

Algorithm for Detection of Ground and Canopy Cover in Micropulse Photon-Counting Lidar Altimeter Data in Preparation for the ICESat-2 Mission

Ute Christina Herzfeld, Brian W. McDonald, Bruce F. Wallin, Thomas A. Neumann, Thorsten Markus, Anita Brenner, and Christopher Field

Abstract—NASA’s Ice, Cloud and Land Elevation Satellite-II (ICESat-2) mission is a decadal survey mission (2016 launch). The mission objectives are to measure land ice elevation, sea ice freeboard, and changes in these variables, as well as to collect measurements over vegetation to facilitate canopy height determination. Two innovative components will characterize the ICESat-2 lidar: 1) collection of elevation data by a multibeam system and 2) application of micropulse lidar (photon-counting) technology. A photon-counting altimeter yields clouds of discrete points, resulting from returns of individual photons, and hence new data analysis techniques are required for elevation determination and association of the returned points to reflectors of interest. The objective of this paper is to derive an algorithm that allows detection of ground under dense canopy and identification of ground and canopy levels in simulated ICESat-2 data, based on airborne observations with a Sigma Space micropulse lidar. The mathematical algorithm uses spatial statistical and discrete mathematical concepts, including radial basis functions, density measures, geometrical anisotropy, eigenvectors, and geostatistical classification parameters and hyperparameters. Validation shows that ground and canopy elevation, and hence canopy height, can be expected to be observable with high accuracy by ICESat-2 for all expected beam energies considered for instrument design (93.01%–99.57% correctly selected points for a beam with

expected return of 0.93 mean signals per shot (msp), and 72.85%–98.68% for 0.48 msp). The algorithm derived here is generally applicable for elevation determination from photon-counting lidar altimeter data collected over forested areas, land ice, sea ice, and land surfaces, as well as for cloud detection.

Index Terms—Algorithms, altimetry, laser measurements, satellites.

I. INTRODUCTION

DETERMINATION of vegetation height of earth’s forests is an essential requirement in the estimation of global and regional biomass and carbon levels. Because of the scale of the problem and the inaccessibility of many of earth’s forested areas, this is best achieved by satellites. NASA’s Ice, Cloud and Land Elevation Satellite (ICESat) mission (2003–2009) has resulted in important new findings in ecology [19], [30]–[33], [35]–[37], [40], [42], [43], in addition to many results in the primary mission objectives in the cryospheric sciences (e.g., [5], [7], [15], [16], [25], [27]–[29], [34], [44]–[46], [49]–[52], see also <http://ICESat.gsfc.nasa.gov/publications>). ICESat ceased operation in 2009. The National Research Council’s “Decadal Survey” [38] identified ICESat-2 as one of its first-tier missions citing the urgent need to observe the rapidly changing cryosphere [39], [47], with launch currently planned for 2016 [1], [2].

Laser altimetry is suited to observe vegetation height and structure because the returned signals include return from the top of the canopy, from within the canopy, and from the ground. Therefore the ICESat-2 mission has an ecosystem science requirement, stated as, “ICESat-2 shall produce elevation measurements that enable independent determination of global vegetation height with a ground track spacing of less than 2 km everywhere over a two-year period.” Based on results from the ICESat mission, which included canopy height estimates with root-mean-square errors of 2–6 m [32], [37], [40], [43], it is expected that the ICESat-2 mission, operating continuously in a 91-day exact repeat orbit, will facilitate derivation of a vegetation height product at 1-km spatial resolution if off-nadir pointing is used to increase the spatial distribution of observations over terrestrial regions. There are, however, different requirements in orbit design and sampling for vegetation science and for the ICESat-2 mission’s primary cryospheric objectives [2], [14], [17], [26]. Hence a different

Manuscript received February 14, 2012; revised July 24, 2012, November 10, 2012, and March 18, 2013; accepted April 2, 2013. The work of U. C. Herzfeld was supported in part by the ICESat-2 Project, NASA Goddard Space Flight Center under Award NNX10AR46G and Award NNX11AH43G S01 to U.C. The work presented here was conducted in collaboration and exchange with the ICESat-2 Project, the ICESat-2 Science Definition Team and Sigma Space Corporation. The Sigma Space Lidar was built under an SBIR under Contract PO FA8650-05C-1817 from the Air Force Research Lab, whom we thank for their permission to use data collected with the sensor for this work.

U. C. Herzfeld is with the Department of Electrical, Computer and Energy Engineering, University of Colorado Boulder, Boulder, CO 80523 USA, and also with the Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO 80309-0449 USA (e-mail: ute.herzfeld@colorado.edu).

B. W. McDonald is with the Department of Electrical, Computer and Energy Engineering, University of Colorado Boulder, Boulder, CO 80523 USA (e-mail: brian.w.mcdonald@colorado.edu).

B. F. Wallin was with the Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO 80309-0449 USA. He is now with the New Mexico Tech, Socorro, NM 87801 USA (e-mail: bwallin@nmt.edu).

T. A. Neumann and T. Markus are with the Cryospheric Sciences Laboratory, Code 615, NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA (e-mail: thomas.neumann@nasa.gov; thorsten.markus-1@nasa.gov).

A. Brenner and C. Field are with the Sigma Space Corporation, Lanham, MD 20706 USA (e-mail: Anita.Brenner@sigmaspace.com; Christopher.Field@sigmaspace.com).

Digital Object Identifier 10.1109/TGRS.2013.2258350

Decadal Survey Mission, Deformation Ecosystem Structure and Dynamics of Ice (DESDynI), was planned to include a lidar specifically designed to measure vegetation height. The importance of the ICESat-2 project in estimation of biomass and carbon levels has increased substantially following the recent cancellation of the lidar component of the DESDynI mission.

Determination of vegetation height from ICESat-2 measurements will be based on the determination of canopy and ground elevations. This is not trivial, because ICESat-2 will use a fundamentally different lidar than was used on ICESat, and identification of ground and canopy in the resulting data requires development of new mathematical methods and algorithms. Two innovative components will characterize the ICESat-2 lidar compared to the ICESat mission: 1) collection of elevation data by a multibeam system and 2) application of micropulse lidar (photon-counting) technology [2]. In comparison with the classic pulse-limited altimeter, a micropulse photon-counting altimeter yields clouds of discrete points, which result from returns of individual photons, and hence new data analysis techniques are required for elevation determination and association of returned points to reflectors of interest including land and sea ice surfaces, ground, tree canopy, water, clouds, and blowing snow.

Identification of tree canopy is especially challenging because of the fuzzy margin of a tree crown, and detection of ground under a possibly dense canopy is difficult because only a small percentage of the originally transmitted photons penetrate the atmosphere and the tree cover, get reflected from the ground, and, after reflection, penetrate tree cover and atmosphere again, before reaching the receiver aboard the satellite. Therefore, the vegetation algorithm development poses a mathematically more difficult problem than the ice algorithm design. Furthermore, a lower number of signal photons may be expected to be received from vegetation, because reflectance of ice surfaces is much higher than that of tree crowns and ground. In addition, forested areas tend to occur in regions with more humid atmospheres and/or higher aerosol densities than glaciers and ice sheets, which tends to further reduce the number of received signal photons.

Similar technologies to those proposed for ICESat-2 have been applied in airborne lidars. ICESat-2 will use a lidar with a center wavelength 532 nm, while ICESat used two frequencies (532 and 1064 nm) facilitated by a frequency-doubling crystal [44]. Micropulse photon-counting technology has been developed and described by [9]–[13], and data from a Sigma Space Corporation airborne sensor that implements this forms the basis of the algorithm development and analysis in this paper. Micropulse single-photon ranging with a multibeam push-broom configuration operating at 532 and 1064 nm has been used by the Slope Imaging Multi-polarization Photon-counting Lidar (SIMPL) altimeter [8], [18].

Spaceborne multibeam lidar technology is being used for topography observations of the moon and Mercury. The lunar orbiter laser altimeter (LOLA) of the Lunar Reconnaissance Orbiter (LRO) (<http://lunar.gsfc.nasa.gov/lola/about.html>, launch date June 18, 2009) uses a multibeam laser generated by propagating a single laser pulse through a

diffractive optical element that splits it into five beams and analysis of the returned pulse. The mercury laser altimeter (MLA) of the MESSENGER mission to Mercury is a pulse-limited 1064-nm four-beam laser, operated at 5 Hz (<http://nssdc.gsfc.nasa.gov/nmc/experimentDisplay.do?id=2004-030A-05>, launch date August 3, 2004). Both these missions employ analysis of laser pulses, whereas ICESat-2 will base altimeter measurements on analysis of ranging to individual photons, which necessitates development of new algorithms for elevation determination. Such an approach is described here.

The objective of this paper is to derive and describe a mathematical algorithm that allows detection of ground and canopy in micropulse photon-counting lidar data, of characteristics similar to those that will be expected from ICESat-2, and to apply these to forest data. So that the most challenging cases can be solved, data were taken from the smithsonian environmental research center (SERC) forest, which has a dense canopy. In Section II, the simulated data and ICESat-2 instrument design cases for vegetated areas are described to aid the reader in understanding the data structure and hence the possibilities for generalizations of the method to similar and other lidar data. The simulation method is not the topic of this paper. In Section III, the mathematical concepts applied in the algorithm are defined; Section IV then provides a description of the numerical algorithm. Section III provides the mathematical fundamentals and the motivation for their application. This section also serves as a reference for applications to other similar data analysis tasks. Section IV can be used by a programmer as a step-by-step guide for implementation of the algorithm and generation of a software for analysis of photon-counting laser altimeter data over forested areas. These sections may serve as a basis for an algorithm theoretical basis document for ICESat-2, but are written with a broad audience in mind. In Section V, the algorithm is applied to analyze simulated data, based on observations over a dense broad-leaf forest, and Section VI is dedicated to algorithm validation. Section VII summarizes, discusses, and concludes this paper.

II. ICESAT-2 INSTRUMENT DESIGN CASES AND DATA DESCRIPTION

A. Micropulse Photon-Counting Lidar Data

The sensor used in the ICESat mission, namely, the Geoscience Laser Altimeter System (GLAS) [44], was a pulse-limited laser altimeter. Elevation determination is based on the analysis of waveforms fitted to the returned signal, the peak of the waveform is associated with geolocation of the “point” (footprint center) from which the signal is returned, and elevation is derived from the two-way travel time associated with the waveform peak. Micropulse photon-counting technology, as pioneered by [9]–[13], is realized in an airborne system built by Sigma Space Corporation (and in other instruments). The Sigma Space system operates at 532 nm wavelength, and the same wavelength will be used for the Advanced Topographic Laser Altimeter System (ATLAS) that is under development for ICESat-2. In this paper, simulations of ICESat-2 ATLAS data based on lidar data collected with the

Sigma Space system will be used for algorithm development. The Sigma Space system yields data in a 3-D point cloud, resulting from 100 beams. The ATLAS system will have discrete sets of individual beams of different strengths. This requires a reduction, resampling, and simulation to generate the expected ATLAS-type data from the collected Sigma Space data. The received ATLAS-type data for a single beam are simulated as summarized in Section II-D; the output is given as a 2-D projection of the cloud of single-photon reflections versus along-track distance of a ground track. Pseudo-ranges are converted to pseudo-elevations in meters. Use of a 2-D projection in the simulated data is appropriate because the ATLAS lidar will assign all received noise and signal photon counts to the central beam axis. We analyze simulated data from beams of different strengths. The multibeam capacity of ICESat-2 is not analyzed here.

B. Design Cases for a Multibeam Sensor for ICESat-2

Designs of a multibeam system discussed for ICESat-2 include a combination of stronger and weaker beams. Science requirements in ice observation have led to the observation requirement of a multibeam system, while energy constraints limit the number of equally strong beams to about 2–4. On one hand, beams with higher transmit energy have a higher probability of penetrating clouds and dense atmosphere and hence yield sufficiently many surface returns for surface identification. The number of surface photons is constrained by both environmental factors (such as atmospheric conditions and surface reflectance) and instrument factors (such as transmit energy or detector efficiency). A lidar system only penetrates thin clouds, but clouds prevail in the Arctic 70%–80% of the time. On the other hand, a larger number of beams are needed to observe the spatial variability of the ice surface, which provides characteristics indicative of ice types, ice dynamics, morphogenesis of sea ice, and other parameters of interest, and improves accuracy of ice elevation mapping and change detection (see [1], [20] and other work cited therein [2]). The combination of these constraints suggests a design that includes beams of different strengths. The two favorites at times of this research were a nine-beam design (with beam strengths 1-2-1; 2-4-2; 1-2-1; i.e., center beam in each row twice as strong as outer beams, yielding corner beams a quarter of the strength of the center beam) and a six-beam design (with strengths 1-4; 1-4; 1-4; i.e., a weak beam and a strong beam in each row). In this paper, we use predictions of ATLAS radiometry over ground and tree canopy to generate simulated datasets from the Sigma Space lidar data collected over vegetated areas. These simulated datasets are used to develop an algorithm to derive both the ground elevation and canopy height.

C. Characteristics of the Forest Type

The dense forests of the SERC, located in eastern North America (38.889° N, 79.559° W), were selected as study sites, because the SERC forests are well characterized through previous work and the relatively dense canopy represents a challenge for ground detection. Hence an algorithm that works

for ground detection in the SERC forests is likely to work for less dense forests. Airborne Sigma Space lidar data were collected there during leaf-on conditions in early October 2009.

The SERC forest contains 3350 trees of 84 recognized species on 16 ha and is situated adjacent to a subestuary of the Chesapeake Bay on the coastal plain near Edgewater, Maryland. The square 16-ha plot is dominated by mature secondary upland forest but is bisected with a section of floodplain forests, both around 120 years since initiation. The upland forest is an example of the “tulip poplar” association with an overstory dominated by tulip poplar (*Liriodendron tulipifera*), several oaks (*Quercus* spp.), beech (*Fagus grandifolia*), and several hickories (*Carya* spp.); a mid-canopy of red maple (*Acer rubrum*) and sour gum (*Nyssa sylvatica*); and an understory composed of American hornbeam (*Carpinus caroliniana*), spicebush (*Lindera benzoin*), and paw-paw (*Asimina triloba*). The flood plain forest is dominated by ashes (*Fraxinus* spp.), sycamore (*Platanus occidentalis*), and American elm (*Ulmus Americana*). Installation of the plot began in September 2007. The forest is rather tall (as high as 40 m) and has a high richness for this part of the temperate zone, with more than 34 species of at least 20.0 cm diameter at breast height (DBH; it is a standard method of expressing the diameter of the trunk or bole of a standing tree. DBH is one of the most common dendrometric measurements). As of November 2009, the tagging and censusing of all woody stems 31 cm DBH in about 9.0 ha of the plot have been completed [41]. At time of the survey with the Sigma Space photon-counting sensor in October 2009, the SERC forest had reached a mature state with a closed canopy cover (over 95% canopy closure) and leaves were still on (Geoffrey Parker, see <http://www.ctfs.si.edu/site/SERC%3A+Smithsonian+Environmental+Research+Center>; 2-10-2012).

D. Simulated ICESat-2 Data Based on Airborne Sigma Space Corporation Data: File Name Conventions

As stated above, our simulation of ICESat-2 data used photon-counting data over the SERC forests collected by a commercial lidar developed by Sigma Space Corporation. The Sigma Space lidar is a high-repetition-rate (kHz) 100-beam scanning photon-counting lidar. The ICESat-2 is expected to have six beams, at fixed angles, and will also use a high repetition rate (10 kHz). Given these differences, the photon cloud collected by the Sigma lidar spans roughly 800 m at the flight altitude used here in the across-track direction and has substantially higher signal photon density than ICESat-2 is expected to achieve. Therefore, the Sigma data must be down-sampled both geometrically and radiometrically to simulate ICESat-2. Observation flights were conducted at dusk, which results in reduction of background photons from sunlight, compared to mid-day ambient light. Elimination of many (but not all) background photons above canopy and below ground was performed by manually applying a prescribed spatially variable range window that includes trees and ground surface. The resulting dataset is referred to as a database of signal-only photons here, though we are unable to perfectly discriminate

between signal and noise photons. Our simulation generates ICESat-2 simulated datasets that vary with respect to noise levels, radius of photon capture, resampling, laser intensity and expected number of mean photons per shot, subarea/flight track, and several realizations of random processes [3], [4].

The Sigma data were scaled to mimic the expected spatial distribution of ICESat-2 data. The combination of the nominal 10-kHz repetition rate of the ICESat-2 laser with a nominal 7 km/s ground speed of the spacecraft yields a footprint on the ground every 70 cm along the flight direction of ICESat-2. Therefore, in order to scale the Sigma data, we selected straight-line segments along the aircraft ground track and defined footprint locations every 70 cm by interpolating along the aircraft track. The footprint of ICESat-2 is expected to be circular, nominally 10 m in diameter, with a Gaussian energy distribution within the footprint. For a given ICESat-2 footprint, the location of the footprint center will be known, but the point of origin of any recorded photon within a footprint will not be known. As such, all received photons are effectively collapsed in space to the footprint center. For the Sigma data, individual photon locations are provided; these locations are therefore collapsed to the defined footprint center in our ICESat-2 simulated data.

For each ICESat-2 simulated footprint, the simulation has three main steps. First, the expected number of returned photons in the footprint is generated using a Poisson-distributed random number with a mean equal to the expected number of signal photons per shot (msp) based on the ATLAS radiometry model for vegetation and the ground under vegetation. ([3], see also A. Martino, AtlasPerformance20100421.xls on the ICESat-2 website). Because of the canopy closure, reflectivities of the ground surface and vegetation are imperfectly known and the atmospheric conditions are not known, so the number of signal photons in a given footprint is determined in large part by ground observation and published values for ground and vegetation reflectivity. As such, the simulation should be considered notional rather than a precise simulation of the SERC forest.

Second, the locations of these signal photons are chosen from the Sigma Space data using a Gaussian-weighted random distribution with a 2-sigma diameter of 10 m to select a radial distance from the footprint center and a uniform random distribution to select an angle with respect to aircraft ground track. This last step allows us to randomly sample the ground within the 10-m swath of ICESat-2 footprints and to simulate the jitter of footprint locations.

Third, since the Sigma Space dataset has a lot more photons than we expect from ICESat-2, we need a metric to determine which Sigma Space photon to select to represent the photon location determined in step 2. We select the closest Sigma Space photon to the photon location from step 2. The region of photon selection is limited to a 1-m radius circle around the desired photon location from step 2, and extends vertically through the Sigma Space data cloud (termed “1-m cap size”). If no photons are found within this cylinder, then none is selected for the simulation. Photon selection is limited to within the cap size for three reasons: 1) to avoid selecting photons that are too far away from the desired location to be

selected by the ICESat-2 instrument. For example, if step 2 yielded a photon location near the edge of the 10-m diameter footprint, and we used a 5-m cap size, a photon 15 m from the footprint center could be chosen; 2) to minimize computer time required to select photons. Given the Sigma Space data density, a large cap size would greatly increase the computational effort. For each photon location selected in step 2, a 10-m cap size would not only exacerbate the problem noted above but also mean searching through thousands of Sigma Space photons to find the one closest to the desired location; and 3) control the average number of signal photons in a given simulation. A smaller cap size is used to simulate ICESat-2 data from source datasets with very large photon density compared to a larger cap size for datasets with relatively sparse photon data. If a 10-m cap size was used to select photons from the Sigma data, the resulting data product would have far more signal photons than can be expected from ICESat-2; this is primarily a consequence of the number of beams (100) in the Sigma lidar, the fast pulse repetition rate, and the relatively slow aircraft ground speed. The radius of photon capture is a cut-off value for the entire simulated dataset.

Step 3 is repeated for each footprint location (from step 2) until the number of desired photons (from step 1) is reached. A Sigma Space photon may be used in one or more simulated footprints owing to the overlapping nature of the ICESat-2 footprints. We tuned the simulation parameter “capsize” so that the average number of signal photons selected over the entire length of the simulated dataset was approximately the same as the average number of signal photons predicted by the ATLAS radiometry model.

Because ICESat-2 assigns all photon locations to the footprint center, a projection of the three-dimensional point locations to 2-D point locations (x_i, z_i) is applied, with x_i distance in meters along a ground track, z_i pseudo-elevation in meters, and $i = 1, \dots, n, n \in \mathcal{N}$ the number of points in the simulated dataset. The total distance along track for the Sigma datasets from SERC is 2500 m. Noise is added in a subsequent step, with noise levels as described in Section II-C.

In the following description, file-name extensions are given in brackets in the order in which they occur in the file names.

Radius of photon capture (cap1), defined as described in step 3 above, is 1 m in all datasets.

Resampling [r0 (no resampling), r1 (resampling)] indicates the photon reuse flag, $r0$ indicates that no photon is used more than once, and $r1$ indicates that a photon can be selected for any footprint even if it was selected for a previous footprint. Effectively, $r0$ results in fewer recorded photons per shot than $r1$.

Laser intensity, quantified as “msp”, is the mean signal photons expected per shot (p4: 0.48 msp, p9: 0.96 msp). Laser intensity is used to characterize the different strengths of the beams considered for the ICESat-2 instrument. Intensity is quantified by a floating point number indicating the mean number of signal photons desired per footprint (per shot). The number of photons selected for any footprint is calculated using a Poisson-distributed random function with this mean. Note that this paper analyzes the two cases of the weak beam and the medium-strong beam of the nine-beam design, as

these are the cases limiting instrument design; there is also the case of the strong beam (1.93 msp), for which no simulations are included in the 99 ICESat-2 type SERC datasets (version 2010).

Subarea/ flight track (sb0-1, sb0-3, sb0-5). In the airborne experiment conducted by Sigma Space Corporation over the SERC forests, data along five tracks were collected and three of those were used to create the simulated datasets.

Randomization instance (s1, s2, s3, s4) refers to a new run of the simulation with a different random seed. Randomization allows us to generate several different simulations for the same ground track that will select different photons. (A random seed is a large integer number that is selected to start a new simulation, mathematically a new realization of a random function. This is commonplace in probability theory. The two realizations only differ in the random part, while all other constraints of the simulation—here flight track, noise level, resampling, laser intensity—remain the same.)

Noise levels [uz2 (lowest), uz3(middle), uz5(highest)]. Random noise is added in the simulations to mimic different atmospheric conditions, typical of nighttime conditions (uz2, 0.5 MHz), clear daytime conditions, as encountered on a crisp winter day (uz3, 2 MHz), and hazy daytime conditions, as encountered on a humid summer day (uz5, 5 MHz) (pers. comm. S. Palm, see also [48]). The existence of solar background noise and atmospheric scattering provides one of the main challenges in detection of returns from vegetation and ground under canopy.

III. MATHEMATICAL CONCEPTS OF THE ALGORITHM

Problems that must be addressed in the determination of canopy height from photon-counting lidar data include the following: fuzzyness of tree crowns; poor signal-to-noise ratios in many observational cases; roughness of the ground; trends in slope of the ground over larger distances; and specific density of trees per unit area which varies with forest type. An approach that mimics the statistical analysis of data collected with a pulse-limited laser, such as the Geoscience laser altimeter system (GLAS) aboard ICESat, and proceeds by matching along-track histograms to Gaussian functions to identify the return will not generally allow the identification of ground and canopy. The photon-counting data point clouds often yield histograms with multiple maxima and minima, which makes identification of a reflector difficult or result in situations where ground is not visible in the histogram at all. To overcome these problems, we develop a different mathematical approach.

A challenge lies in the implementation of an algorithm that facilitates automation of a “soft” solution that selects those regions as canopy and ground that visually appear as such in the cases of the stronger beams or less noise, and, in addition, succeeds at ground and canopy detection even in those cases that cannot be interpreted visually any more. This will be achieved by a combination of a density quantification that uses radial basis functions [see Section (M.2)], and a generalization of the so-called hyperparameter concept of the geostatistical classification method, adapted and applied to

histograms of the density function results. The mathematical approach described here employs spatial statistical methods and discrete mathematics building on concepts similar to those developed by the first author for other signal processing and spatial classification problems [21], [22], [24] and developing new concepts where needed.

Radial basis functions are used as a means to identify centers of dense point clouds over background noise (M.2, M.3). Noting that canopy has a tendency to extend horizontally more than vertically motivates introduction of an anisotropy matrix (M.4) in the identification of density centers (M.3). Ground is in principle a simply connected feature (in the sense of 1-connectedness in mathematical topology), but may appear as disconnected segments of denser areas in the photon data; this is addressed by an anisotropic search as well. Further analysis of density centers and their histograms will lead to the discovery that a method for selecting the most relevant maxima among other maxima is needed. In our problem the relevant density center maxima will be those that identify canopy and ground. To this end, the so-called hyperparameter concept will be applied. The idea of the hyperparameter concept is to capture those items that stand out visually [24] (see M.5 and M.6). Finally, geostatistical classification facilitates identification of ground and canopy centers (M.7). In summary, the computational algorithm developed for simulated ICESat-2 data analysis includes the following components: 1) calculation of spatial density centers, taking into account anisotropy (implemented with help of eigenvectors); 2) analysis of the cumulative distribution function, filtering, and application of the hyperparameter method of geostatistical classification adapted to identify ground and canopy ranges; 3) design and application of density threshold functions for identification of canopy and ground over background Scatterers; 4) optional linear interpolation; and 5) several plotting options and an optional data output for comparative analyses and validation. A step-by-step description of the algorithm is given in Section IV.

Mathematical concepts that have been developed specifically for this algorithm, or may not be generally known, are explained in the following subsections.

- (M.1) Globalization–localization paradigm.
- (M.2) Radial basis function (rbf).
- (M.3) Rbf density.
- (M.4) Geometrical anisotropy.
- (M.5) Geostatistical classification parameters: slope parameter p_1 and significance parameter p_2 .
- (M.6) Hyperparameters.
- (M.7) Application of geostatistical classification ideas to the histogram of the density values.

(M.1) Globalization–Localization Paradigm

A new globalization–localization approach is used to overcome a well-known statistical sampling problem, by disconnecting sampling bases in different steps of the algorithm. The idea here is to treat the following problem, typical of statistical analysis. If the data window (in distance along the flight path) is too small, then not enough photons are available to derive sufficient statistical information to identify ground

under canopy. If the data window is too large, then ground and canopy may not be separable any more. The globalization–localization idea used here is to disconnect the two problems by using a large window (here, an entire dataset) to derive a suite of statistical parameters then in another algorithm step employ a local classification or detection algorithm that utilizes the globally derived parameters. Future ICESat-2 data are expected to be much larger datasets than the simulated data analyzed here, and hence the globalization–localization paradigm can be implemented using a large moving window in the along-track direction combined with smaller windows inside that.

(M.2) Radial Basis Function

An rbf is a real-valued function whose value depends on distance from a center $c \in \mathcal{D}$ for all x in a definition area \mathcal{D}

$$\Phi(x, c) = \Phi(\|x - c\|) \quad (1)$$

with respect to any norm $\|\cdot\|$. In the algorithm, we utilize a Gaussian rbf (letting $r = x - c$ and $s \in \mathcal{R}$)

$$\Phi(r) = e^{-\left(\frac{r}{\sqrt{2s}}\right)^2}. \quad (2)$$

Visualized as a surface in \mathcal{R}^3 , this rbf has the shape of (half) a Gaussian bell curve rotated around the location of a center $c \in \mathcal{R}^2$. In the photon data analysis, we have $c \in \mathcal{R}^3$ and the surface is in \mathcal{R}^4 . More formally, the Gaussian probability density function is

$$f_{\text{normpdf}} = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)^2} \quad (3)$$

with standard deviation σ and mean μ of the population. Replacing $\sigma = s$ and $\mu = 0$ yields

$$\Phi(r) = \sigma\sqrt{2\pi} f_{\text{normpdf}}. \quad (4)$$

(See [6].)

(M.3) Density Centers

Identification of points within tree crowns is motivated by the observation that a tree crown is a diffuse reflector and a volumetric scatterer, but points within the tree crown have a high probability of being located within clusters of other parts of the tree crown, which is a property that does not hold for reflections of ambient light or noise outside of the tree crowns. To identify points located inside clusters or clouds of points with higher density, the *rbf* concept is applied as follows.

For the photon data analysis problem, the definition set \mathcal{D} is the set of all photons (in a track or window). For each point $c \in \mathcal{D}$, a density value $f_d(c)$ is calculated by summing up rbf values for all neighbors within a 15-m radius

$$f_d(c) = \sum_{x \in \mathcal{D}_c} \Phi(\|x - c\|_a) \quad (5)$$

with $\mathcal{D}_c = \{x \in \mathcal{D} : \|x - c\|_2 \leq 15m\}$ the set of all points within a given radius (here: 15 m) from the center point c [note that in this initial distance determination simply the 2-norm (Euclid norm) $\|\cdot\|_2$ is used]. In the rbf, we use a norm

$\|\cdot\|_a$ that takes anisotropy into account. The value of 15 m is chosen because the radius of a cluster of reflectors within a tree crown is usually less than 15 m. This heuristic choice of a value is used in all analyses, but is easily changed. The concept of density centers is illustrated in Fig. 1(a) and (b).

(M.4) Anisotropy Norm

Using an anisotropy norm is motivated by the notion that tree canopy has a tendency to extend more in the horizontal direction than in the vertical direction. When the anisotropy norm is combined with the rbf, points found in a horizontal direction from the center point are weighted higher than points found in a vertical direction. The following algorithm implements a matrix multiplication that is an affine transformation of the density function (the rbf) into a function of ellipsoidal shape. The anisotropy norm is defined as

$$\|v\|_a = \|A(v)\|_2 \quad (6)$$

for any vector $v \in \mathcal{R}^3$, with a transformation matrix

$$A = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (7)$$

This is applied to the density centers c and all their neighboring points in (6) as

$$\|x - c\|_a = \|A(x - c)\|_2. \quad (8)$$

Points of the same rbf value $\Phi(\|x - c\|_a)$ are now located on an ellipsoid with axes (3, 3, 1) around the center point c and (half) Gaussian bell curves along each radial line. The density value $f_d(c)$ then reflects the tendency of tree crowns to connect horizontally into forest canopies. In the 2-D realization of the simulated dataset, a transformation matrix

$$A = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{pmatrix} \quad (9)$$

is used. (The same anisotropy norm is used for ground, as ground continues more in horizontal direction. For terrain with a high topographic relief, the anisotropy matrix A can be set to a different value, or to identity.)

(M.5–M.7) Geostatistical Classification Ideas and Their Application to Histogram Analysis

Several algorithm concepts are inspired by concepts of the geostatistical classification method [22], [24] and modified to solve the lidar data analysis problem. Analysis of the variogram or its generalization, the vario function, is the basis of the geostatistical classification, but some of the principles transfer to any function that is affected by noise and here are applied to the histogram of the data and the histogram of density. More generally, we may consider any positive real-valued discrete function $f(x_i)$, defined for values x_i , $i = 1, n$.

The geostatistical classification proceeds by the analysis of sequences of minima and maxima in the vario function, derivation of parameters from those sequences, construction of a feature vector from the parameters, and classification or class association based on the feature vector. A related problem in

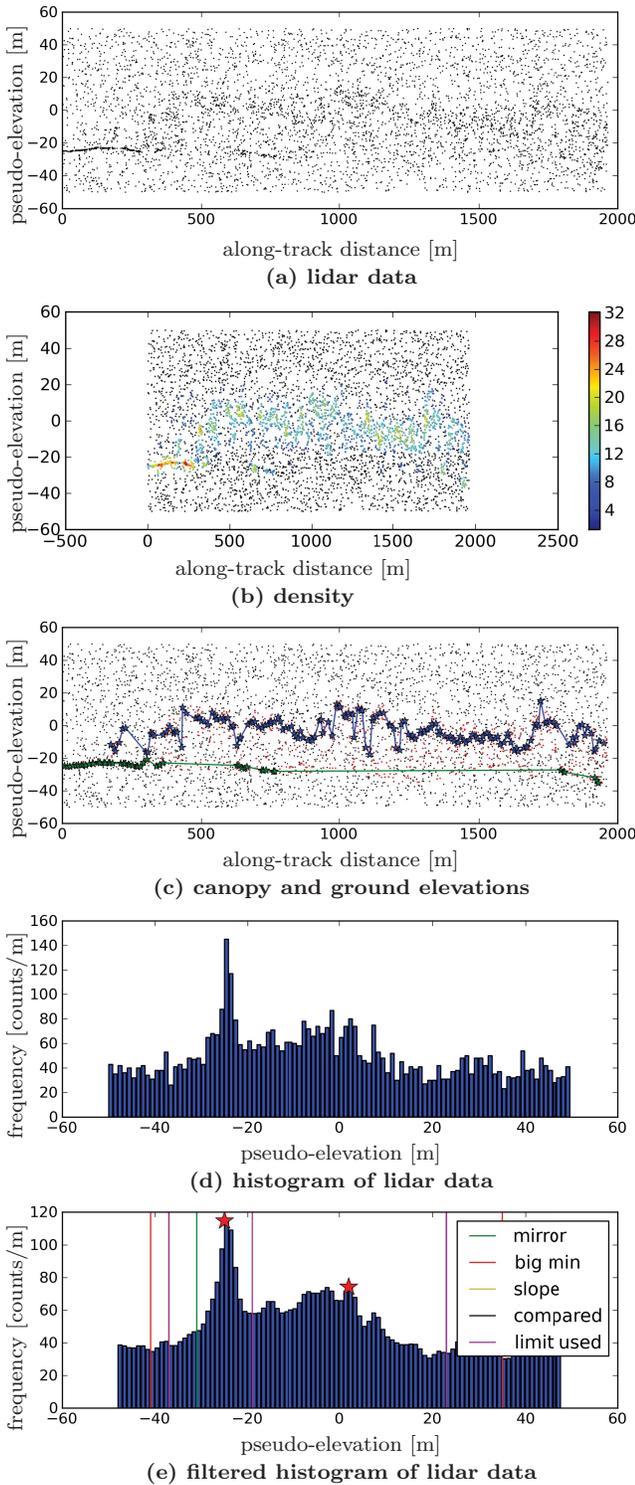


Fig. 1. Analysis concepts. (a) Simulated ICESat-2 lidar data. (b) Density of data in (a), derived from summation of radial basis function values for anisotropic neighborhood. (c) Result of analysis. Stars mark density centers within the two classes of ground and canopy, derived using histograms in (e), and lines are piecewise-linear interpolations of density centers within each class. (d) Histogram of lidar data elevation values shown in (a). (e) Filtered histogram of lidar data elevation values shown in (a). Stars mark the two most significant local maxima, which are determined using the hyperparameter concept and identify highest frequencies in ground and canopy elevations.

signal processing is the analysis of a time series or recording of a time-variable signal, which is often based on the analysis of the minima and maxima of the signal.

(M.5) Geostatistical Classification Parameters

Let $f(x_i)$ be a positive real-valued discrete function defined for values x_i , $i = 1, n$. This function may be a histogram, a variogram, or a vario function. We introduce classification parameters used in the photon classification problem. The mindist parameter is defined as the lag of the first minimum after the first maximum in the function. mindist gives the spacing of parallel features recorded in the function. We further define the significance parameters $p1$ and $p2$ as

$$p1 = \frac{f(x_{\max_1}) - f(x_{\min_1})}{x_{\max_1} - x_{\min_1}} \quad (10)$$

$$p2 = \frac{f(x_{\max_1}) - f(x_{\min_1})}{f(x_{\max_1})} \quad (11)$$

where $p1$ is the slope parameter and $p2$ the relative significance of the first minimum \min_1 after the first maximum \max_1 . In this notation

$$\text{mindist} = x_{\min_1}. \quad (12)$$

Parameters of types $p1$ and $p2$ can be calculated for any max-min sequence, defining

$$pt1(\max_i, \min_j) = \frac{f(x_{\max_i}) - f(x_{\min_j})}{x_{\max_i} - x_{\min_j}} \quad (13)$$

and an analog to (11) for $p2$ -type parameters, for $i \leq j$ and the convention that minimum \min_i always follows maximum \max_i . Note that slope parameters involve distance but $p2$ -type parameters do not.

(M.6) Hyperparameters

A problem typical of the analysis of complex and noisy processes or datasets is that the maxima and minima that tell the “story” of the problem can be identified visually because they stand out but are numerically obscured by noise or by other processes that may interfere with the main process of interest. In the lidar data analysis, we use a robust search algorithm to automatically identify “bigger” max-min sequences and associated generalized parameters, as described in [24]. We determine bigmax, the largest maximum in a group of g maxima, and then bigmin, the smallest minimum in a group of g minima following bigmax. For a fixed groupsize g , a sequence of bigmaxs and bigmins can be determined, and the selected ones are those that survive several increases of the group size. The optimal group size for a given problem can be determined automatically. Here we have applied a criterion to find stable group sizes such that the bigmax–bigmin pair stays the same for three consecutive group sizes. The so-determined parameters bigmax_i , bigmin_i , $i = 1, n$, are termed hypermaxima and hyperminima. For these selected hypermaxima and hyperminima, hyperparameters are defined as generalizations of (10), (11), and (13)

$$pt1(\text{bigmax}_i, \text{bigmin}_j) = \frac{f(x_{\text{bigmax}_i}) - f(x_{\text{bigmin}_j})}{x_{\text{bigmax}_i} - x_{\text{bigmin}_j}} \quad (14)$$

and

$$pt2(\text{bigmax}_i, \text{bigmin}_j) = \frac{f(x_{\text{bigmax}_i}) - f(x_{\text{bigmin}_j})}{f(x_{\text{bigmax}_i})}. \quad (15)$$

(M.7) Application to Histograms of Forest Lidar Data and Density

The hyperparameter concept is applied to identify the two main maxima in the histogram, which represent ground and canopy. It is a necessary piece in the analysis, because even after filtering, many histograms of forest lidar data have several maxima that may be identified as ground or canopy [see Fig. 1(d)]. The geostatistical classification concepts are applied to the histograms of elevation values and to the histograms of density values (see M.3 and M.4).

IV. ALGORITHM STEPS

The algorithm proceeds by the following steps.

- 1) *Import Data:* Data are recorded returns of individual photons, with $P = (x_1, x_2, z)$ is the location of the reflector in three dimensions, $z = z(X) = z(x_1, x_2)$ is the elevation value of a photon in location P , and $X = (x_1, x_2)$ is the projection of the photon's location onto the ground. Data are loaded into the program.
- 2) *Identification of Ground and Canopy Elevation Ranges by Histogram Analysis of Photon Elevation Data:*
 - a) A histogram of the elevation values of received photons is created and grouped by elevation bins. Here, we used 100 elevation bins for a total elevation range of 100 m (bin size 1 m). In the analysis of simulated ICESat-2 datasets, the histograms are built for the entire datasets. In an application of the algorithm to future ICESat-2 data, the analysis will be carried out for moving windows along the satellite ground track.
 - b) The following analysis is based on automated identification of hypermaxima and hyperminima in the histogram (step 2c). Prior to this, high-frequency wiggles in the histogram are smoothed out by the application of a low-pass filter to the histogram created in step (2a) to avoid that the automated algorithm catches on insignificant maxima and minima. As described in [23], this can be achieved by a five-point moving average; formally

$$s_{i_0, \text{fit}} = \alpha_1 s_{i_0-2} + \alpha_2 s_{i_0-1} + \alpha_3 s_{i_0} + \alpha_4 s_{i_0+1} + \alpha_5 s_{i_0+2} \quad (16)$$

with

$$\sum_{j=1,5} \alpha_j = 1 \quad (17)$$

and

$$\alpha_1 = \alpha_5 \quad \text{and} \quad \alpha_2 = \alpha_4 \quad (18)$$

and high values for the central coefficients $\alpha_2, \alpha_3, \alpha_4$ and low values for the outer coefficients α_1, α_5 . A Butterworth filter with

$$\begin{aligned} \alpha &= (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) \\ &= (0.0625, 0.25, 0.375, 0.25, 0.0625) \end{aligned} \quad (19)$$

is selected, because it satisfies the criteria in (16)–(18). A Butterworth filter is a well-known low-distortion low-pass filter and hence reduces high frequencies; when used in the spatial domain

as in our application, it reduces the occurrence of insignificant maxima and minima. Good results have been obtained in applications of this filter in connection with the geostatistical classification method [22], [23]. Here, this filter is applied to histograms, with s_i the histogram value of bin i for $i \in \{1, \dots, m\}$ and m the number of bins. The central index of the filter is indicated by i_0 , and the filtered value $s_{i_0, \text{fit}}$ replaces the central value s_{i_0} . An example is shown in Fig. 1.

- c) In the next step, two hypermaxima are identified (bigmax_1 and bigmax_2). These are the two maxima that stand out visually, and will represent ground and canopy elevation centers (see mathematical concept hyperparameters). For the ground and canopy range detection problem, the hyperparameter location algorithm is adapted from that described in [24] for hyperparameters of vario functions.

For the ground and canopy range detection problem, the following algorithm is used to determine the hypermaxima locations by an iterative process: In the first iteration, group size is $g_1 = 1$, and all local maxima in the histogram are identified and written into an index list. To go from step $(n-1)$ to step n of the iteration, the following is used: given a list of maxima in the index set I_{n-1} , the group size is increased $g_n = g_{n-1} + 1$ and the largest maximum within each group of g_n maxima in the original list is determined and written into Index set I'_n . A maximum is retained in list I_n if it was already in the previous list

$$I_n = I_{n-1} \cap I'_n. \quad (20)$$

Iteration is continued until at most two maxima are left (n_b is the index of the break point of the iteration)

$$|I_{n_b}| < 2. \quad (21)$$

Noting that I_{n_b-1} may contain more than two maxima, the two most significant maxima in I_{n_b-1} are selected, using a param_2-type criterion [see the mathematical concept of the significance parameter p_2 in (M.5)], with the constraint that the final two maxima must be at least eight histogram bins apart. (This corresponds to 8 m in the SERC study and is easily changed.)

The hypermaxima are identified in the histograms in Fig. 2; panel (b) demonstrates that it is necessary to determine the hypermaximum in a series of maxima that remain after Butterworth filtering. After application of this step, two “elevation centers” are identified, bigmax_g and bigmax_c , with $\text{bigmax}_g < \text{bigmax}_c$ and the corresponding x -locations $\text{bigmax}_g(x_g)$ and $\text{bigmax}_c(x_c)$.

- d) The process for determining a canopy elevation range and a ground elevation range, described in this step, is illustrated in Fig. 2(c) and (d).

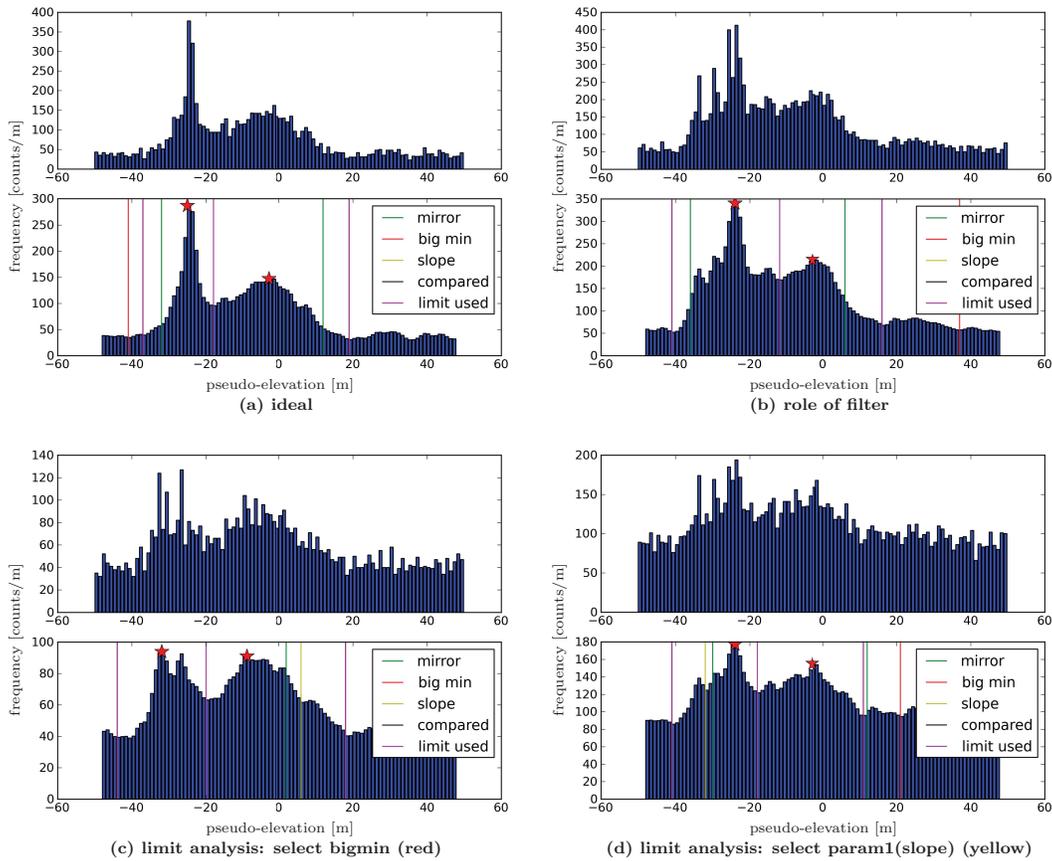


Fig. 2. Histogram analysis. (a) Ideal situation with strong and single maximum for ground (highest) and canopy (second highest). (b) Case where Butterworth filter smooths out outlying maxima. (c) Case where bigmin criterion is used for canopy-range determination. (d) Case where param1 (slope) criterion is used for canopy-range determination. Note: Color bars are plotted in the order they are calculated; earlier lines may be hidden. Color coding is as follows: green: mirror around the selected (starred) maximum; red: parameter bigmin; yellow: parameter param1 (slope); black: compared; magenta: limit used.

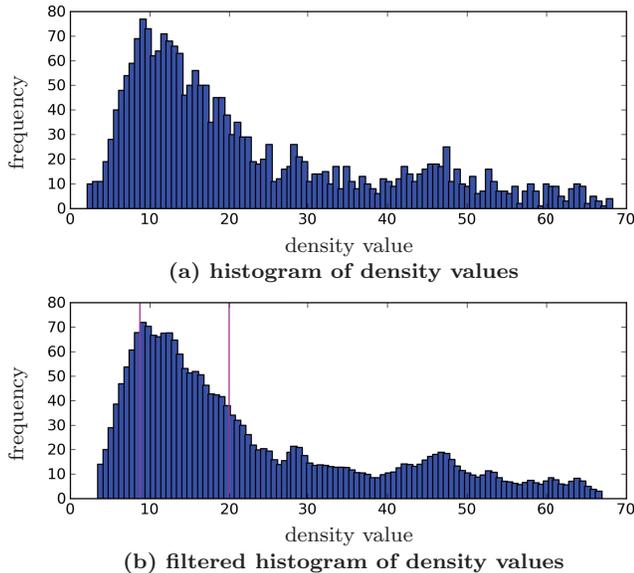


Fig. 3. Threshold analysis, demonstrated for ground detection. The threshold used is the bin associated with 0.5 of the histogram value of the bigmax of density of the ground range dataset. In this example, $0.5 \times 72 = 36$ for the histogram values, the ground threshold then becomes 20. The noise threshold is the bin associated with the bigmax, which is 9 in this case.

Colored lines are used for illustration. First, the minimum $z_{min}(x_0)$ between the ground and canopy centers $bigmax_g$ and $bigmax_c$ is determined. Then

the minimum is mirrored around the ground and canopy center locations, as $x_{green_g} = x_g - (x_0 - x_g)$ and $x_{green_c} = x_c + (x_c - x_0)$. The green lines are placeholders for finding the range values. Three local minima closest to the green lines are identified in I_0 , the one with the lowest minimum is termed $z_{red}(x_{red})$ (this is a hyperminimum) and the one with the steepest slope to the associated hypermaximum is termed $z_{yellow}(x_{yellow})$ (this utilizes a p_1 -type criterion). Finally, the range limits are determined using the slope values from the “red” and “yellow” points to the hypermaxima

$$z_{final}(x_{final}) = z_{red}(x_{red}) \quad (22)$$

if

$$\begin{aligned} &0.8pt1(bigmax_c, z_{yellow}(x_{yellow})) \\ &< pt1(bigmax_c, z_{red}(x_{red})) \\ &< 1.2pt1(bigmax_c, z_{yellow}(x_{yellow})) \end{aligned} \quad (23)$$

and

$$z_{final}(x_{final}) = z_{yellow}(x_{yellow}) \quad (24)$$

otherwise. The elevation range for ground is determined analogously.

3) *Segmentation of the Dataset Into Ground and Canopy Range Sets*: The ground and canopy elevation ranges

determined in step 2 are applied to segment the photon dataset into a canopy range set and a ground range set, and a rest class (elevations higher than canopy range or lower than ground range). It is worth emphasizing that the ground and canopy range sets are not a classification of photons into ground and canopy returns but a segmentation of the global dataset into sets in which ground and canopy can be found.

The next analysis steps are then carried out separately for the ground range set and the canopy range set.

Globalization–localization (M.1). Note that the segmentation algorithm can be applied in a window. For the SERC data, the algorithm steps 1–3 have been applied globally. The following steps (4–9) are applied in a localization. This allows us to use the properties of a larger window, or the whole dataset, for a first identification of elevation points in a likely range, based on the histogram analysis. Then in the second part of the algorithm, different mathematical concepts are applied to identify points that are ground and canopy reflectors.

- 4) *Apply Density Function for Canopy Center Identification*: Density values $f_d(c)$ are calculated as described in M.3, using the rbf (5) for all points in a 15-m radius. For the function evaluation, the distance values transformed according to the anisotropy norm described in M.4 are employed. The sum of all *rbf* values of all neighbors of a point is called the (rbf-)density of that point.
- 5) *Histogram of Photon Density in the Canopy and Ground Region*: A histogram $H(d)$ of the density values d (Step 4) for photon events in the canopy region is calculated in 100 evenly spaced bins and filtered using the Butterworth filter with the same values $\alpha = (0.0625, 0.25, 0.375, 0.25, 0.0625)$ as in step 2b. The maximal histogram value is identified as $H_{\max}(d_m)$ where d_m is the density value for which the maximum occurs.

Then a canopy threshold d_c is set.

Let $H_c = 0.8H_{\max}$ and determine d_c as the density value with $H(d_c) = H_c$ and $d_c \geq d_m$.

For ground threshold d_g , a factor of 0.5 is used. Let $H_g = 0.5H_{\max}$, where d_g is the density value with $H(d_g) = H_g$ and $d_g \geq d_m$. Note that a lower percentage of the histogram’s maximum results in a higher threshold. Fig. 3 illustrates this step for ground detection.

- 6) *Apply the Noise Filter*: The density value d_m for which the largest density count occurs (as defined in step 5) is used as a noise threshold, and points with density less than d_m are rejected. The noise threshold uses a linear function to separate signal area densities and noise area densities. Noise area densities are automatically determined as the lowest uniform density (the areas where these occur do not need to be identified.)
- 7) *Recompute the Density Function*: To eliminate possible high-density noise clusters, the density function (5) is applied a second time, as described above (including anisotropy norm). A high-density point with only noise-type density neighbors will be reassigned a much lower

density value in this second run of density, compared to the first run.

We write d_x for the density value of a point $(z(x), x)$ after the second run of the rbf f_d .

- 8) *“Build line.” Canopy Class Association (More Clearly, Define the Set of Discrete Points That Are in the Canopy Class)*: A point $(z(x), x)$ is identified as a member of the canopy set if all the following hold:
 - a) $d_x > d_c$ (and $d_x > d_g$ for ground);
 - b) $(z(x), x)$ is the point with maximal density in a 10 m along-track interval; and
 - c) a rigidity criterion is satisfied.

The rigidity criterion fixes a maximal elevation difference that is likely to occur among photons reflected from the same tree or neighboring trees, as $|z(x_i) - z(x_{i-1})| \leq rig$ for $z(x)$ elevation in location x .

The rigidity condition may be adjusted to match forest types; for instance, mapping needle trees in sparse stands may require a higher rigidity number than leaf trees in dense forests. The rigidity condition can be relaxed entirely.

- 9) *Ground Detection*: To detect ground under canopy and associate discrete photon points to the ground class, steps 4–8 are repeated using the ground parameters. Canopy and ground lines are illustrated in Fig. 1(c).

V. RESULTS

In this section, we analyze under which conditions the design cases for the beams (see Section II) can be expected to yield useful data for observation of ground and canopy levels in forests. We present results of several case studies, selected from a total of 99 test cases of simulated data. All cases are analyzed with the same algorithm.

In the first case study, typical cases of the medium-strong beam, labeled msp9 (short: p9) are investigated (Fig. 4). This figure demonstrates that the algorithm works for the medium-strong beam, the two options of resampling (without (r0) and with (r1) resampling), and increasing noise levels. The plots show the simulated data in the top panel and the interpretation of ground and canopy by the detection algorithm. Points that are original signal points in SERC forest observations are colored red, while noise points resulting from the simulation are shown in black. The signal-to-noise information was not used in the algorithm, but it aids in visually assessing validity of the algorithm. Information on ground versus canopy or reflections of other items (birds, rocks or other features, atmospheric reflections) are not provided. In this section, visual validation is used; statistical validation is given in Section VI.

In all cases, the level of the canopy is well detected by the algorithm. The canopy assumes similar shapes in all cases in spite of increasing noise and two different sampling strategies. The resampling flag “r1” indicates that resampling is allowed, which increases the signal-to-noise ratio. Cases labeled “r0” (no resampling) constitute a weaker signal, given the same noise level (left column of figure panels, (a), (c), and (e). At the start of the window, no canopy data are identified, and this

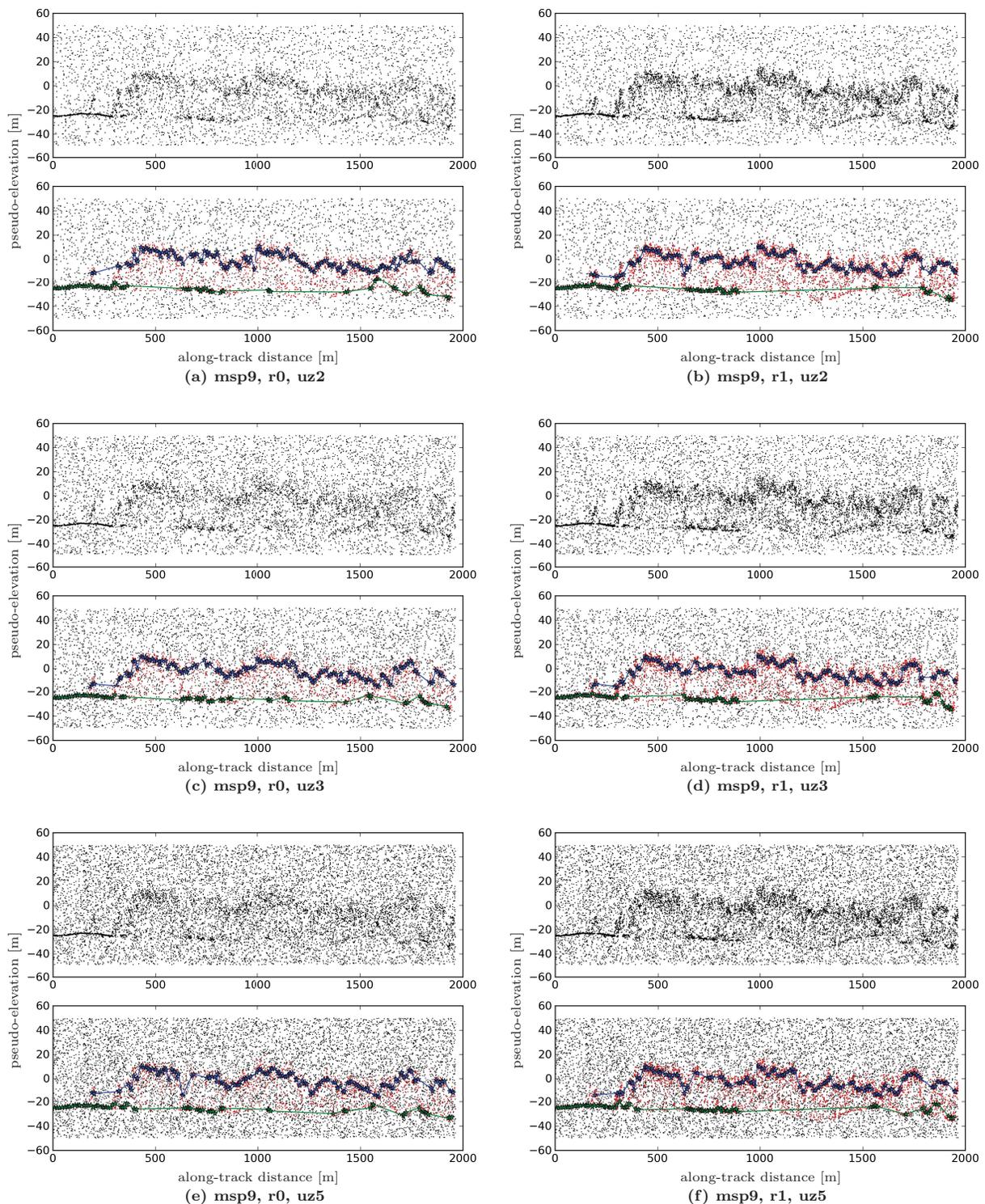


Fig. 4. Data and ground/canopy detection for the stronger beam ($msp9$), without and with resampling (r_0, r_1) (columns) and increasing noise levels (uz_2, uz_3, uz_5) (rows). All examples are SERC forest data, Section I, simulation 2.

matches the visual impression. In the case studies, an entire flight segment of 2500 m is analyzed. For actual satellite or aerial observation datasets, a moving window algorithm will be implemented, which will eliminate edge effects that occur at the start and the end of the shorter segments analyzed here.

Detection of ground level under canopy also works well, but the number of points identified as canopy shows some

variability. The software includes a simple piecewise-linear interpolation option that allows continuation of ground level across large gaps [over 400 m in 4(a), over 600 m in 4(b)]. Even in the worst case of combination of no resampling of beams and the highest noise levels, the detection of ground and canopy succeeds. Since the ground and canopy detection works for both resampling options, science or engineering

TABLE I
SUMMARY OF VALIDATION OF GROUND SET (TOP) AND CANOPY SET (BOTTOM). CORRECTLY IDENTIFIED PHOTONS

Ground												
Case	uz2			uz3			uz5			uzAll		
	Mean	Median	%	Mean	Median	%	Mean	Median	%	Mean	Median	%
p9, r0	0.45	0.00	97.20	0.49	0.00	95.78	0.55	0.00	94.70	0.49	0.00	95.89
p9, r1	0.20	0.00	99.47	0.24	0.00	99.22	0.33	0.00	98.81	0.26	0.00	99.17
p9, rAll	0.33	0.00	98.33	0.36	0.00	97.50	0.44	0.00	96.75	0.38	0.00	97.53
p4, r0	0.89	0.00	90.25	0.93	0.00	89.79	0.82	0.00	85.28	0.88	0.00	88.44
p4, r1	0.31	0.00	98.51	0.38	0.00	98.48	0.38	0.00	98.04	0.36	0.00	98.34
p4, rAll	0.64	0.00	93.79	0.70	0.00	93.52	0.63	0.00	90.75	0.66	0.00	92.68

Canopy												
Case	uz2			uz3			uz5			uzAll		
	Mean	Median	%	Mean	Median	%	Mean	Median	%	Mean	Median	%
p9, r0	0.38	0.00	96.00	0.57	0.00	93.70	0.83	0.00	93.01	0.59	0.00	94.23
p9, r1	0.18	0.00	99.57	0.30	0.00	99.04	0.37	0.00	98.95	0.28	0.00	99.19
p9, rAll	0.28	0.00	97.78	0.43	0.00	96.37	0.60	0.00	95.98	0.44	0.00	96.71
p4, r0	0.94	0.00	85.92	1.29	0.13	82.01	2.26	1.20	72.85	1.50	0.44	80.26
p4, r1	0.25	0.00	98.68	0.35	0.00	98.06	0.53	0.00	97.46	0.38	0.00	98.07
p4, rAll	0.65	0.00	91.39	0.89	0.07	88.89	1.52	0.69	83.40	1.02	0.25	87.89

Mean: mean distance in meters from a point identified as ground/canopy to the nearest signal point in the validation set. Median: median distance in meters from a point identified as ground/canopy to the nearest signal point in the validation set. %: percent of points identified as ground/canopy points that are also in the validation set of signal points. Values are for groups of datasets. p9: all sets associated with medium-strong beams; p4: all sets associated with weak beams. r0, r1: all sets with resampling r0, r1 resp. rAll: all resampling options; uz2,uz3,uz5: low, medium, and high noise levels reps; uzAll: all noise levels.

criteria can be employed for deciding between the two resampling options.

To investigate whether it may be possible to utilize the weakest beams in the ICESat-2 sensor panel and still expect to detect ground under canopy in observations of forests, a case study for the weakest beams (“msp4”) is conducted, using the same algorithm as for stronger beams, without tuning any parameters. Fig. 5 illustrates how the algorithm performs in the worst cases of the weakest beam (msp4) and no resampling (r0), increasing noise. Notably, the algorithm functions for all three noise levels, and automated detection exceeds the possibilities of visual detection of ground and canopy. In comparison to results in Fig. 4, ground can still be detected in sufficiently many locations to derive ground level, but there is a tendency for noise clusters to be misidentified as ground. Canopy heights agree in general among the three cases. Quantifications of these statements will be given in Section VI on validation. The results are encouraging to include the weakest beams in the instrument panel for ICESat-2.

Introduction of a flexible tracking rigidity parameter serves two purposes: 1) options in detection of canopy and ground for weak beams in case of weakly nonstationary ground and canopy levels and 2) a possibility to match forest-type specific characteristics. To explain the application option 1, tracking rigidity can be employed to improve detection of ground and canopy in situations of weak beams and most noise. To give an example of application 2, a forest with wide-standing conifers or pines may result in lidar data that show individual trees, hence a large slope outlining tree shape may be appropriate, and consequently a high rigidity parameter will be helpful. A forest with a dense leaf-tree canopy of homogeneous age typically has a narrow range of crown-top elevations, which is better detected by a lower rigidity parameter. Fig. 6(a) and (b)

illustrate the effect of using two different rigidity parameter values for analysis of the same dataset. In the example, canopy height curves in the lower example (with higher rigidity) have a tendency to pick up points that may be located between tree crowns. The effect of showing individual trees is expected to be more pronounced in analysis of data from wide-standing forests (than in the SERC data, where dense canopies prevail). Fig. 7 shows that the rigidity parameter can be employed to improve ground and canopy detection for the weakest beam (msp4) combined with the no-resampling option (r0), which effectively yields fewer signal photons and the highest noise level (uz5) (see Fig. 5).

VI. VALIDATION

To facilitate algorithm validation, the original signal points are flagged in the simulated ICESat-2-type datasets. (The simulated data contain a column with a 0-1 flag, the value 1 is given for each original data point.) This information was not used in the detection and classification algorithm and can therefore serve for validation of the algorithm. Results of the validation are given in Table I for the following statistical parameters, calculated separately for ground and canopy: 1) percentage of points selected that are signal points; 2) and 3) distance in meters from a point that has been identified as a signal point to the nearest point that is a signal point, given as mean and median of nearest-neighbor distances in 3-D space. Note that the distances in 2, 3 are not elevation errors. Results listed in Table I are summarized from results obtained for all 99 datasets, so that performance for weak beams (msp4), medium-strength beams (msp9), resampling options, and the three noise levels can be analyzed.

For ground, the percentage of correctly selected points is 94.7%–99.47% for all groups of medium-strength (msp9,

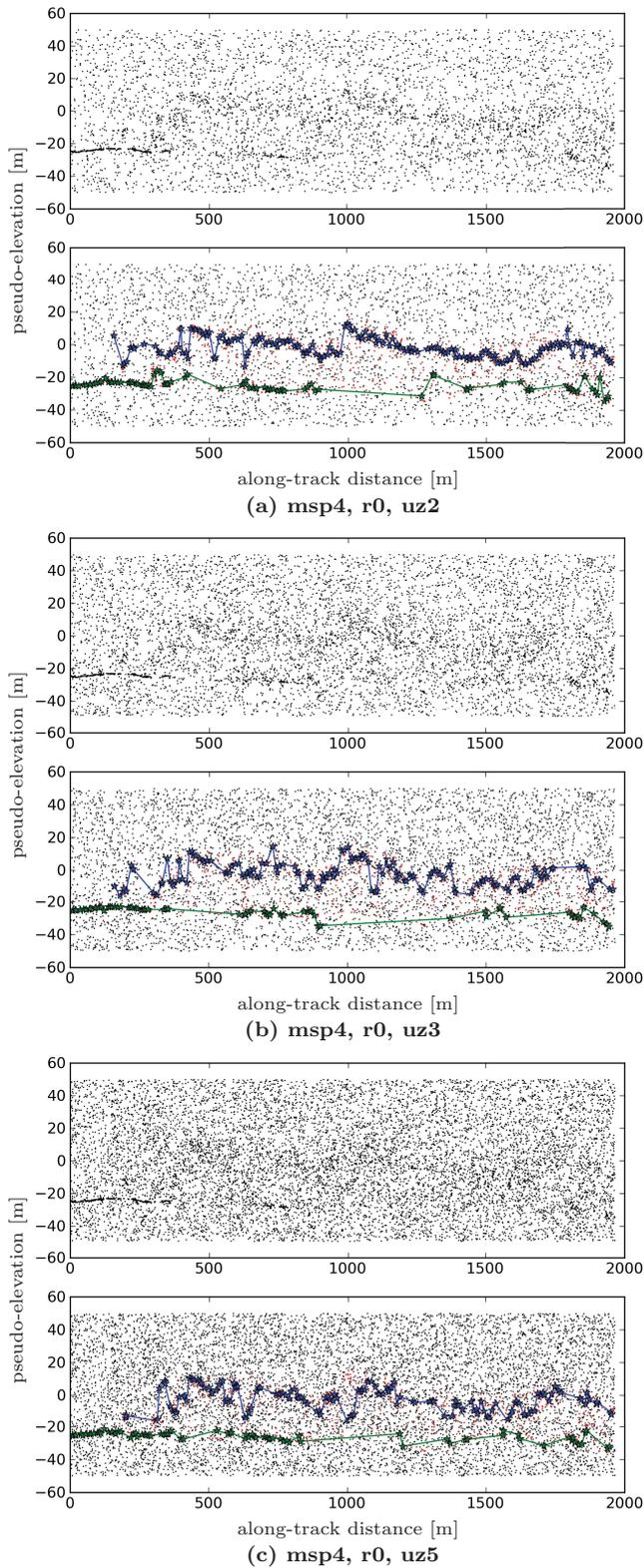


Fig. 5. Data and ground/canopy detection for the weak beam (msp4), without resampling (r0) and increasing noise levels (uz2, uz3, uz5). All examples are SERC forest data, Section I, simulation 1.

short: p9) beams and 85%–98.81% for all groups of weak (msp4, short: p4) beams. The average value over all data sets in a group is 95.89% for (p9, r0) for any noise level, 99.17% for (p9, r1), and 95.53% for all medium-strength (p9) cases.

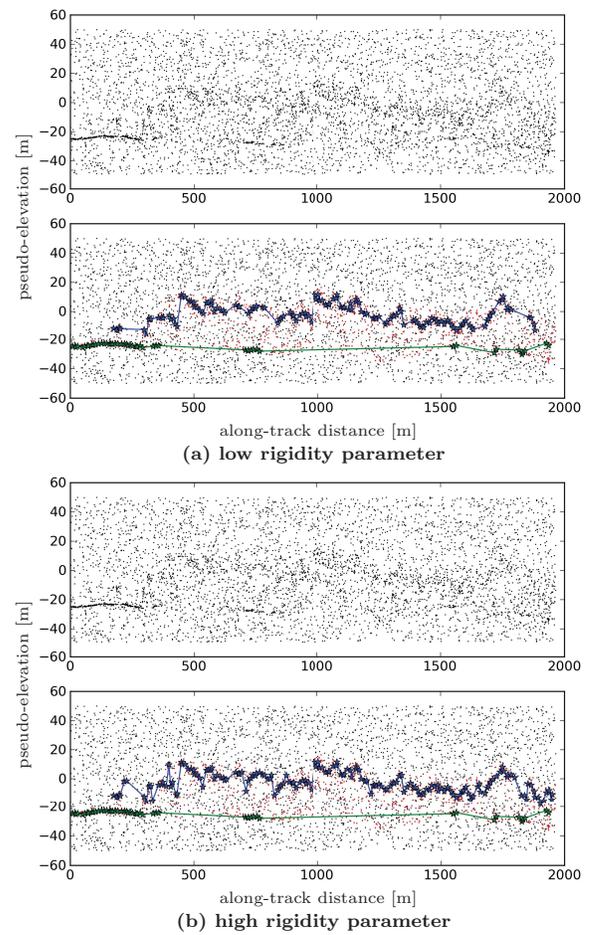


Fig. 6. Experiments using tracking rigidity. Same algorithm, except with lower rigidity parameter (top panel) and higher rigidity parameter (lower panel), applied to the same dataset (msp p4, resampling r0, uz2, v11).

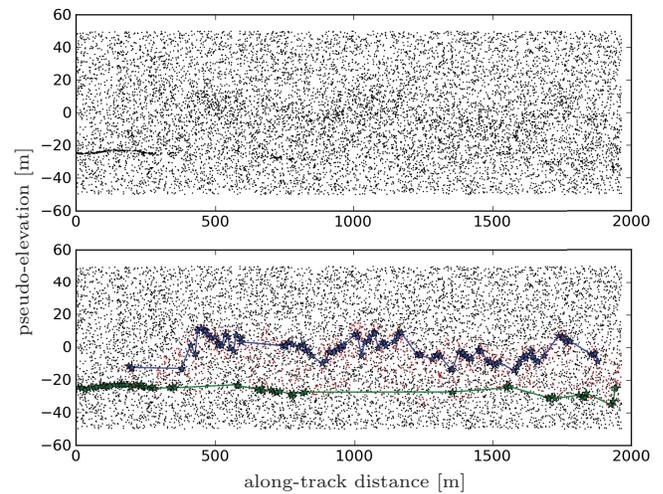


Fig. 7. Application of tracking rigidity to improve detection for weak beams and most noise. Data and ground/canopy detection, msp p4, resampling r0, uz5, v11.

The average value over all data sets in a group is 88.44% for weak beams and no resampling (p4, r0) for any noise level, 98.34% for weak beams and resampling (p4, r1) and 92.68% for all cases of weak beams (p4). The median distance

from a point in the selected set to the nearest neighbor in the signal points set is always zero, and the mean distance is 0.20–0.55 m; the resampling option has a stronger effect than the noise level. The validation demonstrates that the algorithm works very well for detection of ground under canopy. The elevation error is a lot smaller than the distance numbers, but has not been calculated directly, because the piecewise linear interpolation is only included for visualization of the ground and canopy lines, and the objective of the paper is to design a ground detection algorithm, not an interpolation algorithm.

For ground, the percentage of correctly selected points is 93.01%–99.57% for all groups of medium strength (msp9, short: p9) beams and 72.85%–98.68% for all groups of weak (msp4, short: p4) beams. The average value over all datasets in a group is 94.23% for medium-strength beams and no resampling ($p9, r0$) for any noise level; 99.19% for medium-strength beams and resampling ($p9, r1$); and 96.71% for all medium-strength cases ($p9$). The average value over all datasets in a group is 80.26% for weak beams and no resampling ($p4, r0$) for any noise level; 98.07% for weak beams and resampling ($p4, r1$); and 87.89% for all cases of weak beams ($p4$). The median distance from a point in the selected set to the nearest neighbor in the signal points set is zero for all cases of medium-strength beams ($p9$); for all cases of weak beams and lowest nighttime noise levels ($p4, uz2$), it is 0.25 m for the average of all cases of weak beams ($p4$). The mean distance is 0.44 m for the average of all cases of medium-strong beams ($p9$), the lowest for ($p9, r1, uz2$) at 0.18 m, and the highest for ($p9, r0, uz5$) at 0.83 m. The mean distance is 1.02 m for the average of all cases of weak beams ($p4$), the lowest for ($p4, r1, uz2$) at 0.25 m and the highest for ($p4, r0, uz5$) at 2.26 m. In all cases, the resampling option has a stronger effect on the accuracy than the noise level. This is a good result, because the resampling option can be set in the instrument-level detection algorithm, whereas noise from ambient light and atmospheric conditions is an environmental constraint that is corrected for in the data analysis.

The results of the detection algorithm are also very good for canopy detection, with similarly good values as the results for the ground detection. The detection of the canopy is complicated by the facts that the canopy is fuzzy and that the sparse canopy returns have to be extracted from many noise points. These facts may explain a lack of accuracy when only few returns are recorded, as in the case of the weakest beam (msp4).

Even in this hardest case, the average distance from a point in the selected set to the nearest neighbor in the signal points set is 1.02 m, which indicates that the weak beam may be expected to yield useable measurements.

At this point, it should be recalled that the analysis includes the two weaker types of beams considered for ICESat-2, namely, the weak beam and the medium-strength beam, and not the strong beam.

The dataset provided does not identify a ground dataset and a canopy dataset, hence the classification part of the algorithm cannot be validated numerically. Visual inspection of the results indicates that the canopy-class signal points fall in the upper layer and the ground-class identified points fall in

the lower layer, and the continuity of the layers indicates that the classification works correctly. As component of the experimental part of the prelaunch phase ICESat-2 project, validation datasets and instrument test datasets will be collected. To complement future flights with the airborne Multiple Altimeter Beam Experimental Lidar (MABEL), which is a high-altitude (20 km) photon-counting multibeam sensor, validation flights with vegetation lidars of known performance are planned.

VII. SUMMARY, DISCUSSION, OUTLOOK, AND CONCLUSION

In this paper, a set of algorithms has been developed and validated, which allows the detection of ground under dense canopy and identification of ground and canopy levels in simulated ICESat-2-type data. These data constitute a new type of spaceborne lidar altimeter data that will be collected during the ICESat-2 mission with a next-generation multibeam photon-counting lidar altimeter. Data analyzed in this paper are based on airborne observations with a Sigma Space Corporation photon-counting lidar, and simulations vary with respect to the signal strength, noise levels, photon sampling options, and other properties. To consider the mathematically most difficult cases, data stem from dense forests observed during leaf-on conditions and the cases of the two weaker beam types were analyzed. These are: 1) a medium-strength beam with an expected return of 0.93 mean signals per shot (msp9) and 2) a weak beam with 0.48 msp (msp4). The third case is a strong beam with 1.93 msp; this will be used in the ICESat-2 instrument design in any case. The medium-strength beam (msp9) corresponds to the weaker beam in a six-beam design proposed sensor for ICESat-2, whereas an alternative proposed nine-beam design for an ICESat-2 sensor includes four corner beams of strength msp4, four middle beams of strength msp9, and a center beam with a signal rate of 1.93 msp.

A mathematical algorithm was developed using an approach that combines spatial statistical and discrete mathematical concepts, including rbfs, density measures, geometrical anisotropy, eigenvectors and geostatistical classification parameters and hyperparameters. Piecewise linear interpolation was provided as an option to bridge between identified ground points and analogously, canopy centers. The software allows flexibility with respect to output types, which include graphics options and data output for validation and canopy height/ground elevation determination.

Validation using 99 simulated datasets showed that the algorithm works very well and that ground and canopy elevation, and hence canopy height, can be expected to be observable with a high accuracy during the ICESat-2 mission. A result relevant for instrument design is that even the two weaker beam classes considered can be expected to yield useful results for vegetation measurements (93.01%–99.57% correctly selected points for a beam with expected return of 0.93 mean signals per shot [msp9] and 72.85%–98.68% for 0.48 msp [msp4]). The median distance from a point in the selected set to the nearest neighbor in the signal points set is zero for all msp9 cases and for low-noise msp4 cases, and 0.25 m for the average of all msp4 cases. The mean distance

is 0.44 m for the average of all msp9 cases and 1.02 m for the average of all msp4 cases. Notably, this is a 3-D distance error and not an elevation error; the expected elevation error average is lower. In all cases, the option of resampling versus using each detected photon exactly once has a stronger effect on the accuracy than the noise level. Following our analysis, ground and canopy detection, and hence determination of canopy height, is possible in all noise conditions. The resampling option can be set in the instrument-level detection algorithm.

As of Oct 2012, the ATLAS lidar under development for ICESat-2 will have a nominal 41.3-m ($83.3\text{-}\mu R$) diameter field of view where about 85% of the returns will be expected from the central 12.4 m ($25\ \mu R$) containing the transmitter beam footprint (A. Martino, Instrument-Science-Martino.pptx, presented at ICESat-2 Science Definition Team Meeting). To ascertain launch readiness of the mission, signal detection and data analysis algorithms need to be developed at the same time as the sensor is being developed. The algorithm used here for data simulation based on airborne Sigma Space Corporation data employs a sampling from a combination of a 10-m diameter centered on flight track segments and a 1-m diameter cap size, which yields a 12-m possible diameter for photon selection and a 10-m resultant simulated footprint, compared to the 12.4-m central field of view for the Atlas instrument. It is important to notice that, while some parameters best used in ICESat-2 data simulations may change as the ICESat-2 instrument is being refined, the mathematical principles of the surface, canopy, and ground detection algorithm described here remain valid. The MABEL is being developed and further improved as an airborne predecessor whose data will more closely resemble the expected ICESat-2 data.

The analysis presented here does not consider the impact of clouds on the detection of ground and canopy, because the dataset is derived from airborne lidar data collected below cloud levels. In consequence, the results regarding expected retrieval of forest canopy and ground under canopy are valid for cloud-free atmospheric conditions only. Atmospheric conditions in the layer above the forest have been quantified by different noise levels within the layer, which is represented by the simulated data, the noise levels corresponding to nighttime conditions as well as clear winter and hazy summer daytime conditions at mid-latitude. Cloud impact on surface altimetry from a spaceborne 532-nm micropulse photon counting lidar is investigated theoretically in [48]. The effect of clouds on determination of surface and canopy elevations can be studied with MABEL observations made from NASA's high-altitude aircraft ER-2, which can fly at 20000 m above most cloud layers.

The multibeam capacity of the ATLAS instrument for ICESat-2 has not been investigated here. Following the collection of multibeam data with MABEL, this aspect of the ICESat-2 mission can be analyzed. Spatial variability in surface reflectance, other than as recorded by the Sigma Space instrument, has also not been considered. MABEL data also have a response to surface reflectance, which is closer to that expected for the ATLAS instrument, but were not yet collected at time of this analysis.

Because detection of ground and canopy in forested areas presents a technically and mathematically harder problem than

detection of the surface in data collected over land ice and sea ice and most other land surfaces, the algorithm presented here can be expected to be applicable also for land ice, sea ice, as well as land surface detection and elevation determination. As tree canopy may be considered a diffuse reflector, the algorithm may be generalized for other complex and diffuse reflectors, such as rough ice surfaces and atmospheric reflectors such as clouds and blowing snow. In summary, the algorithm derived here can be used as a basis for an algorithm for the analysis of data from the ICESat-2 mission, data from the mission's airborne precursor instrument, MABEL, and for analysis of photon-counting lidar altimeter data in general.

ACKNOWLEDGMENT

The authors would, for helpful discussions, like to thank W. Abdalati, B. Csatho, and the ICESat-2 Science Definition Team (SDT) and members of the ICESat-2 project, especially S. Palm, K. Barbieri, D. Harding, and A. Martino (NASA GSFC) and the vegetation subgroup of the SDT, R. Nelson, A. Neuenschwander, B. Peterson, and M. Lefsky.

REFERENCES

- [1] W. Abdalati, R. Bindshadler, C. Carabajal, B. Csatho, M. DiJoseph, H. A. Fricker, D. Harding, D. Hancock, U. Herzfeld, W. Krabill, R. Kwok, M. Lefsky, T. Markus, A. Marshak, S. Martin, A. Neuenschwander, S. Palm, J. Ranson, R. Schutz, M. Simard, B. Smith, J. Spinhirne, T. Urban, C. Webb, and J. Zwally, "Report of the ad-hoc science definition team for the ice cloud and land elevation satellite-II (ICESat-II)," NASA, Washington, DC, USA, Tech. Rep., 2008.
- [2] W. Abdalati, H. J. Zwally, R. Bindshadler, B. Csatho, S. Farrell, H. Fricker, D. Harding, R. Kwok, M. Lefsky, T. Markus, A. Marshak, T. Neumann, S. Palm, R. Schutz, B. Smith, J. Spinhirn, and C. Webb, "The ICESat-2 laser altimetry mission," *Proc. IEEE*, vol. 98, no. 5, pp. 735–751, May 2010.
- [3] K. Barbieri, A. Brenner, T. Markus, T. Neumann, J. Saba, and D. Yi, "Description of subsampling algorithm for sigma space data," NASA Goddard Space Flight Center, Greenbelt, MD, USA, Tech. Rep., Nov. 2009.
- [4] K. Barbieri, A. Brenner, T. Markus, T. Neumann, J. Saba, D. Yi, and K. Brunt, "Description of ICESat-II simulated data created from sigma space MPL laser data," NASA Goddard Space Flight Center, Greenbelt, MD, USA, Tech. Rep., Aug. 2010.
- [5] A. C. Brenner, J. P. DiMarzio, and H. J. Zwally, "Precision and accuracy of satellite radar and laser altimeter data over the continental ice sheets," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 2, pp. 321–331, Feb. 2007.
- [6] M. D. Buhmann, "Radial basis functions," *Acta Numer.*, vol. 9, pp. 1–38, Jan. 2000.
- [7] C. C. Carabajal and D. J. Harding, "ICESat validation of SRTM C-band digital elevation models," *Geophys. Res. Lett.*, vol. 32, no. 22, p. L22S01, 2005.
- [8] P. Dabney, D. Harding, J. Abshire, T. Huss, G. Jodor, R. Machan, J. Marzouk, K. Roush, C. Shuman, X. Sun, S. Valett, A. Vasilyev, A. Yu, and Y. Zheng, "The slope imaging multi-polarization photon-counting lidar: Development and performance results," in *Proc. IEEE Int. Data Conf. Geosci. Remote Sens. Symp.*, Jul. 2010, pp. 653–656.
- [9] J. J. Degnan, "Unified approach to photon-counting microlaser rangefinders, transponders, and altimeters," *Surv. Geophys.*, vol. 22, no. 5, pp. 431–447, 2001.
- [10] J. J. Degnan, "A conceptual design for a spaceborne 3D imaging lidar," *Elektrotechnik Informationstechnik*, vol. 119, no. 4, pp. 99–106, 2002.
- [11] J. J. Degnan, "Photon-counting multikilohertz microlaser altimeters for airborne and spaceborne topographic measurements," *J. Geodynamics*, vol. 34, nos. 3–4, pp. 503–549, 2002.

- [12] J. J. Degnan, R. Machan, E. Leventhal, D. Lawrence, G. Jodor, and C. Field, "Inflight performance of a second-generation photon-counting 3D imaging lidar," *Proc. SPIE*, vol. 6950, p. 695007, Mar. 2008.
- [13] J. J. Degnan, J. McGarry, T. Zagwodzki, P. Dabney, J. Geiger, R. Chabot, C. Steggerda, J. Marzouk, and A. Chu, "Design and performance of an airborne multikilohertz photon-counting microlaser altimeter," *Int. Archives Photogram. Remote Sens. Spatial Inf. Sci.*, vol. 34, no. 3, pp. 9–16, 2001.
- [14] R. Dubayah and J. Drake, "Lidar remote sensing for forestry," *J. Forestry*, vol. 98, no. 6, pp. 44–46, 2000.
- [15] R. Forsberg and H. Skourup, "Arctic ocean gravity, geoid and sea-ice freeboard heights from ICESat and grace," *Geophys. Res. Lett.*, vol. 32, no. 21, p. L21502, Nov. 2005.
- [16] H. A. Fricker and L. Padman, "Ice shelf grounding zone structure from ICESat laser altimetry," *Geophys. Res. Lett.*, vol. 33, no. 15, p. L15502, 2006.
- [17] F. G. Hall, K. Bergen, J. B. Blair, R. Dubayah, R. Houghton, J. Kellendorfer, M. Lefsky, J. Ranson, H. H. Shugart S. Saatchi, and D. Wickland, "Characterizing 3D vegetation structure from space: Mission requirements," *Remote Sens. Environ.*, vol. 115, no. 11, pp. 2753–2775, 2011.
- [18] D. Harding, P. Dabney, S. Valett, A. Yu, A. Vasilyev, and A. Kelly, "Airborne polarimetric, two-color laser altimeter measurements of lake ice cover: A pathfinder for NASA's ICESat-2 spaceflight mission," in *Proc. IEEE Int. Date Conf. Geosci. Remote Sens. Symp.*, Jul. 2011, pp. 3598–3601.
- [19] D. J. Harding and C. C. Carabajal, "ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure," *Geophys. Res. Lett.*, vol. 32, no. 21, p. L21S10, 2005.
- [20] U. Herzfeld, B. Wallin, and M. Stachura, "Applications of geostatistics in optimal design of satellite altimetry orbits and measurement configurations," *J. Astron. Sci.*, vol. 58, no. 3, pp. 495–511, 2012.
- [21] U. C. Herzfeld, "Vario functions of higher order—Definition and application to characterization of snow surface roughness," *Comput. Geosci.*, vol. 28, no. 5, pp. 641–660, 2002.
- [22] U. C. Herzfeld, "Master of the obscure—Automated geostatistical classification in presence of complex geophysical processes," *Math. Geosci.*, vol. 40, no. 5, pp. 587–618, 2008.
- [23] U. C. Herzfeld and C. A. Higginson, "Automated geostatistical seafloor classification—Principles, parameters, feature vectors, and discrimination criteria," *Comput. Geosci.*, vol. 22, no. 1, pp. 35–41, 1996.
- [24] U. C. Herzfeld, J. A. Maslanik, and M. Sturm, "Geostatistical characterization of snow-depth structures on sea ice near point Barrow, Alaska—A contribution to the AMSR-Ice03 field validation campaign," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3038–3056, Nov. 2006.
- [25] U. C. Herzfeld, P. J. McBride, H. J. Zwally, and J. Dimarzio, "Elevation changes in Pine Island Glacier, Walgreen Coast, Antarctica, based on GLAS (2003) and ERS-1 (1995) altimeter data analyses and glaciological implications," *Int. J. Remote Sens.*, vol. 29, no. 19, pp. 5533–5553, 2008.
- [26] P. Hyde, R. Dubayah, B. Peterson, J. B. Blair, M. Hofton, C. Hunsaker, R. Knox, and W. Walker, "Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems," *Remote Sens. Environ.*, vol. 96, nos. 3–4, pp. 427–437, 2005.
- [27] N. T. Kurtz, T. Markus, D. J. Cavalieri, W. Krabill, J. Sonntag, and J. Miller, "Comparison of ICESat data with airborne laser altimeter measurements over arctic sea ice," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 1913–1924, Jul. 2008.
- [28] R. Kwok, G. F. Cunningham, H. J. Zwally, and D. Yi, "ICESat over arctic sea ice: Interpretation of altimetric and reflectivity profiles," *J. Geophys. Res.*, vol. 111, no. C6, p. C06006, 2006.
- [29] R. Kwok, G. F. Cunningham, H. J. Zwally, and D. Yi, "Ice, cloud, and land elevation satellite (ICESat) over arctic sea ice: Retrieval of freeboard," *J. Geophys. Res.*, vol. 112, no. C12, p. C12013, 2007.
- [30] M. A. Lefsky, "A global forest canopy height map from the moderate resolution imaging spectroradiometer and the geoscience laser altimeter system," *Geosci. Res. Lett.*, vol. 37, no. 15, pp. 933–938, 2010.
- [31] M. A. Lefsky, D. J. Harding, M. Keller, W. B. Cohen C. C. Carabajal, F. D. B. Espirito-Santo, M. O. Hunter, and R. de Oliveira, Jr., "Estimates of forest canopy height and aboveground biomass using ICESat," *Geophys. Res. Lett.*, vol. 32, no. 22, p. L22S02, 2005.
- [32] M. A. Lefsky, M. Keller, Y. Pang, P. de Camargo, and M. O. Hunter, "Revised method for forest canopy height estimation from geoscience laser altimeter system waveforms," *J. Appl. Remote Sens.*, vol. 1, no. 1, p. 013537, 2007.
- [33] M. A. Lefsky and M. McHale, "Volume estimates of trees with complex architecture from terrestrial laser scanning," *J. Appl. Remote Sens.*, vol. 2, no. 1, p. 023521, 2008.
- [34] S. B. Luthcke, D. D. Rowlands, T. A. Williams, and M. Sirota, "Reduction of ICESat systematic geolocation errors and the impact on ice sheet elevation change detection," *Geophys. Res. Lett.*, vol. 32, no. 21, p. L21S05, 2005.
- [35] M. E. Miller, M. A. Lefsky, and Y. Pang, "Optimization of geoscience laser altimeter system waveform metrics to support vegetation measurements," *Remote Sens. Environ.*, vol. 115, no. 2, pp. 298–305, 2011.
- [36] R. Nelson, K. J. Ranson, G. Sun, D. S. Kimes, V. Kharuk, and P. Montesano, "Estimating Siberian timber volume using MODIS and ICESat/GLAS," *Remote Sens. Environ.*, vol. 113, no. 3, pp. 691–701, 2009.
- [37] A. L. Neuenschwander, "Evaluation of waveform deconvolution and decomposition retrieval algorithms for ICESat/GLAS data," *Can. J. Remote Sens.*, vol. 34, no. S2, pp. S240–S246, 2008.
- [38] *Committee on Earth Science, Applications from Space, A Community Assessment, and Strategy for the Future*. Washington, DC, USA: National Academy Press, 2007.
- [39] R. K. Pachauri and A. Reisinger, "Climate Change 2007: Synthesis report. Contribution of working groups I, II and III to the fourth assessment report of the intergovernmental panel on climate change," IPCC, Geneva, Switzerland, Tech. Rep., 2007.
- [40] Y. Pang, M. Lefsky, H. E. Anderson, M. E. Miller, and K. Sherrill, "Validation of the ICESat vegetation product using crown-area-weighted mean height derived using crowd delineation with discrete return lidar data," *Can. J. Remote Sens.*, 34, no. S2, pp. S471–S484, 2008.
- [41] G. Parker. (2012, Feb.). Smithsonian Environmental Research Center, Edgewater, MD, USA [Online]. Available: <http://www.ctfs.si.edu/site/SERC%3A+Smithsonian+Environmental+Research+Center>
- [42] S. Popescu, K. Zhao, A. Neuenschwander, and C. Lin, "Satellite lidar vs. small-footprint airborne lidar: Comparing the accuracy of aboveground biomass estimates and forest structure metrics at footprint level," *Remote Sens. Environ.*, vol. 115, no. 11, pp. 2786–2797, 2011.
- [43] J. A. B. Rosette, P. R. J. North, and J. C. Suarez. "Vegetation height estimates for a mixed temperate forest using satellite laser altimetry," *Int. J. Remote Sens.*, vol. 29, no. 5, pp. 1475–1493, 2008.
- [44] B. E. Schutz, H. J. Zwally, C. A. Shuman, D. Hancock, and J. P. DiMarzio, "Overview of the ICESat mission," *Geophys. Res. Lett.*, vol. 32, no. 21, p. L21S01, 2005.
- [45] C. A. Shuman, H. J. Zwally, B. E. Schutz, A. C. Brenner, J. P. DiMarzio, V. P. Suchdeo, and H. A. Fricker, "ICESat antarctic elevation data: Preliminary precision and accuracy assessment," *Geophys. Res. Lett.*, vol. 33, no. 7, p. L07501, 2006.
- [46] B. E. Smith, C. R. Bentley, and C. F. Raymond, "Recent elevation changes on the ice streams and ridges of the ross embayment from ICESat crossovers," *Geophys. Res. Lett.*, vol. 32, p. L21S09, May 2005.
- [47] S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Avery, M. Tignor, and H. L. Miller, *Climate Change 2007: The Physical Science Basis: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
- [48] Y. Yang, A. Marshak, S. P. Palm, T. Várnai, and W. Wiscombe, "Cloud impact on surface altimetry from a spaceborne 532-nm micro-pulse photon counting lidar: System modeling for cloudy and clear atmospheres," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 12, pp. 4910–4919, Dec. 2011.
- [49] H. J. Zwally, T. Abdalati, K. Herring, J. Larson, J. Saba, and K. Steffen, "Surface melt-induced acceleration of Greenland ice-sheet flow," *Science*, vol. 297, no. 5579, pp. 218–222, 2002.
- [50] H. J. Zwally, M. Giovinetto, J. Li, H. Cornejo, M. Beckley, A. Brenner, J. Saba, and D. Yi, "Mass changes of the Greenland and Antarctic ice sheets and contributions to sea-level rise: 1992–2002," *J. Glaciol.*, vol. 51, no. 175, pp. 509–527, 2005.
- [51] H. J. Zwally, B. Schutz, W. Abdalati, J. Abshire, C. Bentley, A. Brenner, J. Bufton, J. Dezio, D. Hancock, D. Harding, T. Herring, B. Minster, K. Quinn, S. Palm, J. Spinhrn, and R. Thomas, "ICESat's laser measurements of polar ice, atmosphere, ocean, and land," *J. Geodynamics*, vol. 34, nos. 3–4, pp. 405–445, 2002.
- [52] H. J. Zwally, D. Yi, R. Kwok, and Y. Zhao, "ICESat measurements of sea ice freeboard and estimates of sea ice thickness in the Weddell Sea," *J. Geophys. Res.*, vol. 113, no. C2, p. C02S15, 2008.



Ute Christina Herzfeld received the Masters degree in mathematics, evangelical theology, and education (Staatsexamen) from Johannes Gutenberg Universitaet Mainz, Mainz, Germany, in 1983, the Dr. rer. nat. degree in mathematics with applied mathematics and geosciences from Universitaet Mainz, Mainz, in 1986, and the Postgraduate Diploma degree in geomathematics from the Free University of Berlin, Berlin, Germany, in 1986.

She leads the Geomathematics and Remote Sensing Group as a Research Associate Professor with the Department of Electrical, Computer and Energy Engineering, and a Senior Research Associate with the Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA. She is a geomathematician and glaciologist with long-term experience in the analysis of remote-sensing data and extensive field experience in high-latitude glaciers and ice streams, including expeditions to Antarctica, Greenland, and Alaska. His current research interests include the development of algorithms for remote-sensing data analysis, especially for ICESat-2, innovative aerial and satellite observation techniques, study of the current surge of the Bering-Bagley Glacier System (Alaska) and contributions to complexity in sea-ice modeling and sea-level-rise assessments.

Prof. Herzfeld was the recipient of the President's Prize of the International Association for Mathematical Geosciences for outstanding contributions in theoretical and applied geostatistics. She is a member of the IEEE's Geoscience and Remote Sensing Society, the American Geophysical Union, and the International Glaciological Society.



Brian W. McDonald was born in Johannesburg, South Africa, in March 1989. He received the B.S. degree in electrical and computer engineering from the University of Colorado, Boulder, CO, USA, in 2011, where he is currently pursuing the M.S. degree in electrical and computer engineering under the advisory of Prof. Herzfeld.

He joined Dr. Herzfeld's Research Group, University of Colorado, in 2010. He has been a part of numerous projects including field work over Bering Glacier, Alaska. His current research interests include remote sensing of glaciers, sea ice, and land surfaces.



Bruce F. Wallin was born in Laramie, WY, USA, in 1985. He received the B.Sc. degree in applied mathematics from the University of Colorado at Boulder (CU Boulder), CO, USA, in 2009. He currently is pursuing the M.Sc. degree in statistics and operations research with the New Mexico Institute for Mining and Technology, Socorro, NM, USA.

He was a Research Assistant with the Cooperative Institute for Research in Environmental Sciences at CU Boulder from 2009 to 2010. He is currently a Research Assistant with the Institute for Complex Additive Systems Analysis, Socorro, NM, USA. His current research interests include Bayesian statistics applied to remote sensing and environmental sciences.

Mr. Wallin is a member of the American Statistical Association, the Society for Industrial and Applied Mathematics, and the American Geophysical Union.



Thomas A. Neumann received the B.A. degree in geophysical science from the University of Chicago, Chicago, IL, USA, and the Ph.D. degree in geophysics from the University of Washington, Seattle, WA, USA.

He is currently a Cryospheric Scientist at NASA, Houston, TX, USA, who focuses on the development of ICESat-2, the next-generation laser altimeter scheduled for launch in 2016. His current research interests include theoretical and experimental studies of the chemical, physical, and thermodynamic properties of polar snow and ice. His recent work has involved studies using an airborne ICESat-2 simulator to develop surface-finding algorithms and calibrating ICESat altimetry data using ground-based Global Positioning System surveys in Antarctica.



Thorsten Markus received the M.S. and Ph.D. degrees in physics from the University of Bremen, Bremen, Germany, in 1992 and 1995, respectively.

He is currently the Head of the Cryospheric Sciences Branch, Hydrospheric and Biospheric Sciences Laboratory, Goddard Space Flight Center, National Aeronautics and Space Administration, Greenbelt, MD, USA. He is the Project Scientist for ICESat-2. His current research interests include satellite and airborne remote sensing of cryospheric, oceanic, and atmospheric processes.



Anita Brenner received the B.S. degree in aerospace engineering from the University of Cincinnati, Cincinnati, OH, USA, in 1970, and the M.S. degree in electrical engineering with a major in biomedical engineering and a minor in guidance and control from the University of Maryland, College Park, MD, USA, in 1975.

She has been involved in satellite altimetry since 1976, developing algorithms to map the oceans, land, ice sheets, and river changes using radar and laser altimetry. She is currently a Program Manager at Sigma Space Corporation managing programs at the NASA/Goddard Space Flight Center in support of Earth science. She has been supporting the use of radar and laser altimetry for polar ice analysis since 1979, and has over 20 refereed publications using altimetry for studying the ice sheets, oceans, and lakes. She has been a member of the algorithm development teams and calibration and validation teams for NASA and ESA laser and radar altimeters, most recently playing a major role in developing the elevation retrieval algorithms for the Advanced Topographic Laser Altimeter System on the new ICESat-2 mission. She was the recipient of the NASA Group Achievement Awards for supporting the TOPEX, Geosat, and ICESat missions.

Ms. Brenner is a member of Tau Beta Pi, the International Glaciological Society, and the American Geophysical Union.



Christopher Field was born on the Yokosuka Naval Base, Japan, in 1958. He received the B.S. degree in mathematics and physics from Guilford College, Greensboro North Carolina, Greensboro, NC, USA, in 1980, and the M.S. and Ph.D. degrees in electrical and computer engineering from The Johns Hopkins University, Baltimore, MD, USA, in 1987 and 1993, respectively.

He was with the Institute of Energy Conversion, University of Delaware, Delaware, NJ, USA, from 1980 to 1985, engaged in solar energy research. From 1993 to 1995, he was a Post-Doctoral with the Goddard Space Flight Center, Greenbelt, MD, USA, working on lidar systems. From 1995 to 2001, he was an Assistant Professor of electrical engineering with the U.S. Naval Academy, Annapolis, MD. Since 2001, he has been with Sigma Space, Lanham, MD, working on lidar and remote sensing systems.

Dr. Field is a member of the Optical Society of America.