Remote Sensing of Environment 208 (2018) 154–170

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Photon counting LiDAR: An adaptive ground and canopy height retrieval algorithm for ICESat-2 data

S.C. Popescu[a], T. Zhou[a,*], R. Nelson[b], A. Neuenschwander[c], R. Sheridan[a], L. Narine[a], K.M. Walsh[d]

[a] LiDAR Applications for the Study of Ecosystems with Remote Sensing (LASERS) Laboratory, Department of Ecosystem Science and Management, Texas A &M University, College Station, TX 77450, USA
[b] Biospheric Sciences Lab, Code 618, NASA-GSFC, Greenbelt, MD 20771, USA
[c] Geospatial Laser Applications and Measurements, Applied Research Laboratories, University of Texas at Austin, 10000 Burnet Rd, Austin, TX 78758, USA
[d] NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

ARTICLE INFO

Keywords:
ICESat-2
Photon classification
Photon counting LiDAR
ATLAS
MABEL
Canopy height
Terrain elevation

ABSTRACT

The upcoming Ice, Cloud and Land Elevation Satellite-2 (ICESat-2) mission will offer prospects for mapping and monitoring biomass and carbon of terrestrial ecosystems over large areas using photon counting LiDAR data. In this paper, we aim to develop a methodology to derive terrain elevation and vegetation canopy height from test-bed sensor data and further pre-validate the capacity of the mission to meet its science objectives for the ecosystem community. We investigated a novel methodological framework with two essential steps for characterizing terrain and canopy height using Multiple Altimeter Beam Experimental LiDAR (MABEL) data and simulated ICESat-2 data with various vegetation conditions. Our algorithm first implements a multi-level noise filtering approach to minimize noise photons and subsequently classifies the remaining photons into ground and top of canopy using an overlapping moving window method and cubic spline interpolation. Results of noise filtering show that the design of the multi-level filtering process is effective to identify background noise and preserve signal photons in the raw data. Moreover, calibration results using MABEL and simulated ICESat-2 data share similar trends with the retrieved terrain being more accurate than the retrieved canopy height, and the nighttime results being better than corresponding daytime results. Compared to the results of simulated ICESat-2 data, MABEL data achieve lower accuracy for ground and canopy heights in terms of root mean square error (RMSE), which may partly result from the inconsistency between MABEL and reference data. Specifically, simulated ICESat-2 data using 115 various nighttime and daytime scenarios, yield average RMSE values of 1.83 m and 2.80 m for estimated ground elevation, and 2.70 m and 3.59 m for estimated canopy height. Additionally, the accuracy assessment of percentile heights of simulated ICESat-2 data further substantiates the robustness of the methodology from different perspectives. The methodology developed in this study illustrates plausible ways of processing the data that are structurally similar to expected ICESat-2 data and holds the potential to be a benchmark for further method adjustment once genuine ICESat-2 are available.

1. Introduction

Light Detection and Ranging (LiDAR) data such as discrete-return (DR) and full waveform (FW) LiDAR data are increasingly used to characterize the earth's topography, quantify vegetation structure and provide insightful solutions to natural resource inventory and carbon budget characterization (Allouis et al., 2013; Lefsky et al., 2005; Neigh et al., 2013; Popescu, 2007; Zhou et al., 2017; Zolkos et al., 2013). However, the utility of small-footprint LiDAR data over large spatial scales to accurately monitor forest ecosystems remains largely impractical due to their high acquisition cost (Gwenzi et al., 2016; Swatantran et al., 2016) and the limited spatial coverage caused by low operation altitude and high requirements of pulse energy (McGill et al., 2013). The advent of emerging technologies, such as photon-counting LiDAR (PCL), has offered prospects for future spaceborne laser altimeters. In contrast to analog LiDAR, the PCL is unique in that it employs low energy expenditure, increased measurement sensitivity, high repetition rate and space operational altitude. These properties enable PCL to overcome the restriction of spacecraft prime power by generating dense along-track sampling (Zhang and Kerekes, 2014) and ultimately resulting in the large spatial coverage (Wulder et al., 2012).

Due to these advantages of PCL systems, the Advanced Topographic Laser Altimeter System (ATLAS) sensor will be deployed on the upcoming Ice, Cloud and Land Elevation Satellite-2 (ICESat-2) (Markus

* Corresponding author.
E-mail address: tankwin0@tamu.edu (T. Zhou).

https://doi.org/10.1016/j.rse.2018.02.019
Received 12 September 2017; Received in revised form 12 February 2018; Accepted 13 February 2018

et al., 2017). There are two notable features of ATLAS: (1) a multi-beam system that consists of six individual beams (split from a transmitted laser pulse by a diffractive optical element) with three pairs along the track designed to meet the science requirements of detecting the spatial variability of ice surface and monitoring ice dynamics (Herzfeld et al., 2014); for each pair, a weak and strong beam have an energy ratio of approximately 1:4 to compensate for varying surface reflectance; and (2) the micro-pulse photon-counting technology that is capable of efficiently detecting photons reflected back from the earth surface. The ATLAS design allows for dense along-track sampling and large spatial coverage with low energy requirements at high flying altitude (Swatantran et al., 2016). For example, these configurations will generate overlapping footprints on the Earth surface with a diameter of 14 m, spaced at 0.7 m along track. By comparison, the Geoscience Laser Altimeter System (GLAS) aboard the Ice, Cloud and land Elevation Satellite (ICESat) illuminated spots (footprints) of 70 m in diameter, spaced at 170 m intervals. Compared to GLAS, the data to be provided by ATLAS consist of individual geo-located photons with profile configuration instead of waveforms (Markus et al., 2017). In addition, the ATLAS instrument will only operate at one single pulse (532 nm) with 10 kHz laser repetition rate. The dense sampling and extensive spatial coverage will be beneficial to large-scale applications such as sea level change monitoring, forest structural mapping and biomass estimation, improved estimation of Global Digital Terrain Models (GDTM), and reducing uncertainties associated with estimated forest biomass and carbon. Moreover, ICESat-2 will facilitate the production of gridded global products after the three-year mission anticipated lifespan (Neuenschwander and Magruder, 2016) with the potential for increased synergy with other existing remote sensing images, such as Landsat, to further complement ongoing biomass and vegetation mapping efforts.

Ambient noise is generated along the real signal photons since solar background photons can be simultaneously received by the detector. Consequently, an individual photon can be reflected back from targets within the footprint, but the exact corresponding origin location will be unknown (Gwenzi et al., 2016). Much of the noise can be avoided with the nighttime operation of the PCL system when there is less solar background noise. Furthermore, fewer signal photons are expected to reflect off from the vegetation than ice surfaces due to lower reflectance and higher aerosol densities over vegetated areas (Herzfeld et al., 2014).

The test-bed sensors for the upcoming ICESat-2 mission such as the Slope Imaging Multi-Polarization Photon-Counting Lidar (SIMPL) (Dabney et al., 2010) and Multiple Altimeter Beam Experimental Laser (MABEL) (Rosette et al., 2011) have been developed to inform scientists about the potential of future spaceborne laser altimeters to meet various science objectives. Recent studies have demonstrated promising prospects of utilizing these data to discriminate ice and water (Kwok et al., 2016), to retrieve 3-D vegetation structural attributes in a savanna ecosystem (Gwenzi et al., 2016), and to estimate vegetation cover and biomass in conjunction with Landsat 8 data in a dryland ecosystem (Glenn et al., 2016).

A critical task for the ecosystem community is to identify the ground and canopy surface from these photons to meet science objective of determining global canopy heights which hinges upon the ability to detect both the canopy surface and the underlying topography (Neuenschwander and Magruder, 2016). According to the ICESat-2 Algorithm Theoretical Basis Document (ATBD) for land and vegetation, the science objective is to discriminate signal and background photons and generate topography and canopy height products with decent accuracy for biomass estimation. More specifically, the accuracy of ATLAS data for topography and canopy height is expected to be superior to other existing global height products such as the Shuttle Radar Topography Mission (SRTM) whose accuracy ranges from 5 to 10 m, depending upon the amount of topography and vegetation cover over a particular area. Generally, there are two major steps to derive terrain and canopy height from PCL data: (1) noise filtering of raw photons, and (2) canopy and terrain classification of possible signal photons. The performance of noise filtering, on which canopy and terrain classifications depend on, may be of greater significance. A few methods have been developed for noise filtering, such as histogram-based filtering algorithms (Gwenzi et al., 2016; Moussavi et al., 2014), the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) approach (Zhang and Kerekes, 2015) and the Bayesian approach (Wang et al., 2016). Prior to these studies, Magruder et al. (2012) proposed three filtering techniques including canny edge detection, probability distribution and local angle mapping to process MABEL data. Tang et al. (2016) developed a voxel-based spatial filtering method to generate noise-free dataset using the data from the High-Resolution Quantum Lidar System. All of them have been proven effective to some extent while also suffering from some concerns. For example, the histogram-based and DBSCAN methods are prone to error over complex terrain and can potentially lose useful information for subsequent signal and noise photons classification (Tang et al., 2016; Wang et al., 2016). The voxel-based spatial filtering method is mainly oriented for the point cloud dataset with high signal-to-noise ratio (SNR), which may be not suitable for the profile data with low SNR of ICESat-2 data. In addition, these methods were mainly tested on a limited number of MABEL data with accuracy assessment conducted with co-registered reference data (Gwenzi et al., 2016). Undoubtedly, an additional co-registration processes or adjustment will complicate the validation processes and performance evaluation. Furthermore, the efficiency of algorithms for retrieving the canopy height using PCL data over different forest types and noise levels has not been adequately explored.

The overall goal of this paper is to develop a methodological framework to filter noise, classify photons into noise, ground, canopy and top of canopy (TOC) photons and retrieve the terrain and canopy height in various vegetation conditions and noise levels using PCL data such as MABEL and simulated ICESat-2 data, both of which share similar characteristics of the expected data from the ICESat-2 mission. The study is anticipated to advance understanding of ATLAS data, provide insights into the challenges to be expected from data processing and pre-validate the upcoming ICESat-2 mission. The innovative aspects of this study consist of (1) introducing an adaptive methodology to cluster the signal photons and refine parameters for ground and TOC interpolation over diverse vegetation conditions, and (2) building a framework to conduct noise filtering suitable for different possible data scenarios of the upcoming ICESat-2 mission which could render a valuable basis for processing genuine ATLAS data. Ultimately, the methodology and results of this study may potentially allow a more rapid adoption of ICESat-2 data once available by a large scientific community, for a range of ecosystem studies through the probable incorporation of existing remote sensing data.

2. Methods and materials

2.1. Study sites

To prepare the automatic ground and TOC detection algorithms over vegetation areas for the ICESat-2 mission, two test-bed sensor data (MABEL and simulated ICESat-2 data) with multiple vegetation conditions were investigated. For MABEL data, we used two data sets acquired on September 14, 2012, near Hinsdale and Chester in Vermont (VT), and one data set collected on September 21, 2012, near Jacksonville in North Carolina (NC), shown in Fig. 1. The vegetation in Vermont is mainly comprised of sugar maple (*Acer saccharum*), American beech (*Fagus grandifolia*) and hemlock (*Tsuga canadensis*), which belong to the eastern mixed forest type. In contrast, the study site near Jacksonville is mostly deciduous forest covered with Black Willow (*Salix nigra*), Loblolly Pine (*Pinus taeda*) and Eastern Cottonwood (*Populus deltoides*).

Simulated ICESat-2 data were generated from DR LiDAR or FW LiDAR data using the simulator provided by the ICESat-2's NASA