

## PHOTON-COUNTING LASER ALTIMETER DATA FILTERING BASED ON HIERARCHICAL ADAPTIVE FILTER FOR FOREST SCENARIO

Wang Xu<sup>1</sup>, Liang Xinlian<sup>1</sup> \*

<sup>1</sup> State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China-  
(wangxu0920, xinlian.liang) @whu.edu.cn

Commission III, WG III/1

**KEY WORDS:** Hierarchical Adaptive Filtering, ICESat-2, ATL03, Forest, Vegetation.

### ABSTRACT:

Photon-counting laser altimeter data has widely used in forest surveying and mapping with the successfully launch of ICESat-2. However, it is a challenge work that extract signal information from a large amount of noise photons, especially in complex scenario like forest. In this paper, we proposed the hierarchical adaptive filter method to extract signal photons and used this method to filtering ATL03 data in forest scenario. The hierarchical adaptive filter method follows a coarse-to-fine strategy to identify noise photons step by step. The most of noise are removed by filter based on feature of local distance, firstly. Then, the modified adaptive direction filter method is used to remove the noise photons near ground surface and vegetation. Finally, the filter based on continuity of topographic is used to remove the residual noise in the atmosphere. The airborne lidar data in experiment region was used to validate the effectiveness of method. The filter result overall recall is 97.51%, the overall precision is 98.58%, and the F-score is 98.04%. The result show that the hierarchical adaptive method we proposed can extract signal photons in background information effectively and preserve vegetation photons well.

### 1. INTRODUCTION

The forest canopy height is a critical parameter for estimating biomass and carbon stocks (Lefsky, 2010; Li et al., 2015), but to collect forest canopy in a large area is difficult using traditional method like filed and airborne surveying. In 2018, the Ice, Cloud and Land Elevation Satellite-2 (ICESat-2) which carried on Advanced Topographic Laser Altimeter System (ATLAS) has been launched and it can collect forest canopy height globally (Neuenschwander et al., 2020). The ATLAS provides data for estimating forest canopy height, biomass and carbon storage at large scales. Many scholars have used ATLAS data to study forest height estimation methods in many countries (Awadallah et al., 2013; Neuenschwander and Magruder, 2019).

Due to the high detection sensitivity of photon-counting lidar, ATLAS data are easily affected by environmental and instrument factors. Therefore, the first step to use photon-counting lidar data is to separate the signal photons from the background noise. Since the distribution of signal photons are much denser than noise photons in elevation profiles, previous studies proposed many methods for signal photon detection based on photon density. (Brunt et al., 2014) proposed a histogram-based method to detect the ice sheet surface photons. (Xia et al., 2014) proposed a filter method based on the feature of photons local distance and tested it used the Multiple Altimeter Beam Experimental Lidar (MABEL) data where surface covered with vegetation in American. It reported that this method can reach the overall accuracy of 97.6%. The idea of Density-Based Spatial Clustering of Applications with Noise (DBSCAN) method compared the number of points within a given radius for each point and threshold to cluster points into different classes (Ester et al., 1996). Due to that signal photons are distributed along the track, (Zhang et al., 2015) modified the DBSCAN method to use a horizontal ellipse as the search area.

However, due to the influence of terrain slope, the optimal direction of the ellipse search area is not equal to horizontal direction and the direction of ellipse search area is not considered. (Xie et al., 2017) proposed an adaptive directional filter method. The maximum density direction of each photon was considered as the optimal direction of ellipse and was determined by comparing the density in multiple directions. The experiment result reported that the accuracy of adaptive directional filter method was higher than the accuracy of modified DBSCAN method for terrain and vegetation. (Wang et al., 2016) proposed a filter method to derive the probability distribution function of signal and noise photons from k-nearest neighbours (KNN) and classified signal and noise photons based on Bayesian decision theory. (Herzfeld et al., 2017) proposed the density-dimension method that the threshold between density of signal and noise was determined adaptively according to the changing environment to consider the photon's density feature as an additional dimension. (Zhu et al., 2018) proposed a localized statistics-based algorithm that classify signal and noise based on photon density histogram. (Xie et al., 2021) compared seven methods described above, using four types of photon-counting lidar data and four scenarios of land, land ice, sea ice and ocean. (Yang et al., 2021) proposed a filter method using the backward elliptical distance (BED) in forest area. Signal and noise photons were identified by the feature of backward local density. In ICESat-2 algorithm theoretical basis document (ATBD), the differential, regressive, and gaussian adaptive nearest neighbour (DRAGANN) was described which was developed to identify signal photons and remove noise from the ATL03 data (Neuenschwander et al., 2021).

These methods can separate signal photons in background noise to a certain degree. However, since the density of vegetation photons is only slightly larger than that of noise photons,

\* Corresponding author

vegetation photons are easily misclassified as noise photons. This paper proposed an improved hierarchical adaptive filtering method in order to address this issue in forest scenario.

## 2. MATERIAL AND METHODOLOGY

### 2.1 ATL03 Dataset

The experiment region is near Truckee, California, USA. The Global Geolocated Photon Data (ATL03) is collected by ATLAS, which contains information about each photon, including longitude, latitude, ellipsoidal height and time et al. This paper uses ATL03 dataset as experiment data to test the hierarchical adaptive filter method and experiment dataset information is listed as Table 1.

Parameters	Values
Name	ATL03_20201210215731_11870906_005_01
Beam	3L
Time	December 10, 2020
Geolocation	Longitude: 120.232°W ~ 120.224°W Latitude: 39.394°N ~ 39.459°N

Table 1. Experimental dataset description.

### 2.2 ALS Dataset

The airborne laser scanning (ALS) dataset were obtained from the OpenTopography Project in the Preserving Mountains with Forest Management, CA 2020 (Graup, 2021), which is used as reference dataset to evaluate filter's effectiveness. The ALS dataset information is shown in Table 2. Considering possible terrain changes at difference in data collection time, the ATL03 and ALS data with the smallest difference in collection time are selected as far as possible. The collection time difference of two datasets is only one month, and the time is not in the growing period in experimental region, so the topographic changes caused by time difference can be ignored. The ATL03 data and ALS data are shown as Figure 1. The green track is ATL03 dataset and the red region is ALS dataset.

Parameters	Values
Name	Preserving Mountains with Forest Management, CA 2020
Platform	Airborne Platform
Time	November 10, 2020
Geolocation	Longitude: 120.197°W ~ 120.342°W Latitude: 39.390°N ~ 39.480°N
Point Density	20.98 pts/m <sup>2</sup>

Table 2. Reference dataset description.

### 2.3 Data Pre-processing

The noise photon distribution along the track is not uniform, which significantly influences the performance of the density-based filtering method. The photon is homogenized by distribution of the noise photons ratio along the track as proposed in (Wang et al., 2016). In addition, the signal photons on the left and right edges may be considered as noise because the density is lower than that of other regions. So, taking the elevation direction as the axis of symmetry, the photons at the

left and right edge are mirror rotated to supplement the photons at the edge.

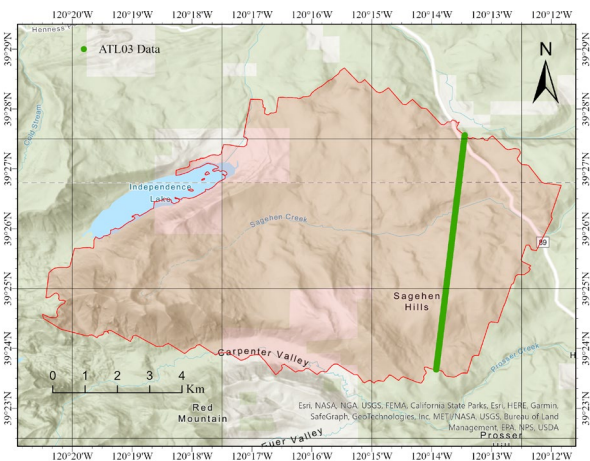


Figure 1. Dataset profile.

### 2.4 Hierarchical Adaptive Filter

This paper proposed a hierarchical adaptive filtering method for forest scenarios. The method first used local feature to remove most of the noise photons. Remaining near-surface noise and cloud and aerosol signal photons was removed by an adaptive density direction filter and an adaptive topographic continuity filter, respectively. The structure of the processing is shown as Figure 2(a) and the hierarchical adaptive filter is detailed in Figure 2(b).

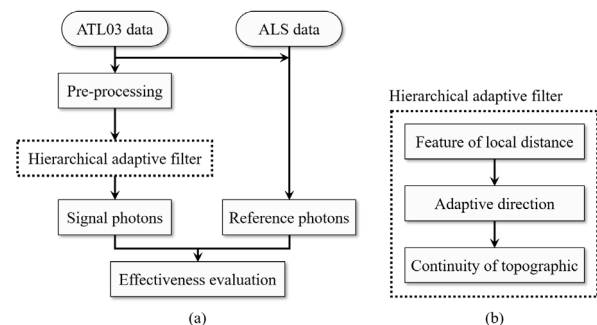


Figure 2. The workflow of this paper.

**2.4.1 Local Feature Filter:** Noisy photon points were first filtered. Since the signal photon is reflected from target objects and the noise is mostly random, the signal photon is supposed to have greater density than the noise photon. Therefore, the maximum distance of signal photon's K nearest neighbour is supposed to be smaller than that of the noise photons.

K is an important parameter that has huge influences on clustering effect. If the value of K is too small, multiple Gaussian distribution will be superimposed on the histogram. The threshold used to distinguish signal and noise will be difficult to selected. If the value of K is too large which eliminates the differences in the spatial distribution of noise and signal photons, there will be only one Gaussian distribution on the histogram (Xia et al., 2014). The selection of the K value can refer to the P value selection in the DRAGANN method (Neuenschwander et al., 2021). In this paper, 200 was selected as K value according to the feature of ATL03 dataset. The

statistics of the maximum distance of KNN was studied which is shown in Figure 3.

The Gaussian fitting was used to determine the threshold to identify signal photons. The signal and noise Gaussian distribution were fitted, respectively, as shown in Figure 3. Due to the spatial distribution of signal photons is much denser than noise photons, the signal and noise Gaussian distribution were on the left and right of the histogram, respectively. The intersection of signal Gaussian and noise Gaussian was selected to be the threshold  $T_1$ . Photons with maximum distance less than  $T_1$  can be roughly regarded as signal photons.

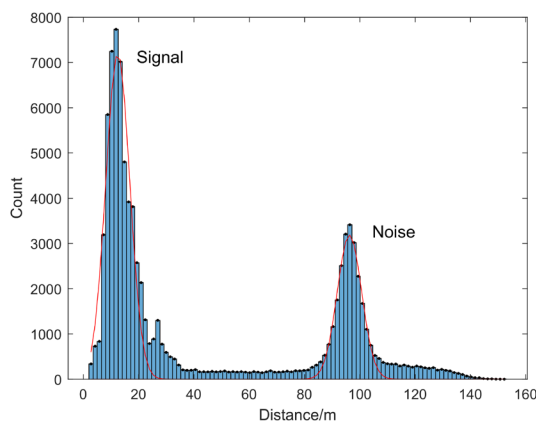


Figure 3. Histogram of the photons KNN distance.

**2.4.2 Optimal Density Direction Filter:** After filtering based on feature of local distance, most of noise photons were removed in ATL03 data. Then, filter based on optimal direction of density is used to filter noise photons near topographic surface.

In previous methods for determining the direction of the photon maximum density, the photon density was calculated at certain angular intervals and the optimal filter direction is the direction of maximum density (Xie et al., 2017; Zhu et al., 2018). However, this method requires a lot of computation and the maximum density direction calculated by this method is not accurate due to the density was calculated at certain angular intervals. Instead, we find the direction of the maximum density for each photon.

The main direction of the neighbourhood point distribution of each photon was evaluated by principal component analysis (PCA). The neighbourhood is circle which radius is determined by the local distance calculated in section 2.4.1. The optimal filter direction, that is, the direction of maximum density is shown in Figure 4. The red dot is coarse signal photons identified by section 2.4.1 and the arrow direction is optimal filter direction. It can be seen that the optimal filtering direction of ground photons is the same as the direction of the surface slope and the optimal filtering direction of vegetation photons is roughly equal to vegetation growth direction.

The neighbourhood was defined as an ellipse with major axis direction equal to the first principal direction calculated by PCA. In the ATL03 dataset, the semi-major axis  $A$  and the semi-minor axis  $B$  of the ellipse were empirical values, which were 10m and 1m, respectively. Counting the number of photons as photon density in the ellipse region for each photon. Then, the threshold  $T_2$  was the right boundary value of the leftmost peak

of the photon density distribution. Photons with density larger than  $T_2$  are signal photons. In this step, conservative threshold  $T_2$  was adopted in our method to prevent vegetation signal photons from being misidentified as noise photons.

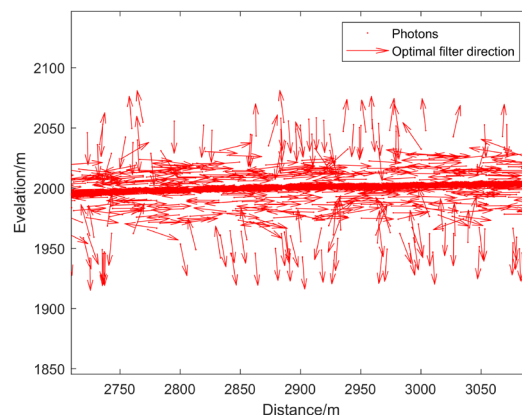


Figure 4. Optimal filter direction. A segment cut from the track.

**2.4.3 Topographic Continuity Filter:** For the purpose of terrain analyses, cloud and aerosol signal photons need to be removed. This paper proposed a method to use a combination of terrain continuity and Gaussian fitting to detect and remove aerosol signal photons. The photons were extracted with a fixed-length sliding window, and then the elevation distribution of these photons was counted. The elevation distribution was expected as Gaussian. The photons whose elevation exceeded three times standard deviation of Gaussian distribution were considered as cloud and aerosol photons and were removed.

## 2.5 Evaluation

Since both ATL03 data and airborne laser data are discrete, for each identified target photon, the reference data points in a neighbour of a photon were searched and the photon classified into ground, vegetation, and noise according to the number and type of reference data points were identified as true data. Then, the recall, precision and F-score were calculated to characterize the filtering method accuracy. Calculated equations of recall, precision and F-score are shown as equation (1).

$$R = \frac{TP}{TP + FN}; P = \frac{TP}{TP + FP}; F = \frac{2P \times R}{P + R} \quad (1)$$

where  $R$  = recall;  $P$  = precision;  $F$  = F-score;  
 $TP$  = Number of correctly identified signal photons  
 $FN$  = Number of signal photons identified as noise  
 $FP$  = Number of noise photons identified as signal

## 3. RESULT

### 3.1 Data Pre-processing

ATL03 photon data is shown as Figure 5. The horizontal axis represents the distance of photons along the track and the vertical axis represents the photons elevation above the WGS84 ellipsoid.

Due to the effect of environmental factors, the distribution of noise photons is not uniform. To deal this issue, the segment

lengths are scaled according to the along-track noise rate to homogenize the photon distribution (Wang et al., 2016).

The flag *bckgrd\_counts\_reduced* recorded in ATL03 dataset *gtxx/bckgrd\_atlas* is defined as number of photons counts in the 50-shot sum after subtracting the number of signal photon events (*gtxx* means ATLAS beam number, including *gt1l*, *gt1r*, *gt2l*, *gt2r*, *gt3l*, *gt3r*). It can be regarded as noise rate recorded at 500Hz along the track. The length of all the segments along the track were scaled by multiplying by the noise rate at the corresponding time and then scaling the track length as a whole back to the original length.

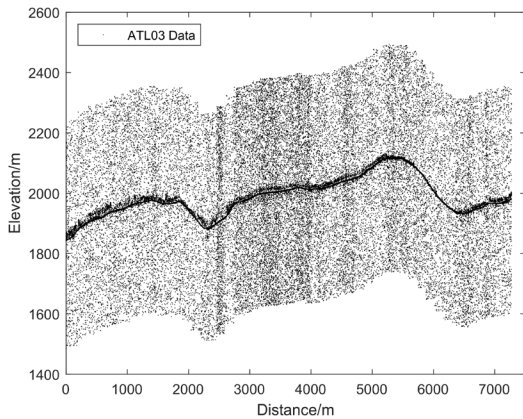


Figure 5. ATL03 photon data.

Then, the photons within 100m of the left and right edge are mirror rotated to supplement the photons at edge. This can prevent the signal photons at the edge from being mistaken for noise photons. The pre-processing result of ATL03 data is shown as Figure 6, where the distribution of photons is uniform.

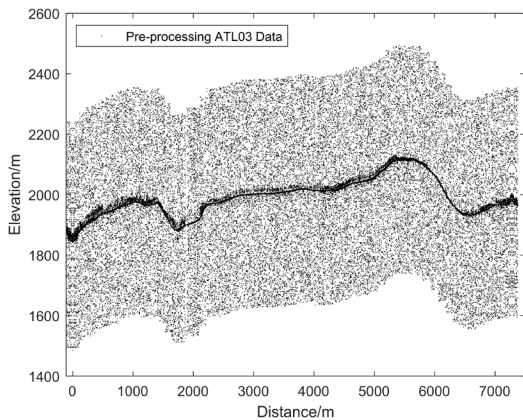


Figure 6. Pre-processing ATL03 photon data.

## 3.2 Hierarchical Adaptive Filter

**3.2.1 Filter based on feature of local distance:** The result of the local distance filtering is shown in Figure 7. The red dot is coarse signal photons and the black dot is noise photons. It can be seen that most of noise photons is removed located above and below the signal photons. The signal photons including topographic surface and vegetation are preserved well. However, the noise photons near the surface still exist. This part of noise is similar to signal photons.

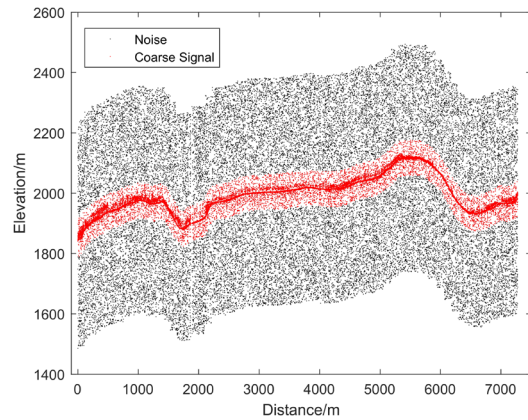


Figure 7. Filter result based on local distance feature.

**3.2.2 Filter based on optimal direction of density:** The result of adaptive direction filter is shown as Figure 8. In order to better show the effectiveness of filtering, only the coarse signal photons identified in section 3.2.1 are shown and signal photons identified by filter based on adaptive direction method are magnified in the figure. The red dot is signal photons identified by this step and the black dot is noise photons. The conservative threshold  $T_2$  was adopted, and this is the reason that vegetation signal photons are well preserved. However, a problem posed by conservative threshold  $T_2$  is obvious that some noise photons in the atmosphere are not identified. This part of photons may be signal from clouds and aerosols and will be removed in the next step.

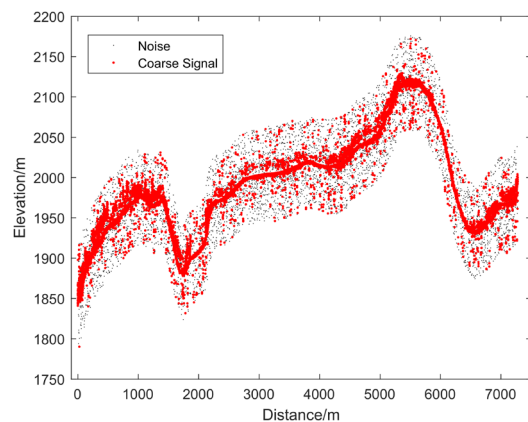
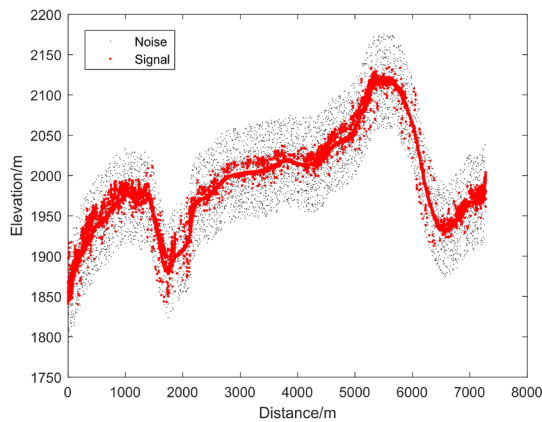


Figure 8. Filter result based on adaptive direction.

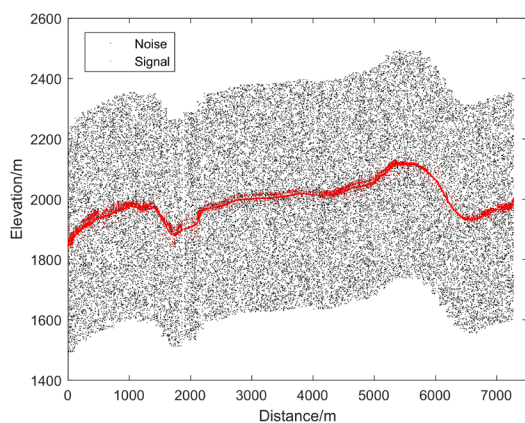
**3.2.3 Filter based on continuity of topographic:** In this experiment, fixed-length sliding window was set to 200m and the sliding step value was set to 50m. The result of filter based on continuity of topographic is shown in Figure 9. The red dot is signal photons identified in this step and the black dot is noise photons.



**Figure 9.** Filter result based on topographic continuity.

It can be seen that the noise photons cannot identified in section 3.2.2 are removed by filter based on continuity of topographic.

The result of hierarchical adaptive filter method is shown in Figure 10. The signal photons are identified from background information and vegetation signal photons were well preserved.

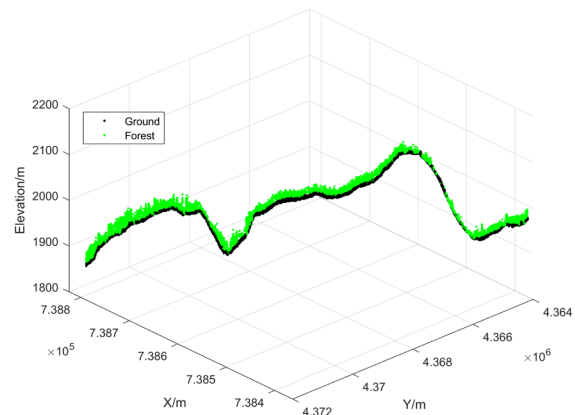


**Figure 10.** Result of hierarchical adaptive filter.

### 3.3 Filter Effectiveness Evaluation

A quantitative evaluation of effectiveness for the filtering method is reported. The extract method of reference signal photons as illustrated in section 2.5 was used. There is reference data extracted shown as Figure 11. The black dots represent ground and the green dots represent vegetation.

The experiment result of hierarchical adaptive filter method was compared with reference data. The overall recall, precision and F-score were calculated to evaluate filter effectiveness. Since photons were not classified into ground and forest, only recall of ground and forest can be calculated, as shown in Table 3.



**Figure 11.** Reference signal photons data.

Indicators	Values (%)		
	Overall	Ground	Forest
Recall	97.51	99.61	88.21
Precision	98.58	\	\
F-score	98.04	\	\

**Table 3.** Filtering indicators of ATL03 data.

## 4. DISCUSSION AND CONCLUSION

In this paper, the hierarchical adaptive filter method is proposed to identify the signal photons from background information in spaceborne photon-counting lidar data. The hierarchical adaptive filter composed of the local feature filter, optimal density direction filter and topographic continuity filter follows a coarse-to-fine strategy to identify noise photons. The local feature filter removed most of noise photons which reduces the computation for subsequent steps. The optimal density direction filter finds the direction of maximum density accurately and adopted the conservative threshold that can preserve vegetation signal photons. The topographic continuity filter removed cloud and aerosol photons to improve the precision of identification of ground and vegetation signals. The hierarchical adaptive filter was designed to identify the signal photons especially vegetation photons with density only slightly larger than that of noise photons from background noise. This method can prevent vegetation photons from being misidentified as noise.

The ATL03 data near Truckee, California, USA which cover with vegetation was selected to test the method. We used ALS data collected at similar time as reference data to evaluate method effectiveness. The filter result overall recall is 97.51%, the overall precision is 98.58%, and the F-score is 98.04%. The result showed that the hierarchical adaptive filter method can identify signal photons from background noise while preserving the vegetation photons on the ground effectively.

## REFERENCES

Awadallah, M., Abbott, A. L., Thomas, V., et al., 2013: Estimating Forest Canopy Height and Biophysical Parameters Using Photon-counting Laser Altimetry. *Proceedings: 13th International Conference on LiDAR Applications for Assessing Forest Ecosystems (SilviLaser 2013)*, Beijing, China, Oct. 2013,

pp. 137-144.

Brunt, K.M., Neumann, T.A., Walsh, K.M., et al., 2014: Determination of Local Slope on the Greenland Ice Sheet Using a Multibeam Photon-Counting Lidar in Preparation for the ICESat-2 Mission. *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 5, pp. 935-939

Ester, M., Kriegel, H.P., Sander, J., et al., 1996: A Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. in *Proc. 2nd Int. Conf. Knowledge Discovery and Data Mining (KDD'96)*, pp. 226–231.

Graup, L., 2021: Preserving Mountains with Forest Management, CA 2020. National Center for Airborne Laser Mapping (NCALM). Distributed by OpenTopography. <https://doi.org/10.5069/G96H4FMX>. Accessed: 2022-01-16.

Herzfeld, U.C., Trantow, T.M., Harding, D., et al. 2017: Surface-Height Determination of Crevassed Glaciers—Mathematical Principles of an Auto-adaptive Density-Dimension Algorithm and Validation Using ICESat-2 Simulator (SIMPL) Data. *IEEE Transactions on Geoscience and Remote Sensing*, 55(4):1874-1896.

Li, L., Guo, Q., Tao, S., et al., 2015: Lidar with Multi-temporal MODIS Provide a Means to Upscale Predictions of Forest Biomass. *ISPRS Journal of Photogrammetry and Remote Sensing*, 102: 198-208.

Neuenschwander, A., Guenther, E., White, J.C., et al., 2020: Validation of ICESat-2 Terrain and Canopy Heights in Boreal Forests. *Remote Sensing of Environment*, 251(112110).

Neuenschwander, A., Magruder L., 2019: Canopy and Terrain Height Retrievals with ICESat-2: A First Look. *Remote Sensing*, Jan;11(14):1721.

Neuenschwander, A., Popescu, S., Nelson, R., et al., 2021: Ice, Cloud, and Land Elevation Satellite 2 (ICESat-2) Algorithm Theoretical Basis Document (ATBD) for Land - Vegetation Along-Track Products (ATL08) Release 005, National Aeronautics and Space Administration. Goddard Space Flight Centre.

Wang, X., Pan, Z.G., Glennie, C., 2016: A Novel Noise Filtering Model for Photon-Counting Laser Altimeter Data. *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 7, pp. 947-951.

Xia, S.B., Wang, C., Xi, X.H., et al., 2014: Point Cloud Filtering and Tree Height Estimation Using Airborne Experiment Data of ICESat-2. *Journal of Remote Sensing*, 18(4): 1199-1207.

Xie, F., Yang, G., Shu, R., et al., 2017: An Adaptive Directional Filter for Photon Counting Lidar Point Cloud Data. *Journal of Infrared and Millimeter Waves*, 36(1):7.

Xie, H., Xu, Q., Ye, D., et al., 2021: A Comparison and Review of Surface Detection Methods Using MBL, MABEL, and ICESat-2 Photon-counting Laser Altimetry Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 7604-7623.

Yang, P., Fu, H., Zhu, J., et al., 2021: An Elliptical Distance Based Photon Point Cloud Filtering Method in Forest

Area. *IEEE Geoscience and Remote Sensing Letters*, 19, 1-5.

Zhang, J., Kerekes, J., Csatho, B., et al., 2015: An Adaptive Density-based Model for Extracting Surface Returns from Photon-counting Laser Altimeter Data. *IEEE Geoscience and Remote Sensing Letters*, 12: 726-730.

Zhu, X., Nie, S., Wang, C., et al., 2018: A Ground Elevation and Vegetation Height Retrieval Algorithm Using Micro-pulse Photon-counting Lidar Data. *Remote Sensing*, 10(12): 1962.