A Progressive Noise Removal Method for ICESat-2 Data Based on Terrain Slope Adaptive Calculation

Zhenyang Hui [1,2,3], Li Zhang [1,2,3], Ying Chen [1], Penggen Cheng [1,2,3,]*, Yuanping Xia [1,2,3], Ting Hui [4], Ziyan Huang [1]

[1] School of Surveying and Geoinformation Engineering, East China University of Technology, Nanchang, China
[2] Jiangxi Key Laboratory of Watershed Ecological Process and Information, East China University of Technology, Nanchang, 330013, China;
[3] Key Laboratory of Mine Environmental Monitoring and Improving around Poyang Lake of Ministry of Natural Resources, East China University of Technology, Nanchang, 330013, China
huizhenyang2008@ecut.edu.cn, 2023110409@ecut.edu.cn, 3032113856@qq.com, pgcheng@ecut.edu.cn, ypxia@ecut.edu.cn, 1273787503@qq.com
[4] College of Management, Guangdong AIB Polytechnic, thui@gdaib.edu.cn

Commission II, WG II/2

Keywords: Noise removal, Signal photon, Terrain slope, Douglas-Peucker

Abstract:

The ICESat-2, the satellite-borne photon-counting laser altimeter, has a wide range of applications. However, the data collected by ICESat-2 are often affected by high levels of noise photons. Therefore, the removal of this noise is a crucial step in processing the ICESat-2 data further. This paper proposes a novel noise removal method based on adaptive terrain slope calculation to address this issue. The method takes advantage of the distinct density distribution differences between signal and noise photons in the vertical dimension. By analyzing filtering windows, the algorithm identifies areas with high-density, low-elevation photons and creates a 50-meter elevation buffer around these points to filter out noise photons that are far from the signal clusters. The Douglas-Peucker algorithm is then used to merge data segments with similar slopes, enabling the adaptive calculation of terrain slopes within local photon regions. Furthermore, clustering based on elliptical density along the primary terrain slope direction is applied to remove discrete noise photons that are in close proximity to signal photons. Lastly, the Local Outlier Factor algorithm is utilized to eliminate residual noise photons located below ground level, in aerial regions, and near tree canopies, effectively separating noise photons from signal photons. To evaluate the effectiveness of the proposed method, experimental data sets from two regions with different geographical characteristics in the United States are selected for testing. The results show an average improvement in F1-score of about 4.6% in gentle terrains and 9.5% in rugged terrains, highlighting the method's superior accuracy and efficiency compared to traditional denoising algorithms.

1. Introduction

The Advanced Topographic Laser Altimeter System (ATLAS), which is installed on ICESat-2, is known for its low energy consumption, high detection sensitivity, and high repetition rates. The data collected by ATLAS has been extensively used in various ecological research fields, such as monitoring polar ice sheets, tracking lake water levels, mapping bathymetry in shallow waters, estimating forest canopy heights, and assessing biomass carbon stocks (Narine et al., 2019; Yuan et al., 2020; Hsu et al., 2021). However, due to the weak signals transmitted and received by the system, factors like atmospheric scattering, solar radiation, and instrumental interferences introduce significant background noise into the recorded point clouds. This noise complicates the accurate identification of signal photons, making effective denoising of photon point clouds essential in processing photon-counting LiDAR data (Zhu et al., 2020).

Current research on denoising photon point clouds can be broadly categorized into three groups: denoising algorithms based on grid image processing, where methods like those used by Magruder et al. (2012) convert profile photon points into two-dimensional images and apply image processing techniques to filter out noise points. Chen et al. (2015) employs the classic active contours method to grid image and applies the Chan-Vese segmentation model to detect potential signal photons.

Denoising algorithms that rely on local statistical parameters, such as the approach by Wang et al. (2016) that uses K-nearest neighbors' probability distribution function to calculate distances for each photon point. Hereafter, he employs Bayesian decision-making for denoising. Xia et al. (2014a) introduces a denoising algorithm based on local distance statistics and applies least squares fitting to determine local curve parameters, achieving satisfied overall accuracy. Zhu et al. (2018) devises a noise photons filtering algorithm based on local statistics with adaptive threshold determination.

Denoising methods based on density-based spatial clustering like the one by Zhang and Kerekes(2014) improves the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm by transitioning from a circular search shape to an elliptical one. Similarly, Zhu et al. (2021) use OPTICS for clustering and sorting to identify structures and remove noisy photons. He et al. (2023) revise a density clustering algorithm with an adaptive mountain slope, achieving better adaptation in forested areas with complex topography.

* Corresponding author

https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-673-2025

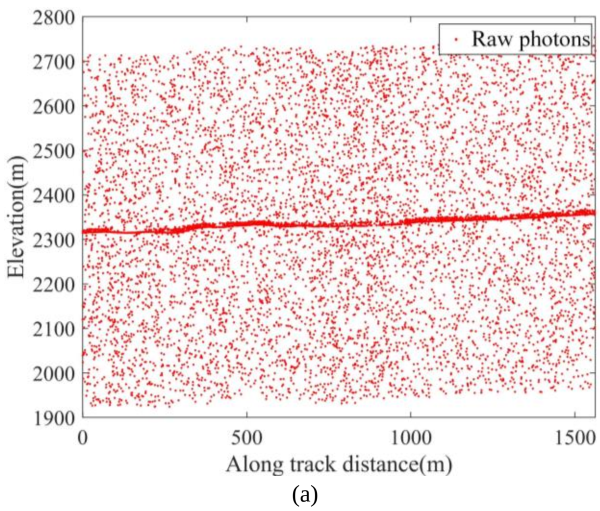
Additionally, with the increasing use of machine learning in data processing, scholars like Lu et al. (2021) have utilized convolutional neural networks for point cloud denoising and classification. Chen et al. (2020) propose a machine learning-based method for detecting potential signal photons from photon-counting LiDAR data.

While these denoising algorithms are effective in removing noise, there are still unresolved issues. For instance, grid-based image processing algorithms may result in information loss; algorithms based on local statistical parameters rely heavily on threshold selection; density-based spatial clustering algorithms are sensitive to terrain variations and parameter settings; and machine learning-based algorithms are influenced by training samples.

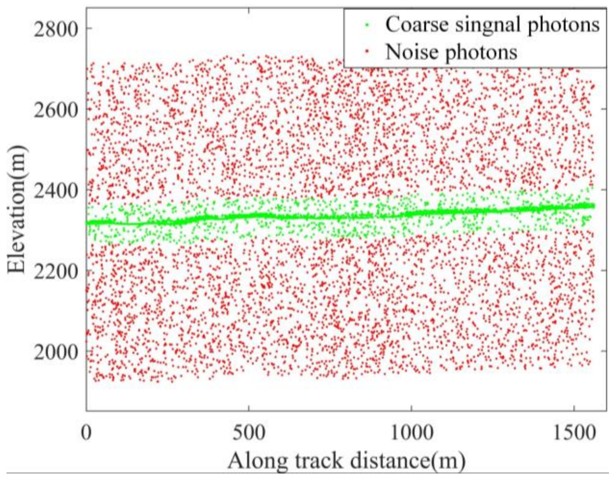
To address these challenges related to terrain slope variations, point cloud density, and parameter sensitivities, this paper proposes a novel terrain slope adaptive denoising algorithm for ICESat-2 LiDAR point clouds. This method involves adapting the primary slope direction to determine optimal clustering parameters for denoising and subsequently conducting precision validation.

2. Methodology

In this methodology, to streamline subsequent data processing, we capitalize on the significant disparity in density distribution between signal and noise photon points along the vertical axis. Initially, small windows are established to identify photon points with high density and low elevation. An elevation buffer zone extending 50 meters above and below the identified photon points is then created to eliminate obvious noise points. Figure 1 showcases the original photon point cloud profile and the results of the initial denoising process. Subsequently, the Douglas-Peucker algorithm is utilized to merge data segments with comparable slopes, facilitating the adaptive calculation of photon point slopes. Next, employing the predominant slope direction, we conduct density-based clustering within elliptical neighborhoods to remove a majority of noise points. Finally, the Local Outlier Factor (LOF) algorithm is employed to eliminate residual noise below ground level, in aerial regions, and near canopy surfaces, effectively segregating noise photons from signal photons. This methodology encompasses three primary steps: slope adaptive calculation based on the Douglas-Peucker algorithm, density-based clustering with adaptive determination of the dominant slope direction, and residual noise removal using the LOF algorithm.



(a)



(b)

Figure 1. Initial denoising with elevation buffer. (a) the original photon point cloud profile. (b) the results of the initial denoising process.

2.1 Adaptive Terrain Slope Calculation Based on the Douglas-Peucker Algorithm

Considering the impact of terrain conditions on signal photon points, the initial phase involves computing terrain slope angles for each segment based on coarse terrain point slopes over consistent along-track distance intervals. This computation is carried out using Equation (1).

$$\theta_i = \tan^{-1}\left(\frac{y_i - y_{i+1}}{x_i - x_{i+1}}\right) \quad (1)$$

where θ_i = i-th terrain slope of each interval

(x_i, y_i), (x_{i+1}, y_{i+1}) = the coordinates of two adjacent terrain points

Subsequently, the Douglas-Peucker algorithm is applied to consolidate similar terrain slopes (Vučetić et al., 2007). The rough terrain points within each interval are linked into a curve, with a virtual straight line drawn between the curve's initial and final points. The residual distance from each remaining point to this straight line is determined using Equations (2-3). A predefined distance threshold *H* is set, and the maximum distance *Dmax* is compared against this threshold. If *Dmax* surpasses *H*, the point furthest from the straight line is retained; otherwise, all points between the two endpoints of the straight line are discarded. The retained points lead to division of the known curve into two segments for further processing through an iterative selection and discarding process until no more points can be discarded. Ultimately, coordinates of curve points meeting specified accuracy criteria are obtained, completing the line simplification process as illustrated in Figure 2.

$$Ax + By + C = 0 \quad (2)$$

$$D_i = \frac{|Ax_i + By_i + C|}{\sqrt{A^2 + B^2}} \quad (3)$$

where D_i = residual distance from point p_i

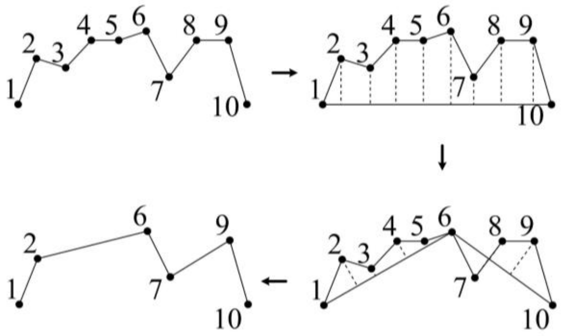


Figure 2. Schematic diagram of Douglas-Peucker algorithm

After curve simplification, the intervals are reassigned based on the simplified points, and the slope angles are recalculated using Equation (1), completing the merging of slope angles. As shown in Figure 3(b), the merged small data segments in Figure 3(a) have the same slope angle.

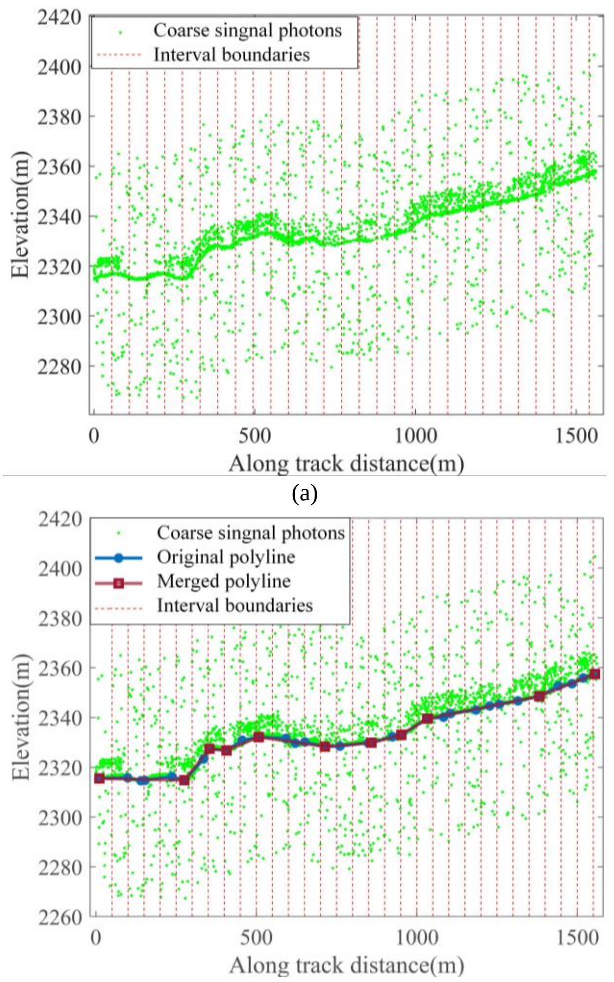


Figure 3. Data segments and merged results. (a) terrain segments with an interval of 50m. (b) Merged terrain segments using the Douglas-Peucker Algorithm

2.2 Adaptive Density Clustering Based on the Calculation of Dominant Slope Direction

Due to the higher density of signal photons along the track direction and the presence of regions with significant surface slopes, we utilize an elliptical neighborhood based on the dominant slope direction for calculating the number of neighboring photon points and performing clustering. The specific steps are outlined below:

(1) Select a point pk randomly from the point set $\{p_i\}$ that has not been visited, and flag it as visited by setting VisitedFlag(pk) = 1.
(2) Determine the number of neighboring points (D_{pk}) within the ellipse in the principal axis direction of pk, store this information in dataset E, and record the indices of neighboring photon points in dataset I. The principal axis direction aligns with the terrain slope, and the calculation formula is provided in Equation (4).
(3) Define a threshold *minpts*. If N_{pk} > *minpts*, classify pk as a signal point and include the indices of pk's neighboring points in the point set {p_neighbors} based on their indices.
(4) Iterate through all points in {p_neighbors}, repeating steps 2 and 3 until all points in {p_neighbors} have been processed.
(5) Stop if all points in $\{p_i\}$ have been visited; otherwise, return to step 1 and continue until clustering is finished.

$$\begin{cases} N_{pk}(p,a,b,\phi) = \left\{ q \middle| (\frac{dx}{a})^2 + (\frac{dy}{b})^2 \le 1 \right\} \\ dx = \left(x_p - x_q\right) * \cos\phi - \left(y_p - y_q\right) * \sin\phi \\ dy = \left(x_p - x_q\right) * \sin\phi + \left(y_p - y_q\right) * \cos\phi \end{cases} \tag{4}$$

where a = half of the major axis of the ellipse
b = half of the minor axis of the ellipse
q = one neighboring point
ϕ = the angle of the major axis direction
x_p, x_q = the x-coordinates of points p and q
y_p, y_q = the y-coordinates of points p and q

Figure 4(a) illustrates the clustering outcomes, while Figure 4(b) presents a close-up view of the signal points derived from the clustering results. These figures illustrate that this clustering approach effectively eliminates noise points.

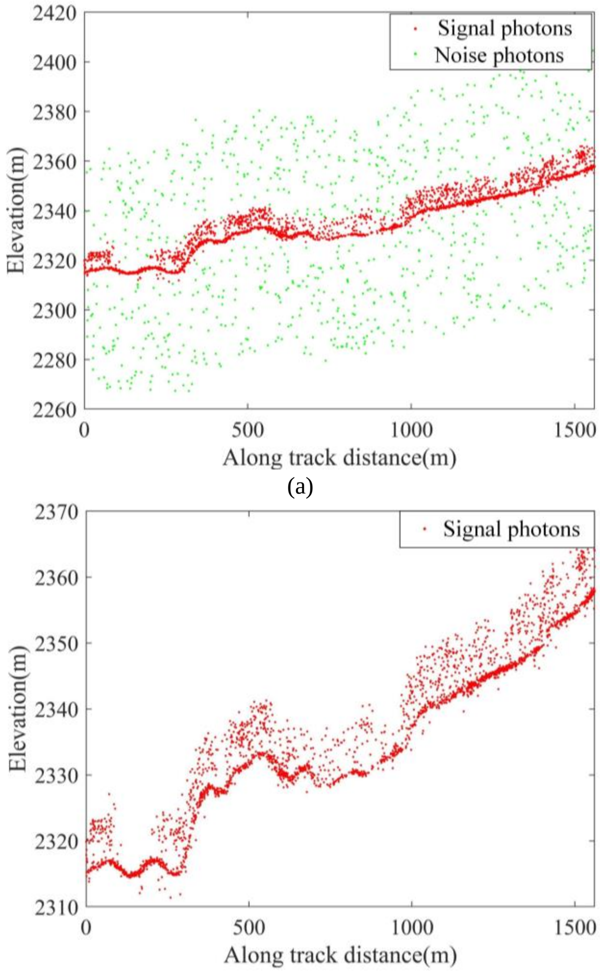


Figure 4. Adaptive density clustering based on the main direction of terrain slope. (a) and (b) are photons before and after noise removal using the adaptive density clustering method.

2.3 Removal of Residual Noise Based on the Local Outlier Factor (LOF) Algorithm

Following the above steps, most noise points have been successfully eliminated. However, despite these efforts, a few noise points may persist below ground level, in the air, or near canopy surfaces within some point clouds due to the application of a uniform global criterion for all data. To address these remaining noise points, this study employs the Local Outlier Factor (LOF) algorithm for further noise removal (Zou et al., 2017).

Initially, for each data point, its reachable distance from other points is computed and arranged from nearest to farthest. The K-nearest neighbors are identified for each data point, and their local reachability density is calculated. Subsequently, the LOF score is evaluated. A higher LOF value indicates increased abnormality while a lower value signifies normality. Finally, a threshold is established where values exceeding it are considered noise and those below it are deemed signals, ultimately leading to denoising results.

Figure 5(a) shows a small number of noise points remaining near the ground, in the air, and on the canopy surface in a portion of the point cloud. Figure 5(b) displays the signal points after removing the residual noise.

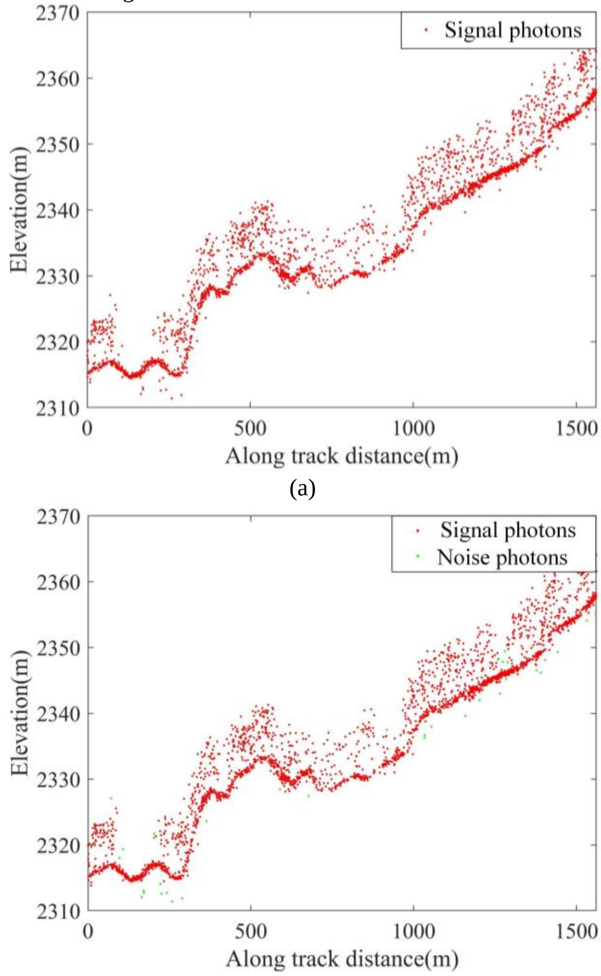


Figure 5. Residual noise removal using the LOF algorithm. (a) and (b) are the results before and after LOF denoising.

3. Experimental results and analysis

3.1 Experimental Datasets

This paper conducts experiments using eight sets of data from two distinct regions in the United States to test a denoising algorithm.Study Area 1, situated at 44.2°N to 45.0°N and 110°W to 111°W within Yellowstone National Park, features elevations ranging from 1847 to 2244 meters and an average canopy height of 14 meters. This area primarily consists of evergreen forests dominated by pine trees, with flat terrain and minimal topographic variations. On the other hand, Study Area 2, located at 35.5°N to 36.0°N and 83°W to 84°W in Great Smoky Mountains National Park, exhibits elevations spanning from 426 to 1978 meters and an average canopy height of 30 meters. The selected region is mainly composed of coniferous forests dominated by fir and hemlock trees, characterized by rugged terrain and significant topographic fluctuations.

3.2 Accuracy Metrics

To quantitatively evaluate the denoising algorithm's accuracy, this study employs four statistical metrics - Recall (R), Precision (P), Accuracy (A), and F1-score (F)- to assess the algorithm's denoising effectiveness as per Equations (5) to (8).

The evaluation criteria for denoising accuracy are based on Digital Terrain Model (DTM) and Digital Surface Model (DSM) data products derived from NEON airborne LiDAR data. Points falling within defined boundaries are classified as signal photons, while those outside are categorized as noise photons (Huang et al., 2022).

$$R = \frac{TP}{TP + FN} \tag{5}$$

$$P = \frac{TP}{TP + FP} \tag{6}$$

$$A = \frac{TP + TN}{TP + TN + FN + FP} \tag{7}$$

$$F = \frac{2P * R}{P + R} \tag{8}$$

where TP = the number of correctly detected signal photons
TN = the number of correctly detected noise photons
FP = the number of noise photons misclassified as signal photons
FN = the number of signal photons misclassified as noise photons

3.3 Experimental Results and Analysis

Three traditional denoising techniques have been selected for comparison: the classic DBSCAN-based denoising method (DBSCAN) (Liu et al., 2024), the approach based on local distance statistics (LDS) (Xia et al., 2014b), and the method of Differential, Regressive, and Gaussian Adaptive Nearest Neighbor filtering (DRAGANN) (Cao et al., 2020).

The DBSCAN-based technique identifies outliers by primarily utilizing density-reachability concepts to group closely positioned points and recognizes outliers as points located in low-density regions. The main idea behind DBSCAN revolves around two key 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 required within this neighborhood to classify an area as dense. Points with a limited number of neighboring points below the minpts threshold are classified as outliers.

The LDS technique identifies noisy photons based on histograms of local distance statistics. Initially, this technique calculates distances between each pair of photons. Then, it computes the sum of neighboring distances for n nearest points for each photon to create a frequency histogram. Generally, signal photons exhibit lower sum values while noisy ones show higher 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 from this frequency histogram, noisy photon points can be identified.

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

In Table 1, the statistical evaluation metrics for each algorithm in flat terrains are presented. While all algorithms can correctly identify noise photon points, the DBSCAN algorithm shows a slightly lower recall rate. Notably, although the LDS algorithm demonstrates higher recall and precision rates, its overall accuracy decreases significantly, leading to decreased F1-scores for all algorithms due to challenges posed by noise points closely resembling signal points within vegetation cover or ground surfaces.

	DBSCAN algorithm				LDS algorithm				DRAGANN algorithm				The proposed method			
	R	P	A	F	R	P	A	F	R	P	A	F	R	P	A	F
data1	0.982	0.873	0.958	0.923	0.983	0.908	0.969	0.944	0.992	0.941	0.981	0.966	0.984	0.978	0.989	0.981
data2	0.837	0.991	0.950	0.907	0.845	0.950	0.945	0.895	0.893	0.986	0.968	0.938	0.872	0.998	0.963	0.930
data3	0.935	0.979	0.963	0.957	0.914	0.997	0.960	0.954	0.953	0.906	0.942	0.929	0.942	0.991	0.971	0.966
data4	0.925	0.998	0.938	0.960	1.000	0.500	0.500	0.667	0.957	0.921	0.909	0.938	0.935	1.000	0.948	0.966
average	0.920	0.960	0.952	0.937	0.936	0.839	0.844	0.865	0.949	0.939	0.950	0.943	0.933	0.992	0.968	0.961

Table 1. Comparison of accuracy indicators for different denoising algorithms in flat areas

The proposed denoising algorithm achieves superior F1-scores in flat terrains compared to other methods, with an average F1-score of 96.1%, outperforming the DBSCAN-based denoising algorithm at 93.7%, the LDS algorithm at 86.5%, and the DRAGANN algorithm at 94.3%. This represents an average improvement of approximately 4.6% in F1-score.

Table 2 displays the statistical evaluation metrics for each algorithm in rugged terrains where F1-scores decrease compared to flat areas, particularly for the DRAGANN algorithm. However, our denoising algorithm excels with an average F1-score of 93.9%, surpassing the DBSCAN-based denoising algorithm at 88.4%, the LDS algorithm at 91.7%, and the DRAGANN algorithm at 72.9%. This signifies an average improvement of around 9.5% in F1-score, highlighting the superiority of our proposed denoising method in rugged terrains compared to traditional approaches.

	DBSCAN algorithm				LDS algorithm				DRAGANN algorithm				The proposed method			
	R	P	A	F	R	P	A	F	R	P	A	F	R	P	A	F
data5	0.891	0.945	0.949	0.917	0.834	0.999	0.940	0.909	0.882	0.522	0.835	0.656	0.883	0.991	0.958	0.934
data6	0.903	0.884	0.948	0.894	0.853	0.979	0.954	0.912	0.887	0.504	0.863	0.643	0.895	0.960	0.962	0.926

data7	0.965	0.909	0.974	0.936	0.843	0.998	0.961	0.914	0.948	0.725	0.934	0.822	0.938	0.968	0.980	0.953
data8	0.980	0.662	0.936	0.789	0.894	0.974	0.974	0.933	0.927	0.693	0.934	0.793	0.933	0.949	0.978	0.941
average	0.935	0.850	0.952	0.884	0.856	0.988	0.957	0.917	0.911	0.611	0.892	0.729	0.912	0.967	0.970	0.939

Table 2. Comparison of accuracy indicators for different denoising algorithms in rugged areas

Overall, based on the results presented in Tables 1 and 2, it can be concluded that our proposed method yields more accurate denoising outcomes than traditional methods in both flat and rugged terrains.

4. Conclusion

This paper introduces a spaceborne LiDAR point cloud noise removal algorithm based on adaptive terrain slope calculation using ICESat-2 data. By considering terrain slope influence and employing the Douglas-Peucker algorithm for segment merging based on similar slopes followed by clustering using elliptical neighborhood densities along dominant slope directions, this approach effectively distinguishes signal photons from noise photons. Experimental findings demonstrate an improved average F1-score by approximately 4.6% in flat areas and by about 9.5% in rugged terrains compared to traditional methods like LDS, DBSCAN-based algorithms, and DRAGANN algorithms. This novel algorithm showcases enhanced performance in removing noise photons near canopy levels and ground surfaces while improving signal point extraction capabilities across different terrains and conditions, affirming its potential for robust point cloud classification and parameter inversion applications.

Acknowledgement

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

References

Chen, B., Pang, Y., A, A.T., 2015. A denoising approach for detection of canopy and ground from ICESat-2's airborne simulator data in Maryland, USA. *Rroc Spie*. 9671, 96711S-96711S-5 .

Cao B.C., Fang Y., Jiang Z.Z., Gao L., Hu H.Y., 2020. Implementation and accuracy evaluation of ICESat-2 ATL08 denoising algorithms. *Bulletin of surveying and mapping*.(05): 25-30.

Hsu, H., Huang, C., Jasinski, M., Li, Y., Gao, H., Yamanokuchi, T., Wang, C., Chang, T., Ren, H., Kuo, C., Tseng, K., 2021. A semi-empirical scheme for bathymetric mapping in shallow water by ICESat-2 and Sentinel-2: A case study in the South China Sea. *Isprs J Photogramm*. 178, 1-19.

He G.H., Wang H., Fang Q., Zhang Y.A., Zhao D.L., Zhang Y.P., 2023. Spaceborne photon counting lidar point cloud denoising method with the adaptive mountain slope. *Journal of Infrared and Millimeter Waves*., 42(02): 250-259.

Huang, J., Xing, Y., Shuai, Y., Zhu, H., 2022. A Novel Noise Filtering Evaluation Criterion of ICESat-2 Signal Photon Data in Forest Environments. *Ieee Geosci Remote S*. 19, 1-5.

Liu, X.F., Wang, Z.H., Yang, W.Z., Chen, S.X., Wang, F.X., Chen, X.W., Xu, W.M., Shu, R., 2024. An Adaptive Signal Photon Detection Method Based on DBSCAN for Photon-Counting Laser Altimeter. *Ieee J-Stars*. 17, 3674-3686.

Lu, D., LI, D., ZHU, X., 2021. Denoising and classification of ICESat-2 photon point cloud based on convolutional neural network. *Journal of Geo-information Science*. 23(11), 2086-2095.

Magruder, L.A., Wharton, M.E.I., Stout, K.D., Neuenschwander, A.L., 2012. Noise filtering techniques for photon-counting LADAR data. *Proceedings of SPIE-The International Society for Optical Engineering*. 8379(2), 24.

Narine, L.L., Popescu, S., Zhou, T., Srinivasan, S., Harbeck, K., 2019. Mapping forest aboveground biomass with a simulated ICESat-2 vegetation canopy product and Landsat data. *Ann for Res*. 62(1), 69-86.

Vučetić, N., Petrović, S., Strunje, A., 2007. Simplification of Lines Containing Fixed Points Using the Douglas-Peucker Algorithm. *Kartografija I Geoinformacije*. 6(7 spec.), 74-81.

Wang, X., Pan, Z., Glennie, C., 2016. A Novel Noise Filtering Model for Photon-Counting Laser Altimeter Data. *Ieee Geosci Remote S*. 13(7), 947-951.

Xia S.B., Wang C., Xi X.H., Luo S.Z., Zeng H.C., 2014a. Point cloud filtering and tree height estimation using airborne experiment data of ICESat-2. *Journal of Remote Sensing*., 18(06): 1199-1207.

Yuan, C., Gong, P., Bai, Y., 2020. Performance Assessment of ICESat-2 Laser Altimeter Data for Water-Level Measurement over Lakes and Reservoirs in China. *Remote Sens-Basel*. 12(5), 770.

Zhang, J.S., Kerekes, J., 2014. An Adaptive Density-Based Model for Extracting Surface Returns From Photon-Counting Laser Altimeter Data. *Ieee Geosci Remote S*. 12(04), 726-730.

Zhu, X., Nie, S., Wang, C., Xi, X., Wang, J., Li, D., Zhou, H., 2021. A Noise Removal Algorithm Based on OPTICS for Photon-Counting LiDAR Data. *Ieee Geosci Remote S*. 18(8), 1471-1475.

Zhu, X., Wang, C., Xi, X., 2020. Research progress of ICESat-2/ATLAS data processing and applications. *Infrared and Laser Engineering*. 49(11), 76-85.

Zhu, X.X., Nie, S., Wang, C., Xi, X.H., Hu, Z.Y., 2018. A Ground Elevation and Vegetation Height Retrieval Algorithm Using Micro-Pulse Photon-Counting Lidar Data. *Remote Sens-Basel*. 10(12), 1962.

Zou, Y., ZHANG, X., SONG, S., 2017. Fast outlier detection algorithm based on local density. *Journal of Computer Applications*. 37(10), 2932-2937.