Ice, Cloud, and land Elevation Satellite-2 (ICESat-2) Project

Algorithm Theoretical Basis Document (ATBD) for Land Ice Along-Track Height Product (ATL06)

Prepared by: Benjamin Smith, University of Washington Applied Physics Lab

with

David Hancock, Kaitlin Harbeck, LeeAnne Roberts, Thomas Neumann, Kelly Brunt, Helen Fricker, Alex Gardner, Matthew Siegfried, Susheel Adusumilli, Beata Csathó, Nicholas Holschuh, Johan Nilsson, and Fernando Paolo

Maintained by: Denis Felikson, NASA Goddard Space Flight Center

This document may be cited as:

Smith, B., D. Hancock, K. Harbeck, L. Roberts, T. Neumann, K. Brunt, H. Fricker, A. Gardner, M. Siegfried, S. Adusumilli, B. Csatho, N. Holschuh, J. Nilsson, F. Paolo, and D. Felikson (2022). *Ice, Cloud, and Land Elevation Satellite (ICESat-2) Project Algorithm Theoretical Basis Document (ATBD) for Land Ice Along-Track Height Product (ATL06), Version 6.* ICESat-2 Project, DOI: 10.5067/VWOKQDYJ7ODB.

National Aeronautics and Space Administration

Goddard Space Flight Center
Greenbelt, Maryland

Abstract

This document describes the theoretical basis of the land ice height processing algorithms and the products that are produced by the ICESat-2 mission. It includes descriptions of the parameters that are provided with each product as well as ancillary geophysical parameters used in the derivation of the products.

CM Foreword

This document is an Ice, Cloud, and land Elevation Satellite-2 (ICESat-2) Project Science Office controlled document. Changes to this document require prior approval of the Science Development Team ATBD Lead or designee. Proposed changes shall be submitted in the ICESat-II Management Information System (MIS) via a Signature Controlled Request (SCoRe), along with supportive material justifying the proposed change.

In this document, a requirement is identified by “shall,” a good practice by “should,” permission by “may” or “can,” expectation by “will,” and descriptive material by “is.”

Questions or comments concerning this document should be addressed to:

ICESat-2 Project Science Office
Mail Stop 615
Goddard Space Flight Center
Greenbelt, Maryland 20771

Preface

This document is the Algorithm Theoretical Basis Document for the TBD processing to be implemented at the ICESat-2 Science Investigator-led Processing System (SIPS). The SIPS supports the ATLAS (Advance Topographic Laser Altimeter System) instrument on the ICESat-2 Spacecraft and encompasses the ATLAS Science Algorithm Software (ASAS) and the Scheduling and Data Management System (SDMS). The science algorithm software will produce Level 0 through Level 4 standard data products as well as the associated product quality assessments and metadata information.

The ICESat-2 Science Development Team, in support of the ICESat-2 Project Science Office (PSO), assumes responsibility for this document and updates it, as required, as algorithms are refined or to meet the needs of the ICESat-2 SIPS. Reviews of this document are performed when appropriate and as needed updates to this document are made. Changes to this document will be made by complete revision.

Changes to this document require prior approval of the Change Authority listed on the signature page. Proposed changes shall be submitted to the ICESat-2 PSO, along with supportive material justifying the proposed change.

Questions or comments concerning this document should be addressed to:

ICESat-2 Project Science Office
Mail Stop 615
Goddard Space Flight Center
Greenbelt, Maryland 20771

Review/Approval Page

Prepared by:

Benjamin Smith
Principal Researcher
University of Washington
Applied Physics Lab Polar Science Center
1013 NE 40th Street
Box 355640
Seattle, WA 98105

Reviewed by:

Alex Gardner
NASA Jet Propulsion Laboratory

Maintained by:

Denis Felikson
Research Scientist
NASA Goddard Space Flight Center
Cryospheric Sciences Laboratory
Code 615, Building 33, Room A210
8800 Greenbelt Rd.
Greenbelt, MD 20771

***** Signatures are available on-line at:** https://ipdtdms.gsfc.nasa.gov ***

Change History Log

Revision Level	Description of Change	Date Approved
1.0	Initial Release	
4.0	Minor changes made to document front matter; added brief abstract; clarified description of corrections applied to land-ice height variable (*h_li*).	
5.0	Removed the parameter *h_robust_sprd* from the *ATL06_quality_summary* flag; updated Figures 8-1 and 8-2.	
6.0	No notable changes to algorithm. However, ATL06 is now being generated for all global land regions and not restricted to just land ice.	

List of TBDs/TBRs

Item No.	Location	Summary	Ind./Org.	Due Date

Table of Contents

List of Figures

List of Tables

1 INTRODUCTION

This document describes the theoretical basis and implementation of the level-3 land-ice processing algorithms. It currently includes ATL06, which provides geolocated land-ice surface heights, and ATL11, which provides time series of surface heights. The higher-level products, providing mapped height, and mapped height change will be described in supplements to this document available 2021.

Starting with release 006, the ATL06 data product is being produced for all global land regions, even those outside of land ice. However, the data outside of land ice should be considered experimental. The ATL06 algorithm was designed to retrieve surface heights over land ice and the data quality has been checked only over land ice. Users should refer to the ATL06 Known Issues document, which describes some coarse checks that can be done on the ATL06 surface heights outside of the land ice regions.

The ATL06 product provides the most basic derived values from the ATLAS instrument on ICESat-2: the surface height at a given point on Earth's surface at a given time relative to the WGS-84 ellipsoid. ATL06 provides estimates of the ice-sheet surface height, and ancillary parameters needed to interpret and assess the quality of these height estimates. ATL06 heights represent the mean surface height averaged along 40-m segments of ground track, 20-m apart, for each of ATLAS's six beams. Segments within adjacent beams are aligned to facilitate estimation of the across-track surface slope; they are also aligned from orbit to orbit so that subsequent repeat tracks give height estimates for nearly the same location on the surface, simplifying the estimation of height changes made through repeat-track analysis. Height estimates from ATL06 can also be compared with other geodetic data and used as inputs to higher-level ICESat-2 products, particularly ATL11, 14, and 15.

Higher-level products are based on the height estimates in ATL06. ATL11 provides heights corrected for displacements between the reference tracks and the location of the ATLAS measurements. ATL14 provides gridded height maps for selected epochs during the mission, based on the corrected heights in ATL11. ATL15 provides height-change maps based on the ATL14 height maps and height differences derived from ATL11.

In this document, Section 2 provides an overview of land-ice products and gives a brief summary of the procedures used to derive products

Section 3 describes the algorithm used to generate the products.

Section 4 gives the processing steps and input data required to derive each parameter, and describes the products in detail.

Section 5 gives a detailed procedure for deriving selected parameters

Section 6 describes test data and specific tests that NASA's implementation of the algorithm should pass.

2 BACKGROUND INFORMATION AND OVERVIEW

This section provides a conceptual description of ICESat-2's ice-sheet height measurements and gives a brief description of the derived products.

2.1 Background

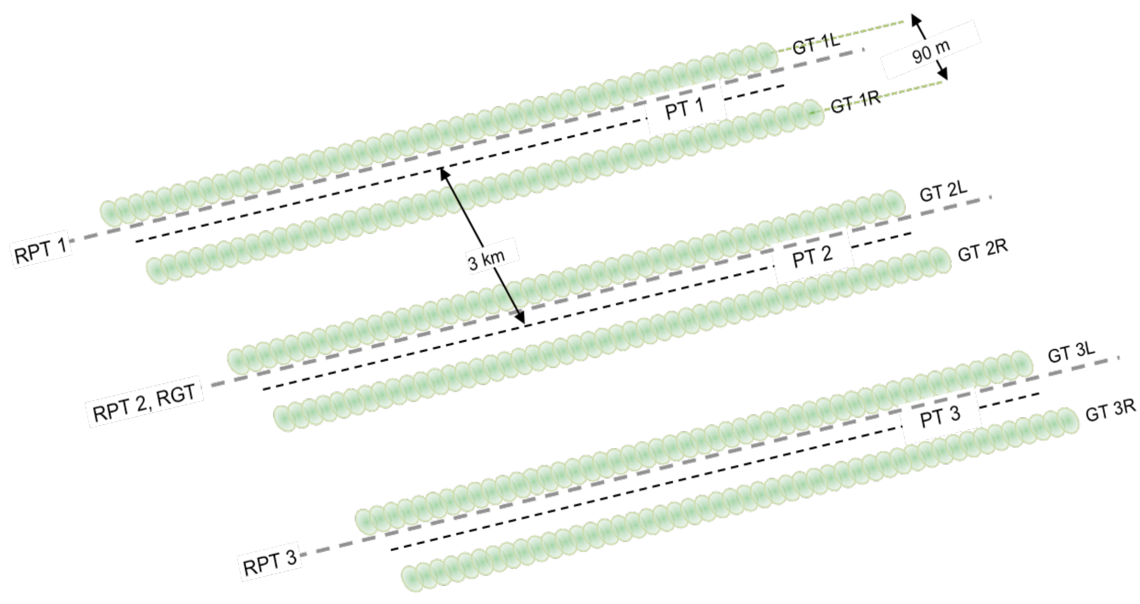
ATLAS on ICESat-2 determines the range between the satellite and the Earth's surface by measuring the two-way time delay of short pulses of laser light that it transmits in six beams. It is different from previous operational ice-sheet altimeters in that uses a photon-counting detector. Previous altimeters (e.g. GLAS on ICESat-1, ATM, and LVIS) have used full-waveform digitizers that received millions or more photons for each transmitted pulse, allowing the receiver to generate a waveform, *i.e.* the return power as a function of time. ATLAS instead records a set of arrival times for individual photons, which are then analyzed to derive surface, vegetation, and cloud properties. Although ATLAS measures much weaker signals than full-waveform altimeters, it has three major design advantages over GLAS:

i) ATLAS has six beams arranged in three pairs (Figure 2-1), so that it samples each of three reference pair tracks with a pair of beams;
ii) ATLAS transmits pulses at 10 kHz, giving approximately one pulse every 0.7 m along track, more than two orders of magnitude finer than the 170-meter along-track of GLAS;
iii) ATLAS's expected pointing control will be better than 90 m RMS, better than the 100-200 m achieved by ICESat-1.

ATLAS's six beams are spread over a small angle so that their projection onto the surface of the earth is a rectangular array with two rows and three columns, with about 3.3 km separation between each column and its neighbors, and 2.5 km between the rows. As ICESat-2 moves along its orbit, the ATLAS beams illuminate six tracks on the Earth's surface; the array is rotated slightly with respect to the satellite's flight direction so that tracks for the fore and aft beams in each column produce pairs of tracks, each pair separated by about 90 m (Figure 2-1). The separation between beams in each pair allows for measurement of the local surface slope in the across-track and along-track direction; this will allow ICESat-2 to make the most precise and detailed repeat estimates of ice-sheet height of any satellite to date.

ATLAS pulses are short, about 1.6 ns long (FWHM), and are transmitted every 0.1 ms (10 kHz); this fast repetition yields footprint centers separated by about 0.7 m in the along-track direction. Each pulse illuminates an approximately circular area on the ground ~17 m in diameter. ATLAS's strong beams detect at most 12 reflected photons from each transmitted pulse. Great care is taken to detect only photons with the same wavelength as the transmitted laser pulse and to limit the field of view of the detectors to a region slightly larger than the illuminated "footprint" of each beam; therefore, ground-return photon events (PEs, meaning photons that are detected) may readily be distinguished from solar background PEs because they are clustered in time, while background PEs are distributed evenly in time and arrive much less frequently.

The high (~45-meter RMS) accuracy of ICESat-2's pointing control means that pairs for consecutive repeats of each RPT (Reference Pair Track) are likely to overlap. The fine along-track sampling and the multi-beam capability allow height products to be defined for segments that are consistent in along-track position for repeated measurements along the same RPT.



Schematic drawing showing the pattern made by ATLAS's 6-beam configuration on the ground, for a track running from lower left to upper right. The RPTs (Reference Pair Tracks, dashed lines) are defined in advance of launch; the central RPT follows the RGT (Reference Ground Track, matching the nadir track of the predicted orbit). The Ground Tracks are the tracks actually measured by ATLAS (GT1L, GT1R, etc, shown by green footprints). Measured Pair Tracks (PTs) are defined by the centers of the pairs of GTs, and deviate slightly from the RPTs because of inaccuracies in repeat-track pointing. The separation of GTs in each pair in this figure is greatly exaggerated relative to the separation of the PTs.

Figure 2-1. ICESat-2 repeat-track schematic

Further processing of ATL06 heights will produce heights corrected for surface slope and curvature that give the estimated time-varying height for selected points on the RPTs and at track-to-track crossover points (ATL11). These shape-corrected heights will be processed further to give i) height maps for selected time intervals (semi-annual or annual, ATL14) and ii) annual height-change maps for the Antarctic and Greenland ice sheets (ATL15)

2.2 Physical Basis of Measurements

2.2.1 Height retrieval over approximately planar surfaces

Light from the ATLAS lasers reaches the earth's surface as flat disks of down-traveling photons, approximately 50 cm in vertical extent, and spread over about 17 m horizontally. On land ice, photons are scattered once, or many times, by snow and ice grains, into every direction, including towards the satellite; a tiny fraction return to the ATLAS telescope's focal plane, and a few of these are counted by the detector electronics and recorded as Photon Events (PEs). Over the vast majority of the earth's land ice, the surface is smooth, with small (single-degree) variations in surface slopes at scales less than a few hundred meters. This allows us to approximate the surface profiles measured by ATLAS with short linear segments. We aggregate PEs received by ATLAS into 50% overlapping along-track segments of a fixed length (40 m), whose centers are 20 m apart. We then fit these PEs with sloping line segments; for each segment, we estimate both the along-track slope and the height at the center of the segment. When both beams in a pair provide height measurements, we also calculate the across-track slope for the pair. Any height variation not captured by this fitting process will be treated as surface roughness.

The time variation in surface height is determined by fitting a simple spatial function to the heights from multiple repeat measurements, and using this function to correct the measurements for the height variations caused by spatial sampling of sloped and curving surfaces. This function is fit to the subset of the repeat measurements that we assess to be of the highest quality, but corrected height estimates are provided for all available repeats, and data-quality metrics are provided to allow users to decide which heights to use.

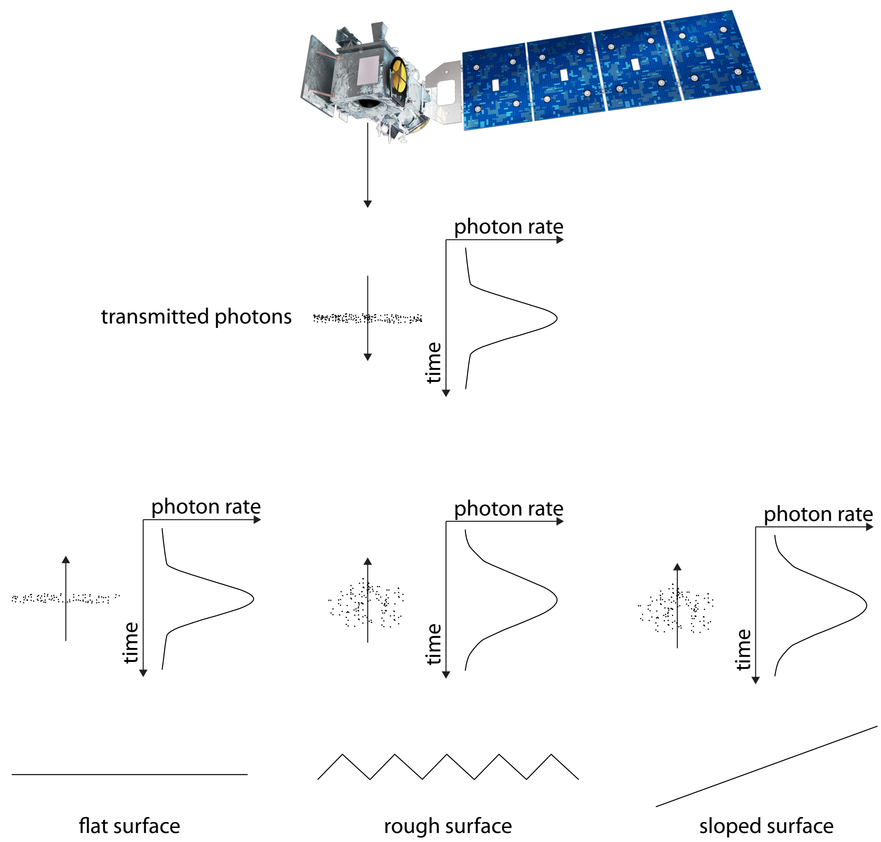
2.2.2 Effects of surface slope and roughness

Figure 2-2 shows how slope and roughness contribute to the shape of the return pulse. For many areas of glaciers, the ground may be treated as a rough planar surface, and the laser pulse as having a Gaussian distribution in space, with intensity falling to $1/e^2$ of its peak value over a distance W/2. The laser pulses also have an approximate Gaussian distribution in time, with standard deviation σ_{tx}. If the incident beam is not parallel to the surface normal, photons from the edge of the footprint farthest from the satellite will be delayed relative to photons from the edge nearest the satellite. At the same time, a rough surface will yield early photons and late photons, further spreading the returned photons. If the angle between the beam and the surface normal is φ, and the surface height within the footprint has a Gaussian distribution with RMS deviation *R* relative to the plane of the surface, then the measured temporal distribution of the returned photons will be Gaussian as well (Yi & Bentley, 1999), with a temporal standard deviation equal to the quadratic sum of the spreads due to the transmitted pulse, the surface slope, and the roughness:

$$\sigma_R = \left[\sigma_{tx}^2 + \left(\frac{2\sigma_{beam}}{c} tan\varphi\right)^2 + \left(\frac{2R}{c}\right)^2\right]^{1/2} \qquad 1$$

For ATLAS, σ_{beam} is expected to be around 4.25 m (one quarter of W), and σ_{tx} around 0.68 ns, corresponding to a FWHM (Full Width at Half Maximum) of 1.6 ns, so spreading due to sloping surfaces will be smaller than the transmit-pulse duration for slopes up to approximately 1.3 degrees.

Figure 2-2. Schematic of returns from different surface types



Top: Transmitted photon distribution. Middle: expected return photon distribution from a flat surface, a rough surface, and a sloping surface. Bottom: surface types.

Surface roughness on a 17-m scale is likely to be small except in heavily crevassed glacier margins and in heavily channeled ablation zones. Although analysis of the return pulse shape does not allow us to distinguish the effects of roughness from those of slope, the geometry of ATLAS's tracks, with pairs of beams separated by 90 m, allows estimates of the across-track slope at scales modestly larger than a single footprint, while the along-track component of the slope can be estimated from the along-track sequence of heights.

2.2.3 Distinguishing return PEs and background PEs

At the same time as signal photons are received by the ATLAS detector, background photons from sunlight are continually entering the telescope. Most of these are eliminated by filters that allow only photons with wavelengths close to the laser wavelengths through, but some pass these filters, and their timing is also recorded. The time distribution of the returned signal photons depends on the geometry and reflectance of the ice surface, and on scattering and attenuation in the atmosphere. We distinguish signal PEs from background PEs by their clustering in time.

Sunlight scattered from bright (*i.e.* snow-covered) surfaces will produce detected PEs at rates up to around 12 MHz. For comparison, a return with as few as three PEs distributed over one half meter of range produces a brief return rate of 900 MHz. Signal returns are also distinct from the background because they are spatially contiguous, so that PEs will be clustered in time in a consistent way from one shot to the next.

2.3 Potential Errors

Errors in ATLAS land-ice products can come from a variety of sources:

1) Sampling error: ATLAS height estimates are based on a random sampling of the surface height distribution;
2) Background noise: Random-noise PEs are mixed with the signal PEs, so sampled PEs will include random outliers;
3) Complex topography: The along-track linear fit and across-track polynomial fit do not always resolve complex surface topography.
4) Misidentified PEs: The ATL03 product will not always identify the correct PEs as signal PEs;
5) First-photon bias: This is an error inherent to photon-counting detectors that results in a high bias in the mean detected PE height that depends on signal strength;
6) Atmospheric forward scattering: Photons traveling downward through a cloudy atmosphere may be scattered through small angles but still be reflected by the surface within the ATLAS field of view; these will be delayed, producing an apparently lower surface;
7) Subsurface scattering: Photons may be scattered many times within ice or snow before returning to the detector; these will be delayed, producing a surface estimate with a low bias.

These errors are each treated in a different way during the ATL06 processing:

1) and 2) are treated as random errors, and their effects are quantified in the error estimates associated with the products.

3) and 4) will produce relatively large errors, and will need to be addressed with consistency checks on the data during the generation of higher-level products.

5) will be corrected routinely during ATL06 processing (see Section 3.0).

6) and 7) require information about cloud structure and ice-surface conditions that will not be available at the time of processing of ATL06. Correcting for these errors remains an active avenue for research.

2.4 Land-ice Level-3 products: ATL06: Land-Ice Height

The ATL06 product provides surface height estimates organized by reference -pair track (RPT), in a format designed to facilitate comparison between different repeat measurements on the same RPT. It also combines information from the two beams in each PT to give across-track slope estimates. A variety of parameters are provided that indicate the quality of the surface-height estimates and the signal and noise levels associated with the measurement. Note that in cycles 1 and 2 of the mission, ICESat-2 did not point at the RPTS, and ICESat2's pairs are offset by up to 2 km from the RPT locations. The first cycle that was collected over the RPTS was the third.

We define ATL06 heights based on fits of a linear model to ATL03 height data from short (40 m) segments of the ground track, centered on reference points spaced at 20-m intervals along-track. We refer to height estimates for these short segments as "segment heights", and segment's horizontal location is that of the reference point, displaced in a direction perpendicular to the RGT to match the GT offset. The choice of 40 m for the segment length provides data from slightly more than two independent (non-overlapping) ATL03 heights (based on 17-m footprints) for the along-track slope estimate, so that this component of the slope can be eliminated as a cause of vertical scatter in the PE height distribution. The spacing between reference points is 20 m, so that each segment overlaps its neighbors by 50%. Defining overlapping segments in this way increases the chances that a segment will overlap a locally smooth area within a crevasse field, potentially improving elevation-rate recovery in these areas.

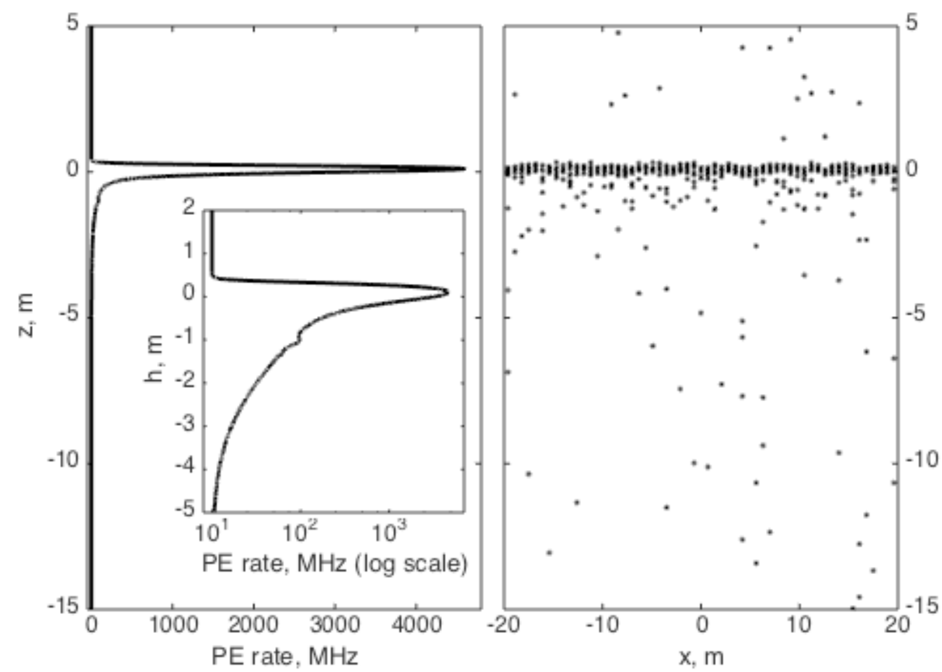
We use the same along-track sampling for both beams in each beam pair, and, for each cycle, use the same reference point each time we calculate a segment height. This allows for direct comparison between segment heights from the same RPT, without the need to interpolate in the along-track direction. The ATL03 PE used for each segment can be determined by associating the */gtxx/land_ice_segments/segment_id* parameter in ATL06 with the */gtxx/geolocation/segment_id* parameter in ATL03: segment *m* in ATL06 includes PEs from ATL03 segments *m-1* and *m* (here xx represents the ATLAS beam, with gt1l and gt1r providing the left and right beams for pair 1).

A minimal representation of the data is given in datasets in the ATL06 product in the */gtxx/land_ice_segments* groups. In these groups, we give the latitude, longitude, height, slope, vertical error estimate, and a quality flag for each segment. This represents the minimum set of parameters needed by most users; a wide variety of parameters describing the segment fit, the input data, and the environmental conditions for the data are available in the subgroups within the *gtxx* groups.

3 ALGORITHM THEORY: DERIVATION OF ATL06 LAND ICE HEIGHT PARAMETERS

In this section, we describe the ATL06 height derivation from lower-level ATLAS data (primarily the PE heights, locations, and times provided by ATL03). This process provides height estimates and segment geolocations for a set of points (called reference points) spaced every 20 m along each of ATLAS's pair tracks. One height is calculated for each beam in each pair, for each reference point, for each cycle of ICESat-2's orbit.

Figure 3-1. Surface return shape

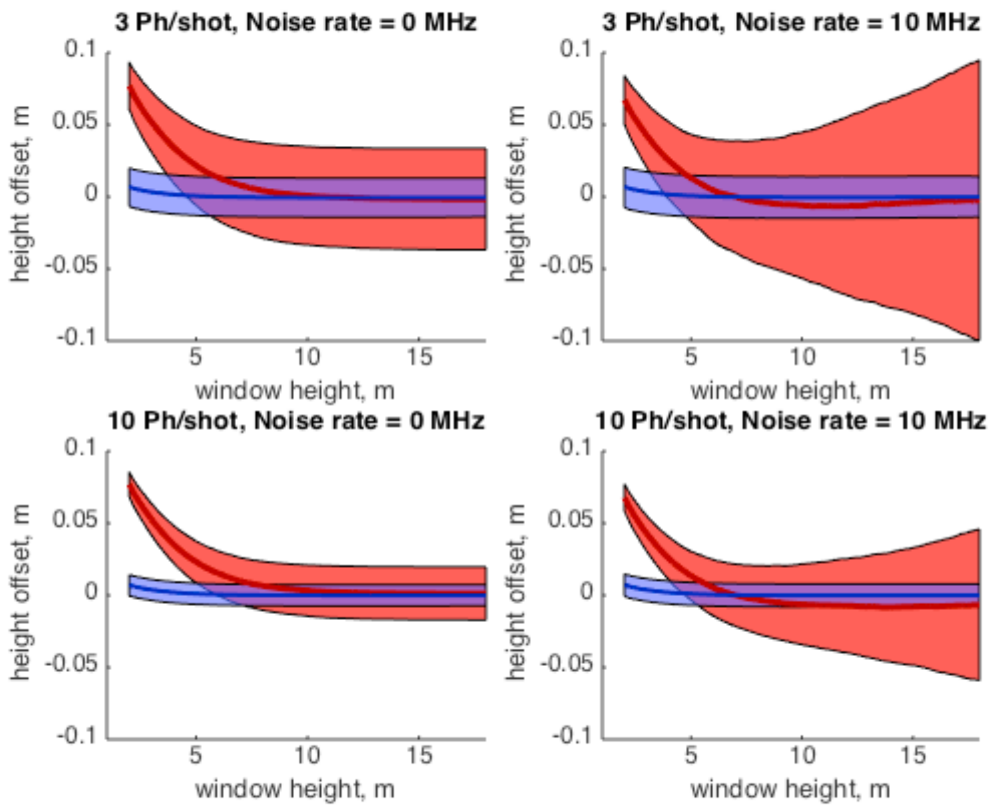


Left: power distribution for a strong beam transmit pulse, expressed as a function of height above the surface, based on the mean of 3000 waveforms measured from an ATLAS prototype laser, with a background noise rate of 10 MHz. Measured waveforms have been smoothed, and noisy portions of the waveform at the beginning and end were replaced by a smooth decay function. Inset: Power distribution on a log scale to better show the falloff in power as a function of time. Right: Simulated PE heights for a 40 meter section of flat ground track, based on the power distribution at left.

1.1 Representation of the surface

Figure 3-1 shows the expected surface-return power as a function of height above the surface, based on waveforms measured from a prototype ATLAS laser, for sunlit ice-sheet conditions with a background PE rate of 10 MHz, and a random set of photon heights generated based on this waveform for a 40-meter along-track segment. The return has a sharp peak in power at the ground, but it is asymmetric, with a leading edge (on the +z side) that is sharper than the trailing edge (on the –z side), and with a long 'tail' of energy on the –z side caused by a slow decay in laser power at the end of the pulse. This produces a dense collection of PEs at the surface height, with scattered PEs above and below, some of which come from the sun and some of which come from the tail of the waveform.

Figure 3-2. Mean and median height biases



Mean (red) and median (blue) heights for 1000 random collections of PEs for 40 meters of along-track data over a flat surface, for weak and strong beams, for two different background noise rates. Solid lines show the average height offset relative to the full-waveform median and mean, colored patches show the 1-σ range.

One way to characterize the surface height for this segment would be to calculate the mean of all PE heights within a pre-determined height range (the 'surface window'). For simplicity, one might choose a large surface window of 10-20 m to ensure the capture of all return PEs. However, this choice would lead to significant noise and potential bias in the estimated surface heights. The noise would come about because the mean of a distribution of heights is sensitive to the extreme values of the distribution, so the photons at the edge of the distribution would produce sampling errors in the recovered heights. The bias could come about if the shape of the transmit pulse were to change over time, because of temperature changes or because of aging of the lasers. If this were to happen, the mean recovered surface height could change even if the true surface height did not, again because the mean is sensitive to outlying data. Figure 3-2 shows the expected bias and scatter magnitudes as a function of the width of the surface window for the means of 1000 random collections of PEs based on the waveform in Figure 3-1. Selecting a small surface window results in a narrow (2 cm or less) scatter of values around the mean, because the range of PE heights in the window is small. However, this leads to a 7-8 cm bias in the surface height, because the tail of the distribution is cut off. Selecting a large surface window leads to a small bias, but, particularly when background noise is large, it leads to scatter in the surface heights, potentially as large as ±10 cm.

We ameliorate this problem in two ways: First, we use an iterative process to select a small surface window that includes the majority of the signal PEs but few background PEs. Second, we express the surface height as the median of the PE heights within the surface window. We select the median instead of the mean because it is less sensitive to sampling error for distributions containing a uniform, ‘background’ component. Median height offsets shown in Figure 3-1 have a spread of less than 2 cm, have maximum biases less than 7 mm, and are nearly independent of the surface-window height. This represents a large improvement in accuracy and precision over the mean, and further processing (discussed in 3.5) can correct for the remaining bias in the median heights.

In the course of processing photon-counting data, we frequently need to estimate the spread of a distribution of PE heights. For other types of data, we might choose to make this estimate based on the standard deviation of the sample of heights, but because our measurements contain a mixture of signal and noise PEs, the standard deviation often overestimates the spread of the data. Instead, we generally use the RDE (Robust Dispersion Estimator), which is equal to half the difference between the 16^{th} and the 84^{th} percentiles of a distribution. For Gaussian-distributed data, this statistic is approximately equal to the standard deviation, and for data containing a mixture of a large fraction of signal and a small fraction of noise, it can give an estimate of the spread of the signal that is relatively insensitive to the noise. In some cases, we use a version of this statistic that estimates the spread of the signal component of a distribution that contains a mixture signal (Gaussian- or near-Gaussian-distributed) PEs and background (uniformly distributed) PEs. In these cases, we estimate the 50^{th} and 75^{th} percentiles of the signal component and scale the difference between these percentiles based on the expected width of these percentiles for a Gaussian distribution. We refer to this measure as “robust spread including background” and describe its implementation in section 5.

3.1.1 Land-ice height definition

The land-ice height is defined as estimated surface height of the segment center for each reference point, using median-based statistics. We calculate this the sum of the least-squares height fit, the first-photon-bias median correction, and the pulse-truncation median correction. Height increment values on the product allow removal of the corrections and calculation of the segment mean height, and first-photon-bias and pulse-truncation corrections appropriate to the segment mean.

3.2 Outline of processing

The outline of the process is as follows for each cycle for each along-track point. First, heights and along-track slopes are calculated for each beam in each pair:

1. PEs from the current cycle falling into the along-track bin for the along-track point are collected (3.3)
2. The heights and surface windows are iteratively refined (3.3.5.2)
3. Corrections and error estimates are calculated based on the edited PEs. (3.4, 3.5, 3.6)

Once these steps are complete, based on the height values for the two beams,

4. The across-track slope is calculated (3.7)

Each of these steps is described in turn below.

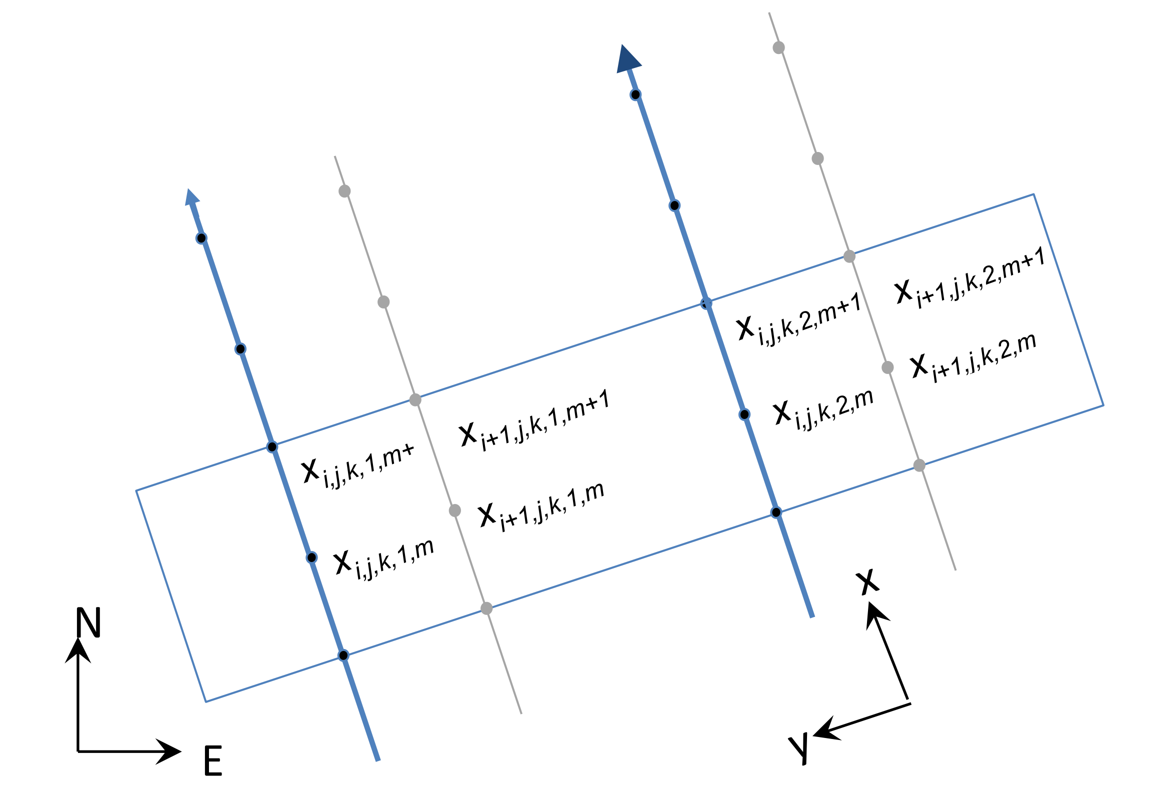
3.3 PE selection

ATL03 provides PE locations and timings for each beam. The first step in ATL06 processing is to select groups of PEs that determine the segment height at each along-track point. Processing is only carried out if the ATL03 *podppd_flag* indicates that the PE geolocation was of high quality for all pulses in the segment, otherwise the segment is skipped.

3.3.1 Along-track segments

Our height- and height-change schemes rely on dividing the data into repeatable along-track segments. We define these segments relative to the pre-defined RGT (see ATL06 Appendix A for definitions related to the ICESat-2 ground and reference tracks) and use them to select groups of PEs for each beam and each pass, and to define local coordinates relative to the RGT. We define a set of reference points, spaced every 20 m in the along-track coordinate x along the RGT, which specify the locations of the height estimates reported in ATL06. One set of reference points is defined for each RPT (Reference Pair Track). An ATL06 segment of data includes all PEs whose x coordinates are within approximately 20 m of that of a given reference point, for a total length of 40 m, so that each segment overlaps its neighbors by 50%. Each individual segment is fit with a least-squares model that gives the slope and height of the segment (Figure 3-3 and Section 3.1.2.4), and height corrections are derived based on the residuals to this model.

Figure 3-3. Reference point numbering schematic



Schematic drawing of two subsequent orbits in geographic (*N,E*) and along-track (*x,y*) coordinates. The rectangle shows the area of data included in a multiple-cycle repeat for reference point *m*.

Along-track segments are designated by five subscripts (Figure 3-3):

-i, the cycle number, numbered from the start of the mission;

-j, the track number, numbered consecutively within the cycle;

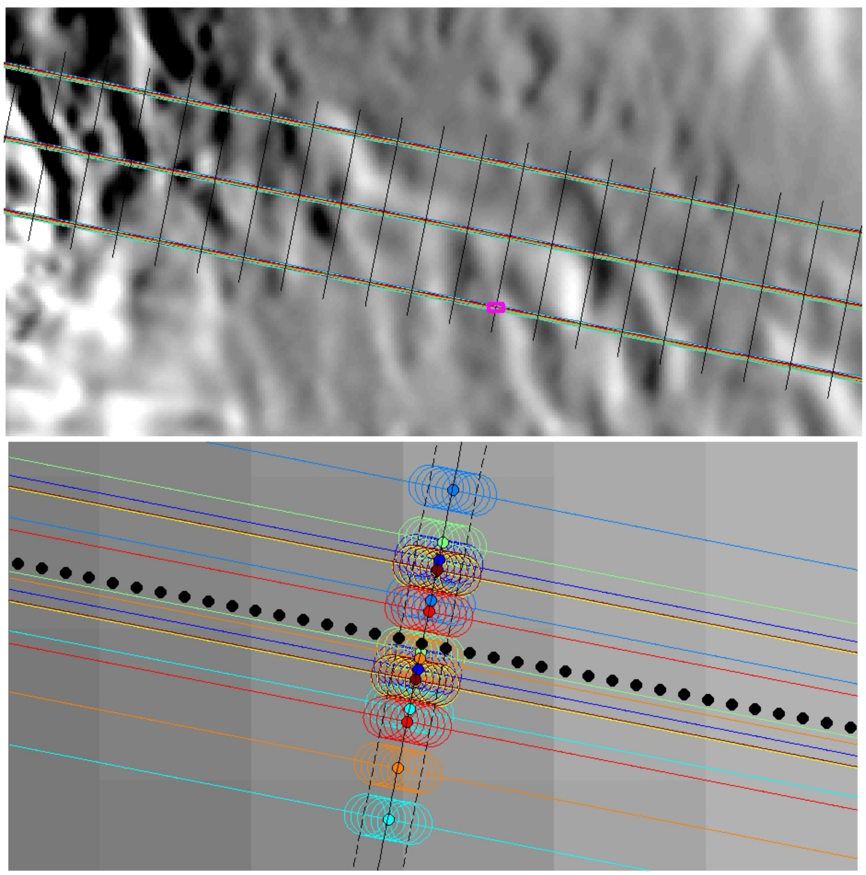
-k, the pair number, numbered from left to right across the satellite swath;

-l, the beam number within the pair, numbered from left to right;

-m, the reference point number, counted from the equator crossing of the RGT.

An along-track repeat measurement for a segment is made up of segments with the same *j*, *k*, and *m*, meaning that the track, the pair, and the along-track coordinates of the measurements are the same. Each cycle, *i*, contributes measurements from two beams, with different *l* values, to the repeat; these different measurements allow the across-track slope to be constrained independently from the height change, and the along-track segment fitting procedure allows us to correct for the along-track slope. Both ATL03 and ATL06 use this segment numbering scheme; however, ATL06 segments are 40 m long and overlap their neighbors by 50%, while ATL03 segments are 20 m long and are disjoint. ATL06 segments are defined as including PE from pairs of adjacent ATL03 segments, and are numbered to match the second of the two, so that ATL06 segment *m* includes ATL03 segments *m* and *m-1*.

Figure 3-4. Example PE selection



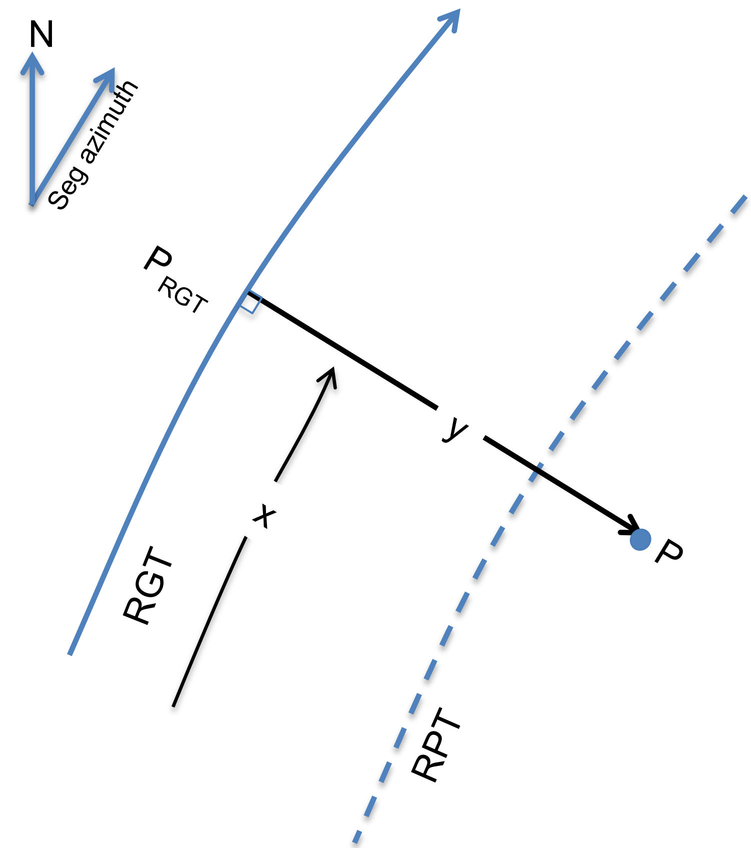
Selecting PEs for a reference point. Top: GT locations for eight simulated repeat measurement of track 188 (colored lines). Black lines are plotted every 2 km in the along-track coordinate x. Bottom: selected footprint locations for a reference point on PT 3 (circles, every 10th shown). Lines and circles are color coded by repeat. Solid points show reference-point locations, dashed lines show the 40-m along-track extent of the segments, filled circles show segment centers. Background image from (Scambos and others, 2007)

3.3.2 Local Coordinate Systems

To select the PEs associated with each reference point, the height data are grouped in local coordinates. The local coordinate system is defined in the ATL03G ATBD. Briefly, the coordinate system is defined separately for each RGT with an x coordinate that follows the RGT, starting at its equator crossing going north. The y coordinate is measured perpendicular to the x coordinate and is positive to the left. Thus, the x coordinate runs from zero to around forty thousand km for each track, the y coordinate runs from approximately -3.3 km for the right beam pair to approximately 3.3 km for the left beam pair, although its values may be larger if ATLAS is pointed off nadir.

To calculate along-track coordinates for any point $\boldsymbol{P}$ adjacent to an RGT, we define the x coordinate to be equal to the x coordinate of the nearest point on the RGT, $\boldsymbol{P}_{RGT}$. The y coordinate is equal the distance between $\boldsymbol{P}$ and $\boldsymbol{P}_{RGT}$, measured to the left of the along-track direction (Figure 3-5). This calculation is carried out for each PE in ATL03: The x coordinate for each PE is equal to the sum of the ATL03 parameters */geolocation/segment_dist_x* and */heights/dist_ph_along*. The y coordinate is equal to the ATL03 *dist_ph_across* parameter. Our reference points are defined to be equal to the start of the first ATL03 segment, so that ATL06 segment m encompasses all PE from ATL03 segments $m-1$ and m.

Figure 3-5. RGT coordinates



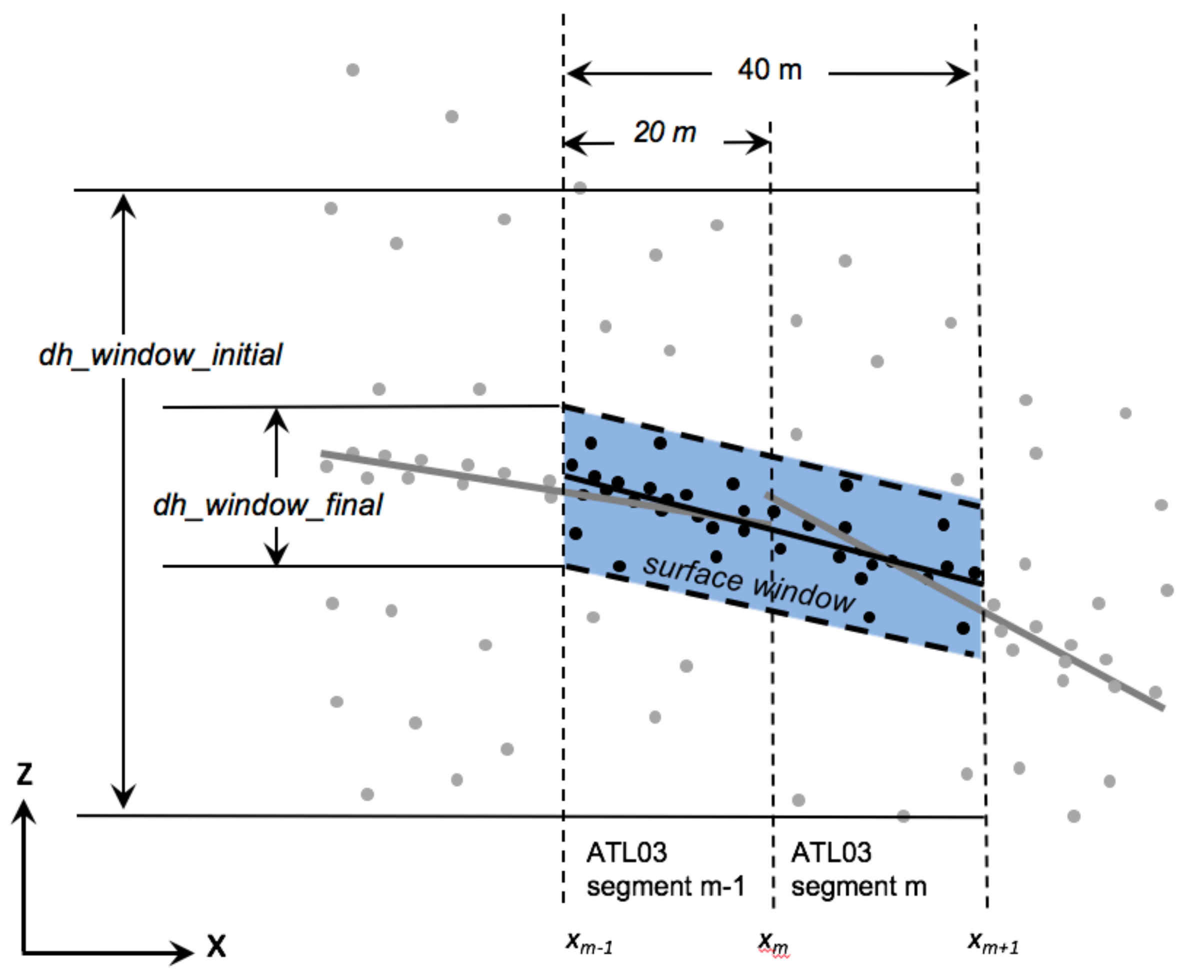
Along-track coordinates for point P relative to its RGT. The track drawn is ascending, in the northern hemisphere, so x is increasing from south to north. Y at P is negative.

The AL06 along-track coordinate for each segment is given by the parameter *x_atc*. The across-track coordinate is given by *y_atc*, and the angle between the along-track vector and local north is given in the parameter *seg_azimuth*. To allow easy referencing between ATL06 and ATL03, we provide the number for the second ATL03 segment in each ATL06 segment in the variable *segment_id*.

3.3.3 Parameters describing selected PEs

ATL06 heights and slopes are estimated by piecewise-linear fits to PEs within each overlapping 40-m segment. Since ATL06 segments are 40-meters long and overlap by 50%, we can collect the photons for each segment, *m*, by selecting all ATL03 PE that have *segment_id* equal to *m-1* or *m*.

Figure 3-6 Segment fitting



Along-track segments fit to PE heights (points) as a function of along-track x for segment *m* (black line) and neighboring segments *m+1* and *m-1* (gray lines). PEs selected during ATL06 processing for reference point *m* are shown in black.

The initial PE selection is shown in Figure 3-6. ATL03 data give a ground-finding confidence flag that indicates whether each PE was detected high confidence (SNR > 100, flag value of 4), medium (100 < SNR < 40, flag value of 3) low confidence (SNR < 40, yet still passes threshold test, flag value of 2), or is included because it falls within 10m of the detected surface (flag value of 1).

An initial surface window is valid if it contains at least 10 PE, and if the along-track distance between the first and last PE is greater than 20 m. This ensures that there are enough PE to determine both the height and slope of the segment. We define three possible sources for signal-selection data:

1. ATL03 confident PE (*signal_selection_source=0)*: PE with *confidence_flag* values > 1 (low or better confidence)
2. All ATL03 detected PE (*signal_selection_source=1)*: PE with *confidence_flag* flag values ≥ 1 (including low or better, and pad PE).
3. A backup signal-finding algorithm (*signal_selection_source=2)*

3.3.3.1 Setting the surface window based on ATL03 flagged PE.

If sources 1 or 2 define a valid surface window, we calculate the slope of that window using an initial least-squares fit to *h* as a function of *x* for the flagged PE. Based on the slope of this window, we calculate *sigma_expected* using equation 1, and calculate the robust spread of the residuals for the flagged PE (correcting for the background PE rate), *r_flagged*. If ATL03 confident PE define a window (case 1), the minimum surface window size, w_{min}, is set to 3 m, and if ATL03 confident PE do not define a window but the combination of ATL03 detected and pad PE do (case 2), w_{min} is set to 10 m. The initial surface window, *w_surface_window_initial* is then set to *max(*$\mathrm{w}_{\mathrm{min}}$, 6 *sigma_expected*, 6 *r_flagged*). The residuals for all of the segment PE are then calculated, and PE with residuals within ±*w_surface_window_initial/2* are selected and passed on to the iterative along-track fitting.

3.3.3.2 Setting the surface window using the backup signal-finding algorithm

If any ATL03 PE are detected but they do not define a window or if no ATL03 PE are present, a backup algorithm is used. First, if any ATL03-flagged PE are present, the along-track slope of the initial window is set to zero, its width is set to 10 m, and it is centered vertically on the mean height of the flagged PE. If the PE within this window fail the along-track-spread test or the ten-PE test, then PE within 40 m along track of the reference point are examined to find the 10-meter-high by 80-meter-long window, centered on the reference point, containing the largest number of PE. Typically, there will be a range of center heights whose PE counts are not significantly different from the maximum; if the maximum count is C_{max}, then any window with a count greater than C_{max}-$\mathrm{C}_{max}^{1/2}$ will be included. The initial window will extend from 5 m below the minimum of these centers to 5 m above the top of these centers, and its length is set to 40 m. If this best window does not contain a good distribution of PE (i.e. more than 10 PE, with a horizontal spread greater than 20 m), the segment is considered invalid. If $\mathrm{C}_{\mathrm{max}}$ is less than 16 (the number of PE that would be detected in an 80-meter long window with a signal strength of 10 PE/40 m, minus one standard deviation), no PE are selected, and the signal selection is marked as invalid.

Table 3-1 *signal_selection_source* values

Value	Meaning
0	Signal selection succeeded using ATL03 confident-or-better flagged PE
1	Signal selection failed using ATL03 confident-or-better flagged PE but succeeded using all flagged ATL03 PE
2	Signal selection failed using all flagged ATL03 PE, but succeeded using the backup algorithm
3	All signal-finding strategies failed.

The *signal_selection_source* parameter describes the success or failure of each step in this process, and Table 3-1 describes the meaning of each value. For each signal-selection algorithm that was attempted, the *signal_selection_status_confident, signal_selection_status_all,* and *signal_selection_status_backup* parameters in the *segment_quality* group give details of the success or failure of each part of the algorithm. The *signal_selection_source* parameter is provided for all segments (successful or not) in the *segment_quality* group, and is provided for segments for which at least one pair has an elevation in the *fit_statistics* subgroup.

Table 3-2 Status parameters for signal-selection algorithms

Signal_selection_status_confident	
0	Signal selection succeeded using ATL03 low-or-better confidence PEs
1	Signal selection using ATL03 low-or-better confidence PEs failed the 20-meter-spread test
2	Signal selection using ATL03 low-or-better confidence PEs failed the 10-photon-count test
3	Signal selection using ATL03 low-or-better confidence PEs failed both tests

Signal_selection_status_all	
0	Signal selection succeeded using all ATL03 flagged PEs (or algorithm not attempted)
1	Signal selection using all ATL03 flagged PEs failed the 20-meter-spread test
2	Signal selection using all ATL03 flagged PEs failed the 10-photon-count test
3	Signal selection using all ATL03 flagged PEs failed both tests
Signal_selection_status_backup	
0	Signal selection succeeded using the backup signal finder after centering the window on flagged PE (or backup signal finder not attempted)
1	Signal selection succeeded using the backup signal finder after searching for the strongest-signal window using four adjacent ATL03 segments
2	Signal selection using the backup signal finder failed the 20-meter-spread test
3	Signal selection using the backup signal finder failed the 10-photon-count test
4	Signal selection using the backup signal finder failed both tests

The final, refined window is described in the *fit_statistics* subgroups. The height of the window is given as *dh_window_final,* and the number of pulses that might contribute PE to the ATL06 segment is given in the *n_seg_pulses* parameter. Note that not all of the pulses in the segment necessarily contribute to the received PEs if the signal strength is low. We calculate *n_seg_pulses* based on the speed of the nadir point, v_{nadir}, of the spacecraft along the ground track, the pulse repetition frequency, and the nominal 40-m length of the ATL06 segment: $N_{seg_{pulses}} = PRF \times 40\ m/v_{nadir}$. This parameter has non-integer values, because it is intended to represent the expected number of pulses in each segment. There is no straightforward way to determine exactly which pulses might have targeted a particular ground segment.

3.3.4 Handling of invalid segments

Segments must pass a series of tests before their elevations are reported in the ATL06 *gtxx*/*land_ice_segments* groups. The signal selection routines must return at least 10 PE, spread over at least 20 m. Fitting does not proceed if these criteria are not met. For segments that continue to the surface window refinement routine, after the surface window refinement is complete, the final PE count and surface-window height are checked against the *snr_significance* parameter, to ensure that the probability of the measured signal-to-noise ration resulting from a random signal selection is small. Only segments with *snr_significance <0.05* (indicating that, given a random-noise input, the algorithm would converge to the calculated SNR less than 5% of the time) proceed to the next stage.

These criteria allow a significant number of low-quality segment heights to be reported in ATL06. This intended for the benefit of users who need to measure surface heights under marginal conditions. To help other users remove these segments, the *land_ice_segments/ATL06_quality_summary* parameter gives a synopsis of the parameters relevant to segment quality (Table 4-3), any one of which could indicate unusable data. The subset of segments with *ATL06_quality_summary* = 0 are unlikely to contain blunders due to signal-finding errors. This choice of parameters may reject useful elevations collected over rough, strongly sloping, or low-reflectivity surfaces and under clouds so obtain more height estimates, users may need to examine additional parameters in ATL06, or regenerate a similar flag for themselves based on a less-stringent set of parameters.

A variety of data flags are available to indicate why a particular segment does not have a reported height parameter. In many cases, the strong-beam segment in a pair will have a reported height, and the weak beam will not; in these cases, a full record is available for the weak-beam segment, providing all parameters up to the step where the fitting process failed. In cases where neither the strong nor the weak beam returned a surface height, the *segment_quality* group provides the *signal_selection_source* parameter, which will show a value of 3 if all signal-selection strategies failed. Only in cases where both segments passed the signal-selection tests but did not pass the *snr_signifcance < 0.05* test will there be an entry in *segment_quality* and no entry in the remainder of the ATL06 records.

Users wishing to apply more- or less-stringent criteria to the data than those described above can examine the refined surface window width *fit_statistics*/*w_surface_window_final,* the signal-to-noise ratio, *fit_statistics*/*snr*, the range-based-error parameter, *land_ice_segments*/*h_li_sigma* and the uncorrected reflectance, *r_eff*, to ensure that they are within expected ranges.

3.3.5 Surface-window refinement and least-squares height estimate

The ATL06 ground-finding algorithm refines the ATL03 surface detection estimate by iterative fitting of the initially-selected ATL03 PEs with the along-track segment model, rejecting PEs with large residuals to the model at each step (3.3.5.2). After the iterations are terminated, the final model height, based on this fit, *h_mean,* is used as an input to the next stage of the algorithm, in which the model residuals are used to derive corrections to the model height.

3.3.5.1 Least-squares fitting

For each segment, we first calculate a least-squares best-fitting segment to the initially selected ATL03 PEs, then use an iterative procedure based on the least-squares fit to refine this window.

Each time we perform the least-squares fit, we construct a design matrix, $\mathbf{G_0}$, from the vector $\boldsymbol{x}$, of along-track coordinates for the selected PEs:

$$\mathbf{G}_0 = [1\ \boldsymbol{x}] \tag{2}$$

The segment height and along-track slope are calculated based on $\mathbf{G_0}$ and the vector of ATL03 heights, $\boldsymbol{h}$, as:

$$[h_{fit}, \frac{dh}{dx}] = (\mathbf{G}_0^T\mathbf{G}_0)^{-1}\boldsymbol{G}_0^T\boldsymbol{h} \tag{3}$$

The residuals to this model are then calculated:

$$r_o = h - \mathbf{G_0}[h_{fit}, \frac{dh}{dx}] \tag{4}$$

3.3.5.2 Iterative ground-window refinement

The initial surface window height may be as large as 20 meters from top to bottom, larger in rough terrain or when the signal-to-noise ratio is small. This means that it may include many noise PEs mixed with the signal PEs. If included in the calculation, these will lead to large random errors in the surface slope and height. We can increase the proportion of signal PEs by shrinking the surface window, but need to avoid shrinking it so much that we lose signal PEs. To do this, we seek to find a window centered on the median height of the surface-return PEs, whose height is three times the spread of the surface PE height residuals. Because the spread and the median of the surface PEs are not initially known, we use an iterative procedure to shrink the size of the surface window, estimating the median and spread at each step.

We have two ways of calculating a value for the spread of the surface return, which we combine as part of our calculation of the width of the surface window. The first is to predict the RMS spread of the surface return using an initial estimate of the surface-slope vector and Equation 1 to give *h_expected_RMS,* assuming zero roughness. The second is to calculate it based on the spread of the residuals to the current model, σ_o. In low-signal-to-noise conditions, we include a correction for the background signal level in this calculation (described in 3.11). Since either of these might provide a good estimate of the spread of the surface PEs we take the maximum of these two values as our spread estimate. To avoid excessive trimming, we eliminate PEs only if their residual magnitude is greater than the maximum of 1.5 m and three times our spread estimate.

We initialize the iterative procedure with the PE selection described in the previous two sections. In cases where the signal selection was initialized with flagged PE (*signal_selection_source=0* *or 1)*, the iterative ground-window refinement is forced to use only PE included in the initial selection. In all other cases, iterations after the first may include PE that were not included in the initial selection, so the window may expand or migrate as iterations progress. In either case the PE that might be selected are the *selectable* PE.

At each step, we

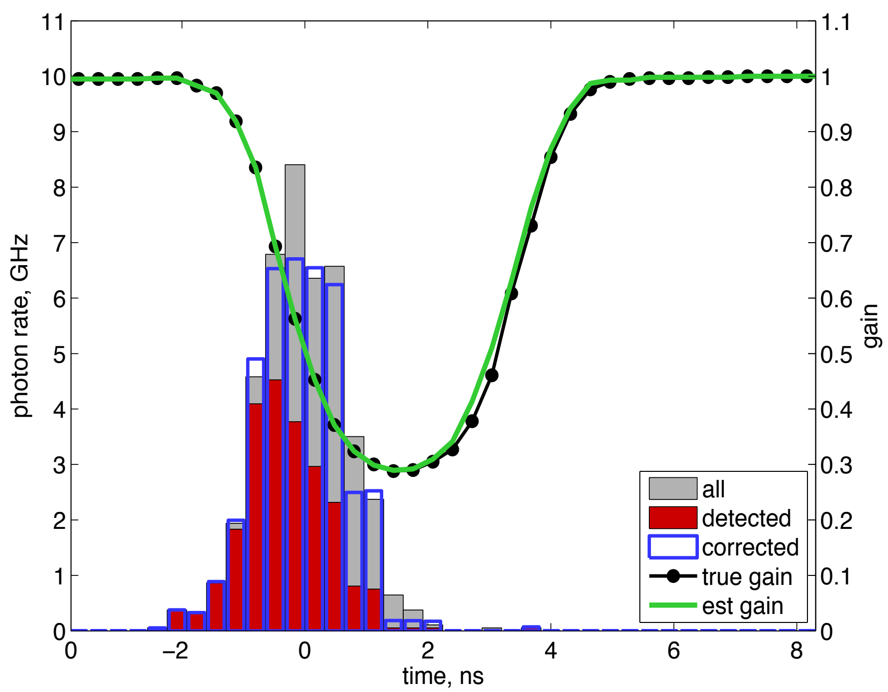
a) Perform a least-squares fit to the currently selected PEs using equation 3, giving a current model estimate, [*h_mean, dh/dx*] and residuals to the model, *r*.
b) Calculate the median and background-corrected RDE (see 3.11) of the distribution of the residuals for the selected PEs, r_{med} and σ_o, and update *h_expected_RMS* based on the current *dh/dx* estimate. The residual spread (σ_o) is limited to a maximum value of 5 m.
c) Calculate the residuals of all of the *selectable* PEs to the current model estimate, *r.*
d) Select PEs from among the *selectable* PEs for which $|r-r_{med}|$ < *H_window/2,* where *H_window*= max(6 σ_o, 6 *h_expected_RMS*, 0.75 *H_window_last,* 3 m).

The iterations are terminated if no further PEs are eliminated in a given step. If a given iteration eliminates PE such that the selected PE no longer define a window, then that step is reversed, and the iterations are terminated. The inclusion of 0.75 *H_window_last* as the minimum size of the window in each step of the calculation attempts to ensure that the calculation does not converge too fast to a spurious value of *h_mean.*

The window width after the final step is reported as *w_surface_window_final*, and the number of PEs in the window is reported as *n_fit_photons*. The final slope of the along-track segment is reported as *dh_fit_dx*. The median residual to the along-track fit is given in the parameter *med_r_fit*, and is used to convert between a mean-based height estimate for each segment and a median-based estimate.

3.4 First-Photon Bias

Figure 3-7. First-photon bias correction



Simulated rates of photon arrivals at the detector (gray) and of detected photons (red) for a strong beam over a flat surface (at 0 ns). The first-photon bias correction gives a corrected histogram (blue outline) and an estimate of the effective detector gain (green). The actual effective gain of the detector (black) is shown for comparison.

The first-photon bias (FPB) results from an inherent problem with the photon-counting detectors selected for ATLAS. For a short time, t_{dead}, after an individual pixel of each detector detects a photon, it cannot detect another. This means that photons early in a ground return are more likely to be detected than those later on, and, for a symmetric return-photon distribution, the mean surface height estimate is biased upwards, an effect that is largest for more intense pulses and for pulses from flat surfaces where the return energy is concentrated in a short period of time. Note that for ATLAS's asymmetric transmit pulse, the first-photon bias may result in either positive or negative height errors, because for small roughness values, the FPB suppresses detection the early, intense part of the waveform, while the tail of the waveform is unaffected, resulting in a negative height bias. For larger roughness values, FPB affects the tail and the peak more equally, and the bias becomes positive. For clarity, we will describe modeling results using a simulated symmetric Gaussian transmit pulse, but the corrections provided on the ATL06 products may have either positive or negative signs.

For ATLAS, t_{dead} is quite short, at approximately 3.2 ns, and there are multiple pixels in each detector (16 for the strong beams, 4 for the weak), to which photons are assigned at random as they reach the detector, resulting in fewer photons reaching each pixel while it is inactive. Despite this, up to several cm of bias may be observed for flat bright surfaces. Figure 3-7 shows simulated instantaneous photon rates for photons incident on the detector, and of detected photons for returns from a flat, smooth surface for a strong spot, under moderately saturated conditions (1.2 photons per pixel per pulse), aggregated over 40 m. Background PE are not included in the simulation, but their effect is likely to be minor, because their contribution to the total PE count is, in strong-signal conditions, a small fraction of the total, and the correction is negligible if the signal is not strong.

We have found that we can generate a correction for the first-photon bias based on a model of the detector for PEs aggregated over a 40-m ground-track segment. In this algorithm, we generate a histogram representing the distribution of heights around the ground return for the segment, as represented by the histogram of PE residuals to the best-fitting sloping segment model. We then estimate the effective gain of the detector, a function that represents the probability that a photon would have been detected if it reached the detector. We use this function to correct the received histogram to an estimate of the histogram of all the photons, detected and undetected. Statistics of this histogram are used to improve estimates of the surface height.

Using the residuals to the best-fitting segment in this calculation assumes that each pulse experiences the same distribution of photon-arrival times, shifted in time by the along-track surface slope, so that a typical distribution can be found by correcting for the along-track slope. If the surface slope or the reflectance has strong variations within a segment this assumption will fail, but for segments where the correction is large (i.e., in the interior of the ice sheets), it should not introduce large errors because ice-sheet surfaces are typically very homogeneous.

3.4.1 Mathematical Description for the First-Photon Bias

The photon distribution incident on the detectors is written as a function of t_i-t_g, where t_i is the PE time and t_{gi} is the time of the ground return. In practice, this is calculated as t_i-t_{gi}=$-r\ c/2$, where r is the height residual to the best-fitting segment. We can express the histogram over N PE times as:

$$N(t;\ t_i - t_{gi}) = \sum_{i=1:N} \sum_{t_i - t_{gi} \in (t, t+\Delta t]} 1 \tag{5}$$

Only some of these photons are detected: After a photon hits a detector, that detector cannot detect another photon until it becomes active, after receiving no photons for a time t_{dead}. This can be expressed by a function giving the status of each pixel for each pulse at time t:

$$A(t, p, pixel) = \begin{cases} 1 & if\ pixel\ \ is\ active\ at\ time\ t\ for\ pulse\ p \\ 0 & if\ pixel\ is\ inactive\ at\ time\ t\ for\ pulse\ p \end{cases} \tag{6}$$

The detected photon distribution is then:

$$N_d\left(t;\ t - t_g\right) = \sum_{i=1:N} \sum_{t_i - t_{gi} \in (t, t+\Delta t]} A(t_i - t_{gi}, pixel_i, P_i) \tag{7}$$

If the photon distribution in $t\text{-}t_g$ is constant over the pulses and over all pixels, then we can write:

$$N_d\left(t - t_g;\ \Delta t\right) = G\left(t - t_g\right) N\left(t - t_g;\ \Delta t\right) \tag{8}$$

Where:

$$G\left(t - t_g\right) = \frac{1}{N_{pulses} N_{pixels}} \sum_{pulses, pixels} A(t - t_g) \tag{9}$$

This function is effectively a gain for this collection of pulses. It ranges between zero, when all pixels are inactive, and one, when all pixels are active. The detector gain is shown by the black line in Figure 3-7. It falls rapidly from unity to about 0.3 during the early part of the surface return, then recovers gradually over a period slightly longer than t_{dead}, about 3.2 ns.

3.4.2 Correction Formulation for the First-Photon Bias

We implement the gain correction based on channel dead-time estimates from ATL03 and a histogram of residual times relative to the best-fitting segment model from 3.3.5.2, truncated by ±*h_window_final*/2. We represent the deadtime for the detector with the mean deadtime for all channels in the detector, and assume that all pixels (and channels) have identical sensitivity. Although the algorithm's function does not depend strongly on the spacing of the histogram bins, our test software has used a bin spacing of 0.05 ns. We express the timing for the correction as a function of time relative to the ground-return time, under the assumption that for an entire segment, the return shape will be consistent relative to the ground-return time:

$$\tau = t - t_g \tag{10}$$

Our strategy in this calculation is to correct an initial histogram of PE arrivals for the effects of detector dead time (G<1) by dividing $N_d(\tau)$ by $G(\tau)$:

$$N_{est}(\tau;\ \Delta t) = \frac{1}{G(\tau)} N_d(\tau, \Delta t) \tag{11}$$

To correct waveforms for the effects of dead time, we can use an *a posteriori* estimate of G(τ) calculated with a simple model of the detector. In this model, we calculate a detected distribution, N_d, as the histogram of PE arrivals relative to the ground bin for a single-segment (40 m) section of track. For each bin in the histogram, we then determine the average number of pixels in the detector that were inactive. This is calculated:

$$P_{dead}(\tau) = \frac{number\ of\ photons\ in\ [\tau - t_{dead}, \tau)}{N_{pix} N_{puses}} \tag{12}$$

The estimated gain is then *1-P_{dead}*. This calculation can be carried out efficiently by convolving the histogram of residuals with a rectangular window of height $1/N_{pix}N_{pulses}$, and shifting the result by the width of the window.

For our simulated example (in Figure 3-7) the recovered gain (green) is approximately equal to the true detector gain; this example is fairly typical of other simulations of this process, where the estimated gain is usually within a few percent of the true gain. There are visible differences between the corrected photon-timing histogram (blue) and the incident photon histogram, but the effects of these variations on the recovered heights are relatively small and have approximately zero bias.

3.4.3 Statistics Derived from the First-Photon-Bias Correction

The output of the gain estimation is a corrected histogram of height differences relative to a reference surface. Statistics of this histogram (e.g. its vertical centroid, its median) can be calculated as they would for the uncorrected PE heights. Since these statistics are calculated on the histogram of uncorrected photon residuals, their values give the correction relative to the mean of the PE heights. Thus, to calculate the corrected mean or median surface height, we add the gain-corrected mean or median of the residuals, respectively, to the uncorrected mean height. Because we expect the transmitted pulse to be skewed, we expect the median height correction to be much larger than the mean height correction.

3.4.3.1 Mean Height Correction

The mean height correction based on the corrected histogram is:

$$fpb\_mean\_corr = \sum \frac{N_{corr,i}}{N_{tot}} dz_i \qquad 13$$

Here dz_i are the bin centers of the histogram of the PE residuals (i.e. the difference between the PE heights and the linear segment fit. The error in the mean correction is found using the error propagation formula for a centroid, assuming that the measured PE counts are Poisson distributed and ignoring the error in the gain estimate. For each bin in the corrected histogram, the corrected count at that bin has an error:

$$\sigma_{N,corr,i} = \frac{N_{0,i}^{1/2}}{G_i} \qquad 14$$

The error in the mean height based on the corrected counts is then:

$$\sigma_{fpb-corr} = \left[\sum \left(\sigma_{N,corr,i} \frac{dz_i - fpb\_corr}{N_{corr,tot}}\right)^2\right]^{1/2} \qquad 15$$

3.4.3.2 Median Height Correction

The median correction and its error are calculated from the CDF (Cumulative Distribution Function) of the corrected histogram as a function of dz:

$$CDF(dz_0) = \sum_{dz_i < dz_0} \frac{N_{corr,i}}{N_{corr,tot}} \qquad 16$$

The median of the corrected histogram is found by interpolating into *dz* as a function of CDF(*dz*) at an abscissa value of 0.5:

$$median\ fpb = CDF^{-1}(0.5) \qquad 17$$

Because CDF is a function of the residuals to the linear segment-fit model, the median calculated in this way gives an offset relative to *h_mean*.

The uncertainty of the median interpolated from the CDF is the slope of the inverse function of CDF*(dz)* with respect to CDF times the statistical uncertainty in the CDF at the median point:

$$\sigma_{med} = \left.\frac{dz}{dCDF}\right]_{CDF=0.5} \sigma_{CDF}(h_{med}) \qquad 18$$

The statistical uncertainty in the CDF achieves half its total variance at the median, so we can calculate its uncertainty at the median as:

$$\sigma_{cdf}(dz_{med}) = \left[\frac{1}{2}\sum \frac{\sigma^2_{N,corr,i}}{N^2_{tot,corr}}\right]^{1/2} \qquad 19$$

We estimate the slope of the CDF based on the 60th and 40th percentiles of dz, calculated from the CDF of dz, and noting that 20% of the residuals should fall within this range. The error in the median correction is then:

$$fpb\_md\_corr\_sigma = \frac{dz_{60} - dz_{40}}{0.2} \sigma_{cdf}(dz_{med}) \qquad 20$$

For both the mean and the median corrections, the error calculated in this way gives the total error in the surface height due to the Poisson sampling in the data. It does not take into account the effects of the along-track distribution of the photons, as the propagated least-squares error (equation 19) does, so the error in the final, corrected height measurement (*h_li_sigma*) is the maximum of *sigma_h_mean* and *fpb_med_corr_sigma*. Note that neither the combined error nor the median error calculated above are rigorous estimates of the error guaranteed to work under all circumstances. However, numerical experiments have shown that these error estimates match the RMS spread of recovered values to within ~10% for numbers of PEs greater than ~20. For smaller numbers of PE, the error estimates may be up to 20% too small.

3.4.3.3 Corrected Return Count

The corrected number of returned photons is calculated:

$$fpb_{N\_photons} = \sum N_{corr} \qquad 21$$

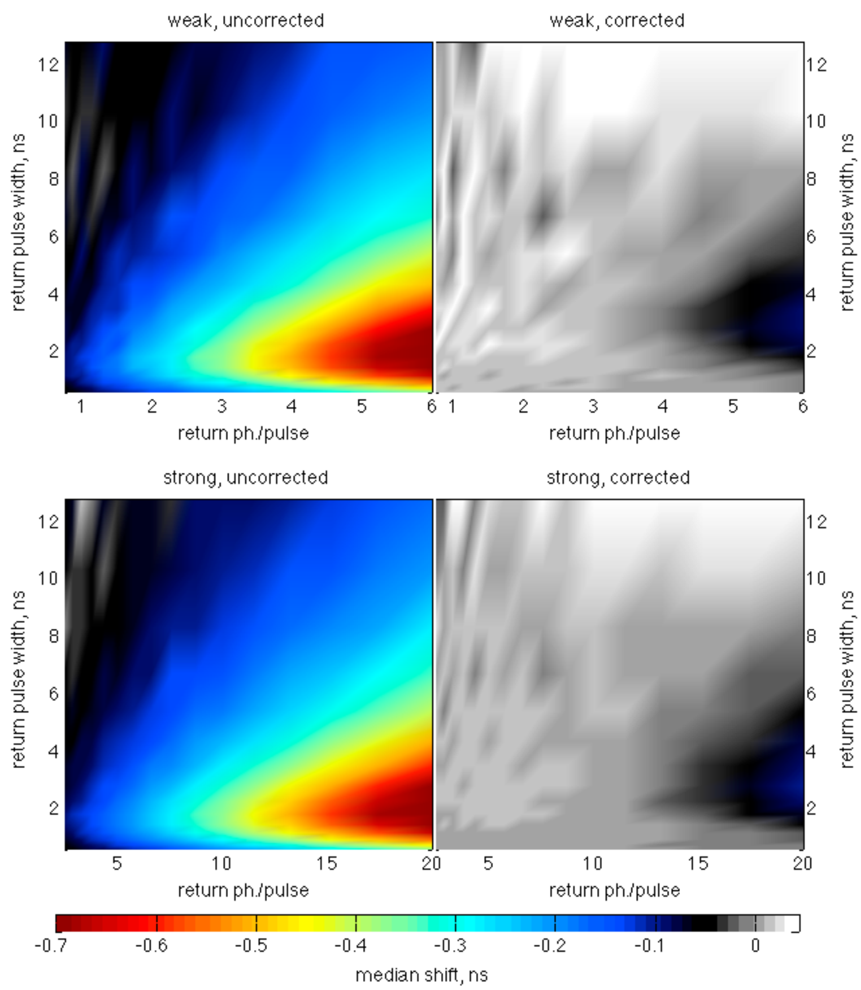
This sum is carried out over the ground window calculated during ground-bin refinement (3.3.5.2). This is similar to the dead-time correction on ATL03.

3.4.3.4 Correction Validity

The correction should provide accurate height and signal-strength corrections as long as there are at least a few active detector pixels during each time increment. If the estimated detector gain for a segment falls below 2/(*N_seg_pulses x n_pixels*), the correction values are set to their invalid value (*NaN*), so that any value that uses these corrections (e.g. *h_li, fpb_n_corr)* will also be marked invalid.

3.4.3.5 Accuracy of the first-photon bias correction

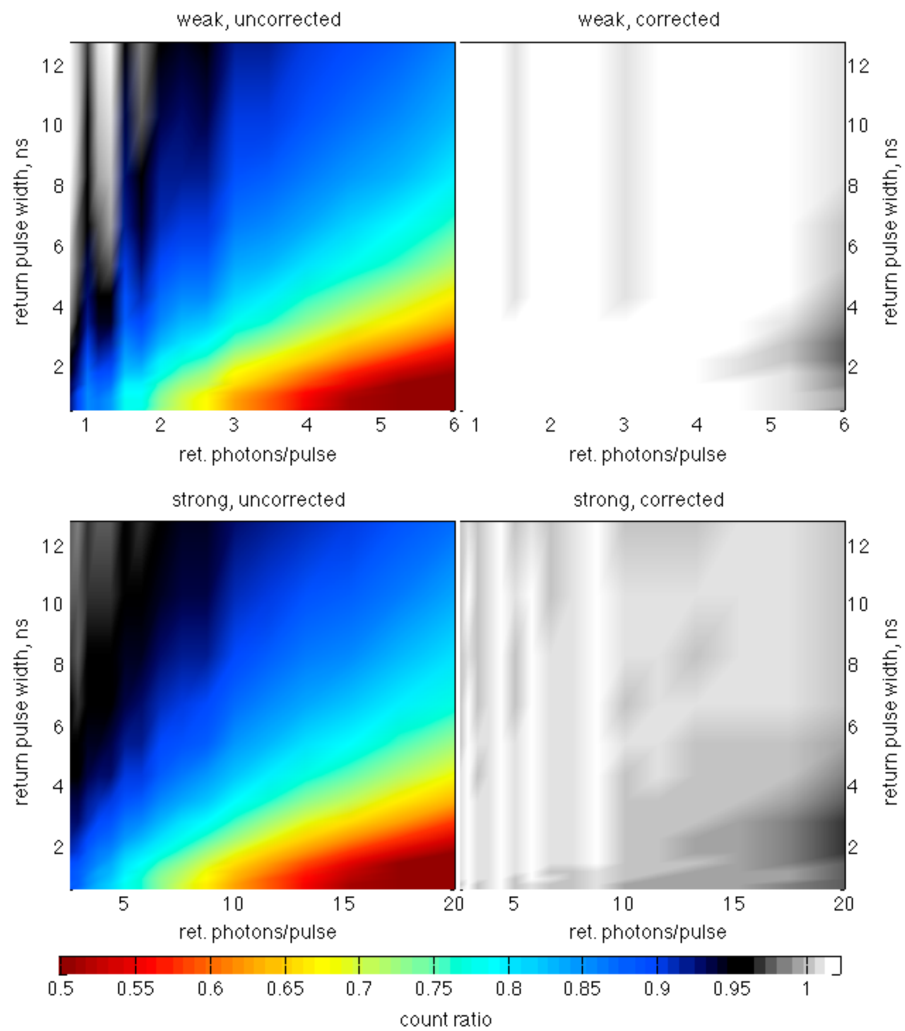
Figure 3-8. Accuracy of first-photon bias correction elevation recovery



Shifts in the median range for weak-beam (top) and strong-beam (bottom) returns for a range of incident photon counts and return-pulse σ values, for a symmetric Gaussian return. At left are the shifts for the uncorrected returns, at right are the shifts after correction. Note that the colorscale runs from bright colors, indicating large biases, to grayscale, indicating small biases.

We assess the potential accuracy of this calculation with a simple simulation of elevation recovery for a strong and a weak ATLAS beam. For each realization of this simulation, we generate random arrival times for a collection of N_{inc} incident return-pulse photons, with standard deviation σ_{inc}. These photons are assigned at random to detector pixels (4 pixels for a weak beam, 16 for a strong beam) and are labeled as detected or undetected based on the detector model described in 3.4 with a dead time of 3.2 ns. Based on these PE times, we then calculate a corrected arrival-time histogram as described in 3.4.2 and calculate statistics for this distribution as described in 3.4.3.

Figure 3-9. Accuracy of first-photon-bias-correction signal strength recovery



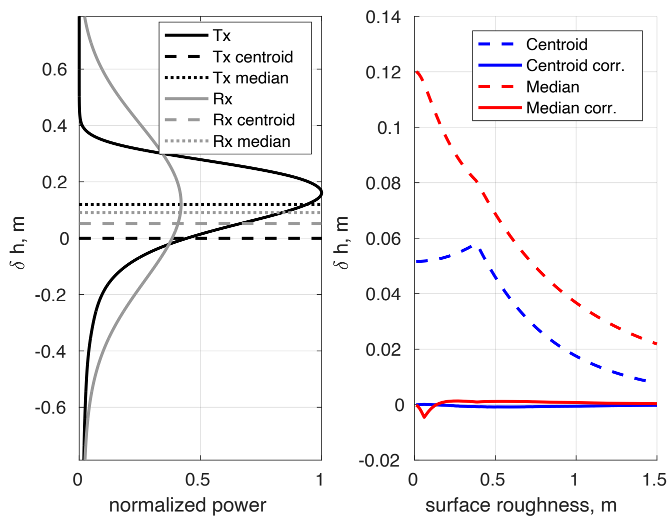
Ratios of the count of weak-beam (top) and strong-beam (bottom) PEs to the number of incident photons for a range of incident photon counts and return-pulse σ values. At left are the ratios for the uncorrected returns, at right are the ratios after correction.

Results of this simulation are shown in Figure 3-8 and Figure 3-9. For the strongest simulated returns, with around two photons per pulse per detector pixel, uncorrected time biases are as large as -0.7 ns, corresponding to positive elevation biases of about 0.1 m. For these returns, only about 60% incident photons are detected. For expected return strengths, of 0.8 photons per pulse per pixel, elevation biases are smaller, around -0.2 ns, and about 85% of incident photons are detected. The largest elevation errors come for return-pulse widths of around 2 ns, and the largest loss of signal photons happens for the smallest pulse widths and the strongest returns.

Applying the correction removes the majority of the bias, both for return times and for signal strengths. Corrected returns have much smaller time biases, accurate to 0.1 ns (1.5 cm) for the strongest (2 photons/pixel/pulse) returns, and 0.02 ns (0.03 cm) for expected (0.8 ph/pixel/pulse) return strengths. Corrected PE counts are within 2% of the incident counts.

3.5 Transmit-pulse shape correction

Figure 3-10. Transmit-pulse-shape correction



Transmit-pulse shape correction example. Left: Transmit (Tx) waveform from a prototype ATLAS laser and a simulated return (Rx) from a rough (0.25 m RMS) surface. The Tx pulse is aligned so that its centroid is at 0 ns (black dashed line), the medians of the Tx and Rx pulses are shown by dotted gray and black lines, respectively, and the centroid of the truncated Rx pulse is shown by a dashed gray line. Right: average bias between the centroid of the Tx pulse and the median and centroid of the windowed Rx pulse, both with (solid) and without (dashed) the transmit-pulse-shape correction applied.

The ATL06 surface-fitting routine and the ATL06 first-photon bias correction both give estimates of the median height of the surface for each segment, relative to the centroid of the transmit pulse, for a 'windowed' collection of photons of limited vertical extent (typically ±1.5 m around the median height). However, the ATL03 PE heights are calculated relative to an estimate of the centroid of the entire transmit pulse. Because the transmitted pulse is not symmetric in time around its centroid, its median is different from its mean, and the centroid of any truncated subset of the photons from this pulse will have a nonzero bias relative to those from the full waveform. This introduces a potential bias in ATL06 height estimates.

The magnitude of the bias depends on three factors: the shape of the 'tail' of the transmitted waveform, the width of the surface window, and the effective surface roughness (i.e. the total broadening introduced by surface slope and roughness). The effects of the tail shape and the surface-window height were described previously (1.1). The effect of increasing effective surface roughness is to increase the scatter in the PEs, producing returns that are closer to symmetrical, as shown for 0.25 m noise in Figure 3-10 (left panel). This larger scatter results in return-waveform medians that have smaller biases than those from a smooth surface, and in smaller biases in the truncated-waveform centroids. Figure 3-10 (right panel) shows the magnitude of biases in return centroids and medians for prototype-laser waveforms, broadened to simulate the effects surface roughness values between 0 and 1.5 meters. For each waveform, we calculated the centroid and median surface height relative to the centroid and median of the transmitted pulse, using a surface window height of a maximum of 3 m and three times the RDE of the returned PEs. The worst of the biases, for the zero-roughness median, is around 15 cm, and biases decrease with increasing roughness. The bias in the centroid is smaller than that of the median, but both are large relative to other expected instrumental biases.

We have found that we can correct for this effect by modeling expected return-pulse shapes and calculating the biases for these shapes, then subtracting the bias from the measured height estimates. The model is based on transmitted-waveform shapes measured periodically during the ICESat-2 orbit using the transmitter-echo-pulse (TEP). Using this TEP waveform and the width of the return, we estimate the extent to which reflection from the sloped, rough surface has broadened the return, and smooth the TEP waveform to broaden it to the same width. We then truncate the broadened synthetic waveform around its mean using the surface window determined in 3.3, then calculate the median and centroid of the broadened, truncated waveform. This gives corrections to the median and mean surface heights.

Note that at the time of writing of this document the relationship between the absolute values of the photon times measured in the TEP and the transmit times of the lasers has not been established. On-orbit calibration exercises and further analysis of pre-launch calibration data should be helpful in this regard, but for now, we take the TEP as a measurement of the shape of the waveform, not the timing of the transmission. Accordingly, we shift the time values on the TEP measurements obtained from ATL03 so that the centroid of the signal photons arrival times is equal to zero, and assume that this shifted TEP represents the transmit pulse.

To estimate the broadened transmit-pulse shape, we begin with an estimate of the transmitted pulse shape derived from ATL03, $P_{tx}(t)$, and $RDE(t_i)$, our estimate of the degree to which the distribution of surface returns, t_i, has been spread by its reflection from a rough or sloping surface:

$$\sigma_s^2 = \max\left((0.01\,ns)^2, RDE(t_i)^2 - RDE\left(P_{tx}(t)\right)^2\right) \qquad 22$$

The $max((0.01\ ns)^2, ...)$ function here is included to ensure that the broadening estimate is positive. From this we generate an estimate of the surface broadening function S(t):

$$S(t) = \exp\left(-\frac{t^2}{2\sigma_s^2}\right) \qquad 23$$

The estimated broadened pulse shape, $P_B(t)$ is the temporal convolution of $P_{tx}(t)$ and S(t):

$$P_B(t) = P_{tx}(t) * S(t) \tag{24}$$

We apply a windowing function, $W_s(t)$, to account for the truncation of the surface return during the ground-bin-selection process:

$$W_s(t) = \begin{cases} 0 & \left|t - mean\left(P_B(t)\right)\right| > h\_window\_final/2 \\ 1 & \left|t - mean\left(P_B(t)\right)\right| \leq h\_window\_final/2 \end{cases} \tag{25}$$

The height correction for the median based on this waveform estimate is then:

$$dh_{tx} = \frac{c}{2} median_t(P_B(t)W_s(t)) \tag{26}$$

Here $median_t()$ represents the temporal median of a function:

$$median_t\left(f(t)\right) \equiv t \text{ such that } \int_{-\infty}^{t} f(t')dt' = \frac{1}{2}\int_{-\infty}^{\infty} f(t')dt' \tag{27}$$

The correction for the mean is identical, but uses the mean instead of the median in equation 26. Figure 3-10 shows that after applying this correction, the remaining bias in the median and mean heights is less than 3 mm. The value calculated in equation 26 is included in the standard surface-height estimate, *h_li*, and is provided in the *tx_median_corr* and *tx_mean_corr* fields in the *bias_correction* parameter subgroup.

3.6 Signal, Noise, and Error Estimates

Before we can calculate the error in the retrieved surface height, we must form estimates of relative contributions of signal and noise PEs to the observed PE count. Under ideal conditions, when the signal level is high and the background count rate is low, few noise PEs will be present among those selected by editing process described above. However, under cloudy conditions when the sun is above the horizon this will often not be true, and it is important that the error estimates reflect the potential presence of background PEs.

3.6.1 Background PE rate

The background PE rate (*bckgrd* in the *geophysical* subgroup) is derived from the ATL03 parameter */bckgrd_atlas/bckgrd_rate*, and is derived from a 50-shot, 200Hz count of PE within the ATLAS signal-finding window, corrected for the number of PE detected by the ATL03 ground-finding algorithm. In general, we expect this parameter to be sufficiently accurate to allow us to predict the number of PE within 10 m of the ground to a precision of better than 10 PE/segment.

The expected background rate, *E_bckgrd*, is also predicted based the solar elevation, assuming a flat, Lambertian surface at the ground. The calculation of this parameter is described in the ATL07 ATBD, section 4.2.3.1. This parameter, when compared against the measured *bckgrd*, is a potential indicator of the surface reflectance and cloud properties.

3.6.2 Signal PE count

The total number of PEs selected in the window, as a function of the number of signal PEs, the background rate, the number of pulses in the window, and the background window height is:

$$N_{tot} = N_{sig} + N_{BG} \tag{28}$$

The number of background PEs in the window has a mean value:

$$N_{BG} = 2\, N_{pulses} h_{window}\, BGR/c \tag{29}$$

Subtracting the two gives an estimate of the number of signal PE, N_{signal}. Because the number of background PE is a Poisson random variable, the calculated N_{signal} may be less than zero in weak-signal conditions. The ratio between the number of signal and noise photons is reported as *fit_statistics/snr*.

To help distinguish high-quality surface returns from returns that are likely a result of signal-finding blunders, we provide the *fit_statistics/snr_significance*, which gives the probability that in the absence of any real ground signal, a segment with at least the observed SNR would be found by the ATL06 signal-selection routine, for the initial range of heights, *h_range_initial* and background rate *bckgrd*. If ATL03 detected photons were used in the signal selection (*signal_selection_source* of 0 or 1, or *signal_selection_status_backup* of 0), *h_range_input* is equal to the range of photon heights. Otherwise it is set to the full range of PE heights provided from ATL03 for the segment. The values of *snr_significance* are calculated from a look-up table based on 1,000,000 realizations of random noise for background-noise values, *bckgrd_table*, between 1 and 10 MHz, and for initial window sizes, *w_table*, between 3 and 80 meters. For each set of random-noise PE, the backup signal-selection algorithm is run to select the input PE for the iterative ground-window refinement routine (3.3.5.2), which is then run to convergence, and the final SNR is recorded. Then, for each value of *bckgrd_table* and *w_table,* the probability of reporting a segment with an SNR value greater than a set of values between -10 and 10, in steps of 0.1, is calculated, and the value is stored in *F_table*. To find *snr_significance* for each segment, we interpolate into *F_table* as a three-dimensional linear function of *h_range_input*, *bckgrd*, and *snr* for that segment.

3.6.3 Per-Photon Errors

Noise PEs are vertically distributed throughout the window with a standard deviation of approximately

$$\sigma_{BG} = 0.287\, h_{window} \tag{30}$$

where the factor 0.287 equals the standard deviation of a uniform random variable on a unit interval.

The signal PEs have an approximate skewed Gaussian distribution, whose width depends on the transmit-pulse duration, the surface roughness, the surface slope, and the footprint width, as described in equation 1, with additional broadening possible due to atmospheric or subsurface scattering. For ice-sheet surfaces and near-vertical beams we assume that the angle between the beam and the surface slope is equal to the magnitude of the surface slope. The total standard deviation of the surface return heights, $\sigma_{photon,est}$ is then:

$$\sigma_{photon,est} = \left(\frac{N_{BG}\sigma_{BG}^2 + N_{signal}\sigma_{signal}^2}{N_{BG} + N_{signal}}\right)^{1/2} \qquad 31$$

With the exception of the surface roughness, all of the quantities needed for this equation are estimated from the data: the slope spreading is estimated from the along-track component of the surface slope and the transmitted pulse width using equation 1, and the background and signal PE counts are estimated from the total number of PEs and the background rate. If we assume the roughness to be zero, and neglect atmospheric and subsurface scattering errors, equation 31 gives a minimum error estimate. An alternate estimate of the per-PE error is the vertical spread of PEs relative to the along-track fit, *h_rms_misfit*. We combine these two estimates by setting our error estimate, σ_{photon}, to the maximum of *h_rms_misfit* and $\sigma_{photon,est}$.

3.6.4 Propagated Height Errors:

Given the established per-PE error, σ_{photon}, the error propagation for the linear fitting equation gives an estimate of the covariance matrix for the fit (Menke, 1989):

$$\mathbf{C}_{\text{fit}} = \left((\mathbf{G}^{\text{T}}\mathbf{G})^{-1}\ \mathbf{G}^{\text{T}}\right)\left((\mathbf{G}^{\text{T}}\mathbf{G})^{-1}\ \mathbf{G}^{T}\right)^{\text{T}}\sigma_{photon}^2 \qquad 32$$

The height error estimate, *sigma_h_mean* is the square root of the upper-left element of $\mathbf{C}_{\text{fit}}$. This error is combined with the sampling error estimated during the first-photon-bias calculation to give the total surface ranging error, *h_li_sigma*. The error in the along-track slope *sigma_dh_fit_dx* , is equal to the square root of the lower-right element of $\mathbf{C}_{\text{fit}}$.

3.6.5 Uncorrected reflectance

The uncorrected reflectance gives the ratio of the measured return energy to the energy expected from a white surface, through a nominal clear atmosphere (Yang and others, 2013). Following the strategy outlined in the ATL09 ATBD, we calculate:

$$r_{eff} = \frac{\pi E_{RX} r^2 F}{N_{seg\_pulses} E_{TX} A T_{opt}} \qquad 33$$

Here E_{RX} is the received energy, r is the range to the surface, A is the telescope area, and T_{opt} is a factor that combines the optical efficiency of the instrument optics and the detector sensitivity. F is a calibration factor that will be determined and maintained as part of the atmospheric science operations. E_{TX} is the transmitted energy per pulse from the ATL03 parameter *tx_pulse_e*. We calculate E_{RX} based on the number of returned PE as:

$$E_{RX} = (fpb_N - N_{BG})\frac{hc}{\lambda} \qquad 34$$

Here fpb_N is the dead-time-corrected segment signal photon count, N_{BG} is the background-photon count (from equation 29), and hc/λ is the energy received per photon. Note that this is the same calculation as equation 4.7 in the ATL09 ATBD, except that we use the ATL06 first-photon-bias-corrected photon count, instead of the correction factor used in ATL09. For an atmospheric transmittance 0.95, we expect to see r_{eff} of about 0.88 over unit-reflectance surfaces.

3.7 Across-track slope calculation

After the iterative editing process is complete, the across-track slope is computed for the pair based on the first-photon-bias-corrected median heights for the two segments:

$$\frac{dh}{dy} = \frac{h_{LI,R} - h_{LI,L}}{y_{ATC,R} - y_{ATC,L}} \qquad 35$$

If only one beam has returned a height, then *across_track_slope* is set to *invalid* for both beams.

3.8 Subsurface-Scattering Bias

The subsurface-scattering, or volume-scattering, bias comes from photons that experience multiple scattering within the snow or ice before returning to the satellite. Ice absorbs green light only weakly, with attenuation lengths of tens of meters or more, but ice grains in firn and air bubbles in ice both scatter green light strongly (Warren and others, 2006). While most photons from an ATLAS pulse are expected to exit the surface of a firn pack within a fraction of a nanosecond, others will likely be delayed significantly, producing a long tail on the histogram of return times. Averaging return times of PEs from this tail with PEs from the surface return leads to a delay in the mean PE return time, and a downward bias in the apparent surface height. The median surface height is modestly less sensitive than the mean, because it less sensitive to outlying data values far from the central peak of the return distribution. This error and its temporal variability is expected to be small for fine-grained snow surfaces such as those found on the Antarctic Plateau and in central Greenland, but it may be more significant in coastal areas where seasonal snow melt leads to large temporal variations in the surface grain size.

The magnitude of the subsurface-scattering bias delay depends in part on the scattering density of the snow and its bulk absorbance, both of which are determined by the density and grain or bubble size close to the surface, and on the impurity content of the snow or ice. Since none of these properties may be known at the time of ATLAS processing, each must be determined independently using external information about the snow, such as meteorological model output or infrared reflectance data.

We do not expect to be able to offer an accurate correction for this effect with our current understanding of the process. This remains an area of active research.

3.9 Atmospheric-Scattering Bias

A second important source of bias in ATLAS height measurements may come from atmospheric scattering of the down-going laser pulse. Scattering by ice particles in the atmosphere redirects much of the light through small angles, often less than about one degree. These photons may fall outside the field of view of the ATLAS detectors, in which case they will be lost and will have no impact on altimetry beyond attenuation of the received pulse, or they may reflect from the surface within the field of view, in which case they may then be detected by ATLAS. However, because their down-going path was longer than the assumed straight down-and-back path assumed in the PRD model, they will give erroneously long ranges, and therefore low surface heights. This effect is increasingly severe for thicker clouds, which scatter more photons, and for clouds closer to the surface, where photons scattered through large angles may still remain in the field of view.

Under cloudy conditions, the received pulse contains a mixture of scattered and unscattered photons, yielding a tail of delayed photons on the downward side of the return pulse; mean and median delays for a segment's aggregate PEs will depend on the relative fraction of the two groups of photons, and the mean path delay per photon. This process has been modeled and found to produce 1-cm level biases on ATLAS height retrievals under most circumstances (Yang and others, 2011) but since the bias may be correlated over large spatial scales it may have a non-negligible impact on continental-scale surface-change retrievals.

As is the case with the subsurface-scattering bias, parameters relating to a possible correction must be determined from datasets external to ATLAS, likely from atmospheric models that give an estimate of the cloud optical depth and the particle size. Potential corrections and data editing strategies for this effect remain an active topic of research.

3.10 Segment geolocation

After ground-window refinement we calculate the final location of the segment. The segment location is defined as the reference-point location plus the across-track unit vector times the mean across-track coordinate of the selected PEs.

To calculate the latitude and longitude of each segment, including the offset between the segment and the reference point, we use the latitude, longitude, and along-track distance provided by ATL03 for the selected PE. We assume that latitude and longitude for the selected PE in the segment are linear functions of along-track distance, and fit a linear function, f_{lat}, to the PE latitudes, and a second linear function, f_{lon}, to the PE longitudes, each as a function of x-x_0. The intercepts of these functions give the segment latitude and longitude.

Geolocation errors in the along- and across-track direction are calculated based on the ATL03 parameters *sigma_geo_AT,* and *sigma_geo_XT* and the radial orbit error, *sigma_geo_r*.

With the surface-slope vector and the geolocation estimate we can calculate the geolocation contribution to the uncertainty in the surface height:

$$\sigma_{geo,h} = \left(sigma_{geo,r}^2 + \left(sigma_{geo,AT} \frac{dh}{dx} \right)^2 + \left(sigma_{geo,XT} \frac{dh}{dy} \right)^2 \right)^{1/2} \tag{36}$$

This value is reported in the *land_ice_segments* group as *sigma_geo_h*, and the contributing *sigma_geo_r, sigma_geo_xt, and sigma_geo_at* are reported in the *ground_track* group.

3.11 Noise-corrected robust estimators of spread

Many of the parameters in this document are based on ordinal statistics. These statistics use the percentiles of a distribution, which are defined based on the cumulative distribution function (CDF) of the distribution. We define the CDF of a discrete sample of values S as:

$$C(x; S) = \frac{the\ number\ of\ values\ in\ S\ that\ are\ less\ than\ x}{the\ number\ of\ values\ in\ S} \tag{37}$$

For a binned distribution (e.g. a histogram or a probability distribution function), $C(x; D(x_o))$, we define the CDF as

$$C\left(x; D(x_0)\right) = \frac{\int_{x_1}^{x} D(x')dx'}{\int_{x_1}^{x_2} D(x')dx'} \tag{38}$$

Here are x_1 and x_2 are the bounds over which the distribution is defined. The percentiles of a distribution are found by calculating the inverse function of the CDF of the distribution:

$$p(r; D) = C^{-1}\left(\frac{r}{100}; D\right) \tag{39}$$

Thus the median of a distribution D is:

$$Median(D) = x\ such\ that\ \ C(x; D) = 0.5 \tag{40}$$

We also define the robust dispersion estimate (RDE) of a distribution as

$$RDE(D) = \frac{p(0.84; D) - p(0.16; D)}{2} \tag{41}$$

This is analogous to the standard deviation of a normal distribution, which is equal to half the difference between its 84th and 16th percentiles, but is less influenced by outlying background values.

In most cases, distributions of ATLAS PEs include a mix of signal and noise PEs. In these cases, the noise PEs and the signal PEs both contribute to the distribution D. We expect the noise PEs are generally uniformly distributed, so we can assume that

$$C(x; D) = \frac{BGR(x - x_1) + \int_{x_1}^{x} D_{signal}(x')dx'}{\int_{x_1}^{x_2} D(x')dx'} \tag{42}$$

Here D_{signal} is the distribution of the signal PEs, and BGR is the background PE rate, in units of x^{-1}. We can solve this for C_{signal}:

$$C(x; D_{signal}, BGR) = \frac{\int_{x_1}^{x} D_{signal}(x')dx'}{N_{signal}} = \frac{\int_{x_1}^{x} D(x')dx' - \frac{BGR(x - x_1)}{N_{total}}}{N_{signal}} \tag{43}$$

Here $N_{total} = \int_{x_1}^{x_2} D(x')dx'$ and $N_{signal} = N_{total} - (x_2 - x_1)BGR$.

Estimating the percentiles of D_{signal} is complicated because C(x; D_{signal}, bckgrd) generally does not have an inverse function in x. However, if we evaluate C(x; D_{signal}, bckgrd) for a set of values, x_i, we can find x_{LT}, the largest value of x_i for which C(x; D_{signal}, bckgrd) < r/100 and x_{GT}, the first value of x_i for which C(x; D_{signal}, bckgrd) >r/100, and interpolate linearly into [x_{LT}, x_{GT}] as a function of [C(x_{LT}; D_{signal}, bckgrd), C(x_{GT}; D_{signal}, bckgrd)] at the point r/100.

The above procedure defines the background-corrected percentiles of a distribution. Based on this we define the noise-corrected median of a distribution, which we designate: median(D; bckgrd). We define the noise-corrected RDE of a distribution somewhat differently from its uncorrected counterpart. For low-noise distributions, the standard deviation of the population can accurately be estimated as half the difference between its 16th and 84th percentiles. In the presence of significant noise, the standard deviation can be estimated more accurately based on the difference between the 25th and 50th percentiles of the distribution, divided by a correction factor of 1.349, equal to the width of the central 50% of a normalized Gaussian distribution.

The surface-window-refinement procedure in section 3.3.5 uses least-squares fitting and the RDE to progressively narrow the surface window. This procedure will not converge under all circumstances. Consider an initial surface window spanning from $-H_i/2$ to $H_i/2$, with noise rate *R* (in PE/m) , containing *s* signal PEs at the center of the window. The normal (non-background-corrected) RDE will find a spread of:

$$\hat{\sigma} = 0.34\, H - \frac{s}{R} \tag{44}$$

If s is small, $\hat{\sigma} \approx 0.34\, H$ so the three-sigma interval will have a width of 2.04 H, and the refinement will not converge. Convergence requires $6\hat{\sigma} < H$, or:

$$s > 1.73HR \tag{45}$$

For a background rate of 10MHz (0.067 PE/m) and a weak beam (three surface PE per pulse), the procedure will converge if H < 26 m. For a strong beam (10 PE per pulse), it will converge if H<86m. The convergence intervals become smaller in proportion to the signal PE count as the surface return is weakened by cloud attenuation or by reduced surface reflectance.

The noise-corrected RDE and median improve on the performance of their uncorrected counterparts, but their performance is limited by the accuracy of the signal-level estimate. The estimate of N_{signal} has an approximate error of $(N_{pulses}\,(HR+s))^{1/2}$ due to the Poisson statistics of the PE. In contrast to the non-robust RDE and median, the process works increasingly well as more shots are aggregated, because N_{signal} increases in proportion to N_{pulses}, while its error increases in proportion to $N_{pulses}^{1/2}$. If we require that $N_{pulses}\, s > a\sigma_n$, we find convergence intervals:

$$H < \frac{N_{pulses}s^2 - a^2 s}{a^2 R} \qquad 46$$

For 10 MHz noise, 3 PE/pulse, and for 57 pulses, this gives $s > 3\sigma_n$ for H < 806 m, implying that the accuracy of the signal-level estimate will not be the limiting factor for any reasonable initial window size.

4 ATL06 DATA PRODUCT DESCRIPTION

Here we describe how the parameters appear in the ATL06 product. The ATL06 parameters are arranged by beam, and within each beam in a number of groups and subgroups. Where parameter descriptions in the ATL06 data dictionary are considered adequate, they are not repeated in this document.

4.1 Data Granules

ATL06 data are provided as HDF5 files. The HDF format allows several datasets of different spatial and temporal resolutions to be included in a file. ATL06 files contain data primarily at the single-segment resolution, divided into different groups to improve the conceptual organization of the files. Each file contains data from a single cycle and a single RGT.

Within each file there are six top-level groups, each corresponding to data from GT: *gt1l, gt1r, gt2l*, etc. The subgroups within these *gtxx* groups are *segment_quality*, *land_ice_segments,* and *residual_histogram.*

In the *segment_quality* group, the data are nearly dense, providing signal-selection and location information for every segment attempted (i.e. those that contain at least one ATL03 PE) in the granule, at the 20-meter along-track segment spacing. Datasets in this group can be used to check the geographic distribution of data gaps in the ATL06 record.

In the *land_ice_segments* group, data are sparse, meaning that values are reported only for those pairs for which adequate signal levels (i.e. more than 10 PE, *snr_significance* > *0.05*) were found for at least one segment: This means that within each pair, every dataset has the same number of values, and that datasets are pre-aligned between pairs, with invalid values (NaNs) posted where the algorithm provided a value for only one beam in a pair. Conversely, if neither beam in a pair successfully obtained a value for *h_li*, that segment is skipped for both beams in the pair. The *segment_id*, timing, and geolocation fields for the valid segments should allow the along-track structure of the data to be reconstructed within these sparse groups. For segments without valid heights that still appear on the product (because the other beam in the pair did contain a valid height) the latitude and longitude are reported for the mean location of all PE for the segment (if any PE are present) or as the location for the valid segment in the pair, displaced by the 90-meter within-pair separation (if no PE are present).

The *residual_histogram* group is at lower resolution than the other groups, giving the distribution of PE relative to the segment heights at a horizontal resolution of 200 m, or around 280 pulses. The *segment_id_list, x_atc_mean, lat_mean,* and *lon_mean* fields in this group all can be used to connect the *residual_histogram* group to the per-segment groups.

In the native format archived at the National Snow and Ice Data Center (NSIDC), each granule (file) of data contains segments from a single pass over a one-degree increment of latitude for a particular RGT, with corresponding data from all six beams. Over most of the globe, ICESat-2 travels in a roughly north-south direction, so each granule will contain approximately 111 km of data for each beam, or approximately 5660 segments. The granules containing the southernmost extent of Antarctica, south of 87S, will contain a considerably longer stretch of data, but because this area will likely be of most interest to researchers investigating continental-scale Antarctic mass balance, the additional coverage will likely be desirable. We expect that because most users will obtain their data through subsetting services provided by the NSIDC, the native granule structure will be of minor importance.

4.2 Segment_quality group

The segment_quality group contains a nearly dense record of the success or failure of the surface-finding strategies, and gives the locations of the ref erence points on the RPTs. It contains a record of the success or failure of the surface-finding strategies, and gives the locations of the reference points on the RPTs.

Locations provided within this group are for the reference points on the pair tracks, not for the segments themselves. This means that both beams in a pair will have the same location (because they are not displaced relative to the reference point), and that the actual segment locations will usually be displaced from the values in *reference_pt_lat* and *reference_pt_lon* in this group by more than 45 m in the across-track direction. The laser beam and spot numbers corresponding to the ground tracks are available in the attributes of the *ground_track* group.

Table 4-1 *Segment_quality* group

Parameter	Units	Description
delta_time	seconds	Elapsed GPS seconds since the reference epoch. Use the metadata attribute *granule_start_seconds* to compute the full GPS time.
segment_id	unitless	segment number corresponding to the second of two ATL03 segments in the ATL06 segment, counted from the RGT equator crossing
reference_pt_lat	degrees	Latitude of the reference segment location on the RPT
reference_pt_lon	degrees	Longitude of the reference segment location on the RPT
record_number	unitless	For those segments that have adequate signal strength, this parameter gives the record for the pair within the other groups in the granule.
signal_selection_source	unitless	Indicates the last algorithm attempted to select the signal for ATL06 fitting, see table Table 3-1.

		A value of 3 indicates that all algorithms failed.

4.2.1 Signal_selection_status subgroup

This subgroup includes the *Signal_selection_status_confident, Signal_selection_status_all,* and *Signal_selection_status_backup* parameters. Their values are described in Table 3-2. Its density structure matches that of the *segment_quality* group.

4.3 *Land_ice_segments* group

The primary set of derived ATL06 parameters are given in the *land_ice_segments* group (Table 4-2). This group contains geolocation, height, and standard error and quality measures for each segment. This group is sparse, meaning that parameters are provided only for pairs of segments for which at least one beam has a valid surface-height measurement. This group contains the *bias correction*, *fit statistics*, *ground_track*, and *geophysical* subgroups, which all have the same sparsity structure as the *land_ice_segments* group.

Table 4-2 *land_ice_segments* group

Parameter	Units	Description	Defined
ATL06_quality_summary	Unitless	Flag indicating: 0: No likely problems identified for the segment 1: One or more likely problems identified for the segment	4.3
delta_time	Seconds	Elapsed GPS seconds since the reference epoch. Use the metadata attribute granule_start_seconds to compute the full gpstime.	Interpolated to the segment center from ATL03
h_li	Meters	Standard land-ice segment height determined by land ice algorithm, corrected for first-photon bias, representing the median-based height of the selected PEs	Equation 47

h_li_sigma	Meters	Propagated error due to sampling error and FPB correction from the land ice algorithm	Equation 48
sigma_geo_h	meters	Total vertical geolocation error due to PPD and POD, including the effects of horizontal geolocation error on the segment vertical error	3.10
latitude	degrees north	Latitude of segment center, WGS84, North=+	3.10
longitude	degrees east	Longitude of segment center, WGS84, East=+	3.10
segment_id	counts	Segment number, counting from the equator. Equal to the *segment_id* for the second of the two 20-m ATL03 segments included in the 40-m ATL06 segment	ATL03

The standard surface height will be given on the ATL06 product as *h_li*. This height is the segment-center height obtained from the along-track slope fit, with the mean-median correction applied so that it represents the median surface height for the segment. By default, *h_li* will be corrected for all height increments in the *geophysical* parameter group except for the ocean tide, the equilibrium tide, and the dynamic atmosphere correction (*dac*); this includes earth, load, and pole tides, and troposphere corrections. Since these parameters are included in the standard ATL03 PE height, no correction is applied at the ATL06 stage. Using the names for product variables:

$$h\_li = h\_mean + fpb\_med\_corr + tx\_med\_corr \tag{47}$$

Tide and troposphere corrections may be removed from *h* by adding the values provided in the ATL06 *geophysical* group. The correction values for the waveform-based corrections are provided in the *bias_correction* group, so that users may convert, for example, from a median-based height estimate to a mean-based estimate.

The errors in the standard land-ice height product are calculated as the maximum of the median error (calculated during the first-photon-bias correction) and the linear-fit error (calculated in 3.6), ignoring errors in the tidal and atmospheric corrections.

$$h\_li\_sigma = \max\left(sigma\_h\_fit,\ fpb\_med\_corr\_sigma\right) \tag{48}$$

This value does not include the effects of geolocation errors on the height estimate, because while the components of *h_li_sigma* should be uncorrelated at the segment-to-segment scale, the geolocation errors are likely to be correlated on much longer scales. The vertical component of the geolocation error, as calculated from the surface-slope vector and the mean horizontal geolocation accuracies of the selected PEs are given in parameter *sigma_geo_h* (see 3.10). The error on a single segment height measurement taken independently of all adjacent measurements should be $(\mathit{h\_li\_sigma}^2 + \mathit{sigma\_geo\_h}^2)^{1/2}$. Averaged over several tens of segments with a consistent surface slope, the error should approach *sigma_geo_h,* but the relative scatter between individual adjacent segments should be *h_li_sigma.*

The geolocation of the segment is given in geographic coordinates by parameters *latitude* and *longitude*. These each represent the horizontal centers of the segments. The corresponding along-track coordinates are given in the *ground_track* group as *x_atc* and *y_atc.*

The *land_ice_segments* group includes the *ATL06_quality_summary* parameter, which indicates the best-quality subset of all ATL06 data. A zero in this parameter implies that no data-quality tests have found a problem with the segment, a one implies that some potential problem has been found. Users who select only segments with zero values for this flag can be relatively certain of obtaining high-quality data, but will likely miss a significant fraction of usable data, particularly in cloudy, rough, or low-surface-reflectance conditions. Table 4-3 gives the parameter values needed for *ATL06_quality_summary* to be reported as zero. The last of these characteristics, the vertical density of photons, helps remove the effects of a common problem where the ATL03 photon selection identifies a cloud top as a likely surface return. In these cases, ATL06 can converge to a large (10+ m) vertical window containing tens of signal photons. Requiring a minimum ratio between the number of photons and the height of the window eliminates most clouds, and eliminates only a few returns from rough or steep surfaces.

Table 4-3 Segment characteristics for *ATL06_quality_summary* to be zero

Characteristic	Threshold	Description
h_li_sigma	< 1 m	Errors in surface height are moderate or better
snr_significance	< 0.02	Surface detection blunders are unlikely
signal_selection_source	<=1	Signal selection must be based on ATL03 photons
n_fit_photons / w_surface_window_final	>1 PE /m for weak beams, > 4 PE/m for strong beams	The vertical density of photons in the final surface window.

4.3.1 geophysical subgroup

The *geophysical* group (Table 4-4) contains tidal and atmospheric corrections that may be added to or removed from *h_li,* and inferred atmospheric properties that may be used to determine whether the elevation of a given segment might be affected by atmospheric forward scattering. Note that the *neutat_delay* parameter and all *tide_* parameters in this group are applied by default except for *tide_ocean* and *dac* (dynamic atmosphere correction).. The sign of the parameters is such that adding the parameter value to *h_IS* removes the correction (for applied corrections) and subtracting the parameter includes the correction (for *tide_ocean*). These parameters are interpolated from the corresponding ATL03 parameters for the 'nominal photons', interpolated as a piecewise linear function of along-track distance to the segment centers. This group is sparse, meaning that parameters are provided only for pairs of segments for which at least one beam has a valid surface-height measurement.

The ocean-tide value (*tide_ocean*) and dynamic atmosphere correction(*dac)* are provided to allow interested users to correct for tides and the inverse-barometer effect over ice shelves. These parameter are not applied because the locations of ice-sheet grounding lines (defining the inland extent of floating ice shelves) are not always precisely known, and may change over time. Different users will want to apply the ocean-tide model to different areas within the grounding zone.

This group also include parameters related to solar background and parameters indicative of the presence or absence of clouds. Some of these parameters are derived from the ATLAS atmospheric channel, and should help identify segments strongly affected by clouds or blowing snow: parameters *cloud_flg_asr* and *cloud_flg_atm* give estimates of the probability of clouds between ATLAS and the ground, based on the apparent surface reflectance and on atmospheric backscatter, respectively. Their values are described in the ATL09 ATBD, and should be evaluated against the standard that cloud optical thickness greater than 0.5 in the lower 3 km of the atmosphere is required to produce a substantial altimetry error. (Yang and others, 2011) . Note that over surfaces other than bright snow (e.g. over blue ice or dirty snow) the *cloud_flg_asr* may indicate clouds when none are present.

Blowing snow has a larger potential to produce altimetry errors, and has been assigned its own flag; the estimated height of a detected blowing-snow layer is given in *bsnow_h*, which is set to zero if no such layer can be detected; the confidence with which a blowing-snow layer can be detected or ruled out is given in *bsnow_conf*. For both flags, cautious users may require a value of 0 or 1 (clear with high/medium confidence) but under sunlit conditions, neither flag may clearly indicate cloud-free conditions. The estimated optical thickness of blowing snow layers, if found, is given in *bsnow_od*.

Table 4-4 *geophysical* subgroup

Parameter	Units	Description	Defined
bckgrd	Hz	Background count rate, derived from the ATL03 50-shot average, interpolated to the segment center.	Interpolated from ATL03
bsnow_conf	unitless	Blowing snow confidence. -3=surface not detected; -2=no surface wind;-1=no scattering layer found; 0=no top layer found; 1=none-little; 2=weak; 3=moderate; 4=moderate-high; 5=high; 6=very high	ATL09
bsnow_od	unitless	Blowing snow layer optical depth	ATL09
bsnow_h	meters	Blowing snow layer top height	ATL09
cloud_flg_asr	counts	Cloud flag (probability) from apparent surface reflectance. 0=clear with high confidence; 1=clear with medium confidence; 2=clear with low confidence; 3=cloudy with low confidence; 4=cloudy with medium confidence; 5=cloudy with high confidence	ATL09
cloud_flg_atm	counts	Number of layers found from the backscatter profile using the DDA layer finder.	ATL09
layer_flag	counts	This flag is a combination of multiple flags (cloud_flag_atm, cloud_flag_asr, and bsnow_con) and takes daytime/nighttime into consideration. A value of 1 means clouds or blowing snow are likely present. A value of 0 indicates the likely absence of clouds or blowing snow.	ATL09

e_bckgrd	Hz	Expected background count rate based on sun angle, surface slope, for unit surface reflectance	Calculated following ATL07
msw_flag	unitless	Multiple Scattering warning flag. The multiple scattering warning flag (ATL09 parameter msw_flag) has values from -1 to 5 where zero means no multiple scattering and 5 the greatest. If no layers were detected, then msw_flag = 0. If blowing snow is detected and its estimated optical depth is greater than or equal to 0.5, then msw_flag = 5. If the blowing snow optical depth is less than 0.5, then msw_flag = 4. If no blowing snow is detected but there are cloud or aerosol layers detected, the msw_flag assumes values of 1 to 3 based on the height of the bottom of the lowest layer: < 1 km, msw_flag = 3; 1-3 km, msw_flag = 2; > 3km, msw_flag = 1. A value of -1 indicates that the signal to noise of the data was too low to reliably ascertain the presence of cloud or blowing snow. We expect values of -1 to occur only during daylight.	ALT09
r_eff	unitless	Effective reflectance, uncorrected for atmospheric effects.	Equation 33
solar_azimuth	degrees_east	The direction, eastwards from north, of the sun vector as seen by an observer at the laser ground spot.	ATL03 solar_azimuth parameter, interpolated to the segment center from the reference photons
solar_elevation	degrees	Solar Angle above or below the plane tangent to the ellipsoid surface at the laser spot. Positive values mean the sun is above the horizon, while negative values mean it is below the horizon. The effect of atmospheric refraction is not included. This is a low-precision value,	ATL03 solar_elevation parameter, interpolated to the segment center from the reference photon

		with approximately TBD degree accuracy.	
tide_earth	meters	Earth tide	Inherited from ATL03
dac	meters	dynamic atmosphere correction	Inherited from ATL03
tide_load	meters	Load Tide	Inherited from ATL03
tide_ocean	meters	Ocean Tide	Inherited from ATL03
tide_pole	meters	Pole Tide	Inherited from ATL03
tide_equilibrium	meters	Equilibrium tide	Inherited from ATL03
neutat_delay_total	meters	Total neutral atmospheric delay correction (wet+dry)	Inherited from ATL03

In some circumstances, the estimated background rate may also give an indication of cloud conditions. The estimated background rate is provided in parameter *bckgrd*, which may be compared with the background rate expected for a unit-reflectance Lambertian surface, with a slope equal to the measured surface slope, *E_bckgrd*. In sunlit conditions, these parameters together allow an estimate of the total sub-satellite reflectance. The effective, uncorrected surface reflectance, *r_eff*, based on first-photon-bias-corrected PE count and the range to the ground, may be compared to these numbers; if *bckgrd* is approximately equal to *e_bckgrd*, the atmosphere and the surface must together have a reflectance close to unity; if *r_eff* is approximately equal to unity, this indicates that the surface below the satellite is likely snow, and likely cloud free; if *bckgrd* is approximately equal to *e_bckgrd* and *r_eff* is small, clouds must be present, and if *bckgrd* is less than *e_bckgrd,* the surface must be dark, and, most likely not snow covered.

Also included in this group are the solar azimuth (*solar_azimuth*) and elevation (*solar_elevation*), used in estimating the expected background rates.

4.3.2 ground_track subgroup

The *ground_track* subgroup (Table 4-5) contains parameters describing the GT and RGT for each segment, as well as angular information about the beams. All the components needed to identify a given segment's orbit number, reference track, pair track, and beam number are given, along with the azimuth and elevation of the beam relative to the ellipsoid surface normal. The orientation of the RPT with respect to local north is given in *seg_azimuth.*

Note that in land-ice products, the ground tracks and pair tracks are numbered separately from the laser beams: the ground tracks are numbered from left to right relative to RGT, and the ground track number is associated with group names within the product: From left to right, they are *gt1l, gt1r, gt2l, gt2r, gt3l, and gt3r*. The laser beams are numbered from left to right relative to the spacecraft flight direction. When the spacecraft is flying with its x axis pointing forwards, the beam numbers are in the same order (beam numbers 1…6 correspond to tracks *gt1l…gt3r*), but when it is in the opposite orientation, the laser-beam numbers are reversed relative to the ground-track numbers (beam numbers 1…6 correspond to tracks *gt3r…gt1l*).

This group is sparse, meaning that parameters are provided only for pairs of segments for which at least one beam has a valid surface-height measurement. Data-set attributes give:

-the reference ground track number

-the correspondence between laser beam numbers and ground tracks

-the cycle number

The RMS accuracy of the horizontal geolocation for the segment is described by the geolocation error ellipse, which is calculated based on the PE-medians of the ATL03 parameters *sigma_geo_xt*, *sigma_geo_at* and *sigma_geo_r*. The along-track and across-track coordinates of the segments are provided by parameters *x_atc* and *y_atc*.

Table 4-5 *ground_track* subgroup

Parameter	Units	Description	Derived
ref_azimuth	degrees	The direction, eastwards from north, of the laser beam vector as seen by an observer at the laser ground spot viewing toward the spacecraft (i.e., the vector from the ground to the spacecraft).	ATL03
ref_coelv	degrees	Coelevation (CE) is direction from vertical of the laser beam as seen by an observer located at the laser ground spot.	ATL03
seg_azimuth	degrees	The azimuth of the pair track, east of local north	3.1.2.2
sigma_geo_at	meters	Geolocation error in the along-track direction	3.10
sigma_geo_xt	meters	Geolocation error in the across-track direction	3.10
sigma_geo_r	meters	Radial orbit error	3.10

x_atc	meters	The along-track x-coordinate of the segment, measured parallel to the RGT, measured from the ascending node of the equatorial crossing of a given RGT	3.1.2.2
y_atc	meters	Along-track y coordinate of the segment, relative to the RGT, measured along the perpendicular to the RGT, positive to the right of the RGT.	3.1.2.2

4.3.3 bias_correction subgroup

The *bias_correction* subgroup (Table 4-6) contains information about the estimated first-photon bias, and the transmit-pulse-shape bias. The standard correction applied in *h_li* is *fpb_med_corr+tx_med_corr*, and its error is *fpb_med_corr_sigma*. The alternate, mean-based correction, is *fpb_mean_corr,* with error *fpb_mean_corr_sigma.* The median-based elevation, without the first-photon-bias correction, may be recovered by subtracting *fpb_med_corr* and adding *med_r_fit.* For example, users who prefer to use the mean statistics instead of the median statistics would use *h_li - fpb_med_corr –tx_med_corr+ fpb_mean_corr +tx_mean_corr* as their height estimate.

The corrected photon count is given as *fpb_n_corr*; this gives an estimate of the number of photons in the surface window as estimated during the FPB correction. The transmit-pulse-shape corrections (*tx_med_corr* and *tx_mean_corr*) are also given.

Table 4-6 *bias_correction* subgroup

Parameter	Units	Description	Derived
fpb_mean_corr	meters	First-photon bias correction to the mean segment height	3.4.3.1
fpb_mean_corr_sigma	meters	Estimated error in *fpb_mean_corr*	3.4.3.1
fpb_med_corr	meters	First-photon-bias correction giving the difference between the mean segment height and the corrected median height	3.4.3.2
fpb_med_corr_sigma	meters	Estimated error in *fpb_med_corr*	3.4.3.2

fpb_n_corr	counts	Estimated window photon count after first-photon-bias correction	3.4.3.3
med_r_fit	meters	Difference between uncorrected mean and median of linear-fit residuals	3.3.5.2
tx_med_corr	meters	Estimate of the difference between the full-waveform transmit-pulse mean and the median of a broadened, truncated waveform consistent with the received pulse	3.5
tx_mean_corr	meters	Estimate of the difference between the full-waveform transmit-pulse mean and the mean of a broadened, truncated waveform consistent with the received pulse	3.5

4.3.4 fit_statistics subgroup

The *fit_statistics* subgroup gives a variety of parameters describing the segment fit and its residuals. These parameters may be used to determine whether a particular segment is potentially usable if it is not identified as problem-free in the *land_ice_segments/ATL06_quality_summary* flag.

Table 4-7 *fit_statistics* subgroup

Parameter	units	Description
dh_fit_dx	unitless	Along-track slope from along-track segment fit
dh_fit_dx_sigma	Unitless	Propagated error in the along-track segment slope
dh_fit_dy	Unitless	Across-track slope from segment fits to weak and strong beams; the same slope is reported for both laser beams in each pair

signal_selection_source	Unitless	Flag describing the source of the information used to select the signal PE. See Table 3-1
signal_selection_source_status	Unitless	Indicates the status of the last signal selection algorithm attempted (see *signal_selection_source*). Values for this flag are given in the sections of Table 3-2.
h_mean	meters	Mean surface height, not corrected for first-photon bias or pulse truncation.
sigma_h_mean	meters	Propagated height error due to PE-height sampling error for height from the along-track fit, not including geolocation-induced error
h_expected_rms	meters	Expected RMS misfit between PE heights and along-track segment fit
h_rms_misfit	meters	RMS misfit between PE heights and along-track segment fit
h_robust_sprd	meters	RDE of misfit between PE heights and the along-track segment fit.
n_seg_pulses	counts (pulse ID)	The number of pulses potentially included in the segment (floating-point number)
n_fit_photons	counts	Number of PEs used in determining h_li after editing
w_surface_window_final	meters	Width of the surface window, top to bottom

snr	unitless	Signal-to-noise ratio in the final refined window
snr_significance	unitless	Probability that signal-finding routine would converge to at least the observed SNR for a random-noise input. Small values indicate a small likelihood of a surface-detection blunder.

4.3.5 *DEM* subgroup

This subgroup (Table 4-8) contains DEM elevations interpolated at the segment centers. It contains only three parameters: the DEM elevation (*dem_h*), the geoid height (*geoid_h*), and the DEM source (*dem_flag*). The best DEMs available in time for the ICESat-2 launch may be significantly better than those available at present (February 2015), but the best current choices are:

- For Antarctica, the REMA DEM : https://www.pgc.umn.edu/data/rema/, filtered to 40-m resolution before interpolation to the ICESat-2 segment centers, with gaps filled with ATL06 data from cycles 1 and 2.
- For the Arctic, the Arctic DEM, based on stereophotogrammetry https://www.pgc.umn.edu/data/arcticdem. The DEM should be filtered to 40-m resolution before interpolation to the ICESat-2 reference points.
- For areas outside the poles, a multi-sensor global DEM, posted at 7.5 arcsec (http://topotools.cr.usgs.gov/gmted_viewer).

This group is sparse, meaning that parameters are provided only for pairs of segments for which at least one beam has a valid surface-height measurement.

Table 4-8 *DEM sub*group

Parameter	Description
dem_h	Height of the DEM, interpolated by cubic-spline interpolation in the DEM coordinate system to the PE location
dem_flag	source for the DEM.1=Antarctic DEM, 2=Arctic DEM, 3=global DEM.
geoid_h	Geoid height, meters

4.4 *residual_histogram* group

This group contains histograms of the residuals between PE heights and the least-squares fit segment heights, at 200-meter along-track resolution. It is intended to allow visualization of the surface-return shapes, and investigation of changes in the return pulse shape or of near-surface scattering, such as that due to dense blowing snow. Each column of the histogram gives the number of PE in a set of bins distributed between -50 and +50 m around the surface. The distribution of these bins is as follows:

From 50 to 20 m below the surface, bins are spaced at 1 m

From 20 m to 10 m below the surface, bins are spaced at 0.5 m

From 10 m to 4 m below the surface, bins are spaced at 0.25 m

From 4 m to 2 m below the surface, bins are spaced at 2 cm

From 2 m below the surface to 2 m above the surface, bins are spaced at 1 cm

From 2 m to 4 m above the surface, bins are spaced at 2 cm

From 4 to 10 m above the surface, bins are spaced at 0.25 m

From 10 to 20 m above the surface, bins are spaced at 0.5 m

From 20 m above the surface to 50 m above the surface, bins are spaced at 1 m.

This distribution of bin edges gives 749 (*N_bins)* vertical bins, with 750 edges. The heights of the bin tops are given in the *bin_top_h* parameter, listed in order from bottom to top. For any bin in the histogram, the bottom elevation is equal to the top of the previous bin, and the elevation of the bottom of the bottom bin is 1 m below its top. The residuals from collections of 10 along-track ATL06 segments are combined into each histogram; because adjacent ATL06 segments overlap by 50%, only those PE within 10 m of each segment center in the along-track direction are included in the histograms. Only those segments with high-quality signals (ATL06_quality_summary =0) are included in the histogram, and a list of the *segment_id* values of included segments is provided in the group (recall that the *segment_id* for a segment corresponds to the second of the two ATL03 segments included in each ATL06 segment). To allow reconstruction of the per-pulse signal levels, the sum of the number of pulses in the valid segments is given for each histogram, and the *background_per_m* parameter is given to indicate the number of background photons expected in each vertical meter of each histogram. The expected number of photons in each histogram bin can be found by multiplying the height difference between the edges of the bin by *background_per_m*. The counts for any histogram bins that are not entirely encompassed by at least one of the two possible telemetry band window ranges are marked as invalid.

Table 4-9 Parameters in the *residual_histogram* group

Parameter	Dimensions	Description
count	N_bins x N_hist	Residual count in 1-cm bins, for PE within 10 (horizontal) m of segment centers for each histogram. Bin-top heights may be found in *residual_histogram/bin_top_h.*
delta_time	1xN_hist	Elapsed GPS seconds since the reference epoch. Use the metadata attribute granule_start_seconds to compute the full gpstime. Calculated from the mean of the *delta_time* for the segments in each histogram bin.
bin_top_h	N_bins	Height of the top of each histogram bin, listed in increasing order. The bottom of each bin is equal to the top of the next-lowest bin, and the bottom of the lowest bin is 1 m below its top
bckgrd_per_m	1xN_hist	Number of background PE expected for each vertical meter of the histogram based on the observed background rate (bckgrd)
segment_id_list	10xN_hist	Segments ids included in each column of the histogram
lat_mean	1x N_hist	Mean latitude of the segments included in the histogram
lon_mean	1x N_hist	Mean longitude of the segments included in the histogram
pulse_count	1xN_hist	Number of pulses potentially included in the histogram (pulses are counted if they are in the central 20 m of each segment, even if no PE from the pulse are selected)
x_atc_mean	1x N_hist	Mean along-track coordinate of the segments included in the histogram.

5 ALGORITHM IMPLEMENTATION: LAND ICE HEIGHT (ATL06/L3A)

This section gives detailed procedures for estimating heights from ATL03 PEs. The procedures are presented as an outline of the steps that need to be programmed to calculate the main parameters from each group; we assume that after interaction with the programming team these outlines will be updated to ensure their accuracy and consistency with the rest of this document.

5.1 Outline of Procedure

The following steps are performed for each along-track reference point:

1. PEs from the current cycle falling into the along-track bin for the along-track point are collected
2. The initial height and along-track slope are estimated for each beam in the pair
3. The heights and surface windows are iteratively refined for each beam in the pair
4. Corrections for subsurface scattering, first-photon bias, median offsets, and error estimates are calculated for each beam based on the edited PEs
5. The across-track slope is calculated

Steps 1-5 are described in the “Processing Procedure” subsection.

5.2 Input Parameters

Steps 1-6 in 5.1.1 can be calculated based on ATL03 inputs. Steps 5 and 6 require information about the background rate, which is provided with the atmospheric data

Table 5-1 lists parameters needed from ATL03 and ATL09 for generation of ATL06.

Individual PE heights, times, IDs, and geolocations are provided by ATL03. A variety of tidal and atmospheric-delay parameters are derived from subsamples of ATL03 fields or by interpolation into data tables used during ATL03 processing. Some ATL03 parameters are provided for every PE (e.g. height and horizontal position). These are averaged over the selected PEs for each segment. Others are provided for ‘reference’ photons spaced approximately every 40 m along track. For these fields, ATL06 values are interpolated as a function of along-track x from the values for the ‘nominal’ photons to the segment centers.

In addition, parameters from the atmospheric channel are used to define the blowing-snow height parameter, the blowing-snow confidence parameter, and the cloud-flag confidence parameter.

The 200-Hz background-rate parameter is used to estimate background rates for each segment, as is the 50-Hz background-rate parameter based on the full atmospheric window. An estimate of the optical depth for the 3 km above the ground and a blowing-snow height estimate and confidence flag are also calculated based on ATL09 parameters.

The transmit-pulse shape is used to correct the truncated means and medians used in estimating the surface shape to reduce potential biases in the recovered surface height.

Table 5-1. Inputs for ATL06

Parameter	Source	Description
podppd flag	*/gtxx/geolocation/podppd_flag*	Flag indicating low/high quality geolocation
Segment_ID	ATL03: /gtxx/geolocation	ATL03 segment ID
Ph_index_beg	ATL03: /gtxx/geolocation	First photon in the segment
Segment_ph_cnt	ATL03: /gtxx/geolocation	Number of PE in each segment
Segment_dist_x	ATL03: /gtxx/geolocation	Along-track distance for each ATL03 segment
Segment_length	ATL03: /gtxx/geolocation	Along-track length of each ATL03 segment.
Velocity_sc	ATL03: /gtxx/geolocation	Spacecraft ground speed
Sigma_across	ATL03: /gtxx/geolocation	across-track component of geolocation error
Sigma_along	ATL03: /gtxx/geolocation	Along-track component of geolocation error
Sigma_h	ATL03: /gtxx/geolocation	Vertical component of geolocation error
Delta_time	ATL03: /gtxx/geolocation	Time for each PE
H_ph	ATL03: /gtxx/heights	WGS-84 PE height
Lat_ph	ATL03:	PE latitude

	/gtxx/heights	
Lon_ph	ATL03: /gtxx/heights	PE longitude
Signal_conf_ph	ATL03: /gtxx/heights	Signal-classification confidence
Ph_id_channel	ATL03: /gtxx/heights	Channel number for each PE
Ph_id_pulse	ATL03: /gtxx/heights	Pulse number for the current PE
Pce_mframe_cnt	ATL03: /gtxx/heights	Major frame number for the current PE
Dist_ph_along	ATL03: /gtxx/heights	Along-track distance relative to the current segment start
Dist_ph_across	ATL03: /gtxx/heights	Along-track distance relative to the RGT
bckgrd_rate	ATL03: /gtxx/bckgrd_atlas	Background rate calculated from the 50-pulse altimetric histogram
delta_time (corresponding to bckgrd_rate)	ATL03: /gtxx/bckgrd_atlas	Time for the first shot in the 50-pulse altimetric histogram
DEM elevation	Standard DEMs	Best-available DEMs (see 4.3.5) interpolated to each segment location
Tide model values	ATL03: /gtxx/geophys_corr	Various tide-model parameters
Tep_hist	ATL03: Atlas_impulse_response/ beam_x/histogram	Transmitter-echo-pulse histogram for the strong/weak spot (should match current spot)

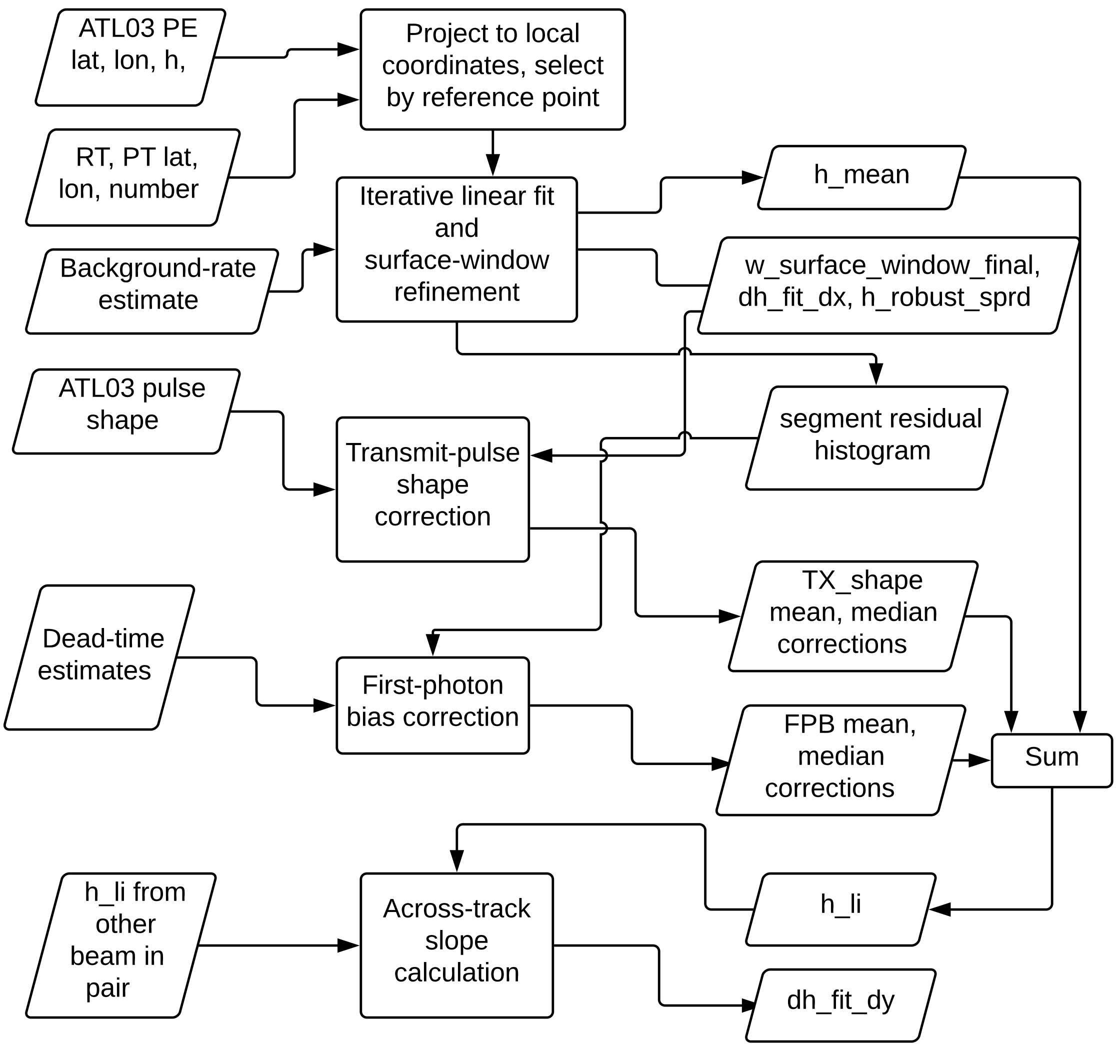
Tep_hist_x	ATL03: Atlas_impulse_response/ beam_x/histogram	Times for transmitter-echo-pulse histogram bins
Tep_bckgrd	ATL03: Atlas_impulse_response/ beam_x/histogram	Transmitter-echo-pulse per-bin background count
Tep_tod	ATL03: Atlas_impulse_response/ beam_x/histogram	Day/time for the TEP measurement used
Channel dead-time estimates	ATL03	dead-time estimates for each channel, from ATL03 parameters /atlas_impulse_response/dead_time
Blowing-snow flag	ATL09	Blowing-snow flag
Blowing-snow confidence	ATL09	Blowing-snow confidence
Cloud flag	ATL09	Cloud flag and confidence

Note that some parameters that are provided for each segment in ATL03 are needed for each PE in ATL06. For example, the along-track distance for a PE is the sum of *segment_dist_x* (provided per segment) and *dist_ph_along* (provided for each PE). To allow us to access these fields, we generate an internal *ph_seg_num* variable, based on the ATL03 *geolocation/ph_index_beg* variables, assigning all photons between the *i*-th value of *geolocation/ph_index_beg* and 1 less than the *i+1*-th value a *ph_seg_num* value of *i*. The background rate is provided in ATL03 on a 50-shot sampling interval; we convert this to the per-PE rate by interpolating as a function of *delta_time.*

5.3 Processing Procedure for Parameters

In this section, we give pseudocode for the calculation of ATL06 parameters. The flow chart for this process is summarized in Figure 5-1. The code is made up of several functions that call one another, following the process described in Section 5.1.

Figure 5-1. Flow chart for top-level ATL06 processing



Processing chain showing inputs and outputs for ATL06 fitting routine.

5.4 Top-Level Fitting Routine

This routine calls the other routines in the processing chain to derive the final heights and corrections. It corresponds to all the steps described in 3.2.

Inputs, for each beam, for ATL03 segments *m-1* and *m*:

x_PE: along-track coordinates of the land-ice PEs, meters

y_PE: across-track coordinates of the land-ice PEs, meters

h_PE: heights of the PE, meters

t_PE: times for PE.

Ice_confidence_flag : Confidence with which the PE has been identified as coming from the surface, unitless

bckgrd : estimated background PE rate for the current segment, counts/second

ch_deadtime: Deadtime estimate for each channel

x0_seg : along-track coordinate of the current reference point

bckgrd_rate: 50-shot-resolution background rate, derived from ATL03, interpolated to the center of the segment.

Spacecraft_ground_speed: The speed of the nadir point below the spacecraft as it moves along the geoid.

Podppd_flag: ATL03 flag indicating high or low quality geolocation

Outputs (repeated for left and right beams)

delta_time: time offset with respect to the beginning of the granule

h_li : land-ice height, meters

h_li_sigma : error in the ice-sheet height, meters

h_robust_sprd : ice-sheet residual robust spread, meters

h_rms_misfit : RMS residual for the residual spread, meters

n_fit_photons: The number of photons used to define the segment.

w_surface_window : width of the refined window used to select PEs, meters

h_expected_rms : expected standard deviation of PEs based on surface geometry and signal levels, meters

dh_fit_dx : along-track slope for the segment, unitless

signal_selection parameters *:* parameters indicating how the initial PE were selected

fpb_corr_mean : first-photon bias correction for the mean surface height, meters

fpb_corr_median: first-photon bias correction for the median surface height, meters

tx_median_corr: return-truncation correction to the median-based segment height

tx_mean_corr: return-truncation correction to the mean-based segment height

fpb_n_corr : corrected PE count from the first-photon bias, meters

y_seg_RGT: segment across-track coordinate

lat_seg_center: segment-center latitude *lon_seg_center*: segment-center longitude

tide and *dac* parameters: geophysical parameters that are averaged and passed on from ATL03

SNR: Estimated signal-to-noise ratio for the segment

atl06_quality_summary: Summary parameter indicating whether a problem in the segment fitting was identified

Output for both beams together:

dh_fit_dy : across-track slope, unitless

Internal variable, that is tracked through the fitting procedure:

h_ range_ input: The range of heights provided as an input to the fitting algorithm.

Parameters:

granule_start_time: the starting time of the granule

dx_seg = 40 meters

sigma_beam: sigma value for pulse surface footprint (expected to be equal to 4.25 m)

SNR_F_table: 3-d table giving the probability of finding a segment with the given SNR for noise-only inputs

PRF: Pulse repetition frequency for ATLAS (equal to 10,000 s^{-1})

Procedure:

1. Select PE for the initial fit.

1a. If the *podppd_flag* indicates degraded geolocation for any pulses, skip to the next segment.

1b. For each beam, select PE with ATL03 segment_id of *m* or *m-1*. Set *h_range_input* equal to the difference between the maximum and minimum of the PE heights. Eliminate any photons that are identified by ATL03 as part of the TEP.

1c. Set initial values for the geolocation and time parameters: set *lat_seg_center, lon_seg_center* and *delta_time* to the means of the corresponding reference photon values.

1d. Calculate *n_seg_pulses* based on the spacecraft ground speed, and the lengths of segments *m-1* and *m*: n_seg pulses=(sum of segment lengths * PRF)/*spacecraft_ground_speed*.

1e Based on the *ice_confidence_flag* values (see **PE selection based on ATL03 flags**), and assign values to *signal_selection_source, signal_selection_status_confident,* and *signal_selection_status_all.* If *signal_selection_source* is equal to 0 or 1, set *h_range input* equal to *H_win*.

1f. If both *signal_selection_status_confident* and *signal_selection_status_all* are nonzero, select PE using the **backup PE selection** routine. If *signal_selection_status_backup* is greater than 1, skip fitting for the current beam and reference point, report invalid for *h_mean,* and for *n_fit_photons*. If *signal_selection_status_backup* is equal to 0 set *h_range_input* equal to *H_win*.

Note: If *h_range_input* is not set in 1d or 1e, it remains equal to the value set in 1a: the difference between the maximum and minimum heights of all photons found in segments *m* and *m-1*.

<u>Output values assigned</u>: *signal_selection_source, signal_selection_status_confident, signal_selection_status_all, signal_selection_status_backup.*

<u>Internal values assigned:</u> PE_selection_flag.

2. For each beam, estimate the surface height and slope using the **iterative least-squares fitting** routine. Set *n_fit_photons* to the number of PE in the final selection. If the final selection includes fewer than 10 PE, or if the along-track spread is less than 20 m, or if the final window width is larger than 20 m, report an invalid fit and set *h_mean* to its invalid value (*NaN*) and return.

<u>Output values assigned</u>, for each beam: *n_fit_photons, dh_fit_dx, h_mean, h_rms_misfit, h_robust_sprd, med_r_fit, w_surface_window_final, SNR.*

<u>Internal values assigned</u>, for each beam: *h_mean, r_fit, selected_PE, h_range_input*

3. For each beam, calculate the first-photon bias correction

For each beam, estimate the first-photon bias correction to the mean height, the first-photon-bias corrected median height, and the corrected return-time histogram based on the residuals to the segment heights calculated in step 3.

3a. Run the first-photon-bias-correction routine on PE flagged with *selected_PE* (see below)

<u>Internal values assigned:</u> fpb-corrected residual histogram, estimated gain.

<u>Output values assigned</u> for each beam: *fpb_mean_corr, fpb_mean_corr_sigma, fpb_median_corr, fpb_median_corr_sigma , FPB_N_PE*

4. Calculate the pulse-truncation correction

Based on the *h_robust_sprd* and *w_surface_window_final* values calculated in the last step of the iterative least-squares fit and the *SNR* calculated in step 2, calculate the pulse-truncation correction (See pulse-truncation-correction section).

<u>Output values assigned</u> for each beam: *tx_med_corr, tx_mean_corr*

5. Calculate remaining output parameters

5a. Calculate *h_li*:

$$h\_li = h\_mean + fpb\_med\_corr + tx\_med\_corr$$

<u>Output values assigned</u>: *h_li*

5b. Calculate *y_seg_RGT,* equal to the median of all *y_PE_RGT* values.

<u>Output values assigned:</u> *y_seg_RGT*

5c. Calculate *seg_time, lat_seg_center* and *lon_seg_center* by regressing (respectively) *time_PE*, *lat_PE* and *lon*_PE as a function of *x_PE* to *x0_seg* for selected PE. For those segments for which fitting has failed, but for which the other beam in the pair has a valid segment, report the latitude and longitude of the valid segment, displaced by 90 m to the left or right in the across-track direction (depending on which segment is valid).

<u>Output values assigned:</u> *seg_time, lat_seg_center, lon_seg_center, delta_time*

5d. Estimate the final cross-track slope, equal to the difference between the *h_li* values divided by the difference between the *y_seg_RGT* values for the two beams.

<u>Output values assigned:</u> *dh_fit_dy*

5e. Calculate error estimates for each beam.

i. For each segment, calculate *h_expected_RMS* based on the footprint size, the along-track track slope, and the transmit pulse duration (equation 1):

$$h\_expected\_RMS = sqrt((dh\_fit\_dx\ \ sigma\_beam)^\wedge 2 + (c/2\ sigma\_xmit\ )^\wedge 2)$$

ii. Add the effects of background noise to *sigma_expected* to calculate *sigma_PE_est.*

$$sigma\_PE\_est= ((N\_signal\ h\_expected\_RMS^2 + N\_noise(0.287\ H\_win)^2)/N\_tot)^{1/2}$$

iii. Calculate linear-fit-model errors. Multiply *h_mean_sigma_unit* and *dh_fit_dx_sigma_unit* by *max(sigma_PE_est, h_rms_misfit)* to obtain *h_mean_sigma* and *dh_fit_dx_sigma*.

<u>Output values assigned:</u> *sigma_h_mean, sigma_dh_fit_dx, sigma_PE_est, h_rms_misfit.*

5f. Set *h_li_sigma* equal to the maximum of *sigma_h_mean* and *fpb_med_corr_sigma.*

<u>Output values assigned</u> for each beam: *h_li_sigma.*

5g. Calculate the uncorrected reflectance, based on the first-photon-bias-corrected total PE count. Equation given in 3.4.3.3.

<u>Output values assigned</u>, for each beam: *r_eff*

5h. Calculate *SNR_significance*, by interpolating into the *SNR_F_table* as a linear function of the table parameters *BGR, SNR*, and *w_surface_window_initial*.

<u>Output value assigned</u>: *SNR_significance*

5i: calculate atl06_quality_summary: *atl06_quality_summary* is zero unless *h_li_sigma* > 1 m or *SNR_significance > 0.02* or *N_fit_photons/w_surface_window_final < 4 (for strong beams)* or <1 (for weak beams) or *signal_selection_source* > 1.

5j: Calculate pass-through parameters: For tide parameters, error parameters, and the *dac*, calculate ATL06 values from the average values for the ATL03 segments.

5k: Calculate systematic error estimates: Based on geolocation error estimates and surface slope, calculate *h_li_sigma_systematic* based on equation 36.

5.5 Signal selection based on ATL03 flags

Inputs, from one beam only, for each PE

x_PE: along-track coordinates of the land-ice PE for the current segment

h_PE: height of PE for the current segment

Ice_confidence_flag: ATL03 classification of the land-ice PE. 0=undetected, 1=PE in the pad region, but not identified as signal PE, 2=low confidence, 3=medium confidence, 4=high confidence.

Input, one per segment:

x0: the along-track location of the segment center.

BGR: the interpolated background PE rate for the segment.

Parameters:

Sigma_beam: The one-sigma expected horizontal spread of the photons on the ground. Equal to 4.25 m (pre-launch estimate)

Sigma_xmit: The one-sigma temporal duration of the transmit pulse.

Outputs:

PE_selection: binary flag, one per input PE, showing whether to use that PE in the initial fit.

Signal_selection_source: parameter indicating the how the signal was selected. See Table 3-1 for values.

signal_selection_status_confident: parameter indicating the success/failure of signal selection using low-or-better confidence PEs.

signal_selection_status_all: parameter indicating the success/failure of signal selection using all flagged PEs.

H_win: Height of the window around the best-fitting line used to select PE.

Procedure:

1. If the inputs are empty (no PE are in the along-track window), set *signal_selection_source* to 3, set *signal_selection_status_confident* to 3, set *signal_selection_status_all* to 3 set *signal_selection_status_backup* to 4, and return.

2. Check if the confidently detected PE are adequate to define an initial segment.

2a. Set *PE_selection* to true for all PE with *Ice_confidence_flag*>=2, to zero for all others

2b: If the difference in *x_PE* between the first and last PE in *PE_selection* is less than 20 m set *signal_selection_status_confident* to 1.

2c: If there are fewer than 10 true elements in *PE_selection*, but the spread between the first and last PE in *PE*_selection is greater than 20 m, set *signal_selection_status_confident* to 2.

2d. If there are fewer than 10 true elements in *PE_selection,* and the spread between the first and last PE is less than 20 m, set *signal_selection_status_confident* to 3.

3. Check if the combination of confidently detected PE and the padded PE are adequate to define an initial segment. If *signal_selection_status_confident* is zero, skip this step.

3a. Set *PE_selection* to true for all PE with non-zero *ice_confidence_flag*.

3b: If the difference in *x_PE* between the first and last PE in *PE_selection* is less than 20 m set *signal_selection_status_all* to 1.

3c: If there are fewer than 10 true elements in *PE_selection*, but the spread between the first and last PE in *PE*_selection is greater than 20 m, set *signal_selection_status_all* to 2.

3d. If there are fewer than 10 true elements in *PE_selection,* and the spread between the first and last PE is less than 20 m, set *signal_selection_status_all* to 3.

3e: If *signal_selection_status_all* is equal to zero, set *signal_selection_source* to 1 and proceed to step 4, otherwise set *signal_selection_source* to 2, and return.

4. Calculate the vertical spread of the selected PE, make the selection consistent with a vertical window around a sloping segment.

4a. Calculate the least-squares fit line between (x_PE-x_0) and h_PE for the selected PE. Internal variables set: *along_track_slope*, *seg_center_height.*

4b. Calculate *r_PE,* the residual between the best-fitting line and *h_PE*.

4c. Calculate *sigma_r*, the robust spread (accounting for noise) of *r_PE*, based on the background density, *BG_density,* with *z_min* and *z_max* set to the minimum and maximum values of *r_PE*. See the ***robust_dispersion*** section for description.

4d. Calculate the expected PE spread, *sigma_expected*, based on the current slope estimate:

$$sigma\_expected=[(c/2\ sigma\_xmit)^2 + sigma\_beam^2 along\_track\_slope^2]^{1/2}$$

4e. Calculate *H_win*:

$$H\_win=max(H\_win\_min,\ 6\ sigma\_expected,\ 6\ sigma\_r)$$

4f. Select all PE that have *abs(r_PE) < H_win/2*. Report the number of selected PE as *N_initial*.

5.6 Backup PE-selection routine.

Inputs:

x_PE : along-track coordinates of all PE for the current beam *h_PE*: heights of all PE for the current beam

x0: along-track bin center for the current bin.

Ice_confidence_flag: ATL03 classification of the land-ice PE. 0=undetected, 1=PE in the pad region, but not identified as signal PE, 2=low confidence, 3=medium confidence, 4=high confidence

signal_selection_source: Flag indicating the how the signal was selected. See Table 3-1 for values.

Outputs:

PE_selection: selected PE for the current bin.

signal_selection_source: Flag indicating the how the signal was selected. See Table 3-1 for values, updated based on the results of this algorithm

signal_selection_status_backup flag indicating the success/failure of signal selection using backup selection algorithm

H_win: Vertical extent of the selected window

Internal variables:

Test_window_center: Vector of test window centers

Window_center_height: Estimated window center height

Procedure:

1. Attempt to center the window on any ATL03 flagged PE that are present.

1a. If any padded or detected PE are found, set *w0* to the maximum of 10 m and the difference between the maximum and minimum selected PE heights, and set *PE_selection* to true for all PE that have heights within *5 m* of the median of the selected PE heights. Set *H_win* equal to 10 m.

1b. If the horizontal spread in the PE marked in *PE_selection* is greater than 20 m, and if 10 or more PE are selected, then set *signal_selection_status_backup* to zero, set *signal_selection_source* to 2, and return.

2. Find the 80-m along-track by 10-m vertical bin that contains the largest number of PEs

2a. Select all PE from ATL03 segments *m-2* to *m+1,* inclusive.

2b. Loop over *test_window_center* values between *floor(min(h_PE))*+0.25 and *ceil(max(h_PE))* in 0.5 m steps. For each *test_window_center* value, count the PE in a 10-m (vertical) bin centered on the *test_window_center* value.

2c. Find the maximum of the window counts, *Cmax*, and calculate its uncertainty, *Csigma=sqrt(Cmax)*. If *Cmax* is less than 16, then set *PE_selection* to null (no selected PE) and skip to step 3.

2d. Set *window_center_height* equal to the center of the range of *test_window_center* values that have a count greater than *Cmax-Csigma*. Set *H_win* to the difference between the minimum and maximum of *test_window_center* values that have a count greater than *Cmax-Csigma*, plus 10 m.

2e. Set *PE_selection* to 1 for all PE in ATL03 segments *m*-1 and *m*, with a height within *H_win/2* of *window_center_height.*

3. Evaluate the selection.

3a. Set *signal_selection_status_backup* to 1.

3b: If the difference in *x_PE* between the first and last PE in *PE_selection* is less than 20 m set *signal_selection_status_backup* to 2.

3c: If there are fewer than 10 true elements in *PE_selection*, but the spread between the first and last PE in *PE*_selection is greater than 20 m, set *signal_selection_status_backup* to 3.

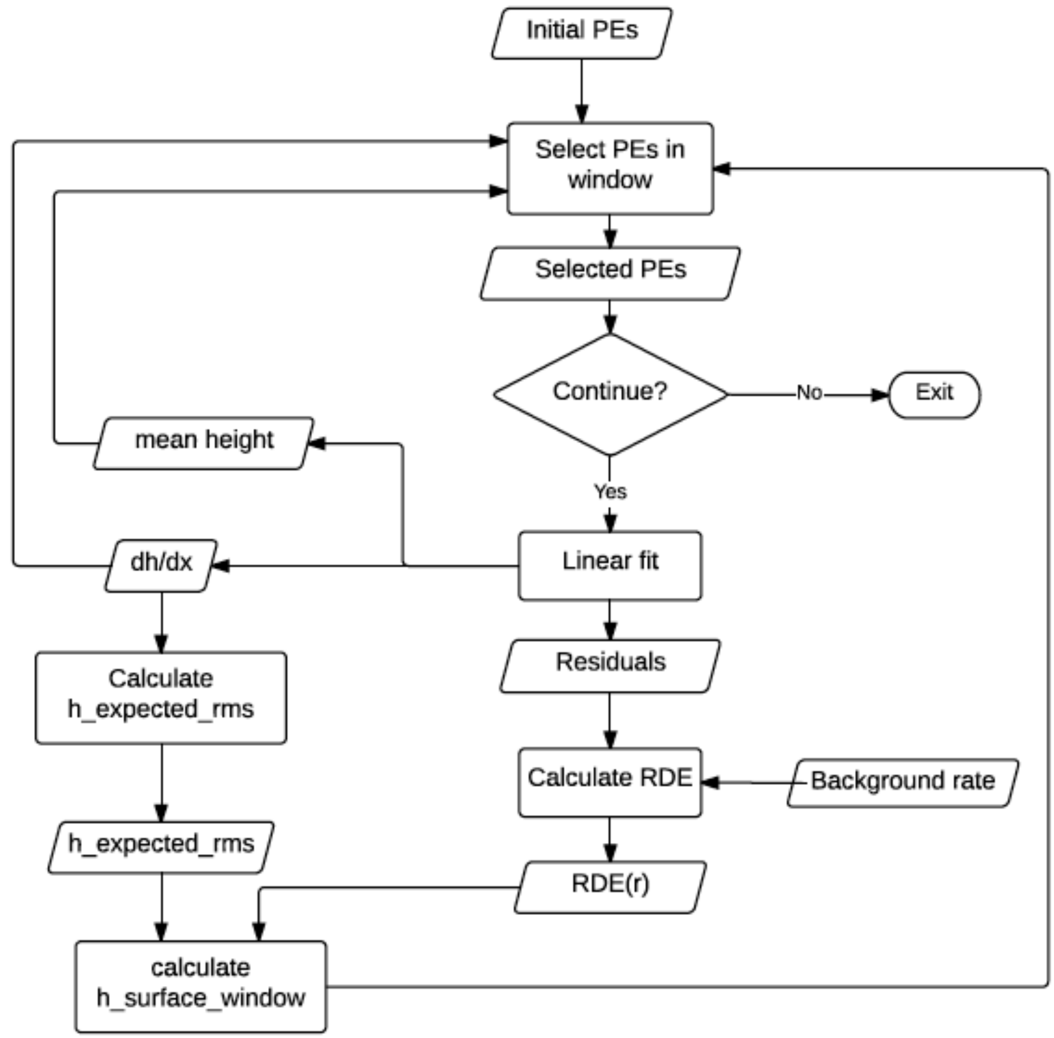
3d. If there are fewer than 10 true elements in *PE_selection,* and the spread between the first and last PE is less than 20 m, set *signal_selection_status_backup* to 4.

3e. If *signal_selection_status_backup* is 1, set *signal_selection_source* to 2, if greater than 1, set *signal_selection_source* to 3.

5.7 Iterative Least-Squares Fitting Routine

This routine performs the iterative least-squares fit to refine the surface window and determine the along-track slope. The process for this step is shown in Figure 5-2.

Figure 5-2. Flow chart for iterative ground fit



Process for iterative ground fit. On exit, all the variables (in slanted parallelograms) are exported. The exit condition cannot happen until after the end of the first iteration.

Inputs:

x_PE : along-track coordinates of PE for the current beam

y_PE : across-track coordinates of PE for the current beam

input_PE_selection: Flag defining the PE selected by the initial selection routine

h_PE: heights of selected PE for the current beam

x0: along-track bin center for the current bin.

bckgrd: Interpolated background-PE rate estimate for the segment

H_win: Initial surface-window height.

signal_selection_source: Flag indicating the source of the initial signal selection

N_it: maximum number of iterations

Parameters:

Sigma_xmit: transmitted pulse duration (seconds)

Sigma_beam: sigma value for pulse surface footprint (expected to be equal to 4.25 m)

L0: Along-track length of the window

N_seg_pulses: Number of pulses in a 40-meter segment (equal to 58 assuming 7 km/s ground-track speed)

H_win_min: Minimum allowed surface window height, equal to 3m.

Outputs:

H_win: the height of the window around the best-fitting segment within which PE are selected.

dh_fit_dx: The along-track slope of the best-fitting segment

h_mean: The mean-based height of the best-fitting segment

PE_fit_flag: A flag indiciating whether a particular PE has been selected based on the segment height and slope and *H_win.*

r0: Residuals to the best-fitting segment

h_mean_sigma_unit: Estimated error in *h_mean* per unit of PE-height error

dh_fit_dx_sigma_unit: Estimated error in *dh_fit_dx* per unit of PE-height error.

N_signal : Estimated number of signal PE

N_BG : Estimated number of background PE

h_robust_sprd : robust spread of residuals

h_rms_misfit: RMS misfit of residuals

SNR: signal-to-noise ratio for window.

Procedure:

1. Initialize the fit.

1a. If *signal_selection_source* is zero or 1, eliminate all PE not marked as 1 in *input_PE_selection,* set *PE_fit_flag* to 1 for all remaining PE.

1b. If *signal_selection_source* is nonzero, Set *PE_fit_flag* to 1 for all PE marked in *input_PE_selection,* zero for all others.

1c. Calculate the vertical noise-photon density:

BG_density = N_seg_pulses median(bckgrd) / (c/2)

2. Iterate the fit.

2a. Check whether enough PE are selected to define a window. If fewer than 10 PE are selected in *PE_fit_flag*, set *H_win, dh_fit_dx, H_mean,* and *r0* to invalid, and return.

2b. Calculate the least-squares linear fit between *h_PE* and *x_PE-x0* for the PE selected in *PE_fit_flag*. The intercept of the fit is *h_mean,* the slope is *dh_fit_dx*. Calculate the residual to this fit for the selected PE, *r0* and for all PE, *r*. If the along-track spread between the first and last selected PE is less than 10 m, fit for the height only, and set the along-track slope estimate to zero.

2c. Calculate *sigma_r*, the robust spread (accounting for noise) of *r0*, based on the background density, *BG_density,* and current window height, *H_win*. The variables input to the ***robust dispersion including a background estimate*** routine are *z=r0*, *zmin=-H_win/2*, *zmax=H_win/2*, *N_BG=H_win BG_density*. If the resulting *sigma_r* is greater than 5 m, set it to 5 m.

2d. Calculate the expected PE spread, *sigma_expected*, based on the current slope estimate:

$$\mathit{sigma\_expected}=[(c/2\ \mathit{sigma\_xmit})^2 + \mathit{sigma\_beam}^2\,\mathit{along\_track\_slope}^2]^{1/2}$$

2e. Save the value of *H_win* in *H_win_previous*, then calculate the window height from *sigma_expected* and *sigma_r*.

H_win=max(H_win_min, 6 sigma_expected, 6 sigma_r, 0.75 H_win_previous)

2f. Save the values of *PE_fit_flag* in *PE_fit_flag_last.*

2g. Select PE within *H_win*/2 of the segment fit.

PE_fit_flag= 1 for PE with *r*<*H_win*/2, 0 for PE with *r* > *H_win*/2

2h. Evaluate the newly selected PE. If there are fewer than 10 selected PE, or if the along-track spread between the first and last PE is less than 20 m, set *PE_fit_flag* to *PE_fit_flag_last, H_win* to *H_win_previous*, and continue to step 3.

2i. If fewer than *N_iterations* have been completed, and if the values for *PE_fit_flag* have changed since the previous iteration, return to step 2a. Otherwise continue to step 3.

3. Propagate the error in the fit parameters assuming unit data errors (see 3.6, with σ_{photon}=1). This gives the unit errors *h_mean_sigma_unit*, *dh_fit_dx_sigma_unit.*

4. Calculate the number of signal and background PE, and the SNR.

N_BG=bckgrd H_win 2/c N_seg_pulses

N_signal=max(0, number of selected PE – N_BG)

SNR=N_signal/N_BG

5. Calculate output error statistics:

h_rms_misfit=RMS misfit of selected PE

h_robust_sprd = *sigma_r* from the last iteration

5.8 Robust dispersion calculation from a collection of points, not including a background estimate

Input:

z: sampled values

Output:

RDE : the robust dispersion estimate for z.

Procedure:

1. Sort z. *zs* is equal to *z*, sorted in ascending order. Let *Nz* equal to the number of elements in *z*.

2. Calculate an abscissa for *zs*,

2a. Generate *ind*, equal to the sequence of integers between 1 and *Nz*.

2b. Calculate *ind_N*, equal to (*ind-0.5)/Nz.*

3. Interpolate the percentiles of *z*. Interpolate the values of *zs* as a function of *ind_N* at values 0.16 and 0.84. Half the difference between these values is *RDE.*

5.9 Robust dispersion calculation from a collection of points, including a background estimate

Inputs:

z: sampled values

zmin, zmax: window from which the values in z are sampled

N_BG: Estimate of the number of background events between *z_min* and *z_max.*

Output:

RDE : the robust dispersion estimate for z.

Parameter:

Scale_factor: equal to *sqrt*(2)(*erfinv(0.5)-erfinv(-0.5))*, where *erfinv()* is the inverse error function, or 1.3490.

Procedure:

1. Estimate the background rate and signal count.

1a. *bckgrd* is equal to *N_BG* divided by the difference between *zmax* and *zmin.*

1b. *N_sig* is equal to the number of elements in *z*, minus *N_BG*.

1c. If *N_sig*<=1, the RDE is equal to (*zmax-zmin*)/*(the number of elements in z)*, and the rest of the calculation is skipped.

2. Sort z. *zs* is equal to *z*, sorted in ascending order. Let *Nz* equal to the number of elements in *z*.

3. Calculate an abscissa for *zs*. Generate *ind*, equal to the sequence of integers between 1 and *Nz,* minus 0.5.

4. Find the indices for the smallest potential percentiles of *z.*

4a. *i0* is equal to the index of the greatest value of *ind* for which $ind < (0.25 N\_sig + (zs-zmin) bckgrd)$.

4b. *i1* is equal to the index of the smallest value of *ind* for which $ind > (0.75 N\_sig + (zs-zmin) bckgrd)$.

5. If $i1 < i0$, reselect *i0* and *i1* to measure spread of the central *N_sig*/2 values of the distribution:

5a: *i0* is equal to the index of the greatest value of *ind* for which $ind < Nz/2 - Nsig/4$.

5b: *i1* is equal to the index of the smallest value of *ind* for which $ind > Nz/2 + Nsig/4$.

6. Calculate *RDE. RDE* is equal to the difference between the *zs* values at *i0* and *i1*, divided by *scale_factor.*

5.10 First- Photon Bias Correction

These routines calculate the first-photon bias for a collection of residual photon heights. Most of the calculation is done as a function of time, and the results are converted back to height at the end of the routine.

Inputs:

r_p: PE heights, corrected for the along-track segment fit, converted to time (multiplied by –2/c)

N_seg_pulses: the number of pulses in the segment

N_px: the number of pixels in the detector.

Outputs:

G_est: the estimated detector gain

N_hist: The uncorrected PE count histogram (in units of PE)

N_PEcorr: the estimated PE count histogram (in units of PE)

t_full: the time vector for the PE count histogram.

FPB_med_corr: the FPB correction to the median height

Sigma_FPB_med_corr: the error estimate for *FPB_med_corr*

FPB_mean_corr: The FPB correction to the mean height

FPB_mean_corr_sigma: the error estimate for *FPB_mean_corr*.

Fpb_N_photons: the FPB-corrected estimate of the number of PE in the return.

Parameters:

t_dead : the mean detector dead time for the beam.

N_seg_pulses: the number of pulses in the segment

N_px: the number of pixels in the detector.

dt: duration of a histogram bin.

Procedure:

1. Generate a residual histogram

Convert PE height residuals to time residuals (multiply by -2/c). Generate a histogram of time residuals, *N_hist*, in bins of size *dt.*

2. Calculate the gain from the histogram

P_dead for bin i is the sum over bins i-N_dead to i-1 of *N_hist,* divided by *N_seg_pulses N_px. G_est* is equal to 1- *P_dead, where N_dead* is the deadtime expressed in histogram bins.

3. Check if the correction is valid. If the minimum value for *G_est* is less than 2/(*N_seg_pulses N_px*), set all return values equal to invalid (*NaN*) and return.

4. Calculated the corrected histogram:

N_PEcorr is equal to *N_hist* divided by *G_est.*

5. Calculate height statistics

Calculate the gain-corrected mean and median and their errors for the segment, based on the full gain estimate and the full histogram:

FPB_med_corr: -1/2c times the gain-corrected median time based on *N_PE* and *G_est.* See 5.11.

Sigma_FPB_med_corr: the error estimate for *FPB_med_corr*

FPB_mean_corr: -1/2c times the gain-corrected mean time based on *N_PE* and *G_est.* See 5.12.

FPB_mean_corr_sigma: the error estimate for *FPB_mean_corr*.

Fpb_N_photons: the sum of *N_PEcorr.*

5.11 Gain-corrected median

Inputs:

N: The uncorrected histogram

G: The gain estimate,

x: the abscissa for the bin centers, corresponding to N and G.

Outputs:

x_med: the median of N based on G

sigma_x_med: the error in *x_med*

Procedure:

1. Calculate the corrected histogram:

N_corr is equal to *N* divided by *G*.

2. Calculate the CDF of N_corr

The CDF, *C*, is calculated at the bin centers, and at each bin center, *j*, is equal to the sum of all values of *N_corr* for bin centers $i<j$. *C* is normalized so that its last value is equal to 1.

3. Calculate the 40th, 50th, and 60th percentiles of N_corr

C is treated as a function that increases linearly across each bin, such that the upper edge of the ith bin is greater than the lower edge of the ith bin by N_i. The abscissa for *C* runs from zero at x_1-*dx/2*, to x_m+*dx/2*, where x_1 is the first bin center, x_m is the last bin center, and *dx* is the spacing between bin centers. The 40th, 50th, and 60th percentiles of *N_corr* are calculated by interpolating into the vector of bin edges as a function of *C*. If more than one bin has a CDF within numerical precision of the calculated percentile, report the mean x value of all such bins.

4. Calculate the error in the CDF at the 50th percentile

The error in any value of *N_corr* (*sigma_N_corr*) is the inverse gain value for that bin times the square root of *N* for that bin. *sigma_CDF* for any x is found by calculating the RSS of all *sigma_N_corr* values for bins less than x, and dividing by the sum of *N_corr*.

The value for *sigma_CDF* at the 50th percentile is found by interpolating *sigma_CDF* as a function of *C* at a *C* value of 0.5.

5. calculate *sigma_x_med*

Sigma_x_med is found:

$$sigma\_x\_med = \frac{dz_{60} - dz_{40}}{0.2} \sigma_{cdf}(dz_{med})$$

Here dz_{60} and dz_{40} are the 40th and 60th percentiles of *N_corr* from step 3.

5.12 Gain-corrected mean

Inputs

N: The uncorrected histogram

G: The gain estimate

x: the abscissa for the bin centers, corresponding to N and G.

Outputs:

x_mean: the mean of N based on G

sigma_x_mean: the error in *x_mean*

1. Calculate the corrected histogram:

N_corr is equal to *N* divided by *G*.

2. Calculate the corrected mean:

Calculate the mean:

$$x\_mean = \sum \frac{N_{corr,i}}{N_{tot}} x_i$$

3. Calculate the error in the corrected histogram:

$$\sigma_{N,corr,i} = \frac{N_{0,i}^{1/2}}{G_i}$$

4. Calculate the error in the corrected mean:

$$sigma\_x\_mean = \left[\sum \left(\sigma_{N,corr,i} \frac{x_i - x\_mean}{N_{corr,tot}} \right)^2 \right]^{1/2} \qquad 49$$

5.13 Transmit-pulse-shape correction

This routine uses the most recent estimate of the transmit-pulse shape calculated from the transmitter-echo pulse to calculate median and mean offsets for a windowed, truncated received pulse. This correction depends the shape of the transmit pulse, and on three parameters that are unique to each segment: the estimated width of the return pulse, the refined surface-window height, and the signal-to-noise ratio.

Inputs:

-Transmit-pulse-shape estimate (*t_tx, P_tx*). The time vector, *t_tx* is shifted so that *P_tx* has a zero centroid (see 5.15).

-Received-pulse width estimate (*W_rx*)

-Surface-window time duration (*dt_W*)

-Signal-to-noise ratio estimate within the truncated window (*SNR*)

Outputs:

Height offsets for the mean and median transmit-pulse-shape correction.

Procedure:

This correction works by generating a synthetic return pulse that matches the width of the actual return pulse, and truncating it in the same way that the return pulse has been truncated. The median and the mean of the synthetic pulse are then calculated.

1. Calculate the time by which the received pulse was broadened

The spreading needed to broaden the transmitted pulse to match the received pulse is equal to *W_spread=sqrt(max(0.01e-9^2,W_RX^2 – W_TX^2)).*

2. Generate a synthetic received pulse

2a: Calculate the shape of the expected spread pulse:

The synthetic received pulse is generated by convolving the transmitted pulse with a Gaussian function of with a sigma parameter equal to *W_spread*. The Gaussian should have enough samples to include at least *4*W_spread* worth of samples on either side of its center. The synthetic pulse and its time vector are *N_hist_synthetic* and *t_synthetic*.

2b: Calculate the median of the broadened synthetic pulse:

Calculate the median of the synthetic received pulse, *t_synthetic_med*, and set *t_ctr=t_synthetic_med.*

2c: Normalize the waveform and add an estimated noise signal:

N_hist_synthetic is normalized so that its sum is equal to 1, and a background count of 1/*SNR* (*dt/dt_W*) is added to *N_hist_synthetic*.

3. Calculate the centroid of the synthetic received pulse

To find the centroid of the truncated synthetic waveform, an iterative procedure is used:

3a: Calculate the centroid of the synthetic waveform

t_ctr is set to the centroid of the truncated synthetic received waveform, windowed by *t_ctr - dt_W/2* and *t_ctr +dt_W/2*

3b: Check for convergence and iterate

Unless the current and previous values of *t_ctr* are consistent to within 0.1 mm (0.00067 ns) or if 50 iterations are complete, return to *4a*.

4. Calculate the median of the synthetic received pulse

The median of the synthetic received waveform is calculated the synthetic received waveform from 4b, windowed by *t_ctr -dt_W/2* and *t_ctr +dt_W/2*

5. The corrections for the median and mean heights are equal to c/2 times the median and mean time offsets.

5.14 Residual_histogram calculation

Inputs:

Segment_lat : latitude for each segment center

Segment_lon : longitude for each segment center

*Segment_x_*ATC: along-track (x) coordinate for each segment center

Segment_h_mean: mean-based land-ice height for each segment center

Segment_slope: along-track slope for each segment center

Segment_SNR: SNR values for segment fits

Segment_BGR: Background rate estimate for each segment

N_seg_pulses Number of pulses in each segment (including those contributing no PE to the fit).

x_pe: along-track(x) coordinates for all ATL03 PE in the segment

h_pe: ATL03 surface height for all PE in the segment.

Parameters:

N_hist: Number of groups of segments in the histogram (number of horizontal divisions)

N_bins: Number of vertical bins in the residual histogram

bin_top_h: Tops of the histogram bins, listed from bottom to top

Outputs:

Count: *N_bins* x *N_hist*-element array giving the number of residual photons in each bin (*N_bins* is the vertical dimension, *N_hist* is the horizontal dimension)

bckgrd_per_m: 1xN_hist-vector giving the expected background count per vertical meter in each column of the histogram based on the observed background rate (bckgrd) and the number of segments included in the histogram

Segment_id_list: 10 x *N_hist*-element array list of segment IDs included in the histogram

Lat_mean: N_hist-element list giving the mean latitude of all segments included in each horizontal histogram bin

Lon_mean: N_hist-element list giving the mean longitude of all segments included in each horizontal histogram bin

x_ATC_mean: N_hist-element list giving the mean along-track (x) coordinate of all segments included in each horizontal histogram bin

Procedure

1. Calculate the bin-edge heights. There are N_bins+1 edges. The second through last edges are equal to the input *bin_top_h* values. The first (lowest) edge is 1 m lower than the second (i.e. equal to the first value of *bin_top_h* -1).

2. Group segment centers into 10-segment groups: For each RGT, segments 1-10 would be in the first group, 11-20 in the second, etc.

3. For each group, gather all valid segments that have high-quality surface-height estimates (*ATL06_quality_summary=0*). If any high-quality segments are present, calculate the histogram count. Otherwise, report the histogram count as all zeros, and report *lat_mean*, *lon_mean*, *x_atc_mean*, and *segment_id_list* as invalid.

3a. For each valid segment, calculate the histogram and background count.

3a.1: Gather the PE that have *x_segment* -10 m < *x_pe* <= 10 m.

3a.2: Calculate the residual between the segment and the gathered PE: *r=h-h_mean_segment- (x_pe- segment_x_ATC) × segment_slope.*

3a.3: For each vertical bin in the histogram, count the PE with residuals that fall into the bin

3a.4: For each valid segment, add the expected background count per vertical meter, as estimated from the segment background count to the total background-per-meter (*bckgrd_per_m*) for the segment. The contribution for each segment is: *segment_BGR × N_seg_pulses / 2 /(c/2). [N.B. The factors of 2 in the previous statement cancel, leaving* : *segment_BGR × N_seg_pulses /c.]*

3b. Add the segment histograms together to calculate the 10-segment histogram

3c. Calculate the mean values for latitude, longitude, and *x_ATC* for the segment. List the selected segments in *segment_id_list*

5.15 Transmit-echo-pulse initialization

This calculation centers the transmit-echo-pulse reported by ATL03 on its centroid, after using an iterative edit to distinguish between signal and noise. It should be performed each time a new night-time TEP estimate of the waveform becomes available. The TEP consists of the power (*tep_hist*) and time (*tep_hist_x*) that are input from ATL03. Two TEP histograms are available, obtained for laser spot 1 and 3. The ATL03 *tep_valid_spot* parameter specifies with which TEP histogram is used for each of the ground tracks, and the ATL03 *tep_range_prim* parameter specifies the valid range of times for each TEP histogram.

Inputs:

-*tep_hist_x* : Time for the Transmit-pulse-shape estimate

-*tep_hist*: power (or signal count) for the transmit-pulse-shape estimate

The time-sampling interval these is *dt_input*. The transmit pulse is sampled so that at least the first 5 ns and the last 10 ns are representative of the background noise for the transmit pulse.

Outputs:

-*t_tx*: time vector for the transmit pulse estimate, shifted such that *P_tx* has a zero centroid

-*P_tx*: Power for the transmit-pulse estimate,

Algorithm:

1. Identify noise-only and signal samples: mark index *noise_samples* as true for the first 5 ns and last 10 ns of samples in *tep_hist.* Set *sig_samples* to the inverse of *noise_samples*

2. Calculate the noise value for the transmit pulse: N_tx = the mean of *tep_hist* for the samples in *noise_samples*. Subtract *N_tx* from *tep_hist* to give *P_tx.*

3. Calculate the centroid of the transmit pulse: T0_tx = sum(P_tx t_tx) /sum(P_tx).* The sum is carried out over the samples in *sig_samples.*

4. Calculate the RDE of the transmit pulse: The width of the transmitted pulse (W_TX) is equal to half the difference between the 84^{th} percentile and the 16^{th} percentile of the portion of *P_tx* in *sig_samples.*

5. Re-establish the noise-only samples: mark *noise_samples* as true for all samples with times more than 6 W_TX away from *T0_tx,* set *sig_samples* to the inverse if *noise_samples.* If *sig_samples* has changed from its previous values, and if fewer than 10 iterations have taken place, return to *1b*.

6. Center the transmit pulse on its centroid: Subtract *T0_tx* from *t_tx_input* to give *t_tx.*

6 TEST DATA AND SOFTWARE REQUIREMENTS

This section describes a very simple test data set that has been derived to verify the performance of the ATL06 surface code.

6.1 ATL06 Test Data Setup

The ATL06 test data are a set of synthetic data generated based on a planar, sloping surface with a slope of 0.02. Separate data sets are generated for surface reflectance values between 1/16 and 1, and for surface roughness values between zero and 2 m. A detector model with a dead time of 3.2 ns is used to simulate the effects of the first-photon bias. For each segment, a full set of ATL06 parameters are generated using the Matlab prototype code, and with the ASAS production code, and the two are compared. Small numerical differences between the codes can produce different results in the early stages of the signal-finding code, so the most valid comparisons between the results of the two codes are for segments with moderate signal strength (reflectance greater than 0.25). We consider the two codes to produce equally valid results when the difference between the results for any parameter is not significantly different from zero, and when the spreads of the two sets of parameters are not significantly different from one another for segments based on the same number of photons with the same surface window size.

7 BROWSE PRODUCTS AND Q/A STATISTICS

7.1 Browse Products

Browse products include two kinds of plots: Data-quality maps, and profile plots.

Data-quality maps are based on the *signal_selection_source* parameter. Each map shows a background image based on the MODIS mosaics of Greenland or Antarctica (Scambos and others, 2007), with color-coded points showing the mean segment location for each kilometer of the beam track, with the color showing the largest bit in signal_selection_source that is set for more than 50% of all segments in that kilometer of data, assuming that for segments with no data, all bits are set. The plots are made separately for the strong and weak beams, because the two beams are, at the granule scale, very close to one another and would otherwise overlap.

Profile plots are generated separately for each beam pair in the granule. Each plot shows the surface height as a function of along-track distance, and the height for each beam in the pair. A second set of axes, aligned with the first, shows the number of PE per segment (*N_fit_photons*) and the height error estimate, *h_li_sigma*.

7.2 Q/A Statistics

Quality assessment statistics are provided for each beam, for each 10-km increment along track. For each increment we provide:

A synopsis of the *signal_selection_source* parameter:

-The fraction of possible segments with *signal_selection_source* equal to zero.

-The fraction of segments with *signal_selection_source* equal to 1.

-The fraction of segments with *signal_selection_source* equal to 2.

-The fraction of segments with *signal_selection_source* equal to 3.

[Add parameters for the entire file]

8 APPENDIX A: GLOSSARY

This appendix defines terms that are used in ATLAS ATBDs, as derived from a document circulated to the SDT, written by Tom Neunann. Some naming conventions are borrowed from **Spots, Channels and Redundancy Assignments** (ICESat-2-ATSYS-TN-0910) by P. Luers. Some conventions are different than those used by the ATLAS team for the purposes of making the data processing and interpretation simpler.

Spots. The ATLAS instrument creates six spots on the ground, three that are weak and three that are strong, where strong is defined as approximately four times brighter than weak. These designations apply to both the laser-illuminated spots and the instrument fields of view. The spots are numbered as shown in Figure 1. At times, the weak spots are leading (when the direction of travel is in the ATLAS +x direction) and at times the strong spots are leading. However, the spot number does not change based on the orientation of ATLAS. The spots are always numbered with 1L on the far left and 3R on the far right of the pattern. Not: beams, footprints.

Laser pulse (pulse for short). Individual pulses of light emitted from the ATLAS laser are called laser pulses. As the pulse passes through the ATLAS transmit optics, this single pulse is split into 6 individual transmit pulses by the diffractive optical element. The 6 pulses travel to the earth's surface (assuming ATLAS is pointed to the earth's surface). Some attributes of a laser pulse are the wavelength, pulse shape and duration. Not: transmit pulse, laser shot, laser fire.

Laser Beam. The sequential laser pulses emitted from the ATLAS instrument that illuminate spots on the earth's surface are called laser beams. ATLAS generates 6 laser beams. The laser beam numbering convention follows the ATLAS instrument convention with strong beams numbered 1, 3, and 5 and weak beams numbered 2, 4, and 6 as shown in the figures. Not: beamlet.

Transmit Pulse. Individual pulses of light emitted from the ICESat-2 observatory are called transmit pulses. The ATLAS instrument generates 6 transmit pulses of light from a single laser pulse. The transmit pulses generate 6 spots where the laser light illuminates the surface of the earth. Some attributes of a given transmit pulse are the wavelength, the shape, and the energy. Some attributes of the 6 transmit pulses may be different. Not: laser fire, shot, laser shot, laser pulse.

Reflected Pulse. Individual transmit pulses reflected off the surface of the earth and viewed by the ATLAS telescope are called reflected pulses. For a given transmit pulse, there may or may not be a reflected pulse. Not: received pulse, returned pulse.

Photon Event. Some of the energy in a reflected pulse passes through the ATLAS receiver optics and electronics. ATLAS detects and time tags some fraction of the photons that make up the reflected pulse, as well as background photons due to sunlight or instrument noise. Any photon that is time tagged by the ATLAS instrument is called a photon event, regardless of source. Not: received photon, detected photon.

Reference Ground Track (RGT). The reference ground track (RGT) is the track on the earth at which a specified unit vector within the observatory is pointed. Under nominal operating conditions, there will be no data collected along the RGT, as the RGT is spanned by GT2L and GT2R (which are not shown in the figures, but are similar to the GTs that are shown). During spacecraft slews or off-pointing, it is possible that ground tracks may intersect the RGT. The precise unit vector has not yet been defined. The ICESat-2 mission has 1387 RGTs, numbered from 0001xx to 1387xx. The last two digits refer to the cycle number. Not: ground tracks, paths, sub-satellite track.

Cycle Number. Over 91 days, each of the 1387 RGTs will be targeted in the polar regions once. In subsequent 91-day periods, these RGTs will be targeted again. The cycle number tracks the number of 91-day periods that have elapsed since the ICESat-2 observatory entered the science orbit. The first 91-day cycle is numbered 01, the second 91-day cycle is 02, and so on. At the end of the first 3 years of operations, we expect the cycle number to be 12. The cycle number will be carried in the mid-latitudes, though the same RGTs will (in general) not be targeted more than once.

Sub-satellite Track (SST). The sub-satellite track (SST) is the time-ordered series of latitude and longitude points at the geodetic nadir of the ICESat-2 observatory. In order to protect the ATLAS detectors from damage due to specular returns, and the natural variation of the position of the observatory with respect to the RGT throughout the orbit, the SST is generally not the same as the RGT. Not: reference ground track, ground track.

Ground Tracks (GT). As ICESat-2 orbits the earth, sequential transmit pulses illuminate six ground tracks on the surface of the earth. The track width is approximately 10m wide. Each ground track is numbered, according to the laser spot number that generates a given ground track. Ground tracks are therefore always numbered with 1L on the far left of the spot pattern and 3R on the far right of the spot pattern. Not: tracks, paths, reference ground tracks, footpaths.

Reference Pair Track (RPT). The reference pair track is the imaginary line half-way between the planned locations of the strong and weak ground tracks that make up a pair. There are three RPTs: RPT1 is spanned by GT1L and GT1R, RPT2 is spanned by GT2L and GT2R (and may be coincident with the RGT at times), RPT3 is spanned by GT3L and GT3R. Note that this is the planned location of the midway point between GTs. We will not know this location very precisely prior to launch. Not: tracks, paths, reference ground tracks, footpaths, pair tracks.

Pair Track (PT). The pair track is the imaginary line half way between the actual locations of the strong and weak ground tracks that make up a pair. There are three PTs: PT1 is spanned by GT1L and GT1R, PT2 is spanned by GT2L and GT2R (and may be coincident with the RGT at times), PT3 is spanned by GT3L and GT3R. Note that this is the actual location of the midway point between GTs, and will be defined by the actual location of the GTs. Not: tracks, paths, reference ground tracks, footpaths, reference pair tracks.

Pairs. When considered together, individual strong and weak ground tracks form a pair. For example, GT2L and GT2R form the central pair of the array. The pairs are numbered 1 through 3: Pair 1 is comprised of GT1L and GT1R, pair 2 is comprised of GT2L and GT2R, and pair 3 is comprised of GT3L and 3R.

Along-track. The direction of travel of the ICESat-2 observatory in the orbit frame is defined as the along-track coordinate, and is denoted as the +x direction. The positive x direction is therefore along the Earth-Centered Earth-Fixed velocity vector of the observatory. Each pair has a unique coordinate system, with the +x direction aligned with the Reference Pair Tracks.

Across-track. The across-track coordinate is y and is positive to the left, with the origins at the Reference Pair Tracks.

Segment. An along-track span (or aggregation) of PE data from a single ground track or other defined track is called a segment. A segment can be measured as a time duration (e.g. from the time of the first PE to the time of the last PE), as a distance (e.g. the distance between the location of the first and last PEs), or as an accumulation of a desired number of photons. Segments can be as short or as long as desired.

Signal Photon. Any photon event that an algorithm determines to be part of the reflected pulse.

Background Photon. Any photon event that is not classified as a signal photon is classified as a background photon. Background photons could be due to noise in the ATLAS instrument (e.g. stray light, or detector dark counts), sunlight, or mis-classified signal photons. Not: noise photon.

h_.** Signal photons will be used by higher-level products to determine height above the WGS-84 reference ellipsoid, using a semi-major axis (equatorial radius) of 6378137m and a flattening of 1/298.257223563. This can be abbreviated as ‘ellipsoidal height’ or ‘height above ellipsoid’. These heights are denoted by h; the subscript ** will refer to the specific algorithm used to determine that elevation (e.g. is = ice sheet algorithm, si = sea ice algorithm, etc…). Not: elevation.

Photon Cloud. The collection of all telemetered photon time tags in a given segment is the (or a) photon cloud. Not: point cloud.

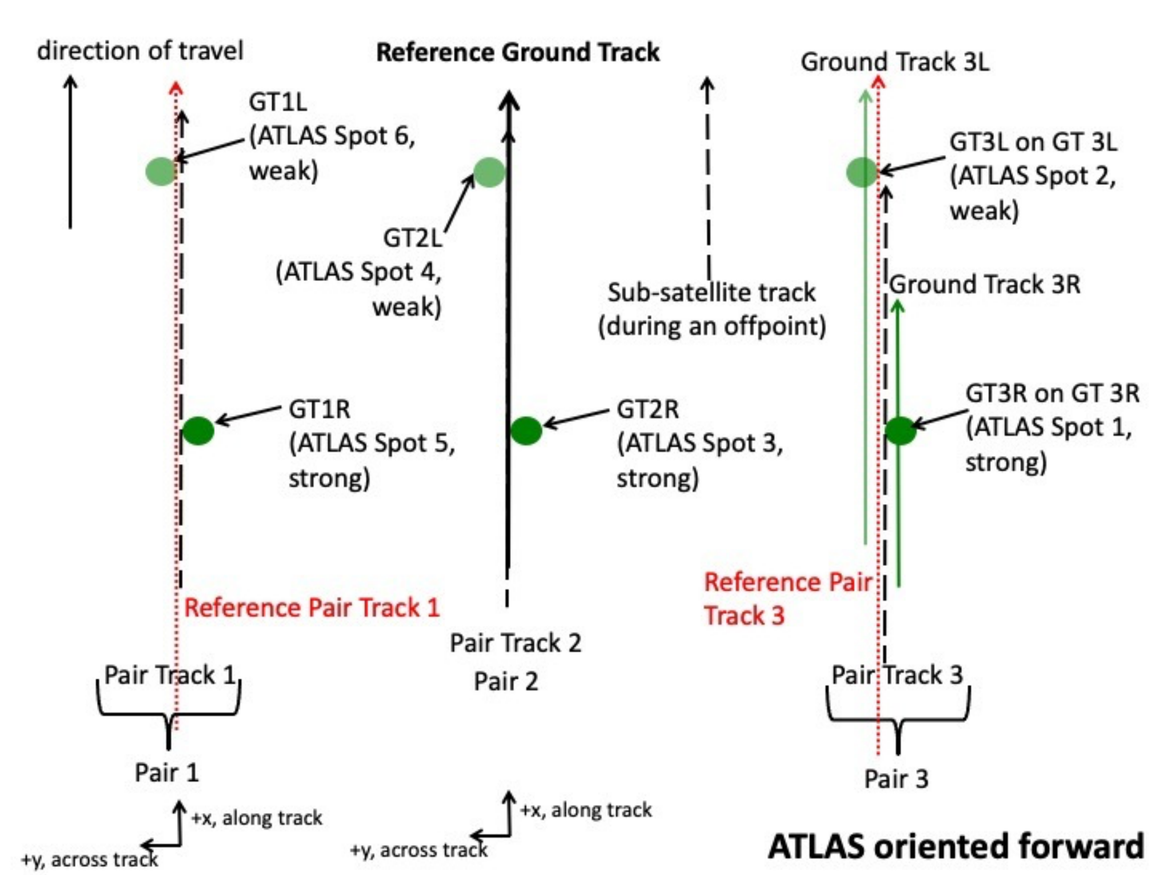
Background Count Rate. The number of background photons in a given time span is the background count rate. Therefore a value of the background count rate requires a segment of PEs and an algorithm to distinguish signal and background photons. Not: Noise rate, background rate.

Noise Count Rate. The rate at which the ATLAS instrument receives photons in the absence of any light entering the ATLAS telescope or receiver optics. The noise count rate includes PEs due to detector dark counts or stray light from within the instrument. Not: noise rate, background rate, background count rate.

Telemetry band. The subset of PEs selected by the science algorithm on board ATLAS to be telemetered to the ground is called the telemetry band. The width of the telemetry band is a function of the signal to noise ratio of the data (calculated by the science algorithm onboard ATLAS), the location on the earth (e.g. ocean, land, sea ice, etc…), and the roughness of the terrain, among other parameters. The widths of telemetry bands are adjustable on-orbit. The telemetry band width is described in Section 7 or the ATLAS Flight Science Receiver Algorithms document. The total volume of telemetred photon events must meet the data volume constraint (currently 577 GBits/day).

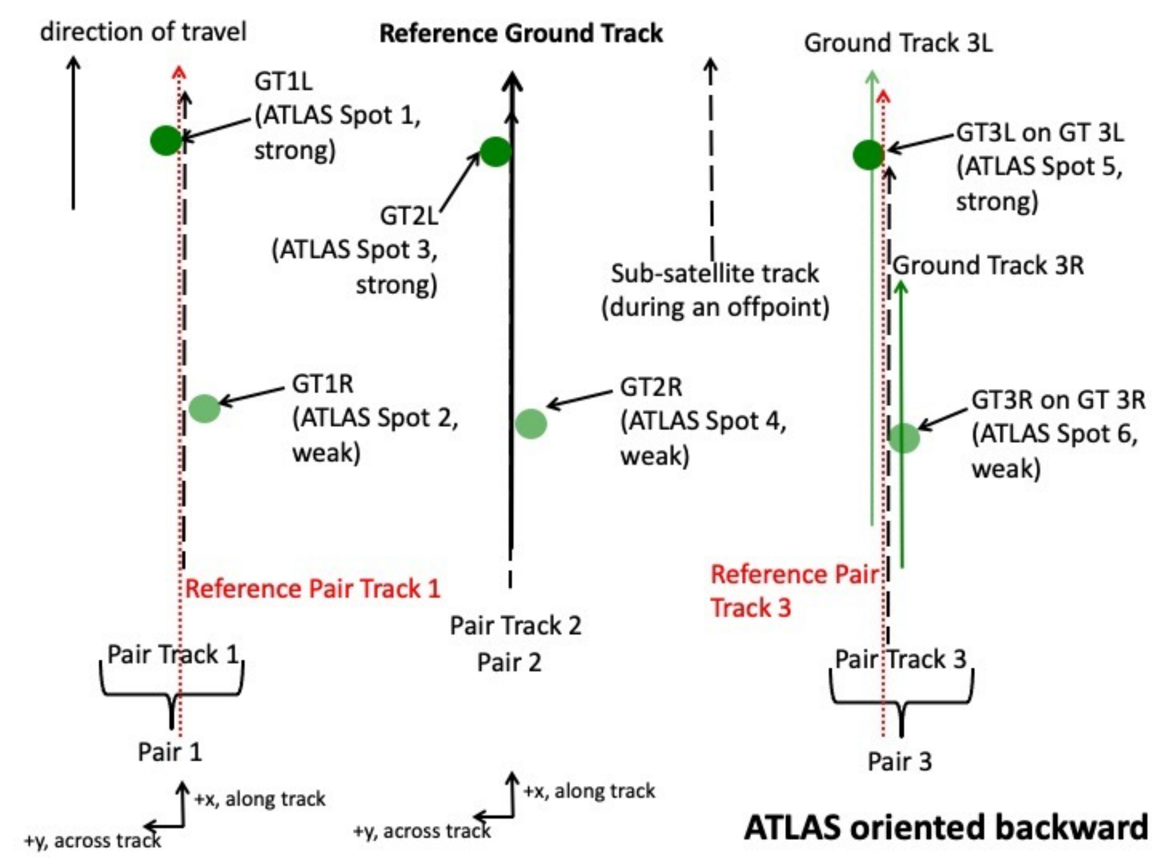
Window, Window Width, Window Duration. A subset of the telemetry band of PEs is called a window. If the vertical extent of a window is defined in terms of distance, the window is said to have a width. If the vertical extent of a window is defined in terms of time, the window is said to have a duration. The window width is always less than or equal to the telemetry band.

Figure 8-1. Spots and tracks, forward flight



Spot and track naming convention with ATLAS oriented in the forward (instrument coordinate +x) direction.

Figure 8-2. Spots and tracks, forward flight



Spot and track naming convention with ATLAS oriented in the backward (instrument coordinate -x) direction.

Glossary/Acronyms

ASAS	ATLAS Science Algorithm Software
ATBD	Algorithm Theoretical Basis Document
ATLAS	ATLAS Advance Topographic Laser Altimeter System
CDF	Cumulative Distribution Function
DEM	Digital Elevation Model
GSFC	Goddard Space Flight Center
GTs	Ground Tracks
ICESat-2	Ice, Cloud, and land Elevation Satellite-2
MABEL	Multiple altimeter Beam Experimental Lidar
MIS	Management Information System
NASA	National Aeronautics and Space Administration
PE	Photon Event
POD	Precision Orbit Determination
PPD	Precision Pointing Determination
PRD	Precise Range Determination
PSO	ICESat-2 Project Science Office
PTs	Pair Tracks
RDE	Robust Dispersion Estimate
RGT	Reference Ground Track
RMS	Root Mean Square
RPTs	Reference Pair Tracks
RT	Real Time
SCoRe	Signature Controlled Request

SIPS	ICESat-2 Science Investigator-led Processing System
TBD	To Be Determined
TL/DR	Too Long/Didn't Read.

References

Bamber, J.L., J.L. Gomez-Dans and J.A. Griggs 2009. A new 1 km digital elevation model of the Antarctic derived from combined satellite radar and laser data - Part 1: Data and methods. *Cryosphere*, **3**(1): 101-111.

Menke, W. 1989. *Geophysical data analysis: discrete inverse theory.* San Diego, CA, Academic Press.

Scambos, T.A., T.M. Haran, M.A. Fahnestock, T.H. Painter and J. Bohlander 2007. MODIS-based Mosaic of Antarctica (MOA) data sets: Continent-wide surface morphology and snow grain size. *Remote Sensing of Environment*, **111**(2-3): 242-257.

Warren, S.G., R.E. Brandt and T.C. Grenfell 2006. Visible and near-ultraviolet absorption spectrum of ice from transmission of solar radiation into snow. *Applied Optics*, **45**(21): 5320-5334.

Yang, Y., A. Marshak, S.P. Palm, T. Varnai and W.J. Wiscombe 2011. Cloud Impact on Surface Altimetry From a Spaceborne 532-nm Micropulse Photon-Counting Lidar: System Modeling for Cloudy and Clear Atmospheres. *Ieee Transactions on Geoscience and Remote Sensing*, **49**(12): 4910-4919.

Yang, Y., A. Marshak, S.P. Palm, Z. Wang and C. Schaaf 2013. Assessment of Cloud Screening With Apparent Surface Reflectance in Support of the ICESat-2 Mission. *Ieee Transactions on Geoscience and Remote Sensing*, **51**(2): 1037-1045.

Yi, D.H. and C.R. Bentley 1999. Geoscience Laser Altimeter System waveform simulation and its applications. *Annals of Glaciology, Vol 29, 1999*, **29**: 279-285.