ICESat-2 data classification and canopy height validation--a case study in the northern region of China

Li He[1,3], Yong Pang[2,3,*], Zhongjun Zhang[1], Xiaojun Liang[2,3], Bowei Chen[4]

[1]School of Artificial Intelligence, Beijing Normal University, Beijing 100875, China
Email: allen_he1996@outlook.com; zzj@bnu.edu.cn

[2]Institute of Forest Resource Information Technique, Chinese Academy of Forestry, Beijing 100091, China
[3]Key Laboratory of Forestry Remote Sensing and Information System, National Forestry and Grassland, Beijing 100091, China
Email: pangy@ifrit.ac.cn; Stanfordlxj@163.com

[4]Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China
Email: rs.cbw@foxmail.com

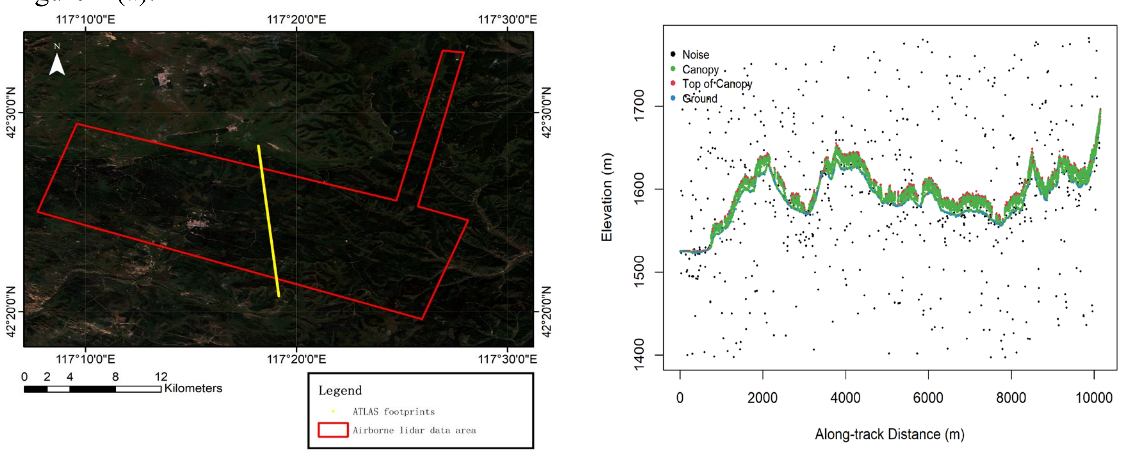
1. Introduction

NASA's ICESat-2 was launched in the fall of 2018 and is collecting a huge amount of data globally. One of ICESat-2 Level-2 products, ATL03, provides time, latitude, longitude, and ellipsoidal height for each photon downlinked from Advanced Topographic Laser Altimeter System (ATLAS) (Neumann et al., 2021). However, noise could seriously affect the observation results, accurate classification of photons is the basis of subsequent study. Some photon classification algorithms have been proposed in the past. Zhang's algorithm based on DBSCAN (2014) and Zhu's algorithm based on OPTICS (2021) use density clustering to distinguish photons. Their methods are affected by parameters and need to be verified in large-scale study area. ATL08, one of ICESat-2 products, is using the Differential, Regressive, and Gaussian Adaptive Nearest Neighbor (DRAGANN) filtering technique by adaptive neighborhood search to identify and remove noise photons (Neuenschwander et al., 2021). According to the distribution features of photons, Chen modified Local Outlier Factor (LOF) algorithm by defining the ellipse search area to filter the noise (Chen et al., 2019a).

Based on Chen's research, we proposed some methods to adapt the algorithm for large scale ATLAS data processing. Using ATL03 data in our study area, we classified photons by our improved algorithm. Then we used signal photons to predict heights of canopy. Results show that our improved algorithm could effectively remove noise photons.

2. Data and Methods

Our study area is located in the Saihanba Forest Farm, Hebei Province, north of China. We used the ATL03 data acquired at the night of June, 2019 by strong beam in this area as the original data for photon classification. The data of airborne lidar was used as reference data (Pang et al., 2021), and the acquisition time of airborne lidar data was the summer of 2018. The location of our study area is shown in Figure 1 (a).



(a) The location of our study area. (b) The classification result of our algorithm.

Figure 1: The location of our study area and the classification result of our algorithm.

Published in: Markus Hollaus, Norbert Pfeifer (Eds.): Proceedings of the SilviLaser Conference 2021, Vienna, Austria, 28–30 September 2021. Technische Universität Wien, 2021. DOI: 10.34726/wim.1861
This paper was peer-reviewed. DOI of this paper: 10.34726/wim.1969

Firstly, based on LOF with Ellipse Searching Area, our methods proposed to adapt the algorithm for large scale ATLAS data processing are as followed:

(1) Coarse Denoising by LOF with Horizontal Ellipse Searching Area. After the signal interval is determined by the elevation histogram, coarse denoising is carried out. The purpose of this method is to use the algorithm to find ground photons, so as to filter out noise photons far away from ground surface. Because the density of ground photons is usually larger than others, and ground photons is always distributed in the lower part of signal photons, it is not hard to extract coarse ground photons by LOF algorithm. Then we can get coarse terrain in the area. In order to avoid local dense photons affecting the extraction of ground photons, the number of domain members can be selected larger in LOF algorithm.

(2) Accurate Denoising by LOF with Rotating Ellipse Searching Area. The purpose of this method is to reduce the influence of terrain slope on LOF algorithm by ellipse searching area adapted to terrain. Photons after coarse denoising are divided into several intervals along the track. We calculate the terrain slope by coarse ground photons in the interval. In each interval, rotating angle of ellipse searching area is equal to the terrain slope. Then, we calculate LOF scores for photons in the interval to classify signal photons and noise. In order to avoid the influence of boundaries on LOF, we set buffers on the left and right sides. Finally, according to the spatial distribution, the signal photons are classified as Top of Canopy photons, Canopy photons and Ground photons.

Secondly, we used airborne lidar data as the reference to analyze accuracy of our algorithm in different scales. We generated DTM and CHM with 1m resolution by airborne data, then extracted corresponding region by UTM coordinates of photons. The length of this region is the same as that of track about 10000 m, and the width is 17m, which is similar to the diameter of footprints. In this region, we calculated mean terrain height per meter along the track in DTM as the reference terrain height, and calculated the canopy heights of ATL data. CHM was used to compute canopy heights from airborne lidar data. Referring to the work by Neuenschwander et al. (2020), we regarded the 98^{th} percentile of the reference heights (RH98) as evaluation metric, and set up nine different scales ranging 20 m to 100 m. ATL_RH98 means 98^{th} percentile of signal photon heights, and ALS_RH98 means 98^{th} percentile of all return heights.

3. Results and Discussion

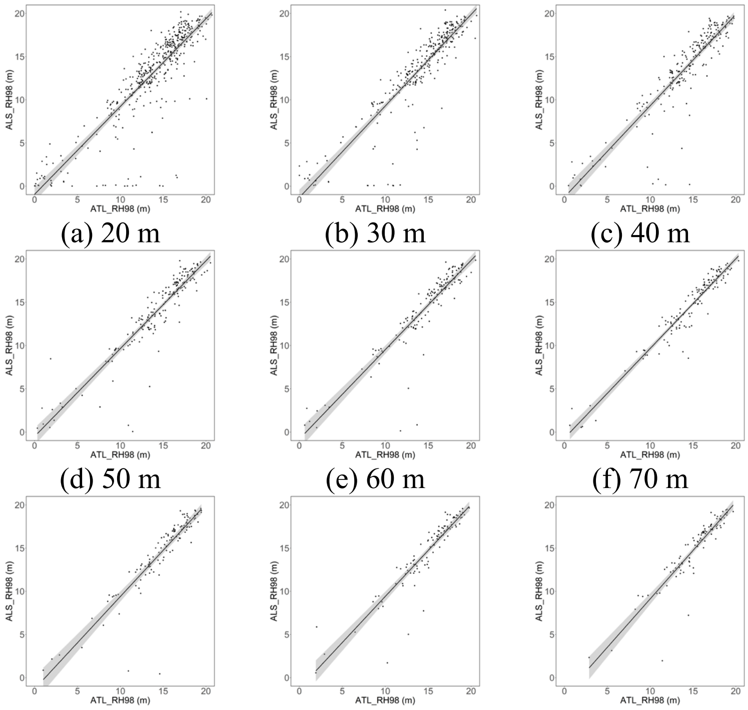
According to the methods mentioned above, the result of classification of our algorithm is shown in Figure 1 (b). We could find that most signal photons can be classified correctly, and our method can adapt to terrain changes and accurately extract signal photons in steep terrain areas. A small number of signal photons close to the ground (higher than the ground) would be misclassified as noise due to their lower local density.

The results of canopy heights comparison are shown in Table 1 and Figure 2. The results show that canopy heights calculated from classified photons have good consistency with airborne lidar data, and we get the minimal Root Mean Squared Error and the maximal R-square at 70 m. When research scale goes beyond 50 m, canopy heights from classified photons are in better consistency with ALS data. At the same scale, canopy heights have smaller RMSE, which shows that our improved algorithm can effectively remove part of noise photons which are difficult to be removed by the original algorithm in Chen's research (Chen et al., 2019b).

The goal of this study is to verify the accuracy of photon classification algorithm by canopy heights. The following two reasons might contribute to those points with large errors of canopy height. Firstly, the noise with large local density are hard to be filtered by our methods and are misclassified as signal photons, resulting in the errors of canopy height. Secondly, the ATL03 data contains photons from land buildings, while the ALS data filters the data of land buildings. This difference will lead to errors. In future study, we will further improve our methods and reduce the errors caused by inconsistent data.

Table 1. The evaluation in the RH98 between our algorithm and airborne lidar data for different scales.

Scales	20 m	30 m	40 m	50 m	60 m	70 m	80 m	90 m	100 m
R^2	0.74	0.80	0.80	0.84	0.83	0.92	0.81	0.85	0.84
RMSE (m)	2.76	2.30	2.22	1.84	1.88	1.18	1.89	1.65	1.54



(a) 20 m (b) 30 m (c) 40 m

(d) 50 m (e) 60 m (f) 70 m

(g) 80 m (h) 90 m (i) 100 m

Figure 2: The comparison results in the RH98 between our algorithm and airborne lidar data for different scales.

4. Conclusions

In this study, we made some improvements to adapt the algorithm for large scale ATLAS data processing. We classified photons as Noise, Top of Canopy photons, Canopy photons and Ground photons in ATL03 data and by comparing canopy heights calculated by classified photons with that calculated by ALS data, the results indicate that our method can effectively separate signal photons and noise in different terrain.

Acknowledgements

This study was funded by the Natural Science Foundation of China (41871278) and National Key Research and Development Program of China (2017YFD0600404).

References

Chen B W, Pang Y, Li Z Y, *et al*, 2019a. Ground and Top of Canopy Extraction From Photon-Counting LiDAR Data Using Local Outlier Factor With Ellipse Searching Area. *IEEE Geoscience and Remote Sensing Letters,* 16(9):1447-1451.

Chen B W, Pang Y, Li Z Y, *et al*, 2019b. Potential of Forest Parameter Estimation Using Metrics from Photon Counting LiDAR Data in Howland Research Forest. *Remote Sensing*, 11(7):856-877.

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 (2020) 112110.

Neuenschwander A, Pitts K, Jelley B, *et al*, 2021. Ice, Cloud, and Land Elevation 1 Satellite 2 (ICESat-2) Algorithm Theoretical Basis Document (ATBD) for Land-Vegetation Along-Track Products (ATL08). National Aeronautics and Space Administration, Goddard Space Flight Center.

Neumann T, Brenner A, Hancock D, *et al*, 2021. Ice, Cloud, and Land Elevation Satellite–2 (ICESat-2) Project: Algorithm Theoretical Basis Document (ATBD) for Global Geolocated Photons (ATL03). National Aeronautics and Space Administration, Goddard Space Flight Center.

Pang Y, Liang X J, Jia W, *et al*, 2021. The comprehensive airborne remote sensing experiment in Saihanba forest farm. *National Remote Sensing Bulletin*, 25(4):904-917.

Zhang J, Kerekes J, Csatho B, *et al,* 2014. A clustering approach for detection of ground in micropulse photon-counting lidar altimeter data. *2014 IEEE Geoscience and Remote Sensing Symposium.*

Zhu X X, Nie S, Wang C, *et al,* 2021. A noise removal algorithm based on OPTICS for photon-counting LiDAR data. *IEEE Geoscience and Remote Sensing Letters,* 18(8):1471-1475.