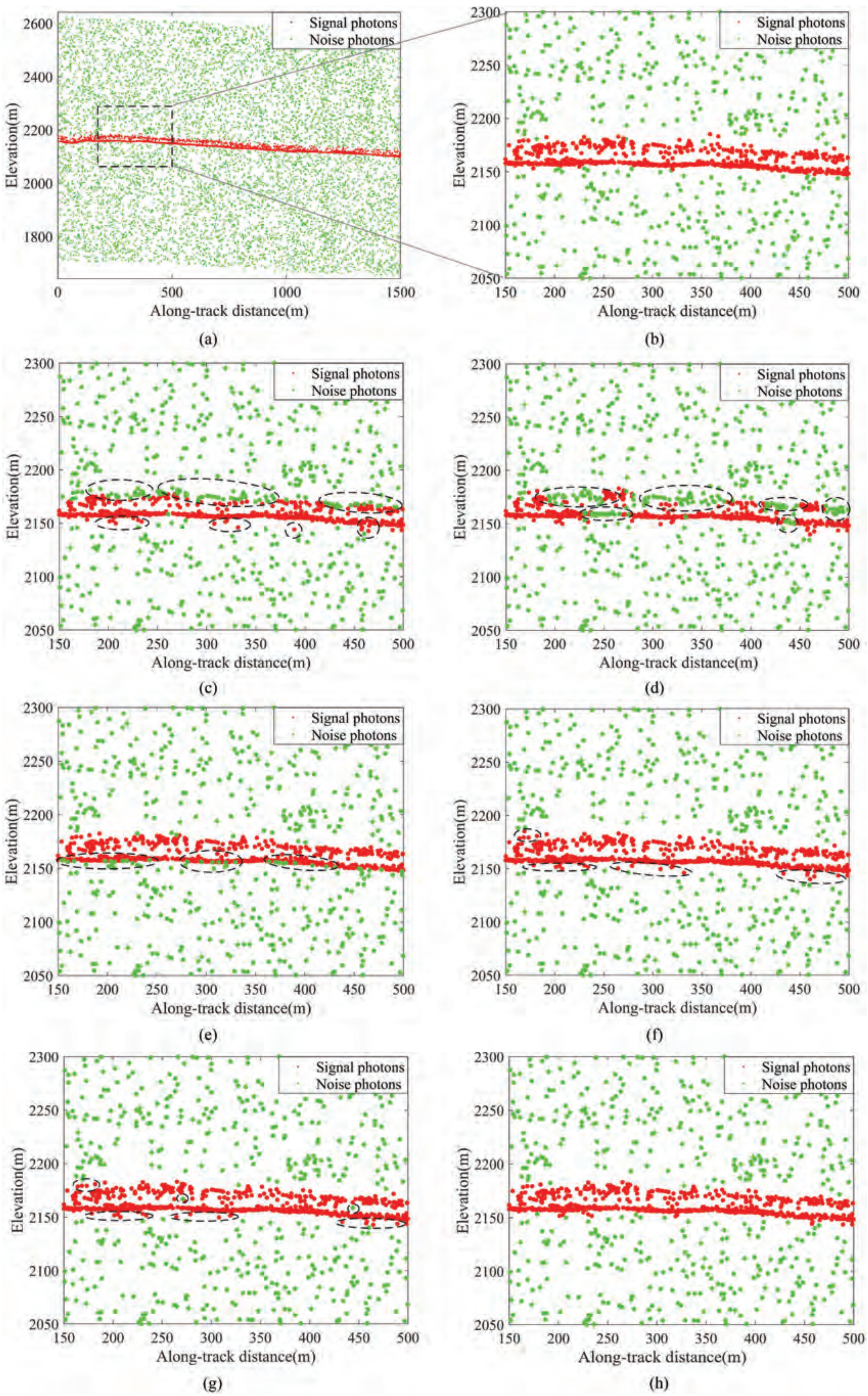


Figure 11. Detailed comparison of denoising results for ICESat-2 photons gathered during daytime located in gentle terrain among these six methods. (a) ICESat-2 photons located in gentle terrain. (b) The referenced denoising result for the selected area. (c) The

classify some lower noise photons as signal photons. Finally, Figure 11h presents the results of noise photon removal using the proposed method which demonstrates significantly more accurate denoising outcomes than the aforementioned methods.

Figure 10c,d depict comparisons of mean F_1 score and accuracy among different methods for datasets in gentle and steep terrains, respectively. For datasets in gentle terrains, where terrain slope changes gradually, all five methods (DBSCAN, DRAGANN, Modified DBSCAN, Modified OPTICS and the proposed method) perform well except for the LDS method. The proposed method demonstrates superior performance in both F_1 score and accuracy. For datasets in steep terrains, where terrain slope changes more drastically, there are variations in F_1 score values among the six methods; however, the proposed method still achieves the highest F_1 score. This trend is also observed in terms of accuracy metrics. Thus, it can be concluded that the proposed method consistently delivers optimal denoising performance regardless of terrain type – whether gentle or steep.

In contrast to datasets collected during daytime, nighttime-collected photons generally exhibit lower levels of noise but also possess lower photon density (Kui et al. 2023; Pan et al. 2024). To assess the proposed method's performance on nighttime photons, a dataset gathered in a flat terrain during nighttime was selected for testing purposes, as illustrated in Figure 12a. From Figures 12c–h, it becomes apparent that due to reduced noise levels among nighttime photons, fewer misclassifications occur compared to their daytime counterparts. Figure 12c illustrates the denoising results obtained using the LDS method, where several photons at lower elevations are mistakenly categorized as noise photons. In Figure 12d, the DBSCAN denoising outcome is presented. As this method utilizes a clustering approach, instances where the clustered results fail to meet the thresholds result in incorrect classifications, as depicted in Figure 12d. Subsequently, Figure 12e showcases the denoising outcome of the DRAGANN method. Similar to Figure 11e, some photons from lower terrain are misclassified as noise photons in addition to all signal photons labeled within the left ellipse. This misclassification is attributed to the DRAGANN method's detection of noise photons based on the assumption of bimodal distribution of elevations for noise and signal photons. By identifying the intersection of two Gaussian curves, noise photons can be eliminated. However, accurately determining this intersection proves challenging. Incorrectly determined thresholds can lead to misclassification of photons across the entire region. Regarding Figure 12f,g, a similar observation can be made as in Figure 11f,g: the Modified DBSCAN and Modified OPTICS methods tend to incorrectly classify lower noise photons. A comparison with the denoising outcomes displayed in Figure 12b reveals that the noise photon removal results by the proposed method (Figure 12h) are much closer to the reference outcome.

In addition to the data collected by ICESat-2 in gentle terrains, Figures 13 and 14 illustrate how different methods perform when processing datasets from steep terrain. Figure 13a shows an area situated in mountainous terrain with varying slopes. A comparison with the reference result in Figure 13b reveals that some photons with higher or lower elevations are incorrectly classified by the LDS method, as depicted in Figure 13c. This is because the LDS method identifies noise photons based on local distance statistical histograms, using a threshold derived from the frequency histogram mean plus t times the standard deviation to distinguish noisy points. However, determining an accurate threshold in rugged terrains poses a challenge, leading to misclassification of photons near signal points as shown in Figure 13c.

Figure 13d presents the result of noise photon removal using the DBSCAN method, where misclassified photons are mixed with signal photons unlike in the LDS method's output. The DRAGANN method performs poorly on this dataset (Figure 13e), erroneously labeling many signal photons as noise due to challenges in discerning noise from signal photons based on intersecting Gaussian curves in local density histograms. Both Figure 13f,g demonstrate that the Modified DBSCAN and Modified OPTICS methods tend to misclassify noise photons that are close to signal photons in steep terrains.

Compared to the other three methods, this study achieved superior denoising results (Figure 13h) by classifying noise photons into three types and gradually removing them based on their characteristics, yielding satisfactory outcomes even in rugged terrains.

Figure 14 displays denoising outcomes of the four methods applied to ICESat-2 data collected during nighttime in steep terrain. A comparison with the reference result (Figure 14b) demonstrates that the LDS, DBSCAN, DRAGANN, Modified DBSCAN and Modified OPTICS

denoising result obtained by the LDS method. (d) The denoising result obtained by the DBSCAN method. (e) The denoising result obtained by the DRAGANN method. (f) The denoising result obtained by the modified DBSCAN method. (g) The denoising result obtained by the modified OPTICS method. (h) The denoising result obtained by the proposed method.

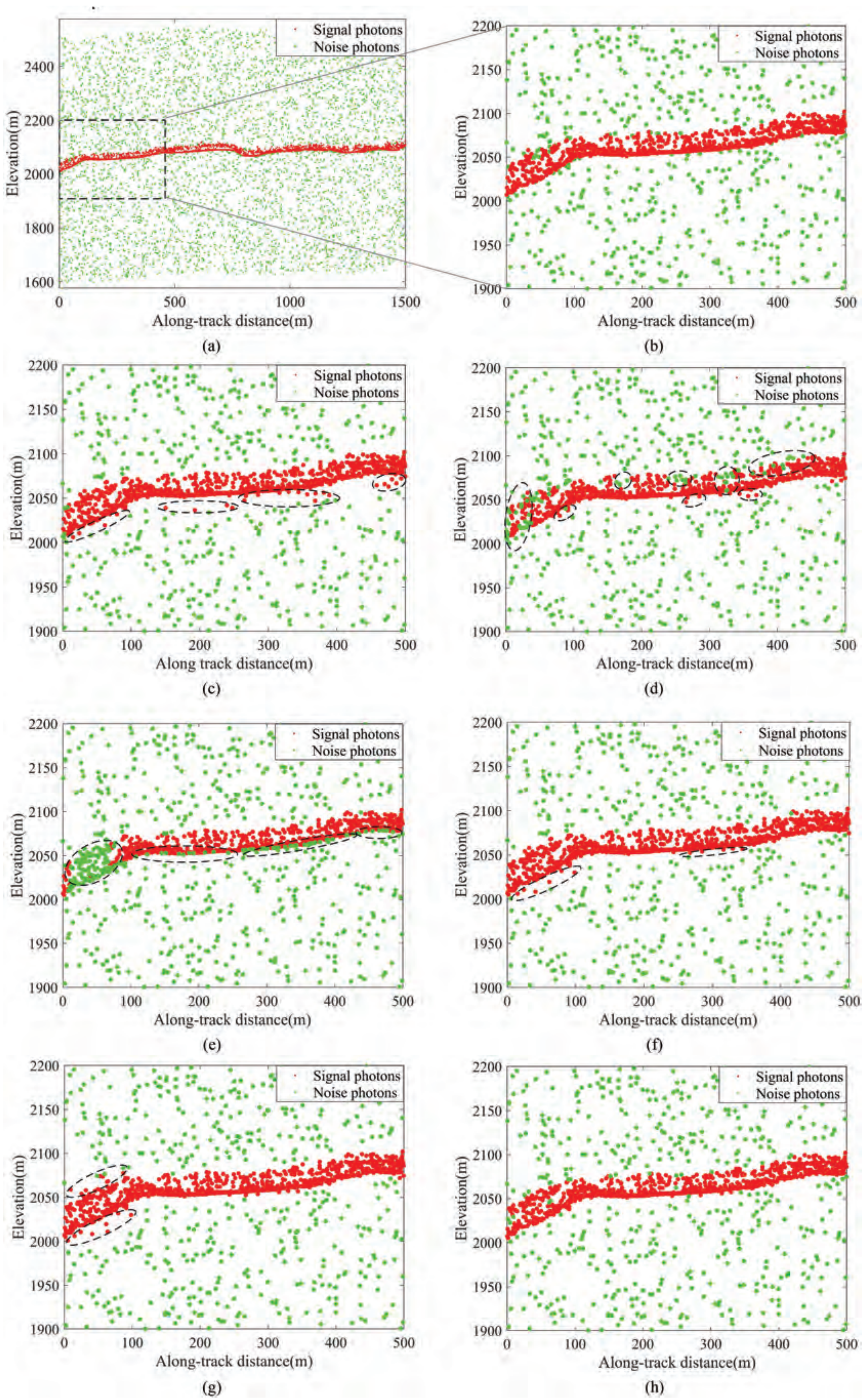


Figure 12. Detailed comparison of denoising results for ICESat-2 photons gathered during nighttime located in gentle terrain among these six methods. (a) ICESat-2 photons located in gentle terrain. (b) The referenced denoising result for the selected area. (c) The

methods perform notably worse than the proposed method in rugged terrains, with numerous misclassified photons shown in Figures 14c–g due to parameter-dependent performance leading to improper classification thresholds.

Conversely, the proposed method adaptively detects and removes noise photons without manual threshold adjustments, enhancing robustness across different terrain types and yielding satisfactory denoising results as depicted in Figure 14h.

5. Discussion

5.1. Necessity analysis towards different steps combination

In this paper, noise photons are categorized into isolated, low-density clustered, and outer clustered types based on their unique spatial distribution characteristics. Each type is then targeted with specific denoising techniques using a multi-step strategy. The denoising process involves three steps: Step 2.1 focuses on removing isolated noise photons through a multi-thresholding strategy, Step 2.2 involves adaptive calculation of terrain slopes to eliminate low-density clustered noise photons, and Step 2.3 deals with removing outer clustered noise photons using the box plot analysis technique.

Results from Section 2.2 show that a significant number of noise photons are removed after the second step of denoising. This prompts the need to evaluate performance without implementing Step 2.1 and consider the impact of swapping the order of Step 2.1 and Step 2.2. To address these questions, various combinations of steps were tested on datasets listed in Table 1, and the outcomes are presented in Table 4.

Analysis of Table 4 reveals that omitting Step 2.1 results in lower accuracy across all four indicators compared to the proposed method, highlighting the importance of this step in the denoising process. While a considerable reduction in noise photons is achieved after Step 2.2, performance significantly deteriorates without the preceding Step 2.1. Swapping the order of Steps 2.1 and 2.2 leads to notably inferior results, particularly in F1-score and accuracy indicators.

The observed decline in performance can be attributed to the structured approach of the proposed method, which systematically addresses each type of noise photon based on its distribution characteristics. By following this sequential removal process tailored to distinct noise types, superior outcomes are achieved compared to directly combining multiple steps without such differentiation.

5.2. Impact of elevation frequency histogram resolution

In Section 2.2, the core photons are identified based on the elevation frequency histogram, which is crucial for calculating terrain slope. Therefore, it is essential to analyze the effect of varying the resolution of the elevation frequency histogram. Generally, a smaller resolution is suitable for flat areas like gentle terrain. This ensures that high-frequency photons are accurately identified as terrain points. Conversely, in steep terrain, a larger resolution is needed to ensure proper selection of high-frequency photons as terrain points. In this study, a resolution of 1 m is used for gentle terrain and 15 m for steep terrain.

To assess the impact of different resolutions of elevation frequency histograms on slope estimation, this study calculates the mean slope deviation (MSD) and root mean square error (RMSE) of slope estimation under various resolutions. The reference terrain slope is calculated using a DTM generated from NEON's airborne LiDAR data, while the estimated terrain slope is calculated using two successive core photons. The results are presented in Table 5.

Table 5 shows that smaller resolutions lead to lower slope deviations for gentle terrains (data1, data2, data3, and data4), but higher deviations for steep terrains (data5, data6, data7, and data8). Conversely, larger resolutions yield opposite results. By using smaller resolutions for gentle terrain and larger resolutions for steep terrain, more accurate slope estimations can be achieved. However, it should be noted that regardless of the resolution used, the mean slope deviation across all eight datasets remains relatively consistent. This suggests that terrain slope estimation using core photons is effective in this study.

5.3. Influence of parameter settings

In this paper, several parameters are involved in the proposed method, namely the filtering window size (S) and the number of neighboring points (N) mentioned in Section 2.1, as well as the minor axis of the ellipse (b) mentioned in Section 2.2. To analyze the impact of these parameter settings on performance, this study tested

denoising result obtained by the LDS method. (d) The denoising result obtained by the DBSCAN method. (e) The denoising result obtained by the DRAGANN method. (f) The denoising result obtained by the modified DBSCAN method. (g) The denoising result obtained by the modified OPTICS method. (h) The denoising result obtained by the proposed method.

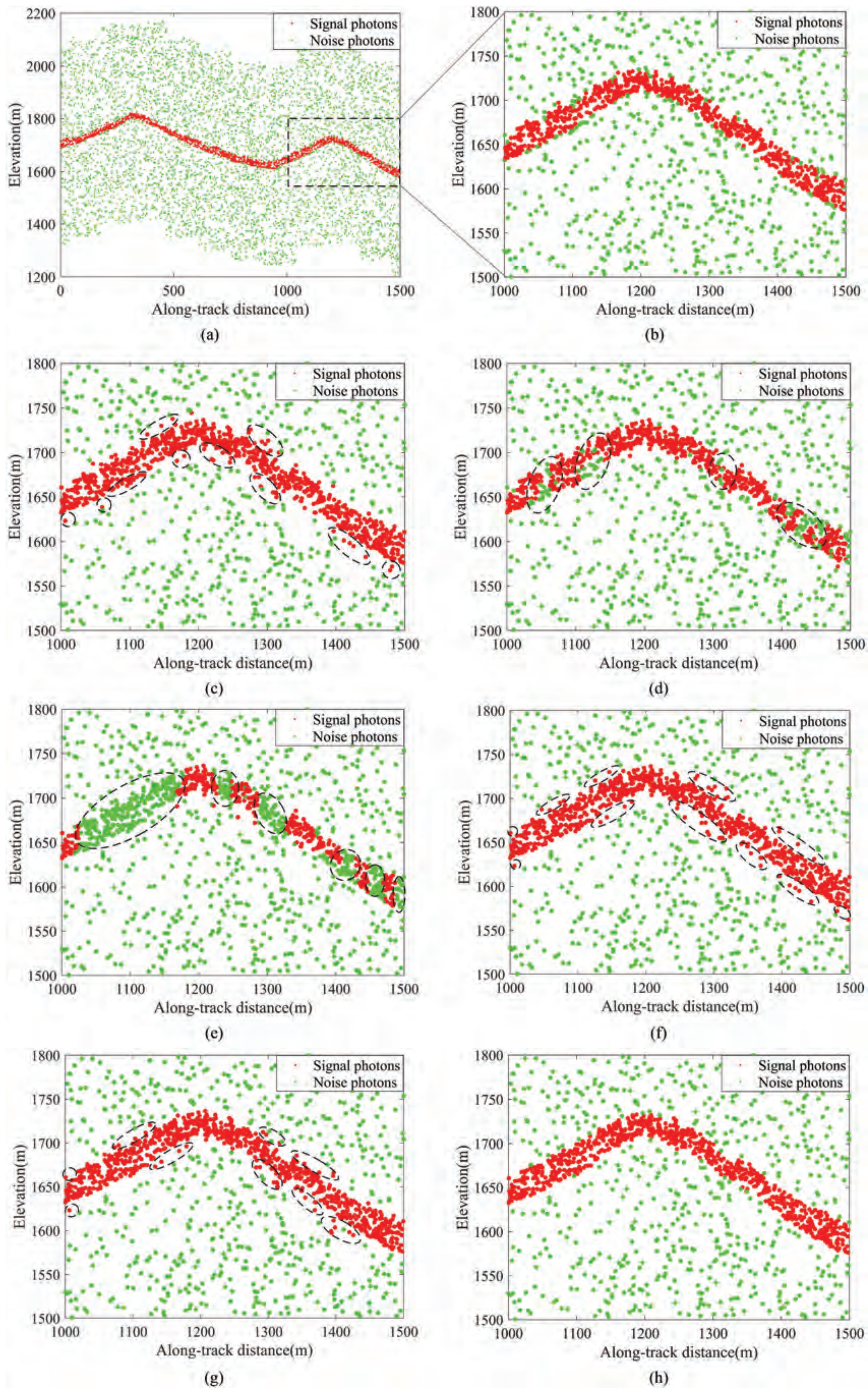


Figure 13. Detailed comparison of denoising results for ICESat-2 photons gathered during daytime located in steep terrain among these six methods. (a) ICESat-2 photons located in steep terrain. (b) The referenced denoising result for the selected area. (c) The

eight datasets listed in Table 1. The accuracy metrics obtained with different parameter settings are presented in Table 6. It is observed that while varying parameter settings may influence the denoising outcome, the calculated accuracy metrics remain relatively stable. In essence, although multiple parameters are involved in this study, their adjustment does not significantly alter denoising performance. This is attributed to the categorization of noise photons into three distinct groups and their gradual removal using a specific denoising technique outlined in this paper. Adopting a multi-step approach allows for incremental removal of noise photons, thereby improving the robustness and efficacy of the proposed method.

5.4. Performance towards the weak beam data

As mentioned in Section 3.2, ICESat-2/ATLAS emits six beams of laser pulses along its orbital path, each containing three pairs. These pairs consist of a strong signal and a weak signal, with an energy ratio of 4:1 between them. The datasets tested in Table 1 correspond to the strong beam data. To evaluate the performance of the proposed method on weak data, this study conducted additional tests using the weak beam data from ICESat-2, as detailed in Table 7.

Given that the energy of the weak beam (40μJ) is only one-quarter that of the strong beam (160μJ), it is evident that the number of signal photons in the strong beam surpasses that in the weak beam, as depicted in Figure 15. Consequently, the signal-to-noise ratio (SNR) of weak beam data is significantly lower compared to that of strong beam data.

The paper also calculated four accuracy indicators for five other methods – LDS, DBSCAN, DRAGANN, Modified DBSCAN, and Modified OPTICS – for comparison purposes. The comparison results are displayed in Figure 16. When analyzing precision (*P*) for weak beam data, the proposed method achieved the higher average precision value of 0.938, outperforming all other methods. Particularly noteworthy is that the DRAGANN method demonstrated the lowest average precision at 0.700, notably inferior to our approach. Precision represents the proportion of true positive predictions among all positive predictions, indicating that our method excels in accurately identifying signal photons within weak beam data.

In terms of recall (*R*), LDS, DBSCAN and Modified DBSCAN showcased similar average recall rates ranging from 0.738 to 0.797. While both DRAGANN and our proposed method achieved average recall rates above 0.850, DBSCAN performed least effectively with an average recall rate of 0.738. This highlights our method's ability to detect a higher number of signal photons within weak beam data.

The *F1* score – a statistical metric evaluating binary classification model performance by considering the harmonic mean of precision and recall – signifies overall model effectiveness. Our method exhibited the highest average *F1* score of 0.900, further validating its superior performance on weak beam data.

Accuracy (*Acc*) directly measures overall model performance, with higher values indicating fewer prediction errors. Our method demonstrated an average accuracy of 0.961, surpassing all other tested methods. This underscores our approach's efficacy in detecting and eliminating noise photons, leading to exceptional signal-noise photon classification outcomes.

5.5. Performance towards different land covers

To evaluate the performance on various land covers, this study selected other eight datasets covering three distinct types: forest, vegetation, and mixture. The forest land cover includes evergreen, deciduous, and mixed forests. Vegetation cover comprises shrubs, crops, woody wetlands, and herbaceous wetlands. Mixture cover includes bare land and building materials. The ICESat-2 data were matched with the National Land Cover Database (NLCD) in ArcGIS to determine the corresponding land cover types. The three land covers used for testing are shown in Figure 17.

The evaluation metrics of the proposed method and five comparative approaches across different land cover types are presented in Figure 18. Subfigures (a)-(c) display the precision, recall, and F1-scores for six methodologies under forest, vegetation, and mixture conditions. Our method maintains a balance between precision and recall across all categories, consistently achieving the highest F1-scores.

For forest land cover, DRAGANN shows poor precision suitability for forested areas. Modified OPTICS has the highest precision (0.975) but insufficient recall (0.888). LDS has higher recall (0.916) but suboptimal F1-score (0.928). DRAGANN has high recall (0.913) but

denoising result obtained by the LDS method. (d) The denoising result obtained by the DBSCAN method. (e) The denoising result obtained by the DRAGANN method. (f) The denoising result obtained by the modified DBSCAN method. (g) The denoising result obtained by the modified OPTICS method. (h) The denoising result obtained by the proposed method.

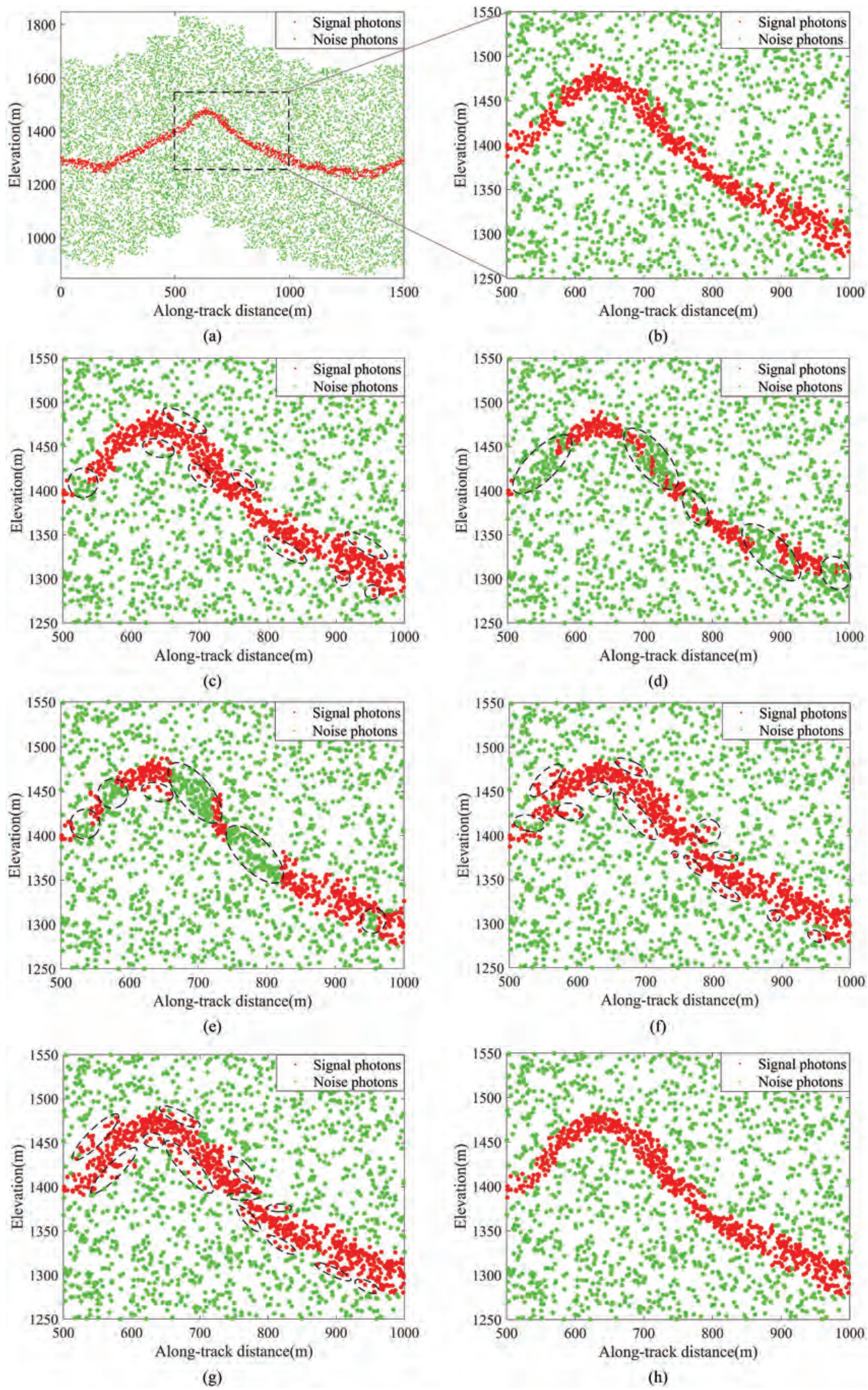


Figure 14. Detailed comparison of denoising results for ICESat-2 photons gathered during nighttime located in steep terrain among these six methods. (a) ICESat-2 photons located in steep terrain. (b) The referenced denoising result for the selected area. (c) The

Table 4. Necessity analysis toward different steps combination.

	P	R	F1	Acc
Without Step 2.1	0.966	0.910	0.937	0.962
Swapping Step 2.1 and Step 2.2	0.710	0.862	0.762	0.866
Step 2.1 +Step 2.2 +Step 2.3	0.978	0.925	0.950	0.969

Table 5. Slope estimation accuracy across varying resolutions of elevation frequency histograms (in degrees).

Resolution	Indicator	data1	data2	data3	data4	data5	data6	data7	data8	mean
1 m	*MSD*	0.762	0.224	0.246	0.246	7.041	15.612	9.116	10.742	5.499
	RMSE	0.600	0.310	0.403	0.358	10.385	11.656	12.398	15.244	6.419
15 m	*MSD*	4.768	7.469	8.738	4.568	4.733	7.193	3.304	3.287	5.508
	RMSE	7.841	14.128	13.820	6.585	5.878	9.324	4.489	3.749	8.227
1 m &	*MSD*	0.762	0.224	0.246	0.246	4.733	7.193	3.304	3.287	2.499
15 m	*RMSE*	0.600	0.310	0.403	0.358	5.878	9.324	4.489	3.749	3.139

Table 6. Accuracy metrics using different parameter settings.

Parameter	value	*P*	*R*	*F1*	*Acc*
S	45	97.796	92.462	94.993	96.884
	55	97.781	92.517	95.016	96.899
	65	97.726	92.473	94.967	96.874
N	40	97.533	92.717	94.998	96.890
	50	97.675	92.601	95.007	96.896
	60	97.811	92.420	94.978	96.879
b	4	89.928	93.624	91.148	95.467
	5	96.069	92.395	94.029	96.374
	6	98.287	91.265	94.550	96.452

Table 7. Detailed information of ICESat-2 weak beam data for the study sites.

Study sites	Datasets	Time	Data type	Terrain slope
YELL	data1	Daytime	gt3l/weak	gentle
	data2	Daytime	gt1r/weak	gentle
	data3	Nighttime	gt1l/weak	gentle
	data4	Nighttime	gt1l/weak	gentle
GRSM	data5	Daytime	gt1r/weak	steep
	data6	Daytime	gt3r/weak	steep
	data7	Nighttime	gt1r/weak	steep
	data8	Nighttime	gt1l/weak	steep

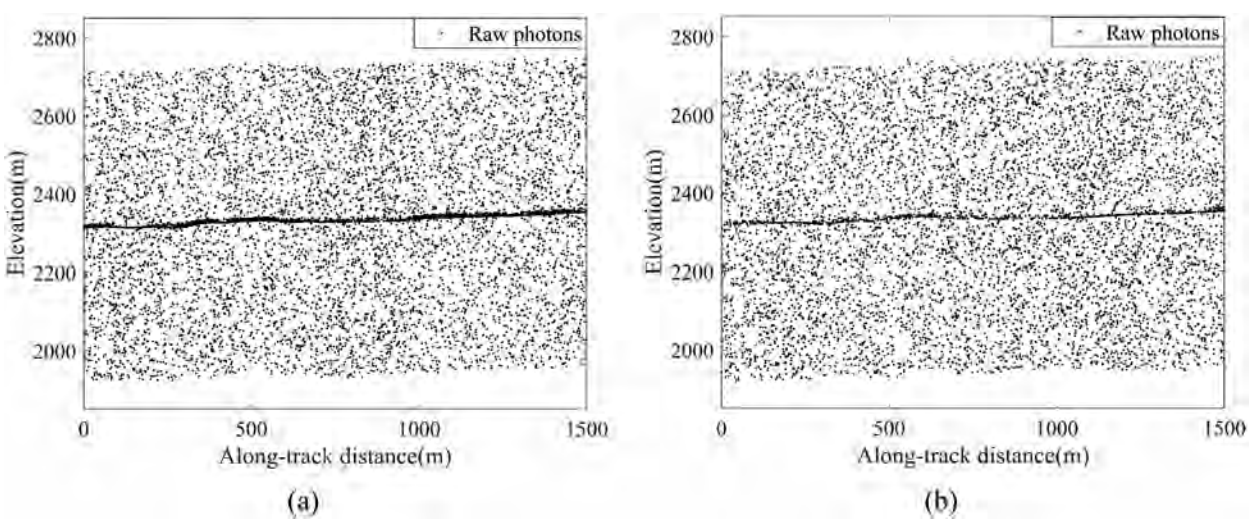
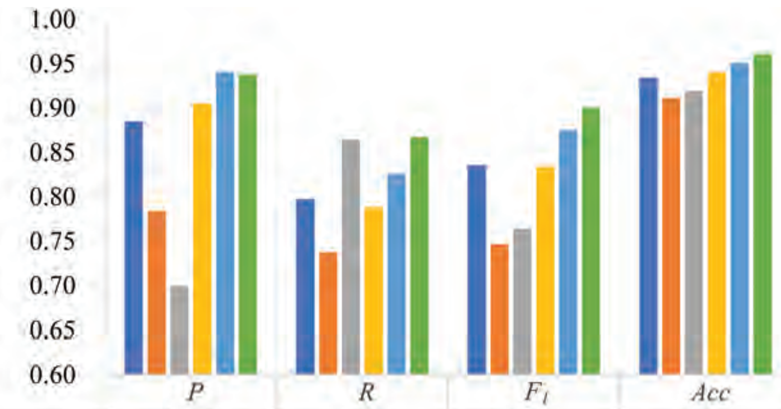


Figure 15. Strong and weak beam data comparison. (a) Strong beam data. (b) Weak beam data.

denoising result obtained by the LDS method. (d) The denoising result obtained by the DBSCAN method. (e) The denoising result obtained by the DRAGANN method. (f) The denoising result obtained by the modified DBSCAN method. (g) The denoising result obtained by the modified OPTICS method. (h) The denoising result obtained by the proposed method.



	P	R	F_1	Acc
LDS	0.885	0.797	0.836	0.935
DBSCAN	0.784	0.738	0.747	0.912
DRAGANN	0.700	0.864	0.764	0.920
Modified DBSCAN	0.905	0.789	0.834	0.941
Modified OPTICS	0.941	0.826	0.876	0.951
The proposed method	0.938	0.867	0.900	0.961

Figure 16. Comparison of mean accuracy indicators among these six methods toward the weak beam data.

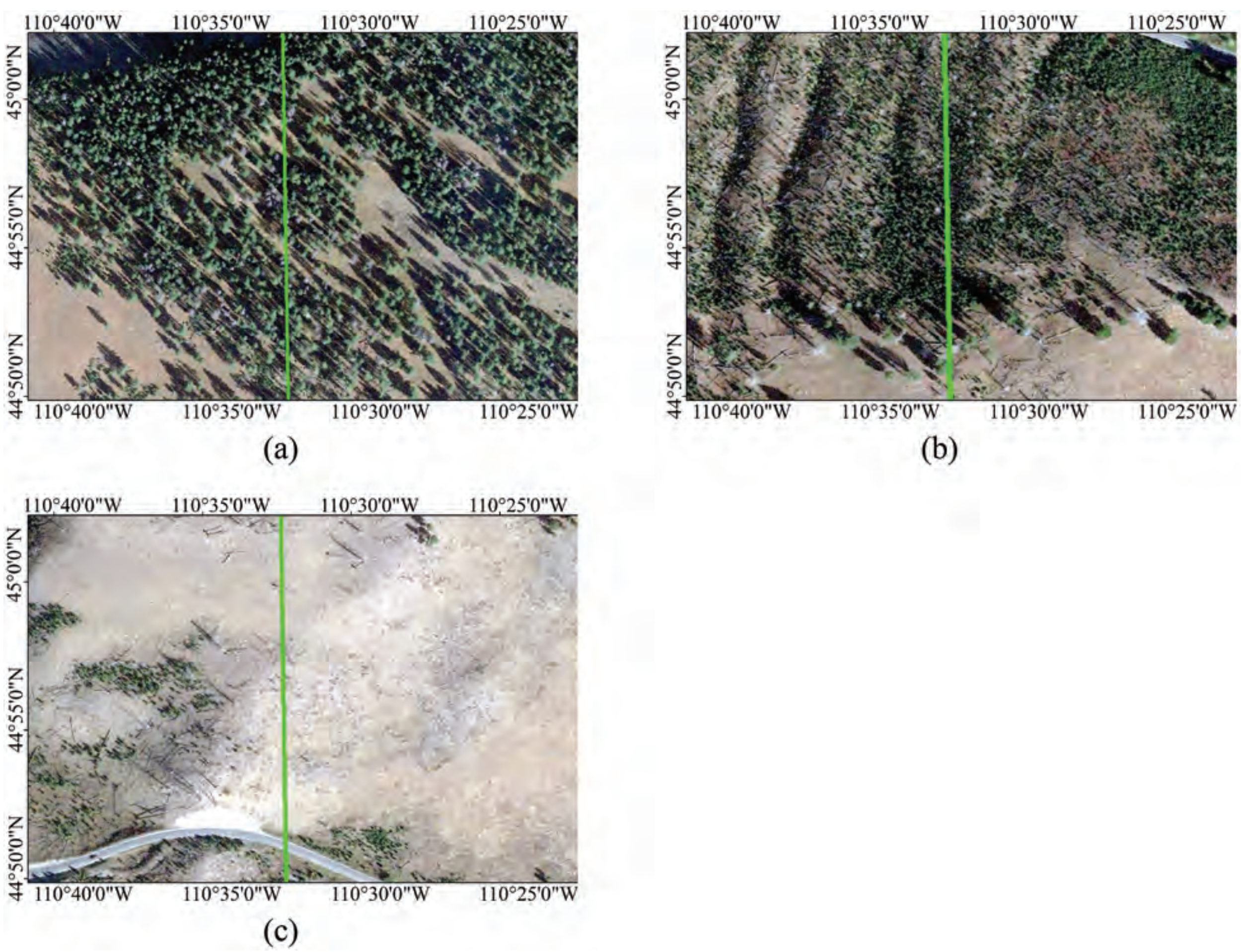


Figure 17. Icesat-2 photons corresponding to different land covers. (a) Forest land cover. (b) Vegetation land cover. (c) Mixture land cover.

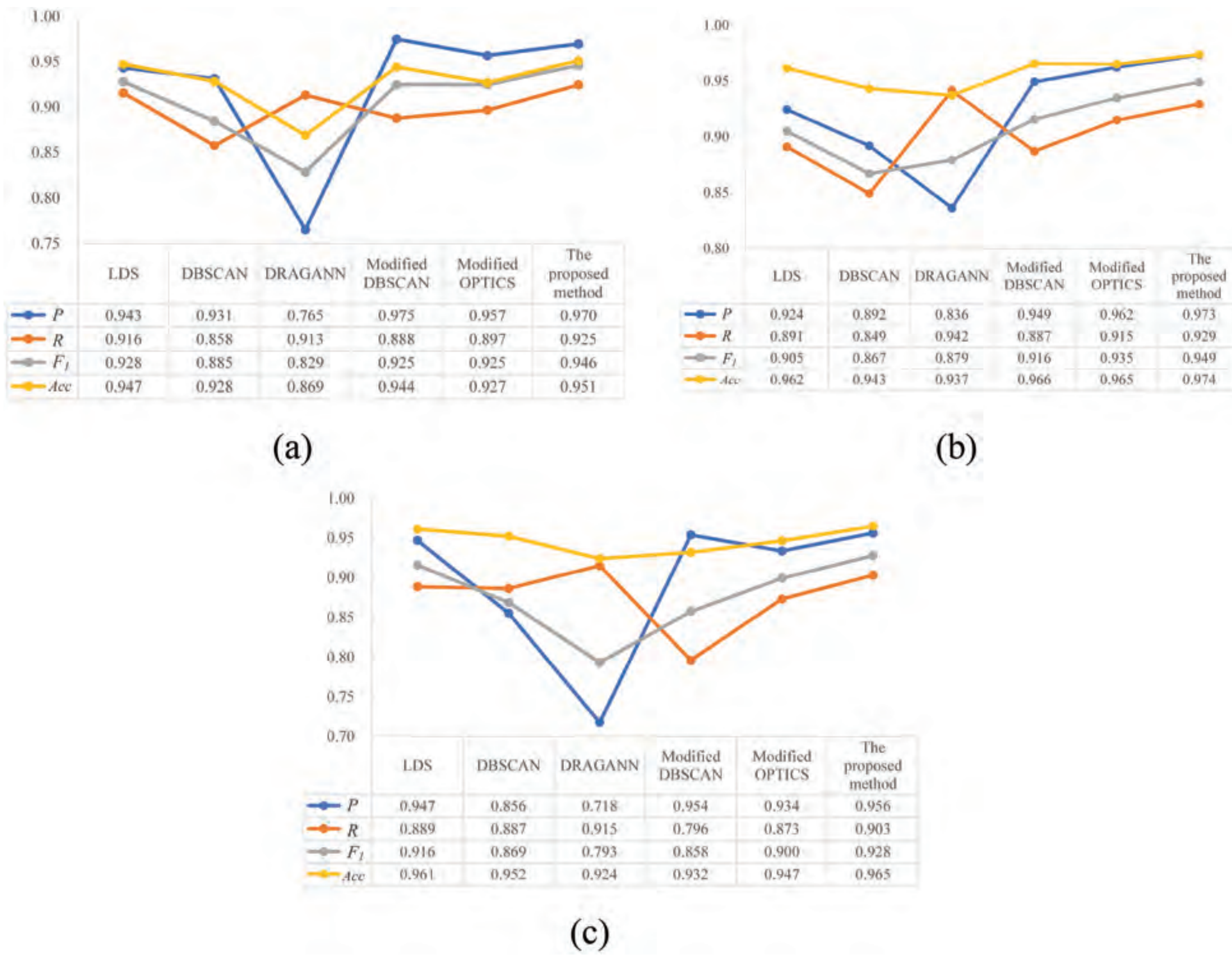


Figure 18. Accuracy indicators comparison toward different land covers. (a) Comparison result for the forest land cover. (b) Comparison result for the vegetation land cover. (c) Comparison result for the mixture land cover.

low precision (0.765). For vegetation land cover, DRAGANN has the highest recall (0.942) but inadequate precision (0.836). Both Modified DBSCAN and Modified OPTICS show improved denoising compared to conventional methods. For mixture land cover, DBSCAN and LDS have balanced precision and recall. Modified DBSCAN has high precision (0.954) but lower recall (0.796).

Overall, our proposed methodology provides robust denoising performance across all three land cover types, with strengths in vegetation and mixture covers.

6. Conclusion

ICESat-2 data has been utilized in various fields. However, ICESat-2 data are affected by a significant amount of noise photons. Therefore, the removal of noise is a critical step in processing ICESat-2 photons. As noise photons exhibit a sparse and random spatial distribution while signal photons tend to cluster closely, this study presents a progressive noise removal approach based on different noise characteristics. The classification of noise photons into isolated noise photons, low-density clustered noise photons, and outer clustered noise photons is done based on their distinct spatial distribution features. Isolated noise photons are identified using a multi-thresholding strategy employing the maximum between-clustering variance algorithm. For the low-density clustered noise photons, an ellipse-based photon counting method is proposed considering that signal photons are generally densely clustered with higher photon densities. The major axis of the ellipse aligns with the local terrain slope, which is dynamically calculated using the Douglas-Peucker algorithm. To address the third type of noise photons, a box plots analysis technique based on local elevation distributions is introduced. Evaluations conducted on datasets with diverse characteristics, including both daytime and nighttime acquisitions, as well as varying land covers, consistently highlight the effectiveness of our method.

The innovative aspects of our method lie in its progressive and type-specific denoising strategy, which effectively overcomes the limitations of uniform denoising methods. This study introduces a novel classification of noise photons into three distinct types based on their unique spatial distribution characteristics, providing deeper insights into the nature of noise. Furthermore, tailored denoising techniques were proposed for each type of noise. For instance, isolated noise photons can be eliminated without parameter tuning through the application of the maximum between-clustering variance algorithm. Additionally, to align the ellipse-based photon counting method with changing terrain slopes, we introduced a novel automatic calculation method based on the Douglas-Peucker algorithm. The automatic denoising of outer clustered noise photons was achieved through the introduction of a box plots analysis technique.

The high efficiency in denoising and preservation of signal photons achieved by our method suggest its potential for broad implementation in ICESat-2 data processing workflows. Future research could investigate combining our approach with machine learning (such as XGBoost, LightGBM) for noise photon classification by integrating the spatial structural features created in this study (such as density, local elevation distribution, and principal axis orientation). To improve geographic interpretability, we will also attempt to use spatial autocorrelation metrics (such as Moran's I) to assess the anomalous character of local photon distributions and create multivariate discriminative models that combine terrain features and spatial structural attributes.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Double Thousand Plan of Jiangxi Province [DHSQT42023002]; Outstanding Young Talents Funding of Jiangxi Province [20232ACB213017]; Funding of National Key Laboratory of Uranium Resources Exploration-Mining and Nuclear Remote Sensing [2024QZ-TD-26]; National Natural Science Foundation of China (NSF) [42161060, 41801325]; Natural Science Foundation of Jiangxi Province [20242BAB25176, 20192BAB217010].

Author contributions

Z.H. conceived the original idea of the study and drafted the manuscript. L.Z. conducted the experiments and made the experimental results analysis. P.C., W.C. and Y.Y. Z. contributed to the revision of the manuscript. All authors have read and agreed to the published version of the manuscript.

Data availability statement

The ICESat-2 data for the study sites in this paper is accessible for download at no cost from the following link: https://nsidc.org/data/atl03. The relevant airborne LiDAR data can be accessed from the following site: https://www.neonscience.org/data-collection/LiDAR. The source codes can be accessed at https://github.com/Smart3DLiDAR/ICESat-2-Photon-Denoising.git.

References

Abdalati, W., H. J. Zwally, R. Bindschadler, B. Csatho, S. L. Farrell, H. A. Fricker, D. Harding, et al. 2010. "The ICESat-2 Laser Altimetry Mission." *Proceedings of the IEEE* 98 (5): 735–751. https://doi.org/10.1109/JPROC.2009.2034765.

Agca, M., A. Yucel, E. Kaya, A. İ. Daloglu, M. Kayalık, M. Yetkin, and F. Yalcın. 2024. "Machine Learning Algorithms for Building Height Estimations Using ICESat-2/ATLAS and Airborne LiDAR Data." *Earth Science Informatics* 17 (6): 5123–5134. https://doi.org/10.1007/s12145-024-01429-w.

Cao, B. C., Y. Fang, L. Gao, H. Y. Hu, and Z. Z. Jiang. 2020a. "Verification of ICESat-2/ATLAS Laser Altimetry Data Accuracy Using Airborne Point Cloud." *Journal of Geomatics Science and Technology* 37 (1): 50–55.

Cao, B. C., Y. Fang, Z. Z. Jiang, L. Gao, and H. Y. Hu. 2020b. "Implementation and Accuracy Evaluation of ICESat-2 ATL08 Denoising Algorithms." *Bulletin of Surveying & Mapping* 5 (01): 25–30. https://doi.org/10.13474/j.cnki.11-2246.2020.0139.

Chen, B., Y. Pang, Z. Li, H. Lu, P. North, J. Rosette, M. Yan. 2020. "Forest Signal Detection for Photon Counting LiDAR Using Random Forest." *Remote Sensing Letters* 11 (1): 37–46. https://doi.org/10.1080/2150704X.2019.1682708.

Chen, B. W., and Y. Pang. 2015. "A Denoising Approach for Detection of Canopy and Ground from ICESat-2's Airborne Simulator Data in Maryland, USA." *International Society for Optics and Photonics* 9671 (01): 96711S. https://doi.org/10.1117/12.2202777.

Chen, B. W., Y. Pang, Z. Y. Li, H. Lu, L. X. Liu, P. R. J. North, and J. A. B. Rosette. 2019. "Ground and Top of Canopy Extraction from Photon-Counting LiDAR Data Using Local Outlier Factor with Ellipse Searching Area." *IEEE Geoscience & Remote Sensing Letters* 16 (9): 1447–1451. https://doi.org/10.1109/LGRS.2019.2899011.

Chen, B. W., Y. Pang, Z. Y. Li, P. North, J. Rosette, G. Q. Sun, J. Suárez, I. Bye, and H. Lu. 2019. "Potential of Forest Parameter Estimation Using Metrics from Photon Counting LiDAR Data in Howland Research Forest." *Remote Sensing* 11 (7): 856. https://doi.org/10.3390/rs11070856.

He, G. H., H. Wang, Q. Fang, Y. A. Zhang, D. L. Zhao, and Y. P. Zhang. 2023. "Spaceborne Photon Counting Lidar Point Cloud Denoising Method with the Adaptive Mountain Slope." *Journal of Infrared and Millimeter Waves* 42 (2): 250–259.

He, X. K., F. P. Gan, Y. Guo, Y. Zhuo, B. K. Yan, J. Bai, N. C. Xing, R. Y. Li, L. T. Dai, and J. Z. Yang. 2024. "A Binned Multilevel

Regression Fitting Method for Monitoring Geladandong Glacier Elevation Changes from ICESat-2 Data." *International Journal of Remote Sensing* 45 (22): 7959–7995. https://doi.org/10.1080/01431161.2024.2383383.

Huang, J. P., Y. Q. Xing, Y. M. Shuai, and H. Z. Zhu. 2022. "A Novel Noise Filtering Evaluation Criterion of ICESat-2 Signal Photon Data in Forest Environments." *IEEE Geoscience & Remote Sensing Letters* 19:1–5. https://doi.org/10.1109/LGRS.2022.3163143.

Huang, J. P., Y. Q. Xing, H. T. You, L. Qin, J. Tian, and J. M. Ma. 2019. "Particle Swarm Optimization-Based Noise Filtering Algorithm for Photon Cloud Data in Forest Area." *Remote Sensing* 11 (8): 980. https://doi.org/10.3390/rs11080980.

Jiao, H. H., J. F. Xie, R. Liu, and F. Mo. 2021. "Discussion on Denoising Method of Photon Counting LiDAR for Satellite Ground Observation." *Spacecraft Recovery and Remote Sensing* 42 (5): 140–150.

Kong, D., and Y. Pang. 2024. "ICESat-2 Data Denoising and Forest Canopy Height Estimation Using Machine Learning." *International Journal of Applied Earth Observation and Geoinformation* 135:104263–104263. https://doi.org/10.1016/j.jag.2024.104263.

Kui, M. Y., Y. N. Xu, J. L. Wang, and F. Cheng. 2023. "Research on the Adaptability of Typical Denoising Algorithms Based on ICESat-2 Data." *Remote Sensing* 15 (15): 3884. https://doi.org/10.3390/rs15153884.

Leng, Z. H., J. Zhang, Y. Ma, J. Y. Zhang, and H. T. Zhu. 2022. "A Novel Bathymetry Signal Photon Extraction Algorithm for Photon-Counting LiDAR Based on Adaptive Elliptical Neighborhood." *International Journal of Applied Earth Observation and Geoinformation* 115:103080. https://doi.org/10.1016/j.jag.2022.103080.

Leng, Z., J. Zhang, Y. Ma, and J. Zhang. 2023. "ICESat-2 Bathymetric Signal Reconstruction Method Based on a Deep Learning Model with Active–Passive Data Fusion." *Remote Sensing* 15 (2): 460. https://doi.org/10.3390/rs15020460.

Li, Y., G. Chen, Z. Han, L. Zheng, and F. And Zhang. 2014. "A Hybrid Automatic Thresholding Approach Using Panchromatic Imagery for Rapid Mapping of Landslides." *GIScience and Remote Sensing* 51 (6): 710–730. https://doi.org/10.1080/15481603.2014.972867.

Li, Y. M., Q. H. Guo, B. Wan, H. N. Qin, D. Z. Wang, K. X. Xu, S. L. Song, et al. 2021. "Current Status and Prospect of Three-Dimensional Dynamic Monitoring of Natural Resources Based on LiDAR." *National Remote Sensing Bulletin* 25 (1): 381–402. https://doi.org/10.11834/jrs.20210351.

Lian, W. Q., G. Zhang, H. Cui, Z. W. Chen, S. D. Wei, C. Y. Zhu, and Z. G. Xie. 2022. "Extraction of High-Accuracy Control Points Using ICESat-2 ATL03 in Urban Areas." *International Journal of Applied Earth Observation and Geoinformation* 115:103116. https://doi.org/10.1016/j.jag.2022.103116.

Lin, Y., and A. Knudby. 2023. "Global Automated Extraction of Bathymetric Photons from ICESat-2 Data Based on a PointNet++ Model." *International Journal of Applied Earth Observation and Geoinformation* 124:103512. https://doi.org/10.1016/j.jag.2023.103512.

Liu, J., J. Y. Liu, H. Xie, D. Ye, and P. N. Li. 2023. "A Multi-Level Auto-Adaptive Noise-Filtering Algorithm for Land ICESat-2 Photon-Counting Data." *Remote Sensing* 15 (21): 5176. https://doi.org/10.3390/rs15215176.

Liu, R., X. M. Tang, J. F. Xie, R. J. Ma, F. Mo, and X. M. Yang. 2024. "A Novel Spaceborne Photon-Counting Laser Altimeter Denoising Method Based on Parameter-Adaptive Density Clustering." *GIScience and Remote Sensing* 61 (1): 2326702. https://doi.org/10.1080/15481603.2024.2326702.

Liu, X. F., Z. H. Wang, W. Z. Yang, S. X. Chen, F. X. Wang, X. W. Chen, W. M. Xu, and R. Shu. 2024. "An Adaptive Signal Photon Detection Method Based on DBSCAN for Photon-Counting Laser Altimeter." *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing* 17:3674–3686. https://doi.org/10.1109/JSTARS.2024.3352023.

Liu, Z., Y. Zi, X. Ma, Y. Ma, X. Wu, G. Jiang, H. Liu, F. Cheng, and Y. Zhou. 2024. "VOJA-Net: Vector-Offset Joint Attention Network for ICESat-2 Point Cloud Data Denoising." *IEEE Geoscience & Remote Sensing Letters* 21:1–5. https://doi.org/10.1109/LGRS.2024.3464520.

Lu, D. J., D. Li, X. X. Zhu, S. Nie, G. Q. Zhou, X. Y. Zhang, and C. Yang. 2021. "Denoising and Classification of ICESat-2 Photon Point Cloud Based on Convolutional Neural Network." *Journal of Geo-Information Science* 23 (11): 2086–2095.

Ma, Y., W. Zhang, J. Sun, G. Li, X. H. Wang, S. Li, and N. Xu. 2019. "Photon-Counting Lidar: An Adaptive Signal Detection Method for Different Land Cover Types in Coastal Areas." *Remote Sensing* 11 (4): 471. https://doi.org/10.3390/rs11040471.

Magruder, L. A., and M. E. W. I. Kevin. 2012. "Noise filtering techniques for photon-counting LADAR data." *International Society for Optics and Photonics* 8379 (01): 83790Q. https://doi.org/10.1117/12.919139.

Mansouri, J., M. Jafari, and A. T. Dehkordi. 2024. "Continuous Mapping of Forest Canopy Height Using ICESat-2 Data and a Weighted Kernel Integration of Multi-Temporal Multi-Source Remote Sensing Data Aided by Google Earth Engine." *Environmental Science and Pollution Research* 31 (37): 49757–49779. https://doi.org/10.1007/s11356-024-34415-2.

Neuenschwander, A., L. Duncanson, P. Montesano, D. Minor, E. Guenther, S. Hancock, M. A. Wulder, et al. 2024. "Towards Global Spaceborne Lidar Biomass: Developing and Applying Boreal Forest Biomass Models for ICESat-2 Laser Altimetry Data." *Science of Remote Sensing* 10:100150–100150. https://doi.org/10.1016/j.srs.2024.100150.

Neuenschwander, A., E. Guenther, J. C. White, L. Duncanson, and P. Montesano. 2020. "Validation of ICESat-2 Terrain and Canopy Heights in Boreal Forests." *Remote Sensing of Environment* 251:112110. https://doi.org/10.1016/j.rse.2020.112110.

Neuenschwander, A. L., K. L. Pitts, B. P. Jelley, J. Robbins, B. Klotz, S. C. Popescu, R. F. Nelson, D. Harding, D. Pederson, and R. Sheridan. 2024. "ATLAS/ICESat-2 L3A Land and Vegetation Height, Version 5." https://nsidc.org/data/atl08/versions/5.

Neuenschwander, A., and K. Pitts. 2019. "The ATL08 Land and Vegetation Product for the ICESat-2 Mission." *Remote Sensing of Environment* 221:247–259. https://doi.org/10.1016/j.rse.2018.11.005.

Nie, S., C. Wang, X. H. Xi, S. Z. Luo, G. Y. Li, J. Y. Tian, and H. T. Wang. 2018. "Estimating the Vegetation Canopy Height Using Micro-Pulse Photon-Counting LiDAR Data."

Optics Express 26 (10): A520–A540. https://doi.org/10.1364/OE.26.00A520.

Pan, J. Y., F. Gao, J. L. Wang, J. P. Zhang, Q. W. Liu, and Y. C. Deng. 2024. "A Main Direction-Based Noise Removal Algorithm for ICESat-2 Photon-Counting LiDAR Data." *Journal of Geodesy* 98 (9): 80–80. https://doi.org/10.1007/s00190-024-01887-6.

Qin, W. J., Y. Song, Y. R. Zou, H. T. Zhu, and H. Y. Guan. 2024. "A Novel ICESat-2 Signal Photon Extraction Method Based on Convolutional Neural Network." *Remote Sensing* 16 (1): 203. https://doi.org/10.3390/rs16010203.

Tian, X., and J. Shan. 2023. "Detection of Signal and Ground Photons from ICESat-2 ATL03 Data." *IEEE Transactions on Geoscience & Remote Sensing* 61:1–14. https://doi.org/10.1109/TGRS.2022.3232053.

Varvia, P., S. Saarela, M. Maltamo, P. Packalen, T. Gobakken, E. Næsset, G. Ståhl, and L. Korhonen. 2024. "Estimation of Boreal Forest Biomass from ICESat-2 Data Using Hierarchical Hybrid Inference." *Remote Sensing of Environment* 311:114249. https://doi.org/10.1016/j.rse.2024.114249.

Vučetić, N., S. Petrović, and A. Strunje. 2007. "Simplification of Lines Containing Fixed Points Using the Douglas-Peucker Algorithm." *Kartografija i geoinformacije* 6 (7): 74–81.

Wang, R. Q., Y. G. Lu, D. S. Lu, and G. Y. Li. 2024. "Improving Extraction of Forest Canopy Height Through Reprocessing ICESat-2 ATLAS and GEDI Data in Sparsely Forested Plain Regions." *GIScience and Remote Sensing* 61 (1): 2396807. https://doi.org/10.1080/15481603.2024.2396807.

Wang, X., Z. G. Pan, and C. Glennie. 2016. "A Novel Noise Filtering Model for Photon-Counting Laser Altimeter Data." *IEEE Geoscience & Remote Sensing Letters* 13 (7): 947–951. https://doi.org/10.1109/LGRS.2016.2555308.

Xia, S. B., C. Wang, X. H. Xi, S. Z. Luo, and H. C. Zeng. 2014. "Point Cloud Filtering and Tree Height Estimation Using Airborne Experiment Data of ICESat-2." *National Remote Sensing Bulletin* 18 (6): 1199–1207. https://doi.org/10.11834/jrs.20144029.

Xie, F., G. Yang, R. Shu, and M. Li. 2017. "An Adaptive Directional Filter for Photon Counting Lidar Point Cloud Data." *Journal of Infrared and Millimeter Waves* 36 (1): 107–113.

Xie, Y., S. Shi, L. Xun, and P. Wang. 2023. "A Multitemporal Index for the Automatic Identification of Winter Wheat Based on Sentinel-2 Imagery Time Series." *GIScience and Remote Sensing* 60 (1): 2262833. https://doi.org/10.1080/15481603.2023.2262833.

Xu, X. Q., W. Wang, D. Huang, X. G. Hu, and W. R. Fu. 2023. "Assessment of Interannual and Seasonal Glacier Mass Changes in the Karakoram During 2018–2022 Using ICESat-2 Data." *Journal of Hydrology* 626:130223. https://doi.org/10.1016/j.jhydrol.2023.130223.

Ye, M. Y., C. B. Yang, X. Q. Zhang, S. Li, X. R. Peng, Y. Y. Li, and T. Y. Chen. 2024. "Shallow Water Bathymetry Inversion Based on Machine Learning Using ICESat-2 and Sentinel-2 Data." *Remote Sensing* 16 (23): 4603–4603. https://doi.org/10.3390/rs16234603.

Zhang, B. C., Y. X. Liu, Z. P. Dong, J. Li, Y. L. Chen, Q. H. Tang, G. A. Huang, and J. L. Tao. 2024. "An Optimal Denoising Method for Spaceborne Photon-Counting LiDAR Based on a Multiscale Quadtree." *Remote Sensing* 16 (13): 2475–2475. https://doi.org/10.3390/rs16132475.

Zhang, C., R. Wang, B. Yang, L. Yang, Y. Yang, F. Liu, and K. Xiong. 2025. "A Direction-Adaptive DBSCAN-Based Method for Denoising ICESat-2 Photon Point Clouds in Forested Environments." *Forests* 16 (3): 524. https://doi.org/10.3390/f16030524.

Zhang, G., S. Xing, Q. Xu, P. Li, and D. Wang. 2023. "A Pre-Pruning Quadtree Isolation Method with Changing Threshold for ICESat-2 Bathymetric Photon Extraction." *Remote Sensing* 15 (6): 16296. https://doi.org/10.3390/rs15061629.

Zhang, G., S. Xing, Q. Xu, F. Zhang, M. Dai, and D. Wang. 2023. "An Automatic Algorithm to Extract Nearshore Bathymetric Photons Using Pre-Pruning Quadtree Isolation for ICESat-2 Data." *IEEE Geoscience & Remote Sensing Letters* 20:1–5. https://doi.org/10.1109/LGRS.2023.3263469.

Zhang, G., Q. Xu, S. Xing, P. Li, X. Zhang, D. Wang, and M. Dai. 2021. "A Noise-Removal Algorithm without Input Parameters Based on Quadtree Isolation for Photon-Counting LiDAR." *IEEE Geoscience & Remote Sensing Letters* 19:1–5. https://doi.org/10.1109/LGRS.2021.3081721.

Zhang, J., and J. Kerekes. 2015. "An Adaptive Density-Based Model for Extracting Surface Returns from Photon-Counting Laser Altimeter Data." *IEEE Geoscience & Remote Sensing Letters* 12 (4): 726–730. https://doi.org/10.1109/LGRS.2014.2360367.

Zhang, Z., X. Liu, Y. Ma, N. Xu, W. Zhang, and S. Li. 2021. "Signal Photon Extraction Method for Weak Beam Data of ICESat-2 Using Information Provided by Strong Beam Data in Mountainous Areas." *Remote Sensing* 13 (5): 863. https://doi.org/10.3390/rs13050863.

Zhang, Z., N. Xu, Y. Ma, X. Liu, W. Zhang, and S. Li. 2020. "Land and Snow-Covered Area Classification Method Based on the Background Noise for Satellite Photon-Counting Laser Altimeters." *Optics Express* 28 (11): 16030–16044. https://doi.org/10.1364/OE.392904.

Zhu, X. X., S. Nie, C. Wang, X. H. Xi, and Z. Y. Hu. 2018. "A Ground Elevation and Vegetation Height Retrieval Algorithm Using Micro-Pulse Photon-Counting Lidar Data." *Remote Sensing* 10 (12): 1962. https://doi.org/10.3390/rs10121962.

Zhu, X. X., S. Nie, C. Wang, X. H. Xi, J. S. Wang, D. Li, and H. Y. Zhou. 2021. "A Noise Removal Algorithm Based on OPTICS for Photon-Counting LiDAR Data." *IEEE Geoscience & Remote Sensing Letters* 18 (8): 1471–1475. https://doi.org/10.1109/LGRS.2020.3003191.

Zhu, X. X., C. Wang, X. H. Xi, S. Nie, X. B. Yang, and D. Li. 2020. "Research Progress of ICESat-2/ATLAS Data Processing and Applications." *Infrared and Laser Engineering* 49 (11): 20200259–20200259. https://doi.org/10.3788/IRLA20200259.