Draft, Feb. 27, 2018

ICE, CLOUD, and Land Height Satellite (ICESat-2) Project

Algorithm Theoretical Basis Document (ATBD) for Ocean Surface Height

Prepared By:
ICESat2 Science Definition Team Ocean Working Group

Contributors **/Code:**
James Morison
Ron Kwok
Ben Smith
Mike Jasinski
David Hancock
John Robbins
Suzanne Dickinson
Tim Urban

National Aeronautics and Space Administration

Goddard Space Flight Center
Greenbelt, Maryland

Suggested contents by releases

Intermediate ATBD (due Feb., 2014)
1.0 Introduction: *Short introduction of what this ATBD will cover. Covers a product, a parameter, ancillary processing, etc.*
2 - Overview and background information: *General description of the algorithm function in easily understandable terms that describes what it does, how it does it and any supporting or background information that makes it clearer. This section should provide information that the public affairs office can easily use for their release of information on ICESat-2/ATLAS.*
3- Open Ocean Products
4- Algorithm Theory: Clearly this is where much of the work over the next two years needs to be done. At this point, Section 4.0 can be very high level.
5 - Algorithm Implementation. This will also be a large part of the future work. At this point, the Input Parameters (Section 4.2) and Output Parameters (Section 4.3) is the most important part. We want to make sure the necessary parameters are being collected by ICESat-2, or from other sources.
ATBD content
All sections should be near final except section 6 on test data. Section 6 provide information as available.

ATBD final
All sections final based on ATLAS designed for algorithm implementation, testing and expected processing for at least the first year of mission. It should identify any expected tuning or calibrations that are expect to occur during the Mission verification/calibration phase (Commissioning phase?)

Abstract

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

CM Foreword

This document is an Ice, Cloud, and Land Height (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:

Thorsten Markus, ICESat-2 Project Scientist
Mail Stop 615
Goddard Space Flight Center
Greenbelt, Maryland 20771

Review/Approval Page

Prepared by:

Jamie
<Enter Position Title Here>
Org/Addresss

Reviewed by: ?????

Bea
<Enter Position Title Here>
<Enter Org/Code Here>

Tom
<Enter Position Title Here>
<Enter Org/Code Here>

Approved by: ????

Thorsten
<Enter Position Title Here>
<Enter Org/Code Here>

??????

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

Change History Log

Revision Level	Description of Change	SCoRe No.	Date Approved
	Initial Release Change 8/21/2014. In section 5.3.4.1 G: instead of: (Delete from ***Yfi*** the trailing points corresponding to the zero padding used to lengthen the received distribution, ***rcvhist***, to the next power of two. This yields the estimate of the surface histogram, ***Y***.) For version dated 02/20/2017 relative to version dated 08/31/2016 we changed sections 4.2.1.4 “A Priori Sea State Bias Estimation” and 5.2.5 “Processing to Characterize Long Wavelength Waves, Dependence of Sample Rate on Long Wave Displacement, and *A Priori* Sea State Bias Estimate” to reflect using binned photon heights instead of harmonic fits to characterize the ocean surface as part of a priori sea state bias estimation Version dated 03/29/2017 relative to version dated 02/20/2017 accept changes from before 09/01/2016 and Re: version 03/29/2017 John Robbins (8/16/2017 email) found a typo in the last paragraph of 5.2.5, which has an underlined title: "Compute photon rate slope correlation using bins with containing 2 points or more. "Compute the covariance of surface slope and photon return rate, binAVG_Xr_4_slope, equal to the sum, iii from 1 to nbins2 …." is changed to “Compute the covariance of surface slope and photon return rate, binCOVslope_Xr, equal to the sum, iii from 1 to nbins2…” Morison 11/14/2017 Re: version 03/29/2017 John Robbins (8/21/2017 email), In Section 2.2.1 there is mention of apparent reflectance, but how do we want to estimate the apparent reflectance). That’s a nut we have yet to crack. It is essentially the subsurface returns issue. We had hoped to get a handle on it with SIMPL or some other dual wavelength instrument run over the open ocean, but we won’t be getting the required data for the ocean any time soon, It		

	may have to be a research topic with ICESat2 itself. Section 5.3.2. has been largely TBD. We now specify 3 types of cal/val and explain that we are seeking support outside of NASA Cryosphere for these activities 1.1.1.1.1.1.1 In the contents page, section 2.2.2 should be and is now listed as Waves and Sea State Bias. Re: version 03/29/2017 John Robbins (9/18/2017 email): Is the parameter, binSlopeBias, actually used anywhere? See the tail end of section 5.2.5. -Initially Morison wasn't sure about this, but recent research results suggest it may be important in understanding sea state bias and its effect on the ocean scans for pointing calibration. How to determine if this is true and how to utilize binSlopeBias are TBD Re: version 03/29/2017 John Robbins (9/18/2017 email): At the tail end of the SSB code, you summarize results for all 2800m segments. Do we want a similar summary in the operational code? No, these lines of code were just for the test case only and for SSB research. 1/16/2018 for section 4.2.1.2 paragraph 2 changed 24 geo-bins to 14 geo-bins 1.1.1.1.1.1.2 Re: version 03/29/2017 John Robbins (10/17/2017 email):, Here is a list of a few items in section 3.1.3 (Corrections to height) that are presently listed as "TBD" that can actually be identified. 3.1.3.1. The geoid being reported on ATL03 comes from the mean tide EGM2008 model. Change implemented 1/31/2018 3.1.3.2. You may want to change the title of that sub-section to "Atmospheric delay corrections," in order to be		

	clear, relative to the title of the next section, which you might want to consider changing. Changed 1/31/208 1.1.1.1.1.1.3 3.1.3.3. The Inverted Barometer Effect (IBE) is only a portion of the total DAC correction. The DAC correction also accounts for wind field effects which are important in coastal areas. These are done together, not treated as separate entities. 1.1.1.1.1.1.4 So, I'm not sure whether you agree that this sub-section should be renamed "Dynamic Atmospheric Correction" to include both the IB and wind field effects. If you agree, then the sentence in the section needs to also reflect that (suggestion, below). The DAC being used comes from AVISO and is called MOG2D. You can reference ATL03, section 6.3.2. MOG2D has a spatial resolution of 1/4 degree and a time resolution of 6 hours. 1.1.1.1.1.1.4.1.1 Here is a straw-man sentence simply offered as a suggestion (too wordy?): Heights are corrected for dynamic atmospheric effect that combines two time varying effects; an inverted barometer (IB) effect and wind field forcing effect. The MOG2D modeled sea surface response to these effects has it's source from AVISO. The model is available as 6-hour, global, grids with 1/4° spatial resolution. Details about MOG2D can be found in ATL03 ATBD, section 6.3.2. 1/31/2018 - Changed text to agree with Markus et al. 2007. i.e., Dynamic Atmospheric Correction and the Inverse Barometer Effect (IBE, time-varying) Heights are corrected for the inverse barometer effect due		

	to the direct application of atmospheric pressure to the sea surface and the dynamic changes forced by wind. ICESat-2 has adopted the utilization of global, empirical, 6-h, AVISO MOG2D, 1/4° × 1/4° grids to be used as a near-real time Inverted Barometer (IB) and Dynamic Atmospheric Correction (DAC)[*Carrère and Lyard*, 2003]. These grids are forced by the European Center for Medium-Range Weather Forecasting (ECMWF) model for the surface pressure and 10-m wind fields. This combined correction typically has amplitude on the order of ±50 cm [*Markus et al.*, 2016]. 1.1.1.1.1.1.5 3.1.3,4. Tidal Corrections 1.1.1.1.1.1.6 You provide a list of these, so here are the specific models (in the order mentioned): 1.1.1.1.1.1.7 Solid earth tides: IERS (2010) conventions (details in ATL03 ATBD section 6.3.3). 1.1.1.1.1.1.8 Ocean load tides: We're using load tides that are inherent in GOT4.8 (ocean tidal model). Details are found in ATL03 ATBD section 6.3.4. 1.1.1.1.1.1.9 Pole Tides: Pole tides include both solid earth and ocean pole tides. These each are computed via IERS (2010) conventions. Details are found in ATL03 ATBD sections 6.3.5 and 6.3.6, respectively. 1.1.1.1.1.1.10 Sea-Surface height corrections: Same as ocean tide corrections? The ocean tide model is GOT4.8. Details in ATL03 ATBD, section 6.3.1. Equilibrium tides: We're using a subroutine attached to GOT4.8 called LPEQMT.F. It is a Fortran routine where fifteen tidal spectral lines from the Cartwright-Tayler-Edden tables are summed. Richard Ray is it's author, and a line appears in section 6.3.1 of the ATL03 ATBD		

	about it. 1/31/2018 Changes made to tide sectipn 1.1.1.1.1.1.11 3.1.3.5. Dynamic topography. I believe that this sub-section is now superfluous, since section 3.1.3.3 now includes these effects. Also, you had characterized this a "static," but it's time-varying. So you may want to enhance 3.1.3.3 and delete 3.1.3.5. 1.1.1.1.1.1.12 There are quite a few other TBDs in Chapter 3, but I lack the expertise to advise on most of them. I hope what I've provided is helpful 1/31/2018 Dynamic Topography section deleted as superfluous 2/1/2018 Accept changes up to 1/31/2018 and reorder some headings in section 5 to be more consistent with ATL07 ATBD and correct table of contents. 2/15/2018 Added material on gridding to sections 3.2 2/19/2018 Significant weeding of redundant and outdated, incomplete sections. Resulted in moderate reorganization reflected in comparison with 02/14/2018 draft. 2/24/2019 Modified gridding in sections 3 to include aggregating DOT histograms. Added Figs X2 and X3 (Figs 12 and 13 after 2/27/2018) showing ATL12 and ATL19 (placeholder) block diagrams 2/26 and 2/27/2018. Corrected figure numbers and equation numbers. Improved formatting, edited and made miscellaneous corrections.		

List of TBDs/TBRs

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

Table of Contents

List of Figures

List of Tables

1.0 INTRODUCTION

This ATBD will cover the retrieval of Sea Surface Height (SSH) from ICESat2 ATLAS laser returns. For the purpose of this early draft, the levels of spatial and temporal averaging to be included are still to be definitively decided, as is the amount of high-level data processing to produce gridded fields of SSH and such things as Dynamic Ocean Topography (equal to SSH minus the geoid).

Other ICESat2 ATBDs describe the products for ice sheets, vegetation, sea ice and inland water, the latter two having close relations with the ocean ATBD. Technical ATBDs include orbit and attitude calculations, corrections for atmospheric path-length delays, and corrections for changes in the surface heights due to tidal effects; these other data are needed to convert ranges into absolute surface heights with respect to the geoid.

This document will address sampling the ice-free world ocean, but not ice-covered or inland waters, because the ICESat-2 sampling scheme is different for the open ocean, generally sampling strong beams only, and because the defining characteristics, (e.g., well developed surface waves) are unique to the open ocean. It will include coastal ocean waters. Because of the increasing seasonal variation in the sea ice cover, the boundary between the open-ocean and ice-covered ocean domains will vary seasonally. This ATBD will share some considerations and features with the sea ice ATBD (surface finding algorithm, concern for tides and the geoid) and the vegetation and inland water ATBD (concern with waves in large bodies of inland water). We consider as input data level 2 photon heights for each of the three strong beams along the satellite track along with the required navigation information. (In certain ocean regions near land or sea ice, the weak beams may be also be active. In these cases the weak beam data over the ocean should be processed the same way as the strong beams.) As output, the processing to be described will produce ATL12/L3A (Appendix A), the height of the open ocean surface at a length scales between 70 m (100 Hz) and 7 km (1 Hz), determined by an adaptive surface finding algorithm. Output will include estimates of height distributions (decile bins), significant wave height, surface slope, and apparent reflectance.

Section 2 provides an overview of the ocean altimetry issues,

Section 3 discusses the ocean products the parameters that reside in each product as well as ancillary geophysical parameters, of interest to science users, which are used in the derivation of these ICESat-2 products.

Section 4 provides a theoretical description of the algorithms used in the derivation of the ocean products.

Section 5 describes the specific implementations of the algorithms that are relevant to the development of the processing code. Included here are both algorithmic details and some software architecture details on throughput optimization and computational loading.

Section 6 provides the processing requirements for data quality monitoring and control. These are provided to users as quality assessments of the individual parameters on each file and to provide criteria for automatic quality control to facilitate timely distribution of the product to the users. Summary statistics or images are provided that allow users to easily evaluate (or Browse) whether the data would be useful and of adequate quality for their research, and as needed to aid in the quick approval or disapproval of products prior to public distribution.

Section 7 describes the testing and validation procedures that are planned

2.0 OVERVIEW AND BACKGROUND INFORMATION

The Advanced Topographic Laser Altimeter System (ATLAS) consists of both LIDAR and altimetry subsystems that will fly on the dedicated platform comprising the mission referred to as ICESat2, the Ice, Cloud, and Land Height Satellite. The following subsections discuss the basic concepts of how ATLAS works and the Level 2 data that we will turn into sea surface height. It then describes the basic properties of the sea surface and how these relate to processing the ICESat2 data.

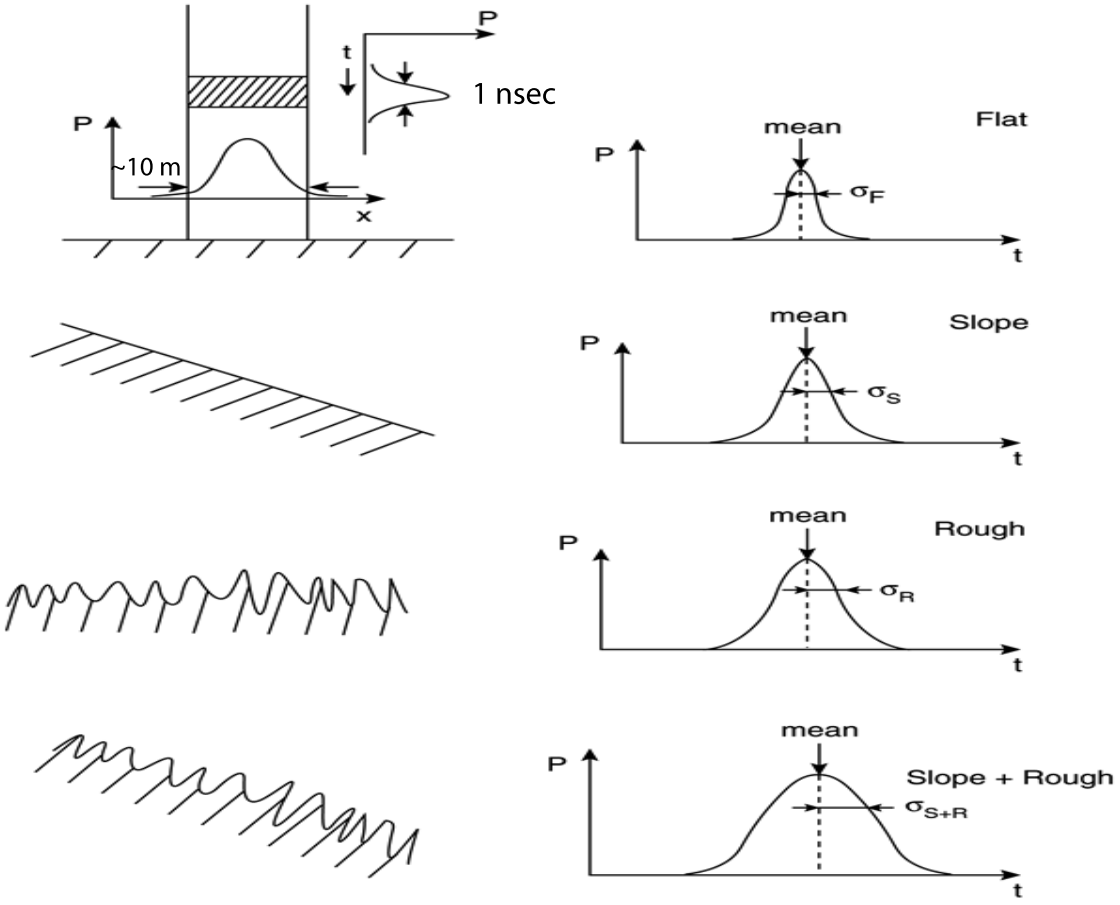


Figure 1 - Characteristics of an idealized distribution of ATLAS return photon arrival times assuming a Gaussian instrument impulse response and a reflective surface with Gaussian roughness. In general surface slope and roughness both broaden the return distribution. For the open ocean the slope effect is likely negligible compared the effect of roughness due to surface waves. An important complication is that surface waves produce a non-Gaussian surface due to their broad troughs and narrow peaks.

The primary purpose of the ATLAS instrument on the ICESat2 mission is to detect

surface height changes. With respect to the ocean, these represent Sea Surface Height (SSH). By way of technical history, the predecessor of ICESat2 ATLAS, the ICESat GLAS instrument, used a laser altimeter to measure the range to the surface. Ranges were determined from the measured time between transmission of the laser pulse and detection of the pulse reflected from the surface and received by the instrument. The GLAS laser footprint diameter on the surface due to beam spreading was nominally 70 m, and the duration of the transmitted pulse was 4 ns.

The ICESat2 ATLAS instrument will sample at a higher rate, 10 kHz with 1-2 ns pulse width, with a smaller intrinsic footprint, ~10 m (Fig. 1), and will rely on detecting the range traveled by individual photons. Experience with MABEL suggests that the instrument impulse response distribution, due largely to the variation in transmit times for the photons of each pulse, will be non-Gaussian with significant skewness. Consequently, four points will be measured on the out going pulse distribution and these will be part of the ICESat2 raw data stream. The distribution of receive times of the surface-reflected photons, and hence photon heights, will be broadened by the distribution of surface heights within the footprint as depicted in Figure 1 for an idealized Gaussian distribution of photon return times or heights. Photon return-times will be digitized in 200 ps (3 cm) range bins. However, the origination time of a photon within a pulse is unknown, and the uncertainty in the time of flight of a single photon will be between 1 and 1.5 ns RMS (30-45 cm flight time) due mostly to the laser pulse width, with smaller contributions from other effects within the instrument. This time uncertainty corresponds to between 15 and 22 cm RMS in range.

Return photon arrivals will be aggregated at a scale depending on the surface type being over flown. Over the open ocean, short segment aggregates will be retrieved at 100 Hz (100 pulses over 70 m of track) and long segment aggregates retrieved up to 1 Hz (10,000 pulses over 7 km corresponding to 25 times the atmospheric sample rate. In contrast, to ensure adequate detection of leads, over sea ice return photons may be taken at the maximum rate of 10 kHz (each pulse every 0.7m).

2.1 Open Ocean Background

Over distances of cm to a few hundred meters, the sea surface is roughened by waves and ocean swell, but over distances of many km, the sea surface is almost flat. Nevertheless, surface slopes and long-wavelength undulations are present, caused by variations in Earth's gravity field represented by the geoid, ocean currents, and variations in atmospheric pressure and seawater density. Satellite radar altimeters have shown remarkable success in measuring sea-surface height and tracking changes in circulation and mean sea level. However, the major satellite radar altimeters, TOPEX/Poseidon/Jason do not go beyond about 62° latitude, and thus miss the sub-Arctic seas and southern parts of the Southern Ocean that are critical to the global overturning circulation and the fate of sea ice in the ice covered Arctic Ocean. ICESat2 over the open ocean will cover this gap between the temperate lower-latitude ocean and the sea ice covered regions of the Arctic

and Antarctic

As discussed above, the distribution of ICESat photon return times, or apparent heights from Level 2 processing, will be determined by the mean SSH, the mean surface slope and a surface roughness within the footprint. Among these, the effect of sea surface slope on the photon height distributions should be small compared to the effect of roughness due to surface waves. For the aggregation scales above, the SSH is the sum of the geoid height (fixed in time and on the order of meters amplitude), the dynamic ocean topography (DOT, on the order of centimeters to tens of centimeters amplitude) associated with mean surface currents, tides (mainly diurnal to semidiurnal periods with amplitudes of up to meters), and sea surface atmospheric pressure (SAP, as much as 10s of cm). As a first approximation, SSH can be estimated as the mean of the distribution of photon heights within a vertical window about the ellipsoid or an approximated canonical sea surface height (e.g., 30 m from the estimated geoid) and over the along track aggregation scales to be decided as above. However, the special and rapidly varying nature of the sea surface

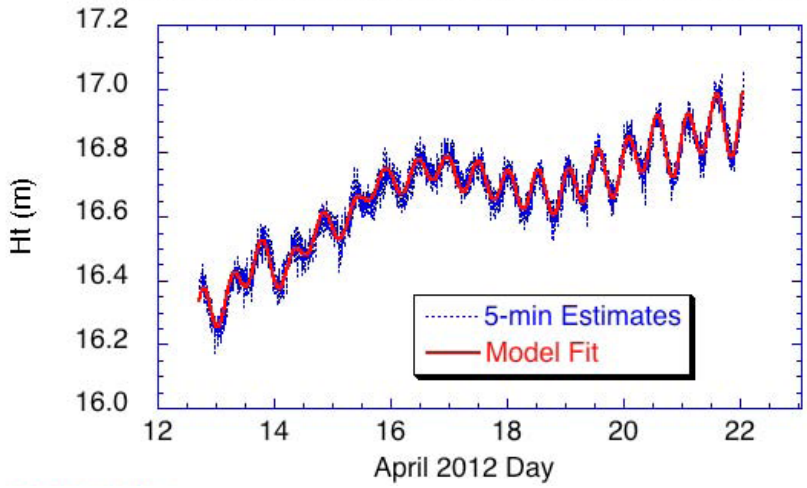


Figure 2. Geodetic height measured with a geodetic grad dual-frequency GPS with PPP processing at a drifting ice camp near the North Pole plotted with an *ad hoc* model of ocean tides and longer variations assumed due to the geoid and DOT variations. The geoid variations over the 60 km drift are estimated to 60 cm, tides 10 cm peak to peak, and DOT only a few cm.

eaningful SSH from ICESat2
ıg in mind that the magnitude
es in DOT and mean SSH, are significantly smaller than most of the other components of SSH.

Tides, Medium-Frequency Changes in Circulation and Aliasing: ICESat2 will complete one orbit of the earth in about 1.5 hours. The repeat period is under discussion but has mainly settled on 91 days. More importantly, many of the most important areas and times to oceanographers are prone to being stormy and overcast. Thus, it is likely that ICESat will only get good surface returns infrequently over large and important regions of the open ocean. Tides and medium to high-frequency changes in circulation will seriously alias into slowly sampled regions, making aggregation into valid long-term means difficult. At some point in ICESat2 ocean processing this will have to be addressed. This can be done by referencing the photon heights to the SSH predicted by a model such as the Ocean Model of Circulation and Tides (OMCT) as used for dealiasing gravity data from the Gravity Recovery and Climate Experiment (GRACE). Essentially, the predicted tidal and circulation contributions to SSH are subtracted from each ICESat2 height observation, the residual is averaged

monthly and similarly averaged model result is added back in. Only the error in the modeled response is aliased. This procedure may be beyond the scope of Level 3 data, but the same ocean model could be used to more narrowly and accurately define the photon height band that needed to be considered in selecting true ocean surface photon returns.

The Geoid, Narrowing the Window on Photon Returns, and Measuring Mean DOT: ICESat2 photon heights will be heavily influenced by the geoid. An example of this for the Arctic Ocean near the North Pole (Fig. 2) shows changes in SSH associated with geoid height variations of about 1 cm per km of horizontal distance. Given the potentially patchy and temporally sparse character of ICESat2 ocean returns, the variations in the geoid from a geoid model (e.g., EGM2008 or an improvement thereon using GOCE and GRACE data) should be extracted from the basic height data to allow meaningful spatial means of SSH to be made. The spatial resolution of these geoid corrections is an important issue affecting the smaller DOT and SSH signals.

Surface Atmospheric Pressure and the Inverse Barometer Effect: Changes in surface atmospheric pressure a short time scales directly cause an inverse deflection of SSH. As with tides, ocean circulation changes, and the geoid, these direct pressure disturbances at the time and place of each ICESat2 photon retrieval should be estimated and removed from estimated SSH before aggregation to avoid spatial and temporal aliasing. This will require bringing atmospheric reanalysis products or direct SAP observations into the processing stream. Knowledge of the likely inverted barometer effect will also help in narrowing the range of return photon heights and thus identifying photons returning from the sea surface.

2.2 The Importance of Waves

2.2.1 Waves and Reflectance

Surface Waves Reflectance, Scattering, and the Sea State Bias: We expect surface waves to have dominant effects on the ICESat2 returns from the open ocean.

Reflectance: Surface waves and wind speed have a critical effect on reflectance. Menzies et al. [*Menzies et al.*, 1998] indicate reflectance is given by:

$$R = W R_f + (1-W) R_s + (1-W R_f) R_u \qquad (1)$$

where R is the total backscatter or retro-reflectance,
W is the fraction of the ocean covered by foam (up to O[1] for winds above about 7 m/s)),
R_f is the reflectance of foam,
R_s is reflectance due to specular reflection from the wave roughened surface, and
R_u is the apparent reflectance due to light that penetrates and comes back up through the surface anywhere it is not reflected downward by surface foam.

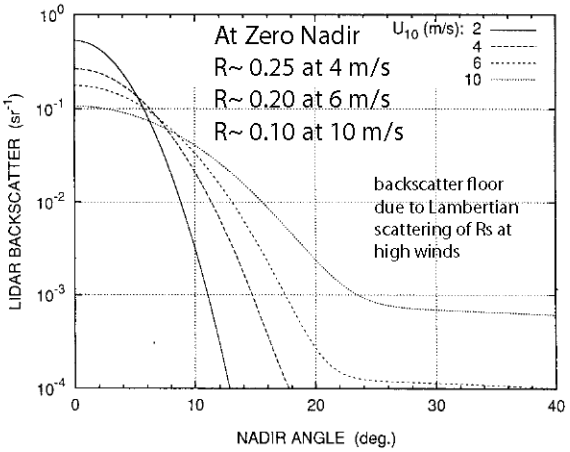


Figure 3. Reflectance as modeled by *Menzies et al.* [1998] as a function of wind speed and nadir angle.

R_f is considered diffuse Lambertian scattering independent of incidence angle. As we would expect, R_s is highly dependent on incidence angle and is the dominant form of backscatter. R_u is Lambertian but it is usually small, about 0.0075 maximum. Menzies et al, model the reflectance owing to R_s and R_f. The main part is an inverse relation between R_s and the variance in surface slope, $R_s \sim 1/ <S^2>$, due to wind generated surface waves. They use a relation between wind speed, U_{10}, and $<S^2>$ by Wu [*Wu*, 1972] based on results of Cox and Munk [*Cox and Munk*, 1954]; $<S^2>$ goes up as the log of wind speed. The model results are given in Figure 3, wherein the strong dependence on nadir angle is apparent. The exception is at high wind speeds where the backscatter from foam starts to become important, flattening the *R* curves for nadir angles greater than 30°. Other results in [*Menzies et al.*, 1998] show good agreement between the model and data from the Lidar In-space Technology Experiment (LITE) shuttle lidar mission of September 1994. The modeled reflectance essentially agrees with the values for reflectance used in the design performance study for ATLAS (Table 1). However, most important with respect to this ATBD, the dependence of sea surface directional reflectance on surface wind stress suggests a method for deriving surface wind speed from ICESat2 measurements of sea surface backscatter. Working solely with ICESat2 observations, we can conceivably

Table 1

Case	Description	Water Reflectance, Lambertian, Green	Water Reflectance, Lambertian, IR
10a	Conical Scan, low wind	0.28	
10b	Conical Scan, medium wind	0.12	
10c	Conical Scan, high wind	0.07	

Modeled Reflectance for Zero Nadir Angle from Figure 4, *Menzies et al.* [1998] Figure 1

Case	Description	Water Reflectance, Green
10a	Wind = 4 m/s	0.25
10b	Wind = 6 m/s	0.20

estimate U_{10} from reflectance. As we discuss below this estimate of U_{10} may be used with SWH, estimated from the variance of the photon height distributions, to calculate the Sea State Bias in ICESat2 SSH measurements.

2.2.2 Waves and Sea State Bias

Sea state bias is a critical issue that has been found to relate to the amplitude of waves and to wind forcing. The SSB problem is fundamental and has received considerable attention with respect to radar altimetry [*Elfouhaily et al.*, 2000; *Gaspar et al.*, 1994]. Improved corrections for SSB in the TOPEX altimeters have been developed [*Chambers et al.*, 2003] relating SSB to wind speed, U_{10}, and significant wave height (SWH, the trough-to-crest height of the highest third of ocean waves) measured with the radar altimeter. Studies using both laser and radar instruments find a fair degree of commonality [*Chapron et al.*, 2000; *Vandemark et al.*, 2005]. Urban and Schutz [*Urban and Schutz*, 2005] compare ICESat-derived SSH with TOPEX/Poseidon SSH and find a negative 10 cm bias in ICESat relative to the radar altimeters. They indicate that the bias is unknown, but that it may be related in part to sea state, and they recognize that the radar altimeter is corrected for SSB using a relation with SWH. Unfortunately, a SWH parameter was not provided with the ICESat GLAS data, a shortcoming we cannot repeat with ICESat2.

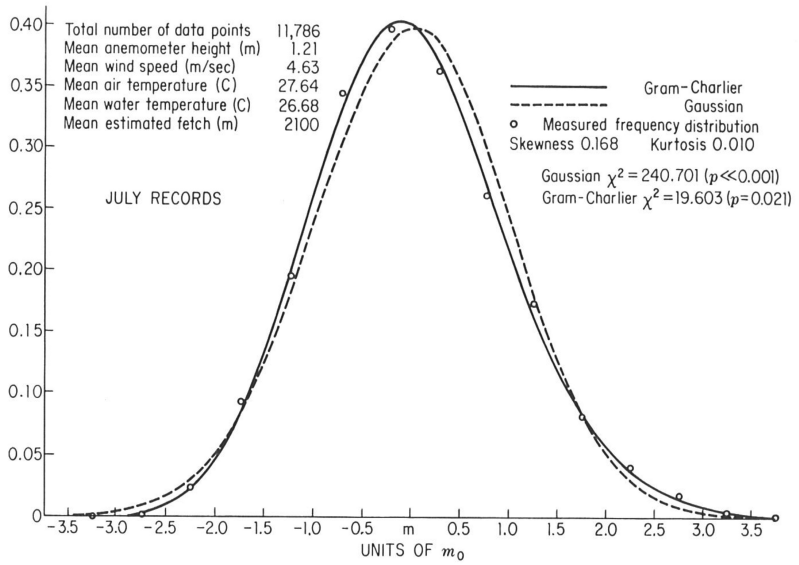


Figure 4. Typical Gram-Charlier distribution sea surface height in a wave environment. The positive skewness and excess kurtosis (Gaussian skewness and excess kurtosis = 0) is due to truncation at low heights and a long tail at high heights. Fig. 7.4-1 of *Kinsman* [1965].

The ICESat ATBD document made the assumption that sea surface heights are Gaussian distributed, so that with a Gaussian transmission pulse, the reflected pulse received by ICESat could be assumed to be Gaussian also. In fact, the ocean surface is not Gaussian and this affects the reflection of either radar or light from the surface. The distribution of surface height in a wave-covered ocean (Fig. 4) is typically given by the Gram-Charlier distribution [Kinsman, 1965]. For this the third and fourth moments of the distribution expressed as skewness and kurtosis (or excess kurtosis = kurtosis-3) are important and indicate truncation of the distribution at low heights with a

long tail at high heights. This reflects the shape of surface waves.

In general, surface waves have trochoidal shape, with narrow, steep sided peaks and relatively broad flat troughs (Fig. 5). This applies particularly to the short-wavelength waves proximally forced by wind. Photons striking the upper portions of these wave are less likely to be returned to ICESat2 than photons striking the lower portions of wave surfaces; peaks are undersampled relative to troughs resulting in a sea state bias (SSB) in estimates of SSH based on the mean of an ICESat2 photon height distribution or, in the case of ICESat, mean arrival time of a return pulse.

Figure 5. Trochoidal shape of surface waves at the limit of steepness, wavelength equal to seven times height. The sharply peaked trochoidal shape is common to all surface waves longer than capillary waves, which have an inverted trochoidal shape. At the steepness limit the waves begin to break. (Figure from http://hyperphysics.phy-astr.gsu.edu/hbase/waves/watwav2.html)

However, commonly the sea surface is composed of swell and wind waves. Swell is the manifestation of large long-wavelength waves that are mature and forced elsewhere. The wavelengths are long enough that in spite of significant amplitude, the slopes are small and the waves are linear and well represented by sine waves. The wind waves, forced by local wind, are shorter (down to capillary wavelengths) and steeper with trochoidal shape. In the extreme, when they reach the limit of steepness (Figure 5), they break and form white caps. As such, these waves can by their shape affect sea state bias in altimetry returns, but their direct bias is small because their amplitudes are small. However, several studies including one done for the development of this ATBD, suggest the character of the short waves varies with position on the larger linear waves and thus can contribute to large SSB.

We have performed a study of the effect of long waves on short waves and their implications for ICESat-2 sea state bias. We have sought a model sea surface suitable for driving the NASA GSFC numerical ATLAS simulator. With this we will be able to test the performance ATLAS over the open ocean and evaluate such things as the sea state bias in mean sea surface height measurements. There are two elements to our artificial sea surface. The first is a realistic representation of the height and surface slope due to waves with wavelengths greater than the 10-m footprint of an ATLAS laser pulse. This captures the bulk of the energy in the wave field. Because the waves are long, a linear model of them can be used. The second component is a measure of the distribution of surface slopes due to the waves with wavelengths less than 10-m laser pulse footprint of ATLAS. The distribution of surface slopes, shifted by the instantaneous slope of the long wave component determines the probability of ATLAS finding a specular return resulting from a

photon laser pulse. Knowing this as a function of surface height and slope due to the long-wave will allow us to predict SSB in the ICESat2 measurements.

We have taken advantage of a code simulating long-wave surface height based on a long-wave spectrum developed by Donelan et al. (1985) to model the long wave component of our ATLAS ocean surface model (AOSM).

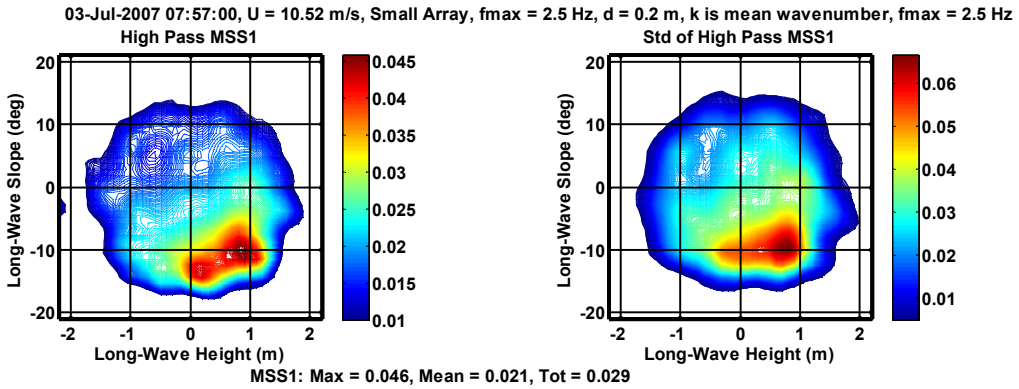


Figure 6. Left – Short wave MSS as a function of long-wave height and slope. Right – the standard deviation of short wave MMS as a function of the same variables.

Our main challenge has been determining the mean square slope (MSS) representing short-wave roughness and its dependence on long-wave height and slope. We had thought that wave wire or some other type of field observation would have been analyzed or at least readily available to determine this relationship, but surprisingly we found no such analysis in the literature. Instead we have analyzed original data from a four-wire wave-gauge array on the University of Miami's ASIS buoy [Will Drennan, personal communication]. The work is described in detail by W. Plant.[*Plant*, 2015a; b] but we provide a brief synopsis here. The ASIS Buoy data were collected in the Deep Ocean Gas Exchange Experiment in the Atlantic Ocean on July 3, 2007 at a wind speed between 10 and 11 m/s. The wave-gauges measure surface heights at orthogonal positions 20 cm apart, so in principle any two gauges measure slope directly down to ~ 2 Hz or wavelength ~ 0.8 m in this case. A single gauge yields a time series of height, and with the dispersion relation, the spectrum of height can be converted to a spectrum of surface slope. The single gauge approach is complicated by the need for the dispersion relation for the short waves to be modified to account for the orbital velocities of the long waves. We have compared the two approaches and found the two-gauge approach yields results inconsistent with the known gauge spacing. However, we can account for the complications to the dispersion relation and produce consistent estimates of the short-wave slope spectra. Spectra for higher frequencies (wave numbers) than 2 Hz are extrapolated using the observed frequency^{-1} spectral slope at 2 Hz. The MSS is obtained by integrating the spectrum out to a maximum frequency (100 Hz), above which increases in MSS are negligible.

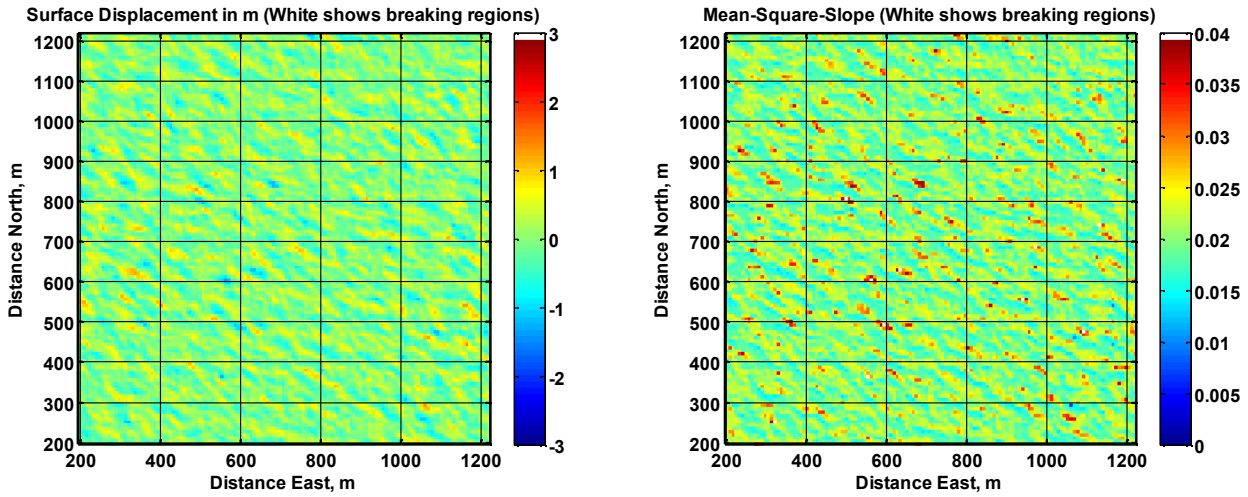


Figure 7. From *Plant* [2015b]. Left – Simulated 2D surface displacement , Right – 2D variation of short-wave mean MSS values that were calculated up to 2 Hz. Downwind direction is from upper right to lower left.

The preliminary result of this analysis has been the dependence of observed short-wave MSS on the slope and height of the long-waves (Fig. 6). MSS is greatest for high heights and negative slopes, i.e., on the upper part of the downwind face of the long-waves. The correlation maps of Figure 6 have been used as a table to specify MSS in a 10-m spot corresponding to the 2D model of long-wave slope and displacement derived from the Donelan spectrum. A realization of long wave height and slope and the corresponding 10-m spot MSS is shown in Figure 7 [*Plant,* 2015a&b].

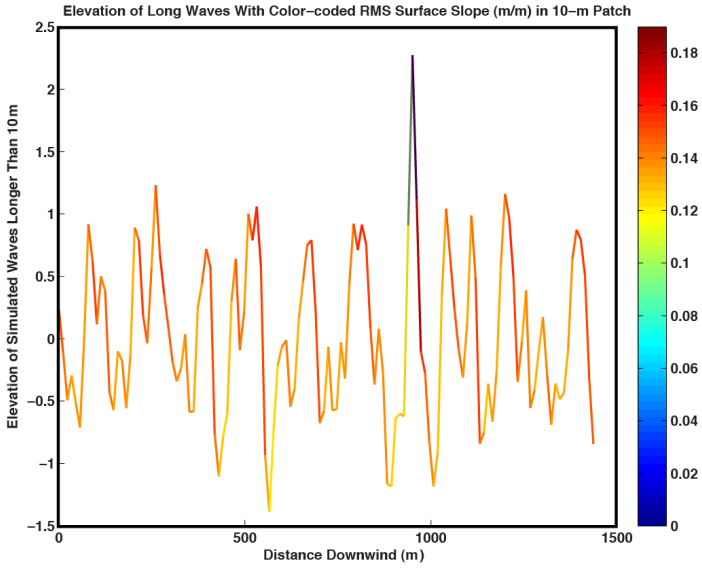


Figure 8. Height due to long-waves versus distance downstream (from upper right corner to lower left in Figure A-left) with color-coded corresponding MSS from Figure A-right.

A 1-D slice in the downwind direction (Figure 8) through the combined data from Figure 7 illustrates how MSS slope tends to increase on the upper parts of the downwind face of waves. It also illustrates in a much-

simplified way how the interplay of the short-wave MSS and long-wave amplitude and slope can affect the probability of specular reflection and consequent sea state bias. For each point along the slice, we assume the surface is normally distributed with a mean equal to the local long-wave slope and a variance equal to the MSS corresponding (Fig. 6) to the long-wave height and slope. The probability of any one photon striking the spot finding a specular point is estimated as the integral of the idealized slope distribution between plus and minus 10^{-5} radians (Fig. 9).

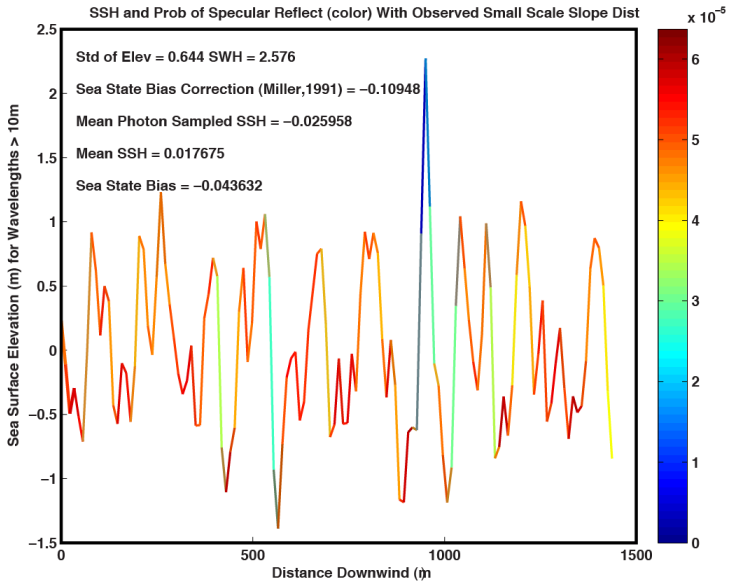


Figure 9. Height due to long-waves versus distance downstream (from upper right corner to lower left in Figure 6-left) with color-coded probability of specular reflection.

The probability of finding a specular point is highest in the toughs and at the peaks, likely because the mean slope is zero. However, the increased MSS in the upper parts of the waves tends to decrease the probability of specular points for higher heights. We know the true mean height along the slice is 1.8 cm. Weighting the samples of height by the probability of a specular point yields a photon sampled mean height of -2.6 cm, a -4.4 cm bias. Interestingly this is about twice the typical sea state bias for radar altimeters for the 10-m wind speed at the time of these ASIS Buoy observations. The bias is not due to nonlinearities on the long-waves. The simulation uses a linear spectrum-based representation of the long waves; there are no trochoidal long waves. When we repeat the probability weighted mean calculation assuming the MSS is uniformly equal to the cut-wise mean MSS, the mean height is only biased -0.7 cm.

There are improvements to be made to this wave analysis, including drawing the MSS values as random values from distributions conditioned on long-wave slope and height and doing the analysis for data at other wind speeds. However, this limited sample suggests that that the systematic distribution of small-scale surface roughness towards the upper parts of the waves produces a negative sea state bias because the wave troughs provide more specular returns than the wave crests. This type of bias is likely compounded by shape-induced biases due to any nonlinearity in the long-waves. As a result is that the SSH observations by ICESat-2 will be non-Gaussian and exhibit sea state bias.

However, further analysis of the wave gauge data taking into account both the Doppler shifting of the small scale wave by long wave orbital velocity of the long waves and the effects of convergence by long-wave orbital velocity on short wave amplitude results in almost the opposite dependence of short wave slope on long wave height and slope and suggests a positive sea state bias. {SEE BILL'S LATEST WRITEUP], Consequently, lacking further data about the dependence of short wave characteristics on long wave position, we seek a method whereby the SSB can be determined tom ICESat-2 data themselves.

2.2.3 ICESat2 Height Statistics and Sea State Bias

Sea State Bias (SSB) is the error in average sea surface height measured by an altimeter due to the dependence of altimeter returns over different parts of waves. The SSB for radar altimeters is negative because the troughs of waves reflect more radar energy than the crests of waves. This effect is averaged over many surface waves encompassed by the typical radar footprint (e.g. 17 km for CryoSat2). Improved corrections for SSB in the TOPEX altimeters have been developed [*Chambers et al.*, 2003] relating SSB to wind speed, U_{10}, and significant wave height (SWH, the trough-to-crest height of the highest third of ocean waves) measured with the radar altimeter. All these studies have measured the SSB empirically by comparing SSH measured by various means or comparing repeat measurements at the same location by a given satellite for differ sea states and wind speeds [*Hausman and Zlotnicki*, 2010].

Studies using both laser and radar instruments find a fair degree of commonality [*Chapron et al.*, 2000; *Vandemark et al.*, 2005]. Urban and Schutz [*Urban and Schutz*, 2005] compare ICESat-derived SSH with TOPEX/Poseidon SSH and find a negative 10 cm bias in ICESat relative to the radar altimeters. They indicate that the bias is unknown, but that it may be related in part to sea state, and they recognize that the radar altimeter is corrected for SSB using a relation with SWH. Unfortunately, a SWH parameter was not provided with the ICESat GLAS data, a shortcoming we cannot repeat with ICESat2.

A Priori Estimation of SSB Using Only Altimeter Data

The small footprint of ICESat2 (17 m for 4σ diameter Gaussian intensity), short distance between pulse/footprint centers (0.7 m) and the essential point sampling at random positions within the footprint of the photon counting lidar, may present an opportunity to capture the shape of the long wave length, energy containing surface waves. If this is true it will be possible to estimate ICESat2 sea state bias solely on the basis of contemporaneous ICESat2 returns without resorting to outside data or even comparison with past ICESat2 returns. The key is evaluating the rate of surface photon returns as a function of surface height. A simple example of this considers an ocean surface with sea surface height, H_{ss}, disturbed by a long sinusoidal surface wave with amplitude *A,* plus random normally distributed perturbations, *N*:

$$Y \quad H_{SS} \quad A\sin\frac{2}{L}x \quad N(0, \quad) \qquad 2$$

If ICESat-2 samples this surface at a constant rate, $\bar{r}$ surface returns per meter, the surface will be uniformly sampled, and there will be no sea state bias. If there is a variation in sample rate, *r*, that is correlated with variations in sea surface height such that:

$$r \quad \bar{r} \quad A\sin\frac{2}{L}x \,, \tag{3}$$

where α is the covariance between *r* and *y* normalized by the variance in *y*:

$$\frac{\mathrm{cov}(ry)}{\sigma_y^2} \,. \tag{4}$$

The surface height estimate, y_e, over a large distance, *X*, is the average of the true height weighted by the sample rate:

$$y_e \quad \frac{1}{X\bar{r}}\int_0^X Yr\,dx \quad \frac{1}{X\bar{r}}\int_0^X (H_{SS} \quad A\sin\frac{2}{L}x \quad N)r\,dx$$

$$\frac{1}{X\bar{r}}\int_0^X (H_{SS} \quad A\sin\frac{2}{L}x \quad N)(\bar{r} \quad A\sin\frac{2}{L}x)dx \tag{5}$$

$$\frac{1}{X\bar{r}}\int_0^X H_{SS}(\bar{r} \quad A\sin\frac{2}{L}x)dx \quad \frac{1}{X\bar{r}}\int_0^X H_{SS}\bar{r}\,dx$$
$$\frac{1}{X\bar{r}}\int_0^X (A\sin\frac{2}{L}x)(\bar{r} \quad A\sin\frac{2}{L}x)dx \quad \frac{1}{X\bar{r}}\int_0^X (A\sin\frac{2}{L}x)(\quad A\sin\frac{2}{L}x)dx$$
$$\frac{1}{X\bar{r}}\int_0^X N(\bar{r} \quad A\sin\frac{2}{L}x)dx \quad 0 \tag{6}$$

For *X* equal to an integral number of wavelengths or very long compared to many wavelengths:

$$y_e \quad \frac{1}{X\bar{r}}\, H_{SS}X\bar{r} \quad \frac{X}{2} \quad A^2 \quad H_{SS} \quad \frac{A^2}{2\bar{r}} \tag{7}$$

So sea state bias, $SSB = y_e - H_{SS}$, is:

$$SSB \quad y_e \quad H_{SS} \quad \frac{A^2}{2\bar{r}} \tag{8}$$

Similarly, Arnold et al. [*Arnold et al.*, 1995] in evaluating Ku-band SSB measured with a combination of radar altimeters and capacitive wave wires mounted on an oil platform in the Gulf of Mexico, calculate electromagnetic bias, ε, due to the variation of backscatter coefficient with surface waves as

$$\frac{\sum_{i=1}^{N} b_i^0 \, \eta_i}{\sum_{i=1}^{N} b_i^0} \tag{9}$$

where η_i is the measured surfaced displacement corresponding $A\sin(2\pi x/L)$ in our idealized example and b_i^0 is the measured backscatter coefficient proportional to the rate of photon returns, r. Our α expressed in the terms of Arnold et al. [1995] is:

$$\frac{\frac{1}{N}\sum_{i=1}^{N} b_i^0 \, \eta_i}{\frac{1}{N}\sum_{i=1}^{N} \eta_i^2}, \tag{10}$$

so their expression for electromagnetic bias is the same as the SSB for the sine wave example:

$$\frac{\sum_{i=1}^{N} b_i^0 \, \eta_i}{\sum_{i=1}^{N} b_i^0} \quad \frac{\sum_{i=1}^{N} \eta_i^2}{\frac{1}{N}\sum_{i=1}^{N} b_i^0} \quad \frac{A^2}{2\bar{r}} \quad SSB \tag{11}$$

Example of SSB Determination for the Uneven Data Distribution of a Photon Detecting Pulsed Lidar Altimeter

Application of 8 is straightforward, and is accurate in cases such as that of *Arnold et al.* [1995], in which the spacing of the data is uniform and the energy of the return is measured and is the metric weighting the sampled surface height. In that application, each radar pulse was powerful and the range and footprint size were small so that every pulse returned a surface height. The energy returned with each pulse and its correlation with height could be expected to aggregate into the SSB over the much larger footprint of satellite radar altimeter.

This is not totally true in the case of the pulsed photon-counting altimeter such as MABEL or ICESat2; the return rate of the height samples is not uniform and the average sea surface height is derived from the histogram of photon heights. The correlation between the rate of surface returns and surface height is only one of several components in the histogram skewness. Other contributors to skewness include the true distribution to surface height and subsurface returns. However, the dominant surface waves typically have long wavelengths and are nearly linear so they can be approximated as sine waves. With this in mind, we experimented with fitting multiple sine waves to the raw history of unevenly spaced photon surface heights and calculating the correlation of observed rate of surface returns with the sine wave fit to yield an estimate of SSB using equation 7.

However, in testing the method on Greenland Sea MABEL data and Airborne Topographic Mapper (ATM) over the Pacific, we found the method was sensitive to finding the correct w t method is ate of ph ese been av s easier to au rmonic m

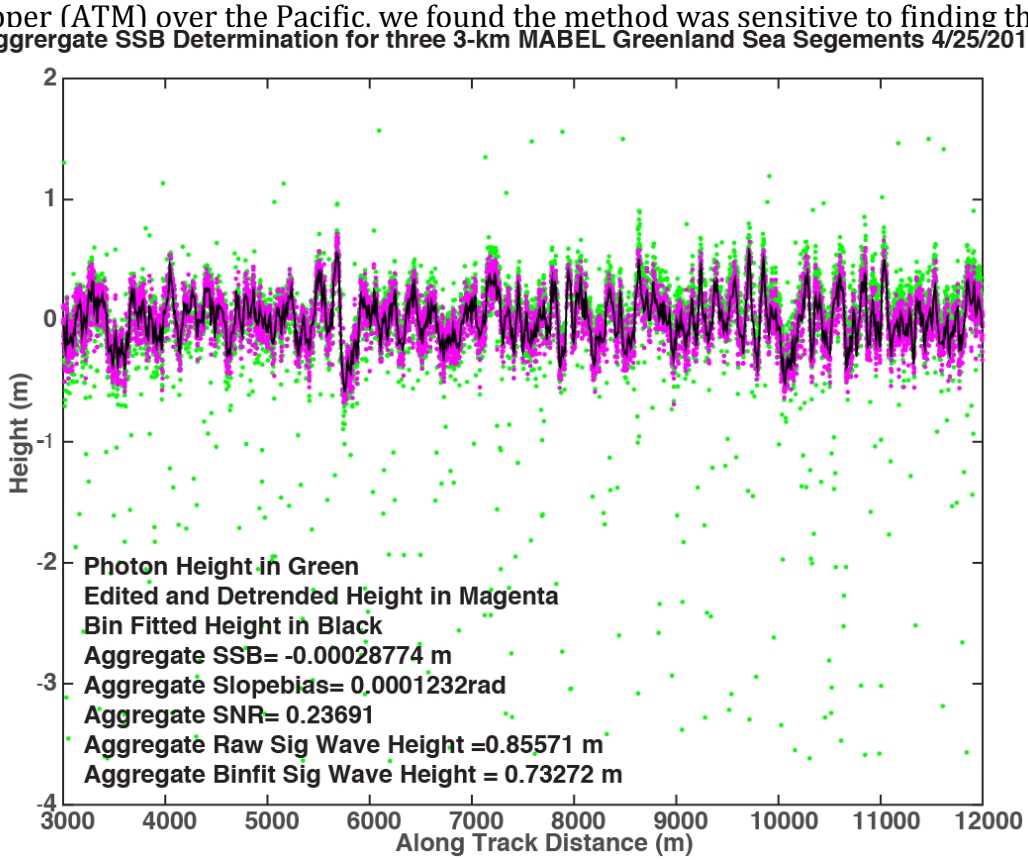


Figure 10. Surface return heights in green from MABEL transit over the Greenland Sea versus distance along track. Three-sigma edited and detrended heights are plotted in green and a10-m bin averages are plotted in black. The bin averages and a priori SSB estimates were computed in three 3000- m segments and aggregated to yield SSB=-0.00029 m.

MABEL data from the Greenland Sea in April 19, 2012 illustrates the application of equation 7. In this example, a 9-km length of MABEL track is broken into 3-km segments. Each 3-km segment is divided into 300 10-m along track bins. Once surface finding selects the surface photon heights, these are edited with two passes of a 3-standard deviation editor, and detrended with along track distance.

The rate of photon returns is computed from the along-track record of sample intervals. The record of sample intervals is given by the difference between successive surface photon along-track locations. For the MABEL data shown here, the sample spacings vary significantly but average a little over one meter. To align the data, the surface height and sample position are interpolated to the center of the sample interval by computing the average of pairs of successive sample heights and position. The resulting records of sample rate are then accumulated in 10-m along track bins and averaged. The bin averaged sample

intervals are then inverted to yield the bin averaged sample rate. Interpolating height to the center of the sample interval and bin-averaging sample interval before inverting avoids an apparent increase in average sample rate in adjacent sample intervals.

The bin averaged sample rates and heights are then used to compute the height rate covariance and the average sample rates used in X8 to determine SSB.

The method is illustrated in Figure 10 with MABEL data gathered April 25, 2012 over the Greenland Sea. Aggregated over three 3000-m intervals, the correlation of photon return rate and height is very low on average, and results in a sea state bias of 0.3 mm, a result that is not surprising given the low energy of the surface waves with a significant wave height (4 x Std of height) of less than a meter for the raw photon returns and the bin-averaged heights.

3.0 OPEN OCEAN PRODUCTS

3.1 Open Ocean Surface Height (ATL12/L3A)

The ATL12 product contains sea surface heights over the ice-free oceans for each of 3 ATLAS strong beams. It provides the most basic data from ICESat-2: the sea surface height at a given point on the open ocean surface at a given time plus parameters needed to assess the quality of the surface height estimates and to interpret and aggregate the estimates over greater distances. These heights can be used in comparisons of ICESat-2 data with other geodetic data and as inputs to higher-level ICESat-2 products, particularly ATL19.

Heights over the sea surface are defined for segments with variable length at variable intervals along the ground track; this is necessary to adapt to the reduced photon counts from the low but variable reflectance of the open ocean surfaces. It also is consistent with surface finding routine used for sea ice covered ocean (ATL07/L3A). We anticipate that requiring a minimum of 100 photon retrievals in the adaptive surface detection algorithm will result in minimum segment lengths of about 70 m. Given the 17-m footprint of the ICESat2 beams, this is equivalent to 4 independent (non-overlapping) spatial samples of the sea surface. We refer to height estimates for these segments as segment heights, as distinct from photon heights in the L2 products.

Heights in marginal ice zones or coastal zones may be defined for shorter variable length segments (e.g., down to 7 meters) sampled at variable intervals along the ground track to adapt to the patches of water between sea ice in the first instance or land in the second instance. These overlap regions will be defined as those that are classified as both sea ice and ocean (marginal ice zone) or land and ocean (coastal zone). The methods of distinguishing open ocean from sea ice and land and setting the ocean segment lengths are TBD.

In purely ocean regions, only strong beams will be active, but in the marginal ice zone and coastal zone overlap regions, the three weak beams will also be active and should be processed identically to the strong beams and output in addition to the strong beam results as part of the ATL 12 ocean product. This is to facilitate a rational merging of the open ocean products (ATL12) and sea ice products (ATL07) and coastal products (ATL??) related to sea surface height.

However, the presence of ocean surface waves in the open ocean far from marginal ice zone and coastal regions, present challenges and an opportunity. The non-Gaussian character of ocean surface waves and the scattering and effective reflectance of the ICESat2 photons likely cause a negative sea state bias (SSB). We anticipate that the higher order moments of the photon height statistics will allow us to better estimate SSB. We will estimate for each ATL12 variable segment at minimum the mean, standard deviation, skewness and kurtosis of the height distribution. These four moments can be determined from the means and standard deviations of a 2-component Gaussian mixture, which will also provide a mixing ratio for the two components. These higher moments make it possible to characterize not only the mean sea surface height, but also the sea state and

likely sea state bias of the height estimate.

However, it is likely that it will be necessary to accumulate large number of surface reflected photons to make meaningful estimates of the higher moments of the height distribution. We estimate that over the open ocean, given the low number of returns from open water (e.g., O[1 photon / pulse]), the minimum sequence length will be 400 pulses (25 Hz) equivalent to about 280 m or fourteen 20-m geo-bins. Even in relatively calm ocean conditions, we will need small multiples of this (e.g., twenty eight 20-m geo-bins = 800-pulses over 560 m at a 12.5 Hz rate) to get enough, O[1000], surface reflected photons for meaningful higher order statistics. Considering that ocean waves can have wavelengths of hundreds of meters, to get meaningful statistics over open water with significant wave activity will likely require accumulating data over several kilometers (e.g., 140 x 20-m geo-bins for 4000 pulses over 2.8 km and a rate of 2.5 Hz). Nevertheless, the 14 geo-bin (400-pulse) increment is especially useful because cloud flags and estimates of the background photon rate will come with the atmospheric products, ATL-09 every 400 pulses. Over the open ocean as opposed to sea ice, higher resolution should not be necessary for tracking purposes. The common number 14 geo-bin (400-pulse) sequences to be pieced together for a common aggregate sequence length is to be decided by experiments with the ATLAS simulator and Mabel data, but it is reasonable to assume that at least ten 14-geo-bin (400-pulse) sequences would commonly be pieced together for on the order of 4,000 signal photons, a 2.5-Hz sequence rate and corresponding track length of 2.8 km.

3.1.1 Height Segment Parameters

3.1.1.1 Geolocation/Time

The location is the mean location of designated signal photons used as input to the surface finding procedure. The time is the mean time of designated signal photons used as input to the surface finding procedure.

3.1.1.2 Height

This is the mean height estimate from the surface detection algorithm. Quality metrics include: Confidence level in the surface height estimate based on the number of photons, the background noise rate, and the analysis from the detection algorithm.

3.1.1.3 Subsurface-scattering corrections

Subsurface-scattering corrections are yet to be determined. They are expected to be minimal for the deep open ocean where the water is clear, but may be significant in near coastal regions where more scattering elements are to be expected in the water. These conditions will be most likely to be apparent as positive skewness in the return height distributions especially as correlated with outside measures of surface color or scattering.

Guided by the treatment of subsurface scattering for the Inland Water ATBD and as we are able to develop relations between photon height distribution skewness and other evidence of scattering, we will derive an appropriate correction for subsurface scattering where applicable and provide it as an output of ATL12.

3.1.1.4 First-photon-bias corrections

Pending further analysis and given the low apparent reflectance of the ocean surface first-photon-bias corrections will likely not be needed for the ocean products. If surface reflectances (e.g. based on Martino's analysis) exceed yet to be determined criteria, first photon bias will be corrected using a method TBD for the ocean.

3.1.1.5 Height Statistics

The photon statistics describe the distribution of the population used in the surface-tracking algorithm. These parameters include the: number of photons, histogram of the population, and, at minimum the mean, variance, and skewness, and kurtosis.

3.1.1.6 Initial Sea State Bias Correction and Surface Wave Properties

As discussed in 2.2.3, ICESat-2 gives us a unique opportunity to make an a priori estimate of sea state bias using only the heights from the surface finding system. The approach is to use the surface photon heights to estimate the surface height and the rate of photon returns in 10-m along-track segments. The correlation of height and return rate normalized by the average photon return rate is the EM or sea state bias. This will be estimated from the surface photon height record before separation of the impulse response in the height histogram and made available as output. The standard deviation of photon heights over the whole segment will be provided as a measure of SWH (=4xStd.) that can be compared with reanalysis products. The standard deviation about the 10-m bin averages will also be provided as an estimate of the energy of short waves that arguably has an important effect on reflectance and sea state bias.

3.1.1.7 Solar Background

The solar background photon rate, *backgr_r_200*, is estimated as an average over 50 laser pulses in ATL03. It is based on the photons included in the altimetric histogram less the photons deemed signal (surface or cloud reflected) photons by ATL03. This is converted to a rate in photons per second by dividing by the total time window reduced by the duration of the window containing signal photons.

For ATL12 we will use the average of the estimated solar background. We will not use the predicted solar background rate based on the solar zenith angle, the solar flux in the receiver etalon's pass band, and the effective aperture of the detectors.

3.1.1.8 Apparent Reflectance

This is based on the comparison of expected photon counts over a sea surface with SWH estimated from photon statistics and photon counts from the actual surface (see 2.2.1.4 and Fig. 2).

3.1.2 Input from IS-2 Products

3.1.2.1 Classified photons (Level 2)

Photons classified as to whether the height is signal, noise, or extended signal. These have a confidence as to type.

3.1.2.2 Atmosphere (Level 3)

Relative/calibrated backscatter, background rates, cloud statistics at 25 Hz.

3.1.3 Corrections to height (based on external inputs)

In anticipation of higher level processing, the standard height products will include a number of corrections applied to the raw height estimates. For example to reduce aliasing problems, corrections for high-frequency and fine spatial scale variations in SSH (e.g., tides and other high frequency ocean circulation changes) that may be inadequately sampled by ICESat2 should be applied before averaging. Estimates of these corrections will be given here. All corrections will be given in terms of height so that to apply a correction, users will add it to the height estimate, and to remove it they will subtract it. Additional corrections that some users may decide to apply will be provided with the product.

3.1.3.1 Geoid adjustment (Static)

Heights are adjusted for local geoid height using mean tide EGM2008 model being reported by ATL03 as taken from the mean-tide EGM2008 model.

3.1.3.2 Atmospheric delay corrections

Heights will be corrected based on an atmospheric model, giving corrections for total delay correction that accounts for wet and dry wet troposphere.
Corrections will be available for the forward-scattering bias, based on available atmospheric-scattering data and an estimate of the attenuation calculated from the apparent surface reflectance.

3.1.3.3 Dynamic Atmospheric Correction and the Inverse Barometer Effect (IBE, time-varying)

Heights are corrected for the inverse barometer effect due to the direct application of atmospheric pressure to the sea surface and the dynamic changes forced by wind. ICESat-2

has adopted the utilization of global, empirical, 6-h, AVISO MOG2D, 1/4° × 1/4° grids to be used as a near-real time Inverted Barometer (IB) and Dynamic Atmospheric Correction (DAC)[*Carrère and Lyard*, 2003]. These grids are forced by the European Center for Medium-Range Weather Forecasting (ECMWF) model for the surface pressure and 10-m wind fields. This combined correction typically has amplitude on the order of ±50 cm [*Markus et al.*, 2016].

3.1.3.4 Tidal corrections (time-varying)

Heights are corrected for:
Solid earth tides (*tide_earth*): Solid earth tides are derived using IERS (2010) conventions are ±30 cm max (details in ATL03 ATBD section 6.3.3)

Ocean load tides (*tide_load*): The local displacement due to ocean loading (-6 to 0 cm) derived from ocean tide model GOT4.8. Details in ATL03 ATBD, section 6.3.1

Pole tide (*tide_pole*): Pole tides include both solid earth and ocean pole tides. These each are computed via IERS (2010) conventions. Details are found in ATL03 ATBD sections 6.3.5 and 6.3.6, respectively.

Ocean pole tide (*tide_oc_pole*): Ocean Pole Tide -Loading due to centrifugal effect of polar motion on the oceans (2 mm, max), subsumed in Pole tide (tide_pole)

Ocean tides (*tide_ocean*): Ocean Tides including diurnal and semi-diurnal (harmonic analysis), and longer period tides (dynamic and self-consistent equilibrium) (±4 m)are from the GOT4.8. Details in ATL03 ATBD, section 6.3.1.

Equilibrium tides (*tide_eq*): Equilibrium long-period tide computed using a subroutine attached to GOT4.8 called LPEQMT.F by Richard Ray. It is a Fortran routine in which fifteen tidal spectral lines from the Cartwright-Tayler-Edden tables are summed. (See section 6.3.1 of the ATL03 ATBD.)

3.1.3.5 Wind and SWH Estimates

Surface winds and significant wave height for forecast/reanalysis products will be taken from an appropriate source such as the ECMWF and with the ATL12 product for comparison for comparison with ICESat-2 derivations of SWH as part of the sea state bias calculation.

3.2 Gridded Sea Surface Height - Open Ocean (ATL19/ L3B)

This product, based entirely on Product ATL12/L3A with no external dependence, contains gridded monthly estimates of sea surface from all IS-2 tracks from the beginning to the end of each month. Below 60°N and above 60°S, the data are mapped on the curvilinear grid of the TOPEX/Poseidon with spacing equivalent to 0.25° of longitude. Above 60°N and below 60°S the grid data are mapped onto a planimetric grid using the SSM/I Polar Stereographic Projection equations at a grid spacing of about 24 km. The exact spacing of the Polar Stereographic grid should be adjusted to match the longitudinal spacing of the TOPEX/Poseidon grid at 60°S and 60°N, and the latitudinal spacing adjust to have an integer number of grid cells between 60°N (S) and the North (South) Pole.

3.2.1 Grid Parameters

3.2.1.1 Sea surface height estimate

With only ATL12 needed as input, this derived product contains the statistical description of the sea surface (mean; standard deviation, skewness, kurtosis) within grid cell. These will be computed in two ways, using aggregation of segment histograms discussed in this section,, and aggregation of segment moments discussed in section 3.2.1.2. The gridding scheme will take as input from ATL12, the histograms of DOT accumulated in segments lying within each grid cell these will be combined by adding the photon counts from each segment in each height bin of each segment in the grid cell.

The average geoid height from each segment will also be averaged, weighted by the number of signal photons in each segment. Similarly, the weighted average a priori sea state bias will be computed.

The four moments of the aggregate histogram will be computed using the 2-Gaussian Maximum Likelihood approach. The sea state bias will be subtracted from the first moment to yield the mean grid cell DOT. The mean SSH can be calculated as the mean DOT plus the weighted average geoid height.

3.2.1.2 Sea surface statistics histogram within grid

We will perform a weighted average of the first through fourth moments of heights in the length segments that go into the grid cell in a given month. Being mindful that each segment going into a grid average may have a different underlying distribution, we will essentially compute the moments of the segment moments. Thus, the moments for the individual segments will be weighted by the number of surface photons in their segments relative to the total number of photons going into the grid cell and then added to yield estimates of the average moments in the grid cell. Where enough segments are included in a grid cell (e.g., more than 20), histograms of the moment moments going into the grid cell can be included as output.

3.2.1.3 Wave statistics within grid

Estimates of SWH, short wave energy and SSB from the *a priori* estimation of sea state bias will also be grid-averaged with appropriate normalization for the number of surface photons in each segment.

3.2.1.4 Sea surface segments (Input)

Sea surface height in surface height segments as described above with the time/location/quality of the sea surface height in the segment.

4.0 ALGORITHM THEORY

In this section, we discuss the following topics:

1. Finding the collection of photon heights representing reflection from the sea surface.
2. Separating the surface wave generated variation in photon heights from other sources of variance.
3. Determining SWH and higher order measures of sea state
4. Formulating the best sea state bias correction

4.1 Introduction

The main purpose of the algorithms developed here is the determination of Sea Surface Height (SSH). There are two main steps to determining sea surface height (SSH) from ICESat-2 photon heights (ATL-03) over the ocean. These are identifying photon heights that likely represent reflection from the sea surface, loosely termed surface finding (Sec. 4.2), and determining the correct statistics of the sea surface height distribution (Sec 4.3). Most importantly we seek the SSH equal to the mean of the height distribution. Because of the motion of the sea surface, surface finding over open water is inherently a search for a distribution of heights representative of the sea surface. Our main product will be the mean of this distribution over time, length and space scales to be determined as part of testing and analysis. In addition other properties of the distribution are useful for assessing the surface wave environment and biases in the SSH determination. Though we focus here on obtaining meaningful statistical distributions of sea surface heights, we have found it possible in tests with MABEL data to create meaningful moving 21-photon means of the heights from the photons designated signal photons by the surface finding routines. This gives a relatively high-resolution time- (or space-) series trace of the sea surface in fair agreement with the analysis using the high-resolution sea ice analysis. This may be used in experimental analyses for surface waves. Figure 11 shows the block diagram for the ATL12 processing to go from ATL03 photon data to along-track histograms and statistics of dynamic ocean topography. Figure 13 illustrates the gridding procedure to go from ATL12 products to ATL19 gridded DOT and SSH.

4.2 ATL12: Finding the surface-reflected photon heights in the photon cloud

Surface finding over the open ocean rests on our limited experience with MABEL data over open water. With few surface returns per pulse and significant wave heights much less than the range of samples, the majority of the of the bins will contain only noise photons so that the median will equal the number of noise photons per bin. Thus the fundamental idea for surface finding is to identify as surface height bins, those bins with counts exceeding the median number of counts. In extreme cases where ICESat-2 encounters significant wave heights approaching the 30-m downlink range of bins, we may have to adopt a slightly

different threshold such as the count value equal to or exceeding 20% of the bin population.

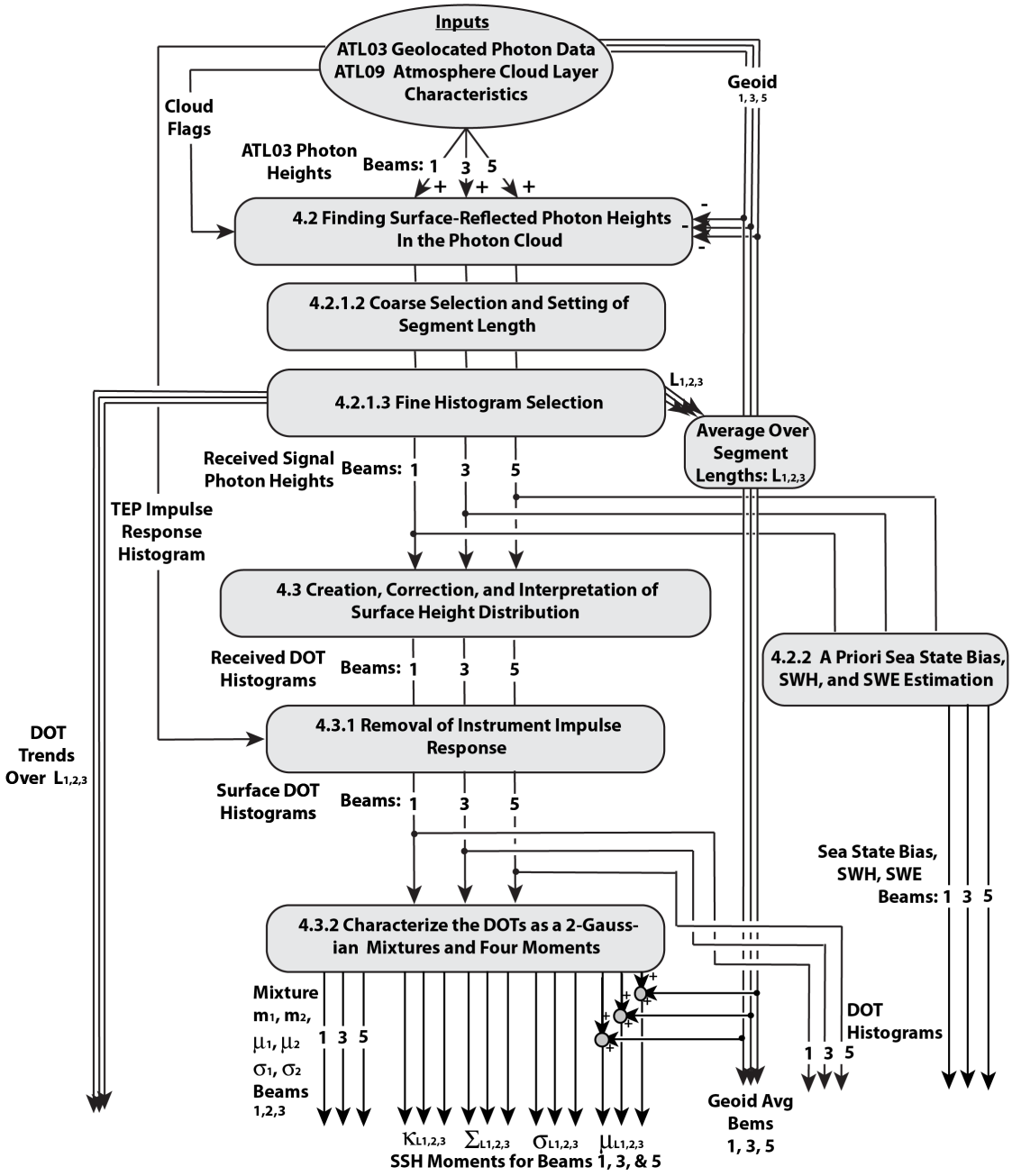


Figure 11. ATL 12 processing block diagram as discussed in Section 3 and 4. μ, σ, Σ, and K denote mean, standard deviation, skewness, and kurtosis respectively.

We anticipate a large number of surface reflected photons may be needed to adequately resolve the higher moments of the sea surface height distribution. Also, fine spatial resolution is probably not required in most open ocean settings. Therefore, we will use a semi-adaptive scheme to accumulate segments long enough to capture up to 10,000 candidate surface photons. This will involve a coarse median-based selection and segment-length setting process (4.2.1.2). This will be followed by a finer scale histogram construction that includes median-based selection of surface photons, a detrending process, and repeat selection (4.2.1.3)

4.2.1 Selection of Signal Photon Heights

4.2.1.1 Input to Selection of Photons:

ATL03 photon heights for each pulse in the ocean ± 15-m downlink band and associated time and geolocation information, plus cloud flag every 400 pulses from ATL09. We assume, that aside from the ± 15-m band about the geoid, no classification as to surface/non-surface photons is made for the ATL03 ocean results.

4.2.1.2 Coarse Selection and Setting of Segment Length

(Note: The process in 4.2.1.2 has not been implemented as of Feb. 2014, but it is based on routines developed for MABEL data and described in 4.2.1.3.)

1. Establish an initial coarse histogram array, H_c, spanning ± 15 m with bin size B_1 (TBD, e.g., 0.01 m) and a data array, A_{coarse}, for up to 10,000 photon heights and associated information (index, geolocation, time) plus noise photon counts. This will be populated with data for the coarse selection of signal photons.
2. Aggregate photon heights over 14 geo-bins (400-pulses) and construct a temporary 14 geo-bin (400-pulse) height histogram spanning ± 15 m with bin size B_1 (TBD, e.g., 0.01 m). This is used to estimate a running total number of signal and noise photons. Note that in our test examples with MABEL data a bin size of 0.05m has been used, but we suggest 0.01 for consistency with later steps in the processing. The 14 geo-bin (400-pulse) ocean segment should be aligned so that the once per 14 geo-bin (400-pulse) cloud flag represents with the cloud conditions during the ocean segment. (it is unclear at this writing what time span an individual cloud flag represents, e.g., average over previous 400 pulses, value at the end of 14 geo-bin(400-pulse)s, etc.)
3. If percentage cloud coverage > *Pc*, (Pc is TBD, e.g., 50%) then proceed to next 14 geo-bin (400-pulse) segment.
4. Photons in the 14 geo-bin (400-pulse) histogram bins with greater than the *Th_Nc* times the median number of photons, N_{median} , are candidate signal photons and photons in the bins with the median number of samples or less are considered noise

photons. *Th_Nc* is TBD, but we have been using *Th_Nc* = 3 in our test examples with MABEL data. Note that with few surface returns per pulse and significant wave heights much less than 30 m, the majority of the of the bins will contain only noise photons so that the median will equal the number of noise photons per bin. In extreme cases where ICESat-2 encounters significant wave heights approaching 30 m, we may have to adopt a slightly different threshold such as the count value equal to or exceeding 20% of the bin population.

5. Add the signal and noise photons from this 14 geo-bin (400-pulse) segment to the running total of candidate signal photons and noise photons, and add all the photons to the coarse histogram, H_c.
6. If the total number of candidate signal photons is greater than or equal to a minimum number, *Th_Ps* (TBD, but e.g., 8,000), of photons or S_{egmax} (e.g., 25) 14 geo-bin (400-pulse) segments have been collected, go on to Fine Selection (Sec 4.2.1.3) with the populated coarse histogram, H_c, ± 15 m with bin size B_1 (TBD, e.g., 0.01 m), and the data array, A_{coarse}. If the total number of signal photons is less than *Th_Ps* , repeat steps 4.2.1.2-2 to 4.2.1.2-2 above and intake another 14 geo-bin (400-pulse) segment.

4.2.1.3 Fine Histogram Selection

1. Considering the coarse histogram array, H_c, perform a N_{bmv} = 1-m/B1 bin (e.g., 101 bins over 1 m) moving bin arithmetic average incremented by 1 bin. Pad the ends of the smoothed histogram with $N_{bmv}/2$ bins (equal the smoothed first and last values) to match the length of the original histogram.

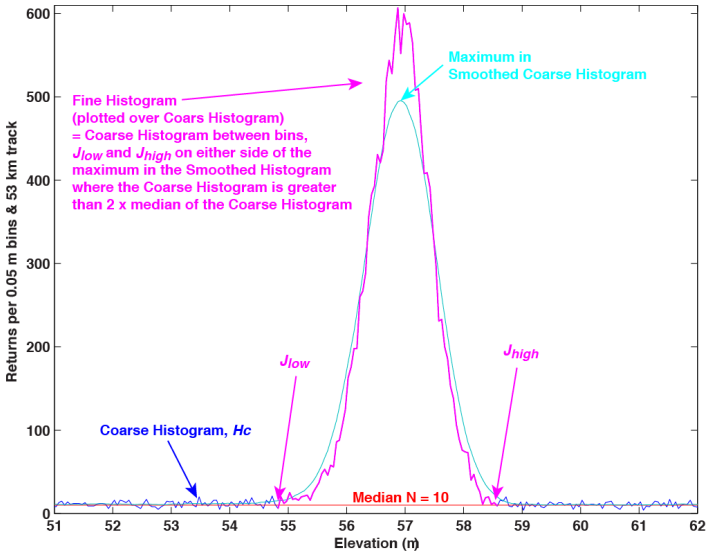


Figure 12. Illustrative coarse, smoothed, and fine histograms from steps 1 and 2 of 4.2.1.3 as applied to 10 minutes of April 2012 MABEL data over Chesapeake Bay. The corresponding figure for detrended data in the second iteration (steps 4-6) are similar.

2. Find the bin limits of the fine histogram as all the coarse histogram bins on either side of the maximum in the 1-m smoothed histogram, from the J_{low} bin position to the J_{high} bin, where the coarse bin photon count is greater than N_f times the coarse histogram median, N_{median}. N_f is TBD, but we have used N_f = 1.
3. Consider the time series of all the heights stored in A_{coarse} that lie

between heights corresponding the J_{low} and the J_{high} bin positions. This array of heights and times represents the first iteration on surface height statistics, but the higher moments are elevated by variations of the mean with time (or distance). Therefore, calculate the least squares linear fit to these versus time (or distance). Store the mean and trend as output variables for the segment and subtract them from the data of A_{coarse} to yield array A_{fine} of detrended photon heights and associated geolocation data.

4. From the heights of A_{fine}, create the detrended coarse histogram, H_d, ± 15 m with bin size B_1 (TBD, e.g., 0.01 m). Then repeat steps1 and 2 substituting H_d for H_c:
5. Considering the detrended coarse histogram array, H_d, perform a N_{bmv} = 1-m/B1 bin (e.g., 101 bins over 1 m) moving bin arithmetic average incremented by 1 bin. Pad the ends of the smoothed histogram with $N_{bmv}/2$ bins (equal the smoothed first and last values) to match the length of the original histogram.
6. Find the bin limits of the detrended fine histogram, H_R, as all the detrended coarse histogram bins on either side of the maximum in the 1-m smoothed detrended coarse histogram, from the J_{low} bin position to the J_{high} bin, where the detrended coarse bin photon count is greater than N_f times the coarse histogram median, N_{median}. N_f is TBD, but we have used N_f = 1.
7. The resulting H_R is the fine histogram of received surface heights to be used in further analysis. The time series of all the heights stored in A_{fine} that lie between heights corresponding the second iteration J_{low} and the J_{high} bin positions will be saved for experimental fine surface and wave analysis (Section 4.2.2).

4.2.2 A Priori Sea State Bias Estimation, SWH, and Short Wave Energy (SWE)

As a side calculation, we will use the spatial record of signal photon heights derived through the process described in Sections 4.2.1.1-4.2.1.3 as an approximation to the true surface height and resolve the structure of surface waves versus along track distance and the variation in photon return rate to estimate sea state bias. The inherently high spatial resolution of ICESat-2 will allow us to measure the variation in surface height over the large energy containing surface waves that lead to sea state bias. We will also know the photon return rate along the waves. Over segments of data equal to 140 geo-bins or 2.8 km, we will determine surface height variation and photon return rate in 10—m along-track bins. From these we will compute the covariance of surface height and photon return rate. Combined with the average photon return rate, the covariance of height and return rate is the sea state bias according to equation (11). We can approximate the significant wave height (SWH) as four times the standard deviation of the bin averaged heights. We can approximate the short wave energy (SWE) as variance of all the photon heights minus the variance of the bin average heights.

4.3 ATL12: Correction and Interpretation of the Surface Height Distribution

There are two main issues in deriving surface statistics from the apparent heights of surface-reflected photons identified in the surface detection phase and manifest in the fine histogram, H_R, of Section 4.2. These are correcting the heights for the instrument impulse response and deriving higher order moments of the corrected surface height distribution.

4.3.1 Separating the Uncertainty due to Instrument impulse response

The received heights of surface reflected photons are the sum of a true height of the surface plus an uncertainty due to the instrument impulse response. The latter is the total of all instrument uncertainty, but it is dominated by uncertainty in the time the individual photons are transmitted. The ATLAS uncertainty in height due to instrument impulse response will have a standard deviation of 15-22 cm. Because the instrument impulse response uncertainty and the true height variation are additive, the received surface photon distribution, H_R, is equal to the convolution of the instrument impulse response height error distribution, H_T, and the true height distribution, H_Y:

$$H_R = H_T * H_Y \tag{12}$$

where * denotes convolution. Consequently, if we want to know the statistics of the surface from ICESat-2 returns, we must deconvolve H_T from H_R to get H_Y.

If we know that H_Y is a Gaussian distribution, we only need concern ourselves computing the mean and standard deviation of the distribution. However, we expect received ocean surface heights to depart from a Gaussian distribution because of a combination the shape of surface waves and the non-uniform sampling due wave-induced variations of specular reflection. In addition to the mean and standard deviation, the third and second moments (skewness, and kurtosis) are also important. If we knew the true sample heights, we could calculate these moments as sample moments, e.g., the n^{th} sample moment for Y_i from the time series of heights:

$$M_n(Y) \quad \frac{1}{m} \sum_{i\ 1}^{m} Y_i \quad \overline{Y}^{\,n} \tag{13}$$

or as the discrete form of the integral over the probability density function,

$$Y \quad M_1 \quad H\ Y\, dY$$

, if we knew *H(Y)*.

However, the instrument impulse response delay, T, and the received height, R, are independent and additive ($R=T+Y$), so in principle it is possible too calculate sample moments for Y from the sample moments of T and R:

$$M_1(Y) \quad M_1(R) \quad M_1(T) \quad \frac{1}{m} \sum_{i\ 1}^{m} R_i \quad \frac{1}{m} \sum_{i\ 1}^{m} T_i \tag{14}$$

$$M_2(Y) \quad M_2(R) \quad M_2(T) \quad \frac{1}{m}\sum_{i=1}^{m} R_i \quad \overline{R}^{\,2} \quad \frac{1}{m}\sum_{i=1}^{m} T_i \quad \overline{T}^{\,2} \tag{15}$$

$$M_3(Y) \quad M_3(R) \quad M_3(T) \quad \frac{1}{m}\sum_{i=1}^{m} R_i \quad \overline{R}^{\,3} \quad \frac{1}{m}\sum_{i=1}^{m} T_i \quad \overline{T}^{\,3} \tag{16}$$

$$M_4(Y) \quad M_4(R) \quad M_4(T) \quad M_2(Y)M_2(T) \quad \frac{1}{m}\sum_{i=1}^{m} R_i \quad \overline{R}^{\,4} \quad \frac{1}{m}\sum_{i=1}^{m} T_i \quad \overline{T}^{\,4} \quad M_2(Y)M_2(T) \tag{17}$$

These sample moments will be calculated for the ICESat-2 heights as a check on more exact methods. However, pending further study, it is TBD whether they will be included in the ATL12 data output. The uncertainty in sample moments is significant [Percival, personal communication] and this uncertainty is compounded with combinations such as (14)-(16). These problems are liable to be particularly important when looking for small variation in skewness, $M_3/M_2^{3/2}$, and excess kurtosis, $M_4/M_2^2 \quad 3$. Optimized approaches for determining the moments of the height distribution require that we separate the effect of instrument impulse response from the received heights to give us the distribution of surface height, H_Y. To do this we must deconvolve H_T from H_R.

Deconvolution can be done in several ways. The Inland Water ATBD (ATL-13) expresses the convolution of (12) for the received histogram as the multiplication of a matrix representing the instrument impulse response histogram times a vector representing the surface height histogram. This matrix is inverted and multiplied times the received histogram to yield the surface histogram. The technique is reportedly sensitive to proper binning to produce a matrix that can be inverted.

Deconvolution can also be done by taking the Fourier transform of (12) and noting that the Fourier transform, $\mathsf{F}(\ )$, of the convolution of two variables, $H_R(k)=\mathsf{F}(H_R)= \mathsf{F}(H_T*H_Y)$ is equal to the product of the Fourier transform of the two variables,

$$H_R(k)=\mathsf{F}(H_R)= \mathsf{F}(H_T*H_Y)= \mathsf{F}(H_T)\ \mathsf{F}(H_Y) \tag{18}$$

where k is the wavenumber in units of m^{-1}, the Fourier space counterpart to height in meters, the histogram independent variable. From (18), we can compute H_Y as an inverse Fourier transform, $\mathsf{F}^{-1}(\ )$ of the ratio of $\mathsf{F}(H_R)$ and $\mathsf{F}(H_T)$,

$$H_Y= \mathsf{F}^{-1}(\mathsf{F}(H_R)/\mathsf{F}(H_T)) \tag{19}$$

The problem with this approach is that there is invariably noise in the received signal histogram, H_R, that produces significant values in $\mathsf{F}(H_R)$ at wavenumbers where $\mathsf{F}(H_T)$ has small values or zeros. This may have something in common with problem of conditioning of the matrix in the matrix inversion technique. In any event, these unrealistic combinations make the inverse Fourier transform unstable. To account for the noise in H_R, we consider

Wiener deconvolution, in which it is assumed the received histogram is contaminated with noise, N, and (12) and (18) become

$$H_R = H_T * H_Y + N \tag{20}$$

and where $H_T(k)$=F(H_T) , $H_Y(k)$=F(H_Y), $H_R(k)$ =F(H_R), and $N(k)$ = F(N),

$$H_R(k) = H_T(k)H_Y(k) + N(k) \tag{21}$$

An estimate $H_Y(k)$ for $H_{Yest}(k)$ of the form

$$H_{Yest}(k) = W(k)\, H_R(\mathrm{k})/H_T(\mathrm{k}) \tag{22}$$

is sought such the expected value of $(H_Y(\mathrm{k})-H_{Yest}(\mathrm{k}))^2$ is a minimum. This is found for

$$W(k) = \mathrm{F}(H_T)^2/[\mathrm{F}(H_T)^2 + \mathrm{F}(Noise)^2/\mathrm{F}(H_Y)^2] \sim \mathrm{F}(H_T)^2/[\mathrm{F}(H_T)^2 + \mathrm{F}(Noise)^2/\mathrm{F}(H_R)^2] \tag{23}$$

and

$$H_{Yest} = \mathrm{F}^{-1}(H_{Yest}(k)) = \mathrm{F}^{-1}(W(k)\mathrm{F}(H_R)/\mathrm{F}(H_T)) \tag{24}$$

For wave numbers where the inverse signal to noise ratio, $\mathrm{F}(Noise)^2/\mathrm{F}(H_R)^2$, is low relative to $H_T(\mathrm{k})$, $W(k)$ goes to one and equation (24) reverts to (19). Where the inverse signal to noise ratio is high (high noise), $W(k)$ goes to zero and the noise is filtered out of the resulting $H_{Yest}(k)$.

4.3.2 Characterizing the Random Sea Surface

The best method for determining sample moments is the method of maximum likelihood (ML). For a distribution of known type $f(x \setminus {}_{1,2,3,4})$, the optimum choice of parameters, ${}_{1,2,3,4}$, for data values $x_1, x_2, x_3 \ldots\ldots x_m$, are those which maximize the likelihood, L,

$$L \quad {}_{1,2,3,4} \setminus x_1, x_2, x_3 \ldots\ldots x_m \qquad \prod_{i\ 1}^{m} f(x_i \setminus {}_{1,2,3,4}) \tag{25}$$

For ICESat-2 sea surface height, we are not sure of the form of the probability density function but commonly the ocean surface has been characterized by the Gram-Charlier distribution for which the first 4 moments are important. Such a distribution can be represented it by mixture of two Gaussian distributions, N_a and N_b

$$f(x_i \setminus {}_{1,2,3,4}) \quad m_a N_a(x_i \setminus {}_{a1,2}) \quad m_b N_b(x_i \setminus {}_{b1,2})$$
$$m_a \quad m_b \quad 1 \tag{26}$$

and the two moments of each distribution plus the mixing ratio, m_a/m_b. Given the surface height histogram, the parameters of the two normal distributions and the mixing ratio can be determined by maximum likelihood.

The aggregate moments of the Gaussian mixture *X* with *n* component Gaussian distributions are functions of the component means, μ_i, and variances, σ_i^2, and the mixing ratio, m_i,

$$E[X_{mix}^j] = \sum_{i=1}^{2} \sum_{k=0}^{j} \frac{j!}{k!(j-k)!} \mu_i^{j-k} m_i E[X_i^k] \tag{27}$$

For example, the aggregate mixture mean is:

$$\mu_{mix} = m_1 \mu_1 + m_2 \mu_2 \tag{28}$$

and the aggregate variance is:

$$\sigma_{mix}^2 = m_1(\sigma_1^2 + \mu_1^2) + m_2(\sigma_2^2 + \mu_2^2) - \mu_{mix}^2 \tag{29}$$

4.3.3 Combined Approaches and Minimum Least Squares

In this revision of the Ocean ATBD the Gaussian Mixture (GM) and Maximum Likelihood (ML) approaches are proposed to analyze ATLAS data over the open ocean. However, there are other approaches being investigated for related ATBDs, Sea-Ice and Inland Water. The Inland Water ATBD describes a matrix inversion method for deconvolution.

The Sea-Ice method combines deconvolution of the source histogram and statistical parameter identification uses a least squared error approach. A table of histograms is developed based on convolution of a known instrument impulse response histogram and normal distributions for various means and standard deviations. Received signal histograms are compared to each of the histograms in the table, and the surface mean and standard deviation are those for the table histogram with the smallest mean squared departure from the received histogram. We have investigated this approach using a 4-dimensional table for the surface mean, standard deviation, skewness, and kurtosis. However, the method was found to be time and memory intensive and seemed to suffer from some ambiguity in choosing the optimum fit to the observed histogram from the 4-dimensional table.

Determination of a Gaussian mixture representation was also done for ICESat/GLAS surface heights. The method was not Maximum Likelihood. A least squares fit of a multiple (most commonly 2) mix of Gaussian distributions to the observed histograms. This approach for fitting the Gaussian mixture should be compared for speed and accuracy to the Expectation Maximization (EM) method of achieving Maximum Likelihood we use here.

4.4 Calculation of Uncertainty

TBD- calculation of uncertainty in SSB correction will be important in matching ICESat2 altimetry to SSH estimates from other altimeters [*Urban and Schutz*, 2005].

4.5 ATL19: Gridding the DOT and SSH

This product, based entirely on Product ATL12/L3A with no external dependence, contains gridded monthly estimates of sea surface from all IS-2 tracks from the beginning to the end of each month. Below 60°N and above 60°S, the data are mapped on the curvilinear grid of the TOPEX/Poseidon with spacing equivalent to 0.25° of longitude. Above 60°N and below

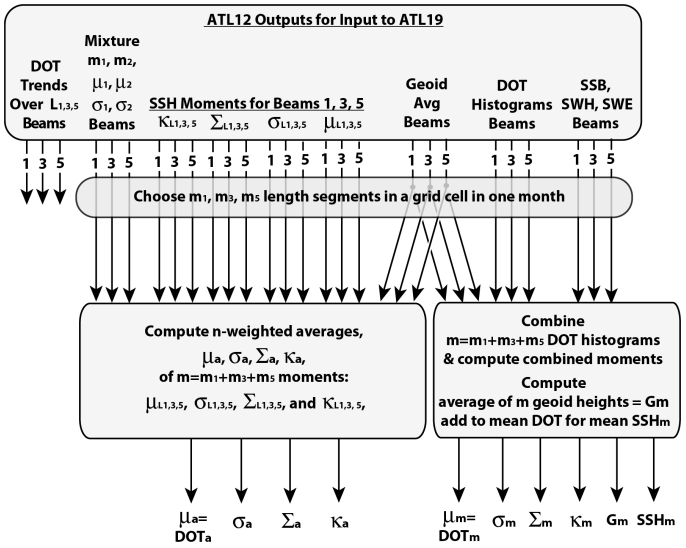


Figure 13. Block diagram for the ATL19 gridding procedure taking ATL12 ocean products as input. μ, σ, Σ, and K denote mean, standard deviation, skewness, and kurtosis respectively.

60°S the grid data are mapped onto a planimetric grid using the SSM/I Polar Stereographic Projection equations at a grid spacing of about 24 km. The exact spacing of the Polar Stereographic grid should be adjusted to match the longitudinal spacing of the TOPEX/Poseidon grid at 60°S and 60°N, and the latitudinal spacing adjust to have an integer number of grid cells between 60°N (S) and the North (South) Pole.

5.0 ALGORITHM IMPLEMENTATION

This section describes the specific implementation of the algorithm for program development.

5.1 Outline of Procedure

1) Finding the surface-reflected photon heights in the photon cloud
   a) Coarse Selection and Setting of Segment Length:
      Use a median-based selection criterion to accumulate a large number of surface-reflected photons.
   b) Fine Histogram Selection:
      Use a median-based selection criterion to derive a preliminary surface height histogram, compute the mean and trend in surface height, remove this trend from all heights, and repeat the median-based selection criterion to derive a final surface height histogram
2) Using the along track signal photon heights from the fine histogram selection step and signal photon return rate to estimate sea state bias and surface wave properties.
3) Correction and Interpretation of the Surface Height Distribution
   a) Separating the Uncertainty due to Instrument impulse response
      Use Wiener deconvolution to derive the surface height histogram with the uncertainty due to the instrument impulse response removed.
   b) Characterizing the Random Sea Surface
      Characterize the surface height histogram as a mixture of two normal distributions and calculate up to 4^{th} moment of the mixture.
   c) Compute the mean, variance, skewness, and kurtosis for the aggregate Gaussian mixture using Equation 17; the final determination of SSH for ATL12 is the mean of the aggregate mixture distribution computed using (28) and the SSH variance is given by (29).

5.2 Input Parameters

5.2.1 Parameters Required from ICESat-2 Data

See Tables 1 and 2. For ATL12, the primary inputs will be photon heights in the downlink band and associated time and geolocation information from ATL03, plus the cloud flag every 400 laser pulses from ATL09. Normally over the open ocean (or sea ice) regions not overlapping with terrestrial or ice sheet regions, the downlink band will be ± 15 m centered about the signal area identified by the ATLAS on-board flight software. In the ocean regions overlapping with land and ice sheet regions, the downlink band will expand to match the band for those regions. Also for ATLAS special operations such as ocean scans and Transmitter Echo Pulse data collection the telemetry band is expanded. For the purposes of ATL12, the EGM2008 geoid supplied with ATL03 (/gtx/geophys_corr/ posted at the 20m along-track rate for each beam in ATL_03) will be used to establish the

narrower ±15-m band within the expanded downlink band of the overlap and special operations regions. Aside from the ± 15-m band about the geoid, no prior classification as to surface/non-surface photons is considered from the ATL03 ocean results. The 14 geo-bin (400-pulse) ocean segment should be aligned so that the once per 14 geo-bin (400-pulse) cloud flag represents the cloud conditions during the ocean segment. (it is unclear to the author at this writing what time span an individual cloud flag represents, e.g., average over previous 400 pulses, value at the end of 14 geo-bin(400-pulse)s, etc?)

Table 1 Input parameters (Source: ATL03)

Label	Description	Symbol
delta_time	Elapsed seconds since first data point in granule	***time_initial***
h_lat	latitude of each received photon	***lat_initial***,
h_lon	longitude of each received photon	***lon_initial***
h_ph	height of each received photon	***ht_initial***
dist_ph_along	Along track distance	
sigma_along	Uncertainty in along-tack distance	
dist_ph_across	Across-track distance	
sigma_across	Uncertainty in across-track distance	
segment_ID	Geo-bin ID	***lat_initial***,
segment_length	Segment length of each geo-bin	***lat_initial***,
h_quality_flag	Flags describing quality of height for use on higher level products	
beam_azimuth	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). When the spacecraft is precisely at the geodetic zenith, the value will be 99999 degrees. 40 Hz.	
beam_coelv	Co-elevation (CE) is direction from vertical of the laser beam as seen by an observer located at the laser ground spot.	
solar_azimuth	The direction, eastwards from north, of the sun vector as seen by an observer at the laser ground spot.	
solar_elevation	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, with approximately TBD degree accuracy.	
bckgrd_rate	Background count rate (*backgrd_atlas/bckgrd_rate*), simply averaged over the segment	
track_type_flag	Flag describing if the track is in normal ice sheet, terrian, ocean scan, target or other non-normal pointing control	
fpb_parm(n)	parameter needed to compute first-photon bias correction to a mean height	TBD
gd_ht	The height of the geoid above the ellipsoid	
geo_motion	Geocenter motion _ difference between center of figure (crust fixed) and center of mass (satellite orbit) _ will likely correct for at least the trend in Z.	

ibe	Inverse Barometer Effect -Ocean Inverted barometer correction (±20 cm)
tide_earth	Solid Earth Tides (±30 cm, max)
tide_eq	Equilibrium long-period tide
tide_load	Load Tide - Local displacement due to Ocean Loading (-6 to 0 cm)
tide_oc_pole	Ocean Pole Tide -Loading due to centrifugal effect of polar motion on the oceans (2 mm, max)
tide_ocean	Ocean Tides including diurnal and semi-diurnal (harmonic analysis), and longer period tides (dynamic and self-consistent equilibrium) (±4 m)
tide_pole	Pole Tide -Rotational deformation due to polar motion (-1.5 to 1.5 cm)
tot_geo	total of the geophysical range corrections applied to compute height.
tropo_dry	Dry Troposphere range correction
tropo_wet	Wet Troposphere range correction

Table 2 Input parameters (Source: ATL09)

Label	Description	Symbol
delta_time	Elapsed GPS seconds since start of the granule. Use the metadata attribute granule_start_seconds to compute full gpstime.	
latitude	Latitude, WGS84, North=+, Lat of segment center	
longitude	Longitude, WGS84, East=+,Lon of segment center	
cloud_asr_prob	Cloud probability based on Apparent Surface Reflectivity	
cloud_flg	Cloud present flag	*Pc*
cloud_flg_conf	cloud probability p=(1-asr/t)*100 flag_values: 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	
erd_flg	Flag that indictes probable atmosphere induced range delay	
asr	Apparent Surface Reflectivity (25Hz)	*asr_25*
asr_q	ASR Quality Flag	*asr_b_25*
bgr	Background count rate, averaged over the segment	*bgr_*25
cab	Calibrated Attenuated Backscatter (CAB) profiles for the RT strong beam or and a sum of the 3 beams covering (nominally) 13 to -1 km at 25 Hz	
cloud_prob	Cloud Probability based on Apparent Surface Reflectivity	

5.2.2 Parameters from Ancillary Sources

Accounting for sea state bias will likely require independent observations of gridded wind speed and direction obtained from other remote sensing or reanalysis product such as provided by NCEP.

5.3 Processing Procedure

5.3.1 Coarse Surface Finding and Setting of Segment Length

We start with 14 geo-bin (400-pulse) segments of ATL03 photon heights (see 5.2.1).

Table 3 Control parameters – coarse surface finding

Parameter	Description	Value
B_c	Bin size of coarse histogram	5 cm
Th_N_c	Number of photons times median required to classify as surface bin	3
Th_Ps	Minimum number of candidate surface photons per segment	TBD
Segmax	Max number of 400-pulse segments in a surface finding segment	TBD
Th_P_c	Max percentage cloud cover	TBD

(A) Establish Working Histogram Array

Establish a temporary **histogram array, H_c,** spanning ± 15 m with **bin size called *Bc*** (TBD, note that in our test examples with MABEL data a bin size of 0.05m has been used, but we may use values as low as 0.01 m)

(B) Establish Counter of Potential Surface Reflected Photons

Establish an **integer variable, N_{good},** counter for accumulating the number of signal photon heights.

(C) Establish Temporary Data Array

Establish the **data array, *HTin***, for up to 10,000 photon heights, geolocation, and time. This will be populated with data for the further selection of signal photons.

(D) Test for Cloud Cover

If **percentage cloud coverage *Pc* > *Th_Pc***, (***TH_Pc*** is TBD, e.g., 50%) then proceed to next 14 geo-bin (400-pulse) segment.

(E) Populate Working Histogram Array

(F) Consider the photon heights in the ±15-m ocean downlink window over 400 laser pulses and use these data to populate the temporary histogram array, $\boldsymbol{H_c}$ of heights. This is used to initially identify candidate signal and noise photons. Test for Surface Photons

---

Photons in the H_c histogram bins with **greater than the $\boldsymbol{Th\_N_c}$ times** the **median number of photons, $\boldsymbol{N_{median}}$,** are candidate signal photons and photons in the bins with the median number of samples or less are considered noise photons. $\boldsymbol{Th\_N_c}$ is TBD, but we have been using $\boldsymbol{Th\_N_c}$ = 3 in our test examples with MABEL data for initial evaluation. Note that with few surface returns per pulse and significant wave heights much less than 30 m, the majority of the of the bins will contain only noise photons so that the median will equal the number of noise photons per bin. In extreme cases where ICESat-2 encounters significant wave heights approaching 30 m, we may have to adopt a slightly different threshold such as the count value equal to or exceeding 20% of the bin population. This will require testing with the ATLAS simulator.

(G) Increase the Count of Surface Photons

Add the number of surface reflected photons to $\boldsymbol{N_{good}}$, the running total of candidate surface photons and noise photons.

(H) Add All Photons to Data Array

Add the all the photons, signal and noise, and their associated geolocation and time information from this 14 geo-bin (400-pulse) segment to the data array ***HTin***.

(I) Test Counter for Sufficient Surface Photons

If the total number of candidate signal photons is greater than or equal to a **minimum number, *Th_Ps*** (TBD, but e.g., 8,000), of photons or if the number, ***Nseg400***, of 14 geo-bin (400-pulse) segments collected reaches $\boldsymbol{S_{egmax}}$ **(e.g., 25)**, go on to 5.2.4 Processing Procedure for Classifying Ocean Surface Photons, Detrending, and Generating a Refined Histogram of Sea Surface Heights with the data array, ***HTin***. The total number of pulses for the segment, ***n_pls_seg***, equals 400 x ***Nseg400***. If the total number of candidate surface reflected photons is less than ***Th_Ps*** , intake another 14 geo-bin (400-pulse) segment and repeat steps 5.2.3.1-9 above.

5.3.2 Processing Procedure for Classifying Ocean Surface Photons, Detrending, and Generating a Refined Histogram of Sea Surface Heights

The procedure based on development of Matlab Program:
ATBDdraftMABEL_reader_PhotonClassifyX2Channel6B_noprint.m.

Table 4 Control parameters – refined surface finding and analysis

Parameter	Description	Value
B_f	Bin size of fine histogram	5 cm
pts2bin	Bins in boxcar smoother	21
Th_P_c	Percentage cloud cover	TBD
ImpulsH1	First value of impulse resp.	TBD
ImpulsH2	Second value of impulse resp.	TBD
ImpulsH3	Third value of impulse resp.	TBD
ImpulsH4	Fourth value of impulse resp.	TBD
ImpulsH5	Fifth value of impulse resp.	TBD

We start with the array of photon height data in ***HTin*** from 5.2.4. These are candidate heights accumulated over a segment of sufficient length to provide an adequate number of signal photons. In view of the low reflectance of the ocean surface this may require as many as 10,000 laser pulses. Following the developmental code used on MABEL we break ***HTin*** into a vector of initial photon heights, ***ht_initial***, a vector of initial times, ***time_initial***, and vectors of initial location, ***lat_initial*** and ***lon_initial***. For ATLAS we would also add a vector of along track distance, ***trackdist_initial***.

(A) Establish or Summon the Bin Edge Vector

Set up for the surface finding histogram, ***N***, by establishing the vector, ***Edges***, (with a length one greater than ***N***) of histogram bin edges with bins ***Bf*** wide between -15 m to +15 m. Also establish the vector, ***Cntrpt***, of bin center points (with length equal to ***N***; for our tests ***Cntrpt*** values equal ***Edges*** values plus ***Bf***/2, deleting the last and largest value). Also establish a vector array, ***BinB***, as long as ***ht_initial*** for bin assignments.

(B) Compute Initial Histogram and Keep Track of Bin Assignment for Each Photon Height

---

Populate the histogram array ***N*** such that each element equals the number of values from ***ht_initial*** that are in the corresponding bin boundaries of ***Edges***. Also, for each value in ***ht_initial***, populate the vector ***BinB*** with the bin number to which ***ht_initial*** is assigned from one to the length, ***LN***, of ***N*** (in Matlab, [N,BinB]=histc(ht_initial,Edges).

(C) Find Median of Initial Histogram

Find the median, ***medianN***, of ***N***.

(D) Smooth Initial Histogram with Boxcar Smoother

- Create a smoothed version, ***smoothN***, of the histogram, ***N***, with a boxcar smoother incremented in one-bin steps over ***pts2bin***, equal to ***nbin*** +1, bins. Here ***nbin*** is chosen as an even number equal to 1/***Bf*** in meters (e.g., for the MABEL test data, ***nbin*** was 20).
- If needed, pad the ends of the smoothed histogram to match length of N. For example, a smoother that starts ***nbin***/2+1 points in from the beginning of the original array and stops at an equal number of points before the end, will be ***nbin*** points shorter than the

original array. In this case add ***nbin***/2 points equal to the first smoothed value to the beginning of the array and add ***nbin***/2 points equal to the last smoothed value to the end of the array.

(E) Find Limits of Valid Histogram

Find the limits of the histogram above (***jlow***) and below (***jhigh***) which the maximum in the 1-m smoothed histogram ***smoothN*** where the value in ***smoothN*** is greater than ***medianN***.

- First find the index, ***imax***, where ***smoothN(imax)*** equals the maximum in ***smoothN***.
- Increment the index, ***i***, downward from ***imax*** until ***smoothN(i)*** is less than ***medianN***. At this point ***jlow*** = ***i***+1.
- Increment the index, ***i***, upnward from ***imax*** until ***smoothN(i)*** is greater than ***medianN***. At this point ***jhigh*** = ***i***-1.

(F) Choose First Cut Signal Photons

The first-cut signal or surface photons are those in the bins of histogram ***N*** between ***jlow*** and ***jhigh***. The noise photons are those from bins below the ***jlow*** bin and above the ***jhigh*** bin. The vector ***BinB*** lists the bin assignment for each photon elevation. Therefore, the indices ***ii*** for which ***BinB(ii)*** is less than or equal to ***jhigh*** and greater than or equal to ***jlow***, are the indices of the surface reflected photon heights in ***ht_initial***.

- Populate the vector of surface photon heights, ***ht_initial_surf***, with the heights at indices ***ii*** in ***ht_initial*** (in Matlab these surface heights would be ht_initial_surf=ht_initial(ii); where ii is the vector of surface photon indices).
- Identify the corresponding track distances, ***trackdist_initial_surf***, in ***trackdist_initial*** at the indices in ***ii***.

(G) Do a Linear Fit of Height Versus Track Distance

We need to do a linear fit to surface heights versus track distance because the trend in surface height contributes substantial variance to the histogram, which can cloud the surface finding algorithm. Therefore, perform a least squares linear fit of the form ***P1*** x ***trackdist_initial_surf***+ ***P0*** to ***ht_initial_surf.***

(H) Subtract Surface Trend

We make a new vector array of detrended heights, ***ht_initial2***, by subtracting the linear surface trend derived above at step 5.2.4 (G) from ***ht_initial***,
i.e., ***ht_initial2***=***ht_initial*** - ***P1*** x ***trackdist_initial***+ ***P0***.

REPEAT SURFACE FINDING AFTER REMOVING TREND

(I) Repeat Steps (B) Through (G) Acting on the Detrended Elevations Starting With: Compute Initial Histogram and Keep Track of Bin Assignment for Each Photon Height
---
Populate the histogram array ***N*** such that each element equals the number of values from ***ht_initial2*** that are in the corresponding bin boundaries of ***Edges***. Also, for each value in ht_initial2, populate the vector BinB with the bin number to which ***ht_initial2*** is assigned from one to the length, ***LN***, of ***N*** [in Matlab, [N,BinB]=histc(ht_initial2,Edges)].

(J) Find Median of Initial Histogram

Find the median, ***medianN***, of ***N***.

(K) Smooth the Initial Histogram with Boxcar Smoother

- Create a smoothed version, ***smoothN***, of the histogram, ***N***, with a boxcar smoother incremented in one-bin steps over ***pts2bin*** , equal to ***nbin*** +1, bins. Here ***nbin*** is chosen as an even number equal to 1/***Bf*** in meters (e.g., for the MABEL data ***nbin*** was 20).
- If needed, pad the ends of the smoothed histogram to match length of N. For example, a smoother that starts ***nbin***/2+1 points in from the beginning of the original array and stops at an equal number of points before the end, will be ***nbin*** points shorter than the original array. In this case add ***nbin***/2 points equal to the first smoothed value to the beginning of the array and add ***nbin***/2 points equal to the last smoothed value to the end of the array.

(L) Find Limits of Valid Histogram

Find the limits of the histogram above (***jlow***)and below (***jhigh***) which the maximum in the 1-m smoothed histogram ***smoothN*** where the value in ***smoothN*** is greater than ***medianN***.

- First find the index, ***imax***, where ***smoothN(imax)*** equals the maximum in ***smoothN***.
- Increment the index, ***i***, downward from ***imax*** until ***smoothN(i)*** is less than ***medianN***. At this point ***jlow*** = ***i***+1.

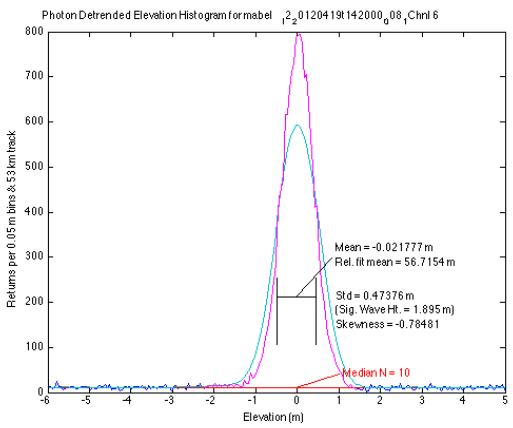


Figure 14. Coarse (blue), smoothed (cyan), and final product (magenta) histograms of detrended, 10 minutes of April 2012 MABEL data over Chesapeake Bay. The magenta histogram is ultimate output of the surface finding phase of the ocean analysis steps 3-6 of 4.2.1.3. The corresponding figure without trend removal is shown in Figure 12.

- Increment the index, ***i***, upnward from ***imax*** until ***smoothN(i)*** is greater than ***medianN***. At this point ***jhigh*** = ***i***-1.

(M) Choose Second-Cut Signal Photons

The second-cut signal or surface photons are those in the bins of histogram ***N*** between ***jlow*** and ***jhigh***. The noise photons are those from bins below the ***jlow*** bin and above the ***jhigh*** bin. The vector ***BinB*** lists the bin assignment for each photon elevation. Therefore, the indices ***ii*** for which ***BinB(ii)*** is less than or equal to ***jhigh*** and greater than or equal to ***jlow***, are the indices of the surface reflected photon heights in ***ht_initial2***.

- Populate the vector of surface photon heights, ***ht_initial2_surf***, with the heights at indices ***ii*** in ***ht_initial2*** (in Matlab these surface heights would be ht_initial2_surf=ht_initial2(ii); where ii is the vector of surface photon indices).
- Identify the corresponding track distances, ***trackdist_initial_surf***, in ***trackdist_initial*** at the indices in ***ii***, [i.e., in Matlab code ***trackdist_initial_surf***,= ***trackdist_initial(ii)***].

ht_initial2_surf is the main product of the surface finding routine. In addition, the second-pass surface indices in ***ii*** are the indices for the time, geolocation, track distance and all other data corresponding to ***ht_initial _surf***. An example of the second pass and the histogram of surface heights for MABEL data are shown in Figure 14.

(N) Calculate Number of Photons, Mean Time, and Geolocation for Segment
---

The number of surface reflected photons in the segment, ***n_photons***, is the length of the index vector ***ii.*** The time, geolocation, and track distance of the surface reflected photons are equal to the values in the vectors ***time_initial***, ***lat_initial***, ***lon_initial***, and ***trackdist_initial*** at the index positions given by ***ii*** (e.g., ***time_initial(ii)***). Compute the segment time, ***t_seg***, as the mean of ***time_initial(ii)***, and the duration of the segment, ***delt_seg***, as the last value of ***time_initial*** - the first value of ***time_initial***. Compute the segment latitude, ***lat_seg***, and segment longitude, ***lon_seg***, as the means of ***lat_initial(ii)*** and ***lon_initial(ii),*** and the segment length, ***length_seg***, as the last value of ***trackdist_initial(ii)*** - the first value of ***trackdist_initial(ii)***.

The total number of photons, ***n_ttl_photon***, equals the length of ***t_seg***, and the number of photons reflected from the surface, ***n_photon_actual***, equals the length of ***ii***. The surface reflected photon rate, r_{surf}, is equal to ***n_photon_actual*** divided by ***delt_seg,*** and the noise photon rate, r_{noise}, is equal to the difference of ***n_ttl_photon*** minus ***n_photon_actual*** divided by ***delt_seg***.

5.3.3 Processing to Characterize Long Wavelength Waves, Dependence of Sample Rate on Long Wave Displacement, and *A Priori* Sea State Bias Estimate

To determine the SSB using Eq. X7 requires that we know the covariance of the surface return rate and the height anomaly associated with the dominant surface waves. To do this we must establish the height variations and photon return rate in evenly spaced 10-m bins along 140-geo-bin segments. Given height data, ***ht_initial2_surf***, unevenly spaced along track at distances ***trackdist_initial_surf,*** we estimate the a priori sea state bias, ***binSSBias*** using the following steps.

Estimating sea state bias over 140-geo-bin segments

At this point, if not earlier, because we want to estimate the properties of waves typically hundreds of meters long, it will be necessary to aggregate 140-geo-bin segments to yield aggregate ***ht_initial2_surf*** and ***trackdist_initial_surf*** with appropriate TBD averaging of the cloud flag and other ancillary information. We assume for the analysis of this section that ***ht_initial2_surf*** and ***trackdist_initial_surf*** are aggregates over 140 geo-bins equivalent to about 2.8 km of ground track corresponding to a 2.5 Hz sample rate or 0.4 sec sample period. These will provide a reasonable compromise between spatial and temporal surface variability to produce a functionally instantaneous representation of the 100-m wavelength and longer surface waves.

Two-passes of eliminating 3-sigma deviates

Considering 140-geo-bin track with along-track distance ***trackdist_initial_surf*** **,** hereafter referred to as *X,* and photon heights ***ht_initial2_surf*** herafter referred to as *Y,* Determine the standard deviation of *Y* over the 140 geo-bins and delete any points and the corresponding values of *X*, where the absolute value of the departure of *Y* from its mean exceeds three standard deviations. Repeat this 3-sigma editing once more.

Detrend *Y* with a linear fit.

Least squares fit $Yfit = aX+b$ to *Y* and then subtract Y_{fit} from *Y* to yield the detrended $Y = Y-aX-b$. Note that because the a priori SSB determination is only concerned with variations in Y, we don't need to track the mean value of Y in this calculation.

Calculate data spacing

First compute the along-track spacing successive positions in *X*, $X_i - X_{i-1}$ or in the parlance of Matlab, yield an array of spacings: *X_spacing* =*X*(2:end)-*X*(1:end-1). Average over negative spacings where $(X_i - X_{i-1})<0$. In our experience with MABEL and ATM data if there is a negative value *X_spacing,* it can be replaced by the average of the previous and succeeding values.

Invert data spacing to make rate

The data rates in units of meter^{-1} are the inverses of the sample interval $SI_i = (X_i - X_{i-1})^{-1.}$ However, if spacing varies significantly and one tries to interpolate between or average two rates together, the result will be biased away from the true average (i.e. $2/(X_{i+1} - X_{i-1})^{-1}$ toward the higher rate. When many individual rates are averaged in this way the result is

greater than the number of photon surface returns divided by the total distance. The problem is virtually eliminated by averaging sample intervals and by interpolating heights and positions to the midpoint between samples rather than interpolating spacing or sample rate back to sample position. Consequently, we next interpolate *X* and *Y* to the middle of sample intervals *Xm by 2*-point averaging, $Xm_i = (X_i + X_{i-1})/2$, and $Ym_i = (Y_i + Y_{i-1})/2$. The result will be series for aligned SI_i, Xm_i, and Ym_i, that have two fewer points than the original *X* and *Y* series associated with loss of the first and last points defining to the first and last sample intervals.

Quantify surface height and slope and return rate in 10-m along-track bins.
Assume 280 10-m bins in the 2.8-km 140-geo-bin interval. Assign each sample interval (*ii*) in the interval-aligned SI_i, Xm_i, and Ym_i, to one of these bins with a bin index *ibin* where *ibin*=round-up (*Xm(ii)*-min(*Xm*))/10). Over all *ii*, compute the sum of each variable in the bis: *xbind(ibin)* = sum of all *Xm(ii), ybind(ibin)*=sum of all *Ym(ii),* and x*rbind(ibin*)= sum of all *SI(ii)* for which *ibin*=round-up (*Xm(ii)*-min(*Xm*))/10). Accumulate the number photon retrievals, *nbind* assigned to each bin.

Find averages, photon return rate and straight-line fits in each 10-m bin
For *ibin* equal to 1 through 280 compute bin averages:
ybind(ibin)=ybind(ibin)/nbind(ibin),
xbind(ibin)=xbind(ibin)/nbind(ibin), and
xrbind(ibin)=xrbind(ibin)/nbind(ibin).
The bin-average sample rate is taken equal to the inverse of the bin averaged sample interval:
$xrbind(ibin) = (xrbind(ibin))^{-1}$

Also compute fits of surface height, *ybinfit*, and slope, *slopebinfit*, at the middle of each bin with at least 3 points. The position, *xbinfit*, will be taken as the center of the bin.
- But first, in the cases with only 1 point in a bin, *ybinfi*t will be set equal to the single height in the bin, and the fitted height slope, *slopebinfit*, will be set to NaN (not a number).
- In the case of two points in a bin, *ybinfi*t will be set equal to *ybind*, the average of the two heights in the bin. Using the along track limits of each bin the temporally first and last along-track distances, *x_first* and *x_last*, of the two points in the bin will be identified. If *x_first* is less than *x_last, sopebinfit* will be equal to the slope between the two Y corresponding to *x_first* and *x_last*. If *x_first* is greater than *x_last* The fitted height slope, *slopebinfit*, will be set to NaN (not a number).
- If the number of points in a bin is greater than or equal to 3, a linear fit with distance will be made to the values of *Y* corresponding to the along-track distance from *x_first* to *x_last*. The *ybinfit* for the bin will be equal to the linear fit evaluated at the center of the bin, *xbinfit* as long as this value is less than one half standard deviation from the simple average of *Y* in the bin, *ybind*. If this is not true, *ybinfit* for the bin will be set to *ybind*.

Compute interval quantities for 2800-m interval
First, delete from consideration any bins with no points at all (i.e. bins for which *nbind*=0). The remaining *nbins* will be termed viable bins.

Compute Sea State Bias using all viable bins with at least one point.
Compute the covariance of surface height and photon return rate, *binCOVHtXr* equal to the sum over viable bins, *iii*, from 1 to *nbins* of (*ybinfit(iii)* *xrbind(iii))/nbins.
Similarly, the average photon return rate, *binAVG_Xr*, equals the average of xrb*ind(iii)* from *iii*=1 to n*bins*.
The bin average sea state bias estimate according to Equation X7 is:

binSSBias equal to *binCOVHtXr/binAVG_Xr*.

Compute photon rate slope correlation using bins containing 2 points or more.
Find bins (indices to be put in vector *jnnan)* where *slopebinfit* is finite (i.e., not equal to NaN). The number of such bins is n*bins2* equal to the length of *jnnan*. Compute the covariance of surface slope and photon return rate, *binCOVslope_Xr,* equal to the sum, *iii* from 1 to *nbins2* of (*slopebinfit*xrbind)/nbins2* over viable bins (indices in *jnnan*).
Similarly, *binAVG_Xr_4_slope* equals the average of the photon rate equal to the sum *xrbind(iii), iii*=1 to n*bins2* in the viable bins (indices in *jnnan*).
The bin average slope bias estimate is an experimental parameter that indicates if photons are being preferentially returned from the backsides or front faces of waves. It is:

binSlopeBias equal to *binCOVslope_Xr/binAVG_Xr_4_slope.*

5.3.4 Processing Procedure for Correction and Interpretation of the Surface Height Distribution

This section, 5.3.4, describes processing steps that analyzes the histogram of received surface photon heights, *N* from Section 5.3.2. The procedure has been tested in a developmental Matlab code (WienerTest_GaussMixandNoise2.m). Matlab executable lines have been replaced by detailed text descriptions.

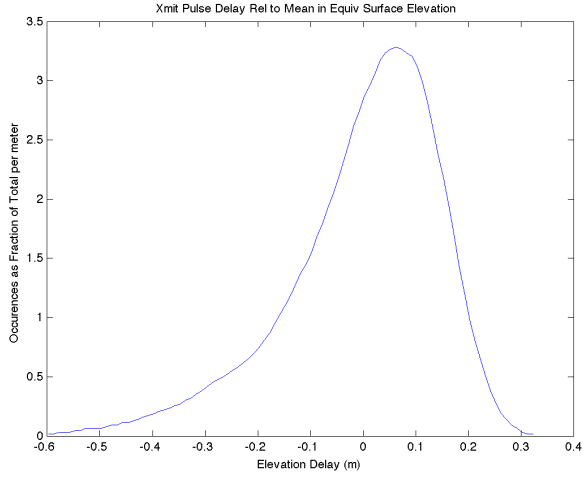


Figure 15. Instrument impulse response histogram (XmitHist00)

As described in Section 5.6, to illustrate and validate the program steps in the developmental code we convolved a real impulse response distribution taken from MABEL data (Fig. XXX) with a synthetic sea surface height, 2-Gaussian mixture distribution with a specified statistical properties (Fig. 16). A representative amount of noise was then added to

this to yield received signal photon height distribution (Fig. 17) with a known true parent surface distribution.

Here parts (A) through (G) go through the analysis process described in part 4.3.1 to determine th ıpulse re) in 5.2.5 (C n Mixture an

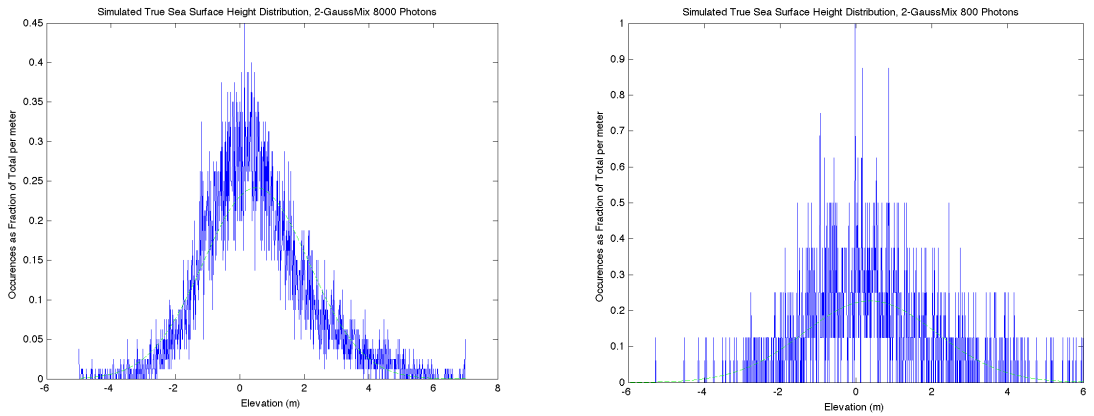


Figure 16. Left - Probability density function of the Synthetic SSH as a Gaussian mixture, 4000 samples from N(0,,1) and 4000 from N(1,2). Right – Same but with 800 photons for which the raw synthetic PDF is fairly noisy, but two peaks and skewness are faintly discernable.

5.3.4.1 Separating the Uncertainty due to Instrument impulse response

Parts

(A) Starting with histogram of received photon heights from the previous section, compute a constant signal to noise ratio (SNRc)

(B) Impulse Response Distribution

(C) Fourier transform of the received histogram

(D) Fourier transform of the instrument impulse response histogram

(E) Compute the Wiener filter

(F) Apply Wiener Deconvolution to Yield the Fourier Transform of the Surface Distribution

(G) Compute the Surface Height Distribution

Code Steps

(A) Noise Ratio Determination

– Start with the histogram of surface reflected photons heights, ***N***, from section (5.3.2 (M)) versus the sea surface height vector, ***sshx,*** with height bin size equal to ***Bf*** (e.g., ***Bf***= 0.01 m). Convert the

histogram to a vector array of probability density function (pdf) values, ***rechist***. The dimensions of ***rechist*** and ***sshx*** are the same.

- Smooth the received pdf with a 12th order, low-pass Butterworth filter with cutoff wavenumber corresponding to 0.1/***binsize***. Run the filter forward and backward over ***rechist*** to yield the smoothed histogram, ***smoothrechist***. Figure 17 shows ***rechist*** in blue and ***snoothrechist*** in red for an 8000 photon sample (left) and 800 photon sample (right)
- Compute a signal to noise ratio (***SNRc***) equal to the standard deviation (std) of ***smoothrechist*** divided by the standard deviation of the difference between the received histogram and the smoothed histogram, ***SNRc***= std(***smoothrechist***)/std((***rechist-smoothrechist***).

(B)

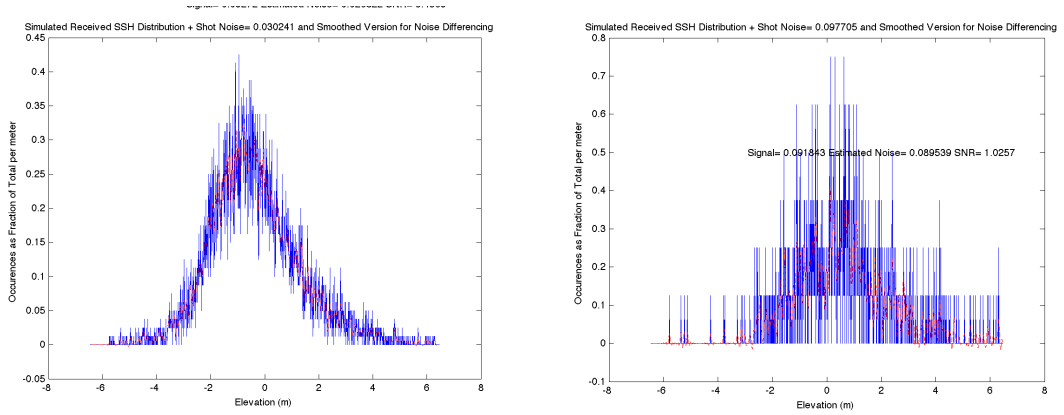


Figure 17. Synthesized received probability density functions (blue) as a convolution of the instrument impulse response PDF (Fig. 15) and the synthesized true SSH PDF (Fig. 16) plus pulse noise. The PDF for 8000 photons is at left and for 800 photons is on the right. These PDFs represent the data as they would come from the surface finding part of the processing. The red line is the smoothed PDF used to estimate the signal energy for the signal to noise ratio of the received data.

Impulse Response Distribution

Obtain the current best estimate for the instrument response distribution, ***XmitHist000***, a vector array of probability density function values for the impulse response. Note that in earlier drafts of this ATBD we referred to the instrument response distribution as XmitHist00. This was the distribution with the bin boundaries shifted so that the mean time was zero. For the purpose of calculating the length of the derived surface distribution it is conceptually useful to interpolate values in the XmitHist00 so that there are an odd number of bins with the origin bin at zero delay (or elevation) and the same bin size (***binsize***) as is used for the received elevation histogram. The method for doing this important step is TBD. For each pulse, four points will be measured on the outgoing pulse distribution and these will be part of the ICESat-2 raw data stream, plus the complete distribution will be downlinked at much less frequent intervals. Some research and

experimentation will be required to find the best method of converting these instrument measures to a representative distribution. To account for the non-Gaussian higher moments of the instrument impulse response may be best to represent it as a mixture of two Gaussian distributions. This would be somewhat analogous to our treatment of the height distribution discussed below. If so, we may want five points (***ImpulsH1, ImpulsH2, ImpulsH3, ImpulsH4, ImpulsH5)*** to be measured on the outgoing pulse to uniquely determine the five variables of a mixture of two Gaussians (2 means, 2 variances, mixing ratio).

(C) Fourier Transform of Received Height Distribution

- Find length of record (***NFFT***) equal to the next power of two greater than the length of ***rcvhist***. This should cover the longest of the impulse response distribution, ***XmitHist000***, and ***rcvhist***.
- Pad the end of ***rcvhist*** with zeros to extend it to length ***NFFT***
- Compute the fast Fourier transform, ***Rf***, of the received height distribution, ***rcvhist***
- Ensure consistent scaling so that energy is preserved (e.g., with Matlab FFT, we multiply the result by the binsize to reflect integration in the space domain
- The highest wavenumber (cycles per meter) is the Nyquist wavenumber equal to half the sampling wavenumber, wvns which equals the inverse of ***binsize*** (***wvns***=1/***binsize***). Using this, we compute the wavenumber vector (***wvn***) corresponding to the Fourier transform of the

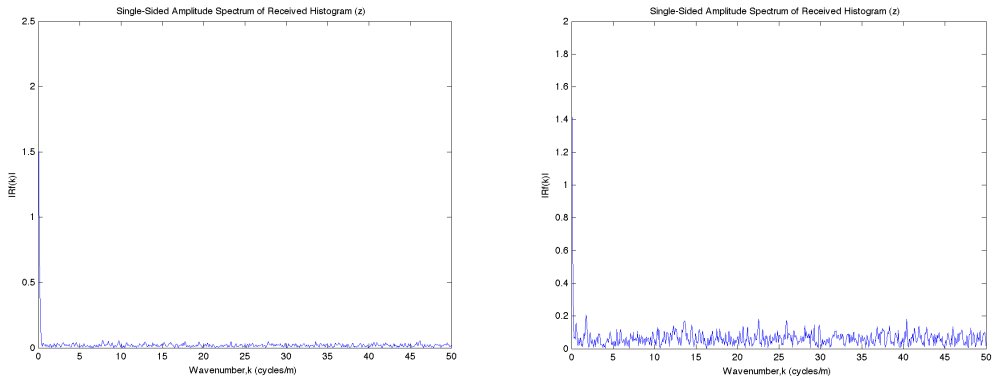


Figure 18. Single-sided amplitude spectra of the received PDFs (Fig. 17). The transform for 8000 photons is on the left and 800 photons is on the right. Note the higher wavenumber variability for 800 photons.

received histogram. Examples of ***Rf*** plotted versus ***wvn*** are shown in Figure 18.

(D) Fourier Transform of Instrument Impulse Response

- Pad the end of ***XmitHist000*** with zeros to extend it to the length ***NFFT***
- Compute the fast Fourier transform, ***Tf***, of the impulse response distribution, ***XmitHist000***

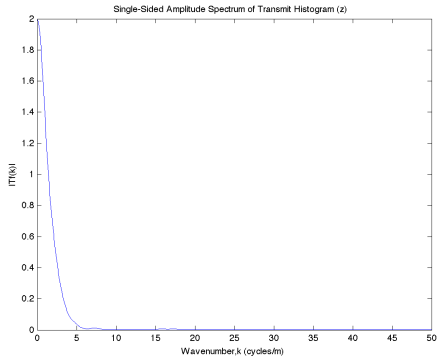


Figure 19. Single-sided amplitude spectrum of the instrument impulse response PDF (Fig. 15).

– Ensure consistent scaling so that energy is preserved (e.g., with Matlab FFT, we multiply the result by the ***binsize*** to reflect integration in the space domain

– Examples of ***Tf*** plotted versus ***wvn*** are shown in Figure 19.

(E) Compute the Wiener Filter

– Compute the Wiener filter, ***WienerF***, equal to ***Tf*** x $\boldsymbol{Tf}_{cc}$ / (***Tf*** x $\boldsymbol{Tf}_{cc}$ +$\boldsymbol{SNRc}^{-2}$), where subscript cc denotes complex conjugate.
– Note that ultimately, we may compute the Wiener filter with a signal to noise variation, ***SNRf***, that varies with wavenumber wavenumber, e.g., ***SNRf*** equal the absolute value of the Fourier transform of the received distribution (***Rf***) divided by the perceived noise (***NoiseRec***), but tests so far using the constant ***SNRc*** have worked well.
– Examples of ***WienerF*** plotted versus ***wvn*** are shown in Figure 20.

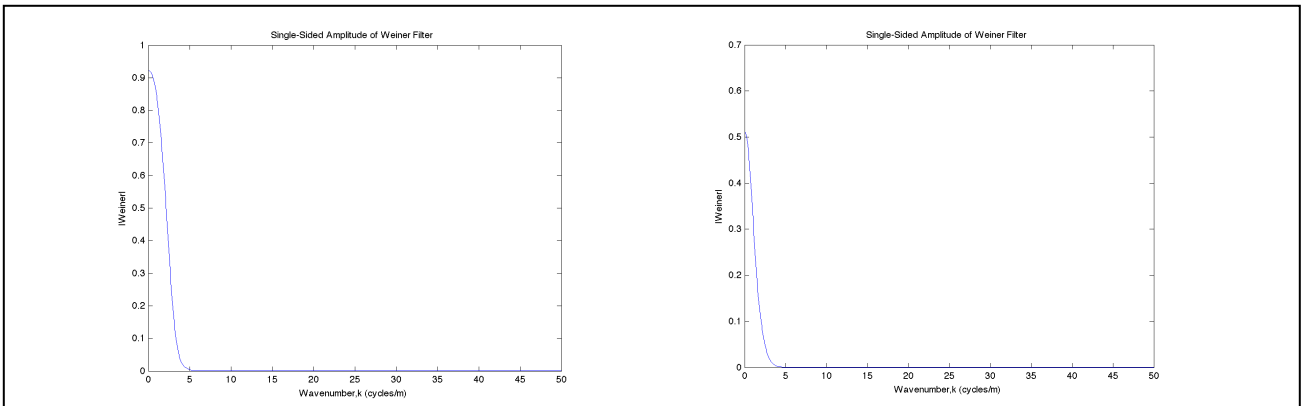


Figure 20. Single-sided amplitude spectra of the Wiener filters computed from the Fourier transform of the instrument impulse response PDFs (Fig. 19) and the signal to noise ratio estimated from the received height PDFs and smoothed versions of them (Fig. 17). Owing to a lower signal to noise ratio, the Wiener filter for 800 photons (right) has a narrower low frequency pass band than the 8000 photon filter (left).

(F) Apply Wiener Deconvolution to Yield the Fourier Transform of the Surface Distribution
---

– Compute the Fourier transform of the surface distribution, ***Yf***, as the Wiener deconvolution of the instrument impulse response distribution and the received height distribution. In Fourier (wavenumber) space this is the Weiner filter, ***WeinerF***, multiplied times the ratio of the transform of the received distribution, Rf, and the transform of the impulse response, ***Tf***, i.e., ***Yf***=***WeinerF*** x ***Rf***/***Tf***.

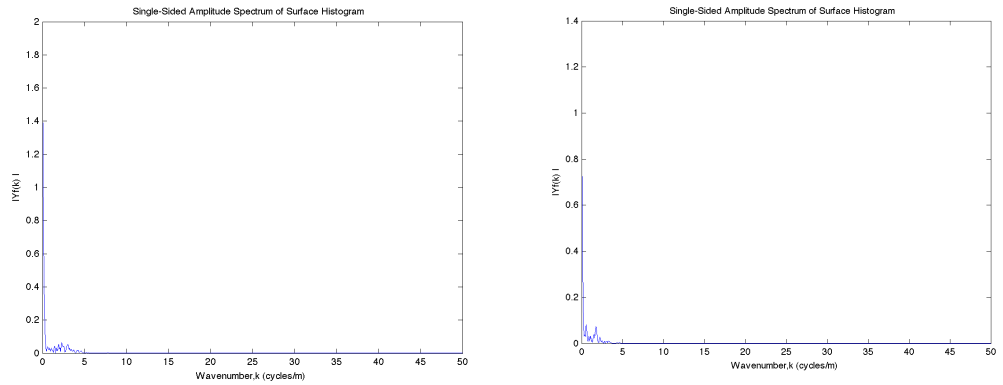
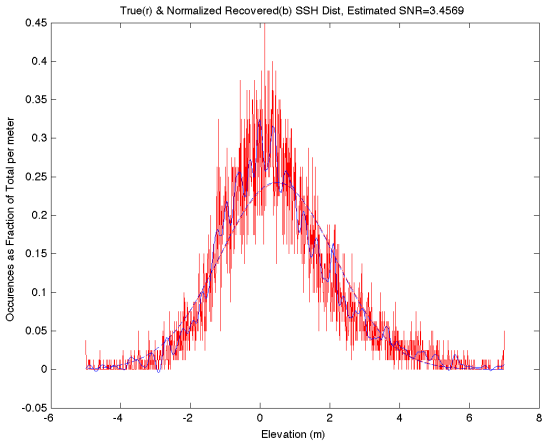


Figure 21. Single-sided amplitude spectra of the true surface spectra estimated as the Wiener filter (Fig. 20) times the Fourier transforms of the received PDFs (Fig. 18) divided by the Fourier transform of the instrument impulse response PDFs (Fig. 19). The result for 8000 photons is on the left and 800 photons is on the right.

– Examples of ***Yf*** plotted versus ***wvn*** are shown in Figure 21.

G) Compute the Surface Height Distribution

- Compute ***Yfi***, equal to the inverse Fourier transform of the output of the Wiener filtering process, ***Yf***. Ensure that the result is normalized to conserve energy and dimensional consistency. For example, the output of the Matlab inverse Fourier transform (ifft.m) must be divided by ***binsize*** for dimensional consistency.
- Compute x-axis derived from ***xrechist*** and characteristics of instrument impulse response. Convolving one finite series of length N with another of length M results in a series of length N+M-1.We know further that if the instrument impulse response histogram has zero mean with the zero bin at index position ***zbin*** this will dictate where the points are added to the true height histogram to lengthen the received height histogram. If the impulse response distribution were symmetric, equal points would be added to each end, but this is not generally true. In fact the minimum equivalent sea surface height error due to the impulse response when added to the minimum sea surface height establishes the minimum received height bin, and the maximum equivalent sea surface height error due to the impulse response when added to the maximum sea surface height establishes the maximum received height bin. Assuming length the length of the impulse response histogram is ***lxmithist***, and the zero point is at index, ***zbin***, ***zbin*-1** points will be added to the beginning of the true surface height histogram and ***lxmithist-zbin*** points will be added to the end of the surface height histogram to produce the convolved received histogram. In processing, we reverse this to trim the length of the x-axis of the received height histogram to the appropriate x-axis for the underlying surface height histogram, essentially deleting the first ***zbin*-1** points and the last ***lxmithist-zbin*** points.
- The resulting length of the derived sea surface height histogram is thus established and used to to delete the trialing points in in the derived sea surface height histogram corresponding to the zero padding used to lengthen the received histogram (rcvhist) to the next power of two. This yields the estimate of the surface histogram, Y. This should have the number of points consistent with deconvolving the impulse response from the received histogram.
- To give the surface height probability density function, ***Y***, normalize the ***Yfi*** to make up for energy lost in the Wiener filter application. The integral of ***Y*** should equal unity, so take ***Y*** equal to ***Yfi*** divided by the integral of ***Yfi***, i.e., ***Y*** = ***Yfi*** / (sum of ***Yfi*** x ***binsize***).
- Examples of ***Y*** and the difference between ***Y*** and the synthetic surface examples that we started with are given in Figure 22 for 8000 photons and Figure 23 for 800 photons.



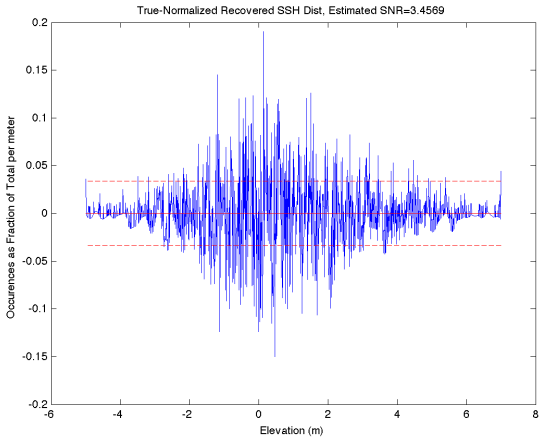


Figure 22. For the 8000 photon realization: On left in red, the PDF of true SSH (Fig. 16) with a Gaussian curve fit (dashed-red). Also in blue the PDF of SSH calculated by the Wiener deconvolution process (inverse transform of Fig. 21) with a Gaussian curve fit (dashed-blue). On the right is the difference between the true SSH and the calculated SSH.

5.3.4.2 Characterizing the Random Sea Surface

This part goes through the analysis process described in part 4.3.2 to determine the true surface distribution, namely: Maximum likelihood analysis to determine the means, variance, and mixing ratio of the two parent normal distributions.

(H) Compute the 2-component Gaussian mixture Equivalent of the Surface Distribution
---

- Here we will use maximum likelihood to compute the 2-component Gaussian mixture corresponding to the observed surface height distribution, ***Y***. The exact method of doing this is TBD. In the development process using Matlab code, we took advantage of the Matlab function gmdistribution.fit to perform the maximum likelihood analysis. Other approaches might be more suitable for ICESat-2 including a least squared error approach developed for ICESat.
- For the development code using the Matlab gmfitdistribution.fit routine, we first created a sea surface height spatial series with the height distribution ***Y***. This is accomplished by first multiplying ***Y*** by 10,000, rounding and deleting zeros to produce an integer distribution, ***YI.*** A series, ***XY***, is assembled by concatenating for each value of ***YI(i)*** for index ***i*** corresponding to height ***sshx(i)***, ***YI***(i) values equal to ***sshx(i)***. Although it is not strictly necessary, the values of ***XY*** are randomly permutated to

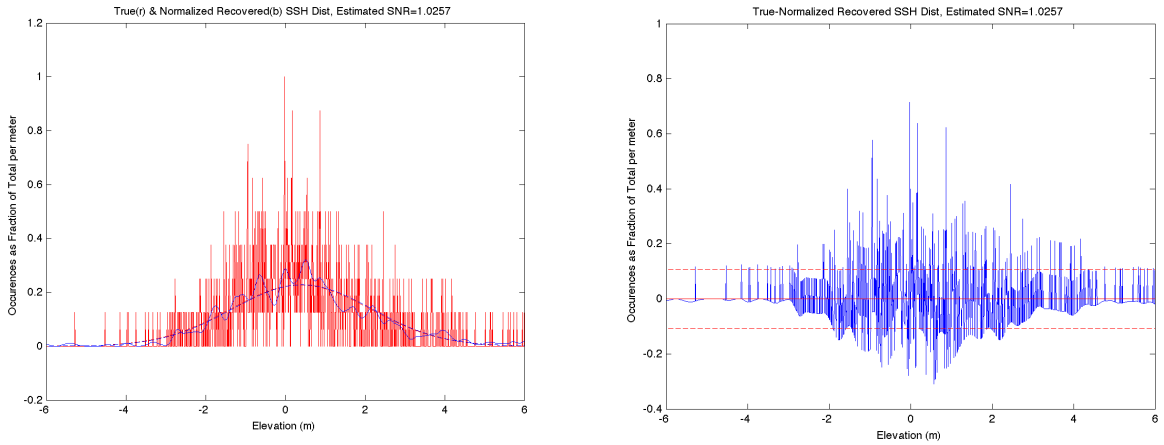


Figure 23. For the 800 photon realization: On left in red, the PDF of true SSH (Fig. 13) with a Gaussian curve fit (dashed-red). Also in blue the PDF of SSH calculated by the Wiener deconvolution process (inverse transform of Fig. 21) with a Gaussian curve fit (dashed-blue). On the right is the difference between the true SSH and the calculated SSH.

produce a realistic distribution of heights with the probability density function ***Y*** on heights given in ***sshx***.

- We then compute the 2-compnent Gaussian mixture variables, of the two-Gaussian mixture using GMfit=gmfitdistribution.fit (***XY'***,2).

– The result is Gaussian mixture distribution with 2 components:

- With example output for 8000 photons
Gaussian mixture distribution with 2 components in 1 dimensions
Component 1:
Mixing proportion, ***m1*** = 0.457026
Mean, ***mu1*** = 1.0488
Standard Deviation, ***Sig1*** = 2.0062
Component 2:
Mixing proportion, ***m2*** = 0.542974
Mean, ***mu2*** = 0.0081
Standard Deviation, ***Sig2*** = 1.0533

- For the 800 photon case the results are:
Gaussian mixture distribution with 2 components in 1 dimensions
Component 1:
Mixing proportion, ***m1*** = 0.533033
Mean, ***mu1*** = 0.8430
Standard Deviation, ***Sig1*** = 2.0559
Component 2:
Mixing proportion, ***m2*** = 0.466967
Mean, ***mu2*** = -0.0553
Standard Deviation, ***Sig2*** = 1.0533

(I) Compute the Aggregate Mean, Variance, Skewness and Kurtosis of Sea Surface Height
---

– Use the five parameters of the 2-Gaussian mixture to compute the aggregate moments, first through fourth of the aggregate mixture. From (27), for a mixture of two Gaussians with a fraction ***m1*** (m_1) from the first Gaussian with mean ***mu1*** (μ_1) and standard deviation ***sig1*** (σ_1) and a fraction ***m2*** (m_2) (note: $m_1+m_2=1$) from the second Gaussian with mean ***mu2*** (μ_2) and standard deviation ***sig2*** (σ_2) the j^{th} moment of the mixture is calculated according to:

$$E\left[X_{mix}^{j}\right] = \sum_{i=1}^{2}\sum_{k=0}^{j}\frac{j!}{k!\,(j-k)!}\,m_i\,E\left[X_i^{j-k}\right]\,E\left[X_{mix}^{k}\right] \qquad (30)$$

$$E\left[X_{mix}^{j}\right] = \sum_{i=1}^{L}\sum_{k=0}^{j}\frac{j!}{k!\,(j-k)!}\,m_i\,E\left[X_i^{j-k}\right]\,E\left[X_{mix}^{k}\right]$$

Note that in computing the moments (30), the histogram data the expectation operation (E[ ]) is not actual executed, because obviously we do not have the sample populations of the mixture components. The expectation operator appears in (30) to indicate the mixture moments and the mixture component moments coming from Section 5.2.5.2 (H) above. These are applied

sequentially using only first and second moments of the mixture components and the mixing proportions to get up to the fourth moment of the mixture.

- The first moment or mean of the mixture is the mean sea surface height, ***SSH***, for the segment to be output in ATL12. From (28), ***SSH*** equal to the mean of the mixture, μ_{mix} is:

$$SSH_{mix} = m_1 \mu_1 + m_2 \mu_2 \quad (31)$$

- The second moment of the mixture is the sea surface height variance, ***SSHvar***, for the segment to be output in ATL12 and used to estimate significant wave height. From (29), the second moment, or squared standard deviation, σ^2_{mix}, of the mixture is:

$$SSHvar_{mix}^2 = m_1 (\sigma_1^2 + \mu_1^2) + m_2 (\sigma_2^2 + \mu_2^2) - \mu_{mix}^2 \quad (32)$$

- Using (30), compute third and fourth moments, we obtain the sea surface height skewness, ***SSHskew***, and excess kurtosis ***SSHkurt***,

$$SSHskew = E[Y^3] / \sigma_{mix}^3 \quad (33)$$

$$SSHkurt = E[Y^4] / \sigma_{mix}^4 - 3 \quad (34)$$

and output these as part of ATL12 to better characterize the sea surface.

5.3.4.3 Conclusions for Section 5.2.5:

Our synthetic true SSH distribution was assumed to be an even mix of heights drawn from two Gaussian distributions, one with a mean of zero and a standard deviation of 1 m, and the other with a mean of 1 m and a standard deviation of 2 m.

The result of the Wiener deconvolution of the synthesized received distribution, made by convolving the true SSH distribution with the instrument impulse response distribution and adding noise, produced what is essentially a smoothed version of the true SSH distribution (Fig. 22 left for 8000 photons and Fig. 23 left for 800 photons). The Gaussian fits to the true and calculated distributions means and standard deviations are virtually identical, but of course, both miss the slight bi-modality and skewness of the true and calculated SSH distributions.

Using the Matlab gmdistribution.fit function on the calculated SSH distribution and assuming mixture distribution with 2 Gaussian components in 1 dimension yields for 8000 photons means of 0.0081 m and 1.0488 m versus the true values of 0 m and 1 m. The Expectation Maximization (EM) of gmdistributin.fit yields standard deviations for the two Gaussian components of 1.1481 m and 2.0062 m versus the true values of 1 m and 2 m. The

gmdistribution.fit value mixture ratios are 0.5430 and 0.4570 versus the true values of 0.50 and 0.50.

For 800 photons the EM process yields means of -0.0553 m and 0.8430 m versus the true values of 0 m and 1 m. The Expectation Maximization of gmdistributin.fit yields standard deviations for the two Gaussian components of 1.0533 m and 2.0559 m versus the true values of 1 m and 2 m. The gmdistribution,fit value mixture ratios are 0.4670 and 0.5330 versus the true values of 0.50 and 0.50.

These results with Wiener convolution and analysis by fitting a 2-component Gaussian Mixtures are promising. Clearly 8000 photons produces much better fidelity than 800 photons. This modeling approach gives us confidence that the analysis procedure is capable of an accurate estimate of surface statistics. Further testing with synthetic data, MABEL data and ATLAS simulator data will be needed fully understand the capabilities of these tools. For example, a more precise test of capability might be to use an idealized synthetic surface histogram that was more nearly perfect, having been derived analytically or drawn as we have here, but from a very large number of samples (e.g., 800,000). Pulse noise would then be added corresponding to various sample sizes. Wiener deconvolution results could then be compared to the perfectly idealized surface histogram and errors attributed solely to the processing procedure, with no error associated with the initial synthesized surface.

In running the Matlab Wiener deconvolution prototype scheme, we found it to be stable over a wide range of imposed noise values and several Gaussian mixtures. For very high assumed signal to noise ratios, it more closely reproduces the true SSH PDF (e.g., Figs. 22 and 23) seemingly accepting as signal what looks to us like noise due to low sample counts. For low signal to noise ratios, the Wiener deconvolution scheme produced a smoothed version of the synthetic SSH distribution, throwing out the noise due to low sample counts. While we think the method of signal to noise estimation used here is reasonable, it must be tested with real data. In whatever way we estimate the signal to noise ratio, the Wiener deconvolution scheme appears to be a robust way of handling it. It does an excellent job of reproducing the overall mean and variance of the true SSH PDF (Figs. 22 and 23).

Fitting a Gaussian mixture model to the data also appears to work well but we should experiment with larger numbers of synthetic photon counts to see what improvements can be made to the accuracy of the component means and variances. We also need to see how the mixture means and higher moments (Eqn. 17) improve with higher photon counts. Because we have treated the Matlab routine gmdistribution.fit as a black box, we also need to learn more about how the EM method works and would be applied in a standalone code. The scheme can be compared to the minimum squared error approach used for GLAS and we can see if that existing routine can be used for ATLAS to derive Gaussian mixtures representing the ocean surface height distribution.

5.3.5 Applying *a priori* SSB Estimate

The *a priori* SSB estimate of Section 5.3.3 can be applied to the mean SSH and DOT (=SSH-EGM2008 Geoid) for comparison to in situ cal/val or other altimeters. The impact of this SSB on the other moments is TBD.

5.3.6 Variance or Uncertainty of Estimates

TBD

5.4 Ancillary Information

5.4.1 Solar Background Photon Rate

See section 3.1.1.5. The solar background rate is taken from ATL03 (*backgrd_atlas/bckgrd_rate*) and simply averaged over the length of the height segment to produce *backdrO*.

5.4.2 Apparent Surface Reflectance (ASR)

5.5 Output Parameters

Table 5 Output to ATL19 (See Appendix A for full product specifications)

Product Label	Description	Symbol
sea_surface_heights		
h	Mean sea surface height	***SSH***
h_var	Variance of best fit probability density function (meters^2)	***SSHvar***
h_skewness	Skewness of photon sea surface height histogram	***SSHskew***
h_kurtosis	Excess kurtosis of sea surface height histogram	***SSHkurt***
mix_m1	Fraction of component 1 in 2-component Gaussian mixture	***m1***
mix_mu1	Mean of component 1 in 2-component Gaussian mixture	***mu1***
mix_sig21	Standard deviation of component 1 in 2-component Gaussian mixture	***Sig1***
mix_m1	Fraction of component 1 in 2-component Gaussian mixture	***m2***
mix_mu1	Mean of component 1 in 2-component Gaussian mixture	***mu2***
mix_sig21	Standard deviation of component 1 in 2-component Gaussian mixture	***Sig2***
t_seg	Mean time of surface photons in segment	***t_seg***
delt_seg	Time duration segment	***t_seg***
lat_seg	Mean latitude of surface photons in segment	***lat_seg***
lon_seg	Mean longitude of surface photons in segment	***lon_seg***
length_seg	Length of segment	***length_seg***

slope_seg	Sea surface slope of segment from coarse fit	***P1***
hbar_seg	Sea surface height of segment from coarse fit	***P0***
n_pulse_seg	Number of laser pulses in segment	**n_pls_seg**

Height_segment_stat		
n_ttl_photon	Number of surface photons found for the segment	***n_ttl_photon***
n_photon_actual	Number of surface photons found for the segment	***n_photon***
Photon_rate	Photon count rate, averaged over the segment	r_{surf}
backgr0_seg	*backgrd_atlas/bckgrd_rate* from ATL03_ averaged over the segment	r_{bkgrd}
PhotonNs_rate	Noise photon count rate, averaged over the segment	r_{noise}

5.6 Synthetic Test Data

The first task in the creation of the developmental software has been the creation of a synthetic data set. This is not part of ATLAS processing per se, but is a good tool for algorithm testing and development.

The procedure is imbedded in a developmental Matlab code (WienerTest_GaussMixandNoise2.m). Here, the percent sign indicates a plane text description of the procedural step and is a non-executable comment in the Matlab code. Matlab executable lines are highlighted in yellow. The program begins with parts (A) through (E) which establish an artificial surface height histogram consisting of a mixture of 2 normal distributions convolved with a real impulse response distribution taken from MABEL data by Ron Kwok. A representative amount of noise is then added to this to synthesize a received surface height with a known underlying true surface distribution.

(A) Establish the instrument impulse response distribution from MABEL data (XmitHist000)

(B) Make a Gaussian mixture representing the ssh distribution (sshhist)

(C) Compute noise (Noise) as the rRMSdifference from the analytic Gaussian mixture and sshist in the tails (+-2sig to +-3 sig)

(D) Convolve the SSH distribution with the instrument impulse response distribution to produce the received distribution (rechist)

(E) Add random Poisson (or pulse) noise due to discrete nature of photon counts representative of what we might expect for a given bin size and total number of photon counts.

% ----------

% Begin Matlab code:

% In this test program we will:
% (A) Establish the instrument impulse response distribution from Mabel data (XmitHist000)

```
% (B) Make a Gaussian mixture representing the ssh distribution (sshhist)
% (C) Compute noise (Noise) as the rms difference from the analytic gaussian mixture
%   and sshist in the tails (+-2sig to +-3 sig)
% (D) Convolve the ssh ditribution with the instrument impulse response distribution to produce
%   the received distribution (rechist)
% (E) Add random Poisson (or pulse) noise due to discrete nature of photon
%  counts representative of what we might expect for a given binsize and
%  total number of photon counts (NptsMix)
%   photons.
%

% ---------------------------------------------------------
%  (A) load and plot Instrument impulse response Histogram
% ---------------------------------------------------------
% ---------------------------------------------------------
cd
/Users/jamie/Documents/ICESat2_SDT/ICESat2_OceanATBD_Development/PhotonSourceDistrib
ution
load XmitHist000
meanXmit00=sum(XmitHist000(:,1).*XmitHist000(:,2))/sum(XmitHist000(:,2))
stdXmit00=sqrt(sum(XmitHist000(:,1).*XmitHist000(:,1).*XmitHist000(:,2))/sum(XmitHist
000(:,2)))

% XmitHist000 is a raw histogram and XmitHist000(:,2) are the total number of photons in each
% 0.01-m bin convert. We  convert it to a probability density function, the integral
% of which is 1.0. Binsize in all histograms/pdfs will be 0.01 m.
% XmitHist000 bins are relative to the mean delay is zero.
binsize=0.01;
totalxmit=sum(XmitHist000(:,2));
XmitHist_fraction=(XmitHist000(:,2)/totalxmit);
XmitHist000(:,2)=XmitHist_fraction/binsize;
checkintegralXMIT=sum(XmitHist000(:,2)*binsize);
meanXmit00=sum(XmitHist000(:,1).*XmitHist000(:,2))/sum(XmitHist000(:,2))
stdXmit00=sqrt(sum(XmitHist000(:,1).*XmitHist000(:,1).*XmitHist000(:,2))/sum(XmitHist
000(:,2)))

% plot
figure
plot(XmitHist000(:,1),XmitHist000(:,2))
title('Xmit Pulse Delay Rel to Mean in Equiv Surface Height')
ylabel('Occurences as Fraction of Total per meter')
xlabel('Height Delay (m)')
```

```
print -dpng XmitHistogram_GM_8k.png

%-----------------------------
% (B) Make a sequence sea surface height conisting of a mix of two normally
% distributed random numbers with Means MuMix, standard deviations SigmaMix
% and percent mixing ratios RatioMix, and number of points total NptsMix
% ---------------------------------------------------------
% ---------------------------------------------------------
MuMix=[1,0]
SigmaMix=[2,1]
RatioMix=[50,50]
NptsMix=8000
  ssh=makeGaussianMix(MuMix,SigmaMix,RatioMix,NptsMix);
meanssh=mean(ssh)
MeanfromMixValues=sum(RatioMix.*MuMix/sum(RatioMix))
stdssh=std(ssh)
StdfromMixValues=sqrt(sum(RatioMix.*(SigmaMix.^2)/sum(RatioMix)))
lssh=length(ssh)

% Compute histogram of ssh using bin size = 0.01 m running from -3 Std to +3
% Std of the ssh mix. The output of hist is a raw histogram and we  convert it to
% a probability density function, the integral of which is 1.0 by dividing by the
% total number of points (=length of ssh) and the binsize
%
Nsigs=3;
sshx=round(meanssh)-Nsigs*round(stdssh):binsize:round(meanssh)+Nsigs*round(stdssh);
sshhist=hist(ssh,sshx);
sshhist_fraction=(sshhist/lssh);
sshhist=sshhist_fraction/binsize;
checkintegralSSH=sum(sshhist*binsize)
%
meansshhist=sum(sshx.*sshhist)/sum(sshhist)
stdsshhist=sqrt(sum((sshx-meansshhist).*(sshx-meansshhist).*sshhist)/sum(sshhist))
totalssh=sum(sshhist)
Nbins=length(sshhist);

% ---------------------------------------------------------
%  (C) Compute noise (Noise) as the rms difference from the analytic gaussian mixture
%  (gaussfit) and sshist in the tails (+-2sig to +-3 sig)
% ---------------------------------------------------------
% ---------------------------------------------------------
```

```
gaussfit=(1/stdssh)*(1/sqrt(2*pi))*exp(-0.5*((sshx-meanssh).^2)/stdssh^2);
NoiseAll=std(sshhist-gaussfit)

% because gaussfit is a pdf we divide histogram values by bin size for
% comparison. Noise values should probably be mutltiplied by bin size for
% computatiuion as noise in histograms ?Noise=binsize*NoiseInTails?
DiffFromNorm=(sshhist)-gaussfit;
DiffFromNormInTails=[DiffFromNorm(1:round(0.5*Nbins/Nsigs)),DiffFromNorm(end-
round(0.5*Nbins/Nsigs))];
NoiseInTails=std(DiffFromNormInTails')

figure
plot(sshx,sshhist, sshx,gaussfit,'--g')
title(['Simulated True Sea Surface Height Distribution, 2-GaussMix ',num2str(NptsMix),'
Photons'])
ylabel('Occurences as Fraction of Total per meter')
xlabel('Height (m)')
print -dpng Sim_SSH_Distriburtion_GM_8k.png


% ------------------------------------------------------------
% (D) Convolve the instrument impulse response and simulated true height histograms
% to get a simulated received height ssh distribution rcvhist
% ------------------------------------------------------------
% ------------------------------------------------------------
% ------------- Matlab Routine conv.m   ------------------
% conv Convolution and polynomial multiplication.
%    C = conv(A, B) convolves vectors A and B.  The resulting vector is
%    length MAX([LENGTH(A)+LENGTH(B)-1,LENGTH(A),LENGTH(B)]). If A and B are
%    vectors of polynomial coefficients, convolving them is equivalent to
%    multiplying the two polynomials.
%
%    C = conv(A, B, SHAPE) returns a subsection of the convolution with size
%    specified by SHAPE:
%      'full'  - (default) returns the full convolution,
%      'same'  - returns the central part of the convolution
%                that is the same size as A.
%      'valid' - returns only those parts of the convolution
%                that are computed without the zero-padded edges.
%                LENGTH(C)is MAX(LENGTH(A)-MAX(0,LENGTH(B)-1),0).
% ------------------------------------------------------------------
% NOTE: conv.m does a raw convolution, i.e., simple sums of products. To
```

```
% replicate a true convolution integral with proper scaling, the results of
% conv.m must be multiplied by the bin size so that the scale of the
% resulting convolution is of the same order as the scale on the two
% histograms convolved.
% Also the length of the convolution should equal length of the ssh
% histogram plus the length Xmit histogram minus 1
%--------------------------------------------------------------------

tic
rcvhist=conv(sshhist,XmitHist000(:,2)')*binsize;
cleanrcvhist=rcvhist;
toc

lrcvhist=length(rcvhist)
lxmithist=length(XmitHist000(:,2))
lsshhist=length(sshhist)
halfwidthrcvhist=(lxmithist+lsshhist-2)/2
if 2*halfwidthrcvhist+1==lrcvhist
   message=['length of rcvhist is consistent']
end

% If the impulse response distribution were symmetric conv.m would pad the length of the sshhist
by (lxmithist-1)/2 at each end. The Xmit histogram is not symmetric. The minimuim x-axis index
of the received histogram should be the minimum index of the true surface height histogram plus
% the minimum negative bin index of the impulse response histogram and the maximu x-axis
% index of the received histogram should be the maximum index of the true histogram plus
% the maximu (positive) bin index of the Xmit histogram.
% Therefore, we find the index of the zero bin,zbin,
% of XmitHist000 (which is the old XmitHist00 histogram interpolated to even cm bin indices)and
% add the bins with index 0 to zbin-1 before the first bin of the ssh histogram and
% add the bins with index zbin+1 to end after the last bin of the ssh histogram

zbin=find(XmitHist000(:,1)==0);
xaddend=XmitHist000(zbin+1:end,1)'+sshx(end)*ones(1,length(XmitHist000(zbin+1:end,1)));
xaddbegin=XmitHist000(1:zbin-1,1)'+sshx(1)*ones(1,length(XmitHist000(1:zbin-1,1)));
xrechist=[xaddbegin,sshx,xaddend];

figure
plot(xrechist,rcvhist)
title('Simulated Convolved Sea Surface Height Distribution')
ylabel('Occurences as Fraction of Total per meter')
xlabel('Height (m)')
```

```
% ----------------------------------------------------------------


% ----------------------------------------------------------------
% ----------------------------------------------------------------
%  (E) Add random Poisson (or pulse) noise due to discrete nature of photon
%   counts representative of what we might expect for a given binsize and
%   total number of photon counts (NptsMix)
% ----------------------------------------------------------------
sumrchhist=sum(rcvhist*binsize)
noisey=rcvhist;
% make Poisson distributed photon counts in histogram to repesent received + pulse noise
for inoise=1:length(rcvhist)
   noisey(inoise)=randpoisson_JM(NptsMix*binsize*cleanrcvhist(inoise));
end

   % NOTE %
   %randpoisson_JM.m is a Matlab subroutine to generate random numbers drawn
   % from a Poisson distribution:

   function X = randpoisson_JM(np)
   x=1:round(np+1)*10;
   p= poisspdf(x,np);
   m=1;
   X = zeros(m,1); % Preallocate memory
   for i = 1:m
      u = rand;
      I = find(u < cumsum(p));
      if isempty(I)
         X(i)=0;
      else
      X(i) = min(I);
      end
   end

   % poisspdf Poisson probability density function.
   %   Y = poisspdf(X,LAMBDA) returns the Poisson probability density
   %    function with parameter LAMBDA at the values in X.
   %    The size of Y is the common size of X and LAMBDA. A scalar input
   %    functions as a constant matrix of the same size as the other input.
   %    Note that the density function is zero unless X is an integer.

Noise=std(noisey/(NptsMix*binsize)-rcvhist);
```

% Convert back to PDF

```
rcvhist=noisey/(NptsMix*binsize);
sumnoiseyrchhist=sum(rcvhist*binsize)

figure
plot(xrechist,rcvhist)
title(['Simulated Received Sea Surface Height Distribution + Pulse Noise= ',num2str(Noise)])
ylabel('Occurences as Fraction of Total per meter')
xlabel('Height (m)')
% ------------------------------------------------------------------
```

% rcvhist is the simulated received histogram at the histogram bins centered at xrechist. It is shown in blue in Figure 17 for 8000 photons (left) and 800 photons (right).

5.6.1 Numerical Computation Considerations

TBD – as needed specific considerations on method of code computation

5.6.2 Programmer/Procedural Considerations

TBD- provide information related to output parameters that were not in the algorithm description

5.6.3 Calibration and Validation

There are three types of open ocean calibration and validation:

1) Direct comparison of sea surface height or dynamic ocean topography with satellite radar altimetry from TOPEX/Poseidon and CryoSat-2. This may be possible to automate so that it can be done by the project personnel using a TBD approach, but funding is being sought outside the ICESat-2 and NASA Cryosphere program for Cal/Val.
2) Comparison of changes in DOT with in situ measurements of dynamic heights from hydrography plus ocean bottom pressure from in situ gauges or GRACE-FO. This will have to be done for the open ocean outside the Cryosphere program.
3) Direct comparison to in situ precision GPS measurements. We are seeking a way to do this as an add on to similar measurements for NASA Oceanography's SWOT cal/val effort.

5.6.4 Quality Control and Diagnostics

TBD

6.0 DATA QUALITY AND BROWSE

This section provides the processing requirements for data quality monitoring to that provide data users a quality assessment of the parameters on each file and to provide criteria to allow nearly automatic distribution of the product to the public. Browse consists of statistics or images that allow users to easily evaluate if the data would be useful and of quality for their research and as needed to aid in the quick approval or disapproval of products prior to public distribution.

6.1 Data Quality Monitor 1

6.1.1 Processing

Description of processing for data Quality monitor and the parameters to compute. Specification of pass and fail values/limits. Description of any visualization of data or quality monitor.

6.1.2 Quality Criteria

Specification of pass, warning and fail values/limits and actions to be taken by operations staff.

6.1.3 Browse

Description of any visualization of data or quality monitor.

6.2 Data Quality Monitor 2

6.2.1 Processing

Description of processing for data Quality monitor and the parameters to compute. Specification of pass and fail values/limits. Description of any visualization of data or quality monitor.

6.2.2 Quality Criteria

Specification of pass, warning and fail values/limits and actions to be taken by operations staff.

6.2.3 Browse

Description of any visualization of data or quality monitor.

7.0 TEST DATA AND RESULTS

7.1 Unit Test Data

7.1.1 Unit test 1 Purpose

Repeat for as many unit tests as needed to test each condition in the algorithm. What functions of the algorithm it will test and algorithm expectation.

7.1.1.1 Unit Test Inputs

Specific value(s) for each input or sets of values to allow simple unit testing that validates the algorithm is implemented as specified.

7.1.1.2 Results

The resulting value(s) for each output parameter based on the inputs

7.2 Simulated Test Data

Repeat as many simulation data sets as needed to test each condition in the algorithm.

7.3 Simulated Data Set 1

7.3.1.1 Source

Description of the data set, what functions of the algorithm it will test and algorithm expectation. Name and source location of the data set
Repeat for as many as simulation tests as needed.

7.3.1.2 Results

Description of the results from the data set being processed by the algorithm. Name and source location of the expected resulting values from processing the data set.

8.0 CONSTRAINTS, LIMITATIONS, AND ASSUMPTIONS

TBD- as needed.

GLOSSARY/ACRONYMS

ATLAS	ATLAS Advance Topographic Laser Altimeter System
GSFC	Goddard Space Flight Center
ICESat-2 MIS	ICESat-2 Management Information System
PSO	ICESat-2 Project Support Office
SIPS	ICESat-2 Science Investigator-led Processing System

TBD

APPENDIX A: ICESat2 Data Products

ICESat-2 Data Products

File ID/Level	Product Name	Concept	Short Description	Frequency
00/0	Telemetry Data	Full rate Along-track with channel info	Raw ATLAS telemetry in Packets with any duplicates removed	Files for each APID for some defined time period
01/1A	Reformatted Telemetry	Full rate Along-track with channel info	Parsed, partially reformatted, time ordered telemetry. Proposed storage format is NCSA HDF5.	Uniform time TBD minutes (1 minute?)
02/1B	Science Unit Converted Telemetry	Full rate Along-track with channel info	Science unit converted time ordered telemetry. Reference Range/Heights determined by ATBD Algorithm using Predict Orbit and s/c pointing. All photon events per channel per pulse. Includes Atmosphere raw profiles.	Uniform time TBD minutes (1 minute?)
03/2A	Global Geolocated Photon Data	Full rate Along-track with channel info	Reference Range/Heights determined by ATBD Algorithm using POD and PPD. All photon events per pulse per beam. Includes POD and PPD vectors. Classification of each photon by several ATBD Algorithms.	Uniform time TBD minutes (1 minute?)
04/2A	Calibrated Backscatter Profiles	3 profiles at 25 Hz rate (based on 400 pulse mean)	Along-track backscatter data at full instrument resolution. The product will include full 532 nm (14 to -1.0 km) calibrated attenuated backscatter profiles at 25 times per second for vertical bins of approximately 30 meters. Also included will be calibration coefficient values for the polar region.	Per orbit
05/2B	Photon Height Histograms	Fixed distances Along-track for each beam	Histograms by prime Classification by several ATBD Algorithms. By beam	Uniform time TBD minutes (30 minutes?)
06/L3	Antarctica Ice Sheet Height / Greenland Ice Sheet Height	Heights calculated with the ice sheet algorithm, as adapted for a dH/dt calculation	Surface heights for each beam, along and across-track slopes calculated for beam pairs. All parameters are calculated for the same along-track increments for each beam and repeat.	There will be TBD files for each ice sheet per orbit

File ID/Level	Product Name	Concept	Short Description	Frequency
07/ L3	Arctic Sea Ice Height/ Antarctic Sea Ice Height	Along-track heights for each beam ~50-100m (uniform sampling); separate Arctic and Antarctic products	Heights of sea ice and open water samples (at TBD length scale) relative to ellipsoid after adjusted for geoidal and tidal variations, and inverted barometer effects. Includes surface roughness from height statistics and apparent reflectance	There will be files for each pole per orbit
08/ L3	Land Water Vegetation Heights	Uniform sampling along-track for each beam pair and variable footpath	Heights of ground including inland water and canopy surface at TBD length scales. Where data permits, include estimates of canopy height, relative canopy cover, canopy height distributions (decile bins), surface roughness, surface slope and aspect, and apparent reflectance. (Inland water > 50 m length -TBD)	Per half (TBD) orbit
09/ L3	ATLAS Atmosphere Cloud Layer Characteristics	Based on 3 profiles at a 25 hz rate. (400 laser pulses are summed for each of the 3 strong beams.)	Cloud and other significant atmosphere layer heights, blowing snow, integrated backscatter, optical depth	Per day
10/ L3	Arctic Sea Ice Freeboard / Antarctic Sea Ice Freeboard	Along-track all beams. Freeboard estimate along-track (per pass); separate Arctic/ Antarctic products	Estimates of freeboard using sea ice heights and available sea surface heights within a ~TBD km length scale; contains statistics of sea surface samples used in the estimates.	There will be files for each polar region per day
11/ L3	Antarctica Ice Sheet H(t) Series/ Greenland Ice Sheet H(t) Series	Height time series for pre-specified points (every 200m) along-track and Crossovers.	Height time series at points on the ice sheet, calculated based on repeat tracks and/or crossovers	There will be files for each ice sheet for each year
12/ L3	Ocean Height	Along-track heights per beam for ocean including coastal areas	Height of the surface10 Hz/700 m (TBD) length scales. Where data permits, include estimates of height distributions (decile bins), surface roughness, surface slope, and apparent reflectance	Per half orbit
13/ L3	Inland Water Height	Along-track height per beam	Along-track inland ground and water height extracted from Land/Water/ Vegetation product. TBD data-derived surface indicator or mask. Includes roughness, slope and aspect.	TBD files Per day

File ID/Level	Product Name	Concept	Short Description	Frequency
14/L4	Antarctica Ice Sheet Gridded/ Greenland Ice Sheet Gridded	Height time series interpolated onto a regular grid for each ice sheet. Series (5-km posting interval)	Height maps of each ice sheet for each year of the mission, based on all available ICESat-2 data.	Per ice sheet per year
15/L4	Antarctica Ice Sheet dh/dt Gridded/ Greenland Ice Sheet dh/dt Gridded	Images of dH/dt for each ice sheet, gridded at 5 km.	Height-change maps of each ice sheet, with error maps, for each mission year and for the whole mission.	Per ice sheet for each year of mission, and for the mission as a whole
16/ L4	ATLAS Atmosphere Weekly	Computed statistics on weekly occurrences of polar cloud and blowing snow	Polar cloud fraction, blowing snow frequency, ground detection frequency	Per polar region Gridded 2 x 2 deg. weekly
17/ L4	ATLAS Atmosphere Monthly	Computed statistics on monthly occurrences of polar cloud and blowing snow	Global cloud fraction, blowing snow and ground detection frequency	Per polar region Gridded 1 x 1 deg. Monthly
18/L4	Land Height/ Canopy Height Gridded	Height model of the ground surface, estimated canopy heights and canopy cover gridded on an annual basis. Final high resolution DEM generated at end of mission	Gridded ground surface heights, canopy height and canopy cover estimates	Products released annually at a coarse resolution (e.g. 0.5 deg. tiles, TBD). End of mission high resolution (~1-2km)
19/ L4	Ocean MSS	Gridded monthly	Gridded ocean height product including coastal areas. TBD grid size. TBD merge with Sea Ice SSH	Monthly
20/ L4	Arctic and Antarctic Gridded Sea Ice Freeboard/	Gridded monthly; separate Arctic and Antarctic products	Gridded sea ice freeboard. (TBD length scale)	Aggregate for entire month for each polar region

File ID/Level	Product Name	Concept	Short Description	Frequency

21/ L4	Arctic Gridded Sea Surface Height within Sea Ice/ Antarctic Gridded Sea Surface Height within Sea Ice	Aggregate for entire month (all sea surface heights within a grid) separate Arctic and Antarctic products	Gridded monthly sea surface height inside the sea ice cover. TBD grid	Aggregate for entire month for each polar region
Experimental	Arctic Sea Ice Thickness / Antarctic Sea Ice Thickness	Per Pass Thickness samples (from 10-100m freeboard means) for every 10 km (TBD) segment (all beams) where leads are available; (per pass)	Sea ice thickness estimates derived from the sea ice freeboard product. External input: snow depth and density for each pass.	There will be files for each polar region per day
Experimental	Arctic Gridded monthly Sea Ice Thickness / Antarctic Gridded monthly Sea Ice Thickness	Aggregate for entire month (all thickness observations within a grid) plus Thickness (corrected for growth)	Gridded sea ice thickness product; centered at mid-month. Include thickness with or without adjustment for ice growth (based on time differences between freeboard observation).	Gridded monthly (all thickness observations within a grid) for each polar region
Experimental	Lake Height	Along reference track per beam in Pan-Arctic basin (>50-60 deg N).	Extracted from Product 08 and 13, for lakes >10 km2, with slope and aspect. Ice on/off flag. TBD water mask developed from existing masks.	Monthly along track product, no pointing
Experimental	Snow Depth	Along reference track per beam for Pan-Arctic basin (>50-60 deg N).	Extracted from Product 08 and 13 along track repeat heights, with slope and aspect. Snow detection flag.	Monthly along track product, no pointing

9.0 REFERENCES

Arnold, D. V., W. K. Melville, R. H. Stewart, J. A. Kong, W. C. Keller, and E. Lamarre (1995), Measurements of electromagnetic bias at Ku and C bands, *J. Geophys. Res.*, *100*(C1), 969-980.

Carrère, L., and F. Lyard (2003), Modeling the barotropic response of the global ocean to atmospheric wind and pressure forcing - comparisons with observations.

Chambers, D. P., S. A. Hayes, J. C. Ries, and T. J. Urban (2003), New TOPEX sea state bias models and their effect on global mean sea level, *J. Geophys. Res.*, *108*(C10), 3305.

Chapron, B., V. Kerbaol, D. Vandemark, and T. Elfouhaily (2000), Importance of peakedness in sea surface slope measurements and applications, *J. Geophys. Res.*, *105*(C7), 17195-17202.

Cox, C., and W. Munk (1954), Measurement of the Roughness of the Sea Surface from Photographs of the Sun?s Glitter, *J. Opt. Soc. Am.*, *44*(11), 838-850.

Elfouhaily, T., D. R. Thompson, B. Chapron, and D. Vandemark (2000), Improved electromagnetic bias theory, *J. Geophys. Res.*, *105*(C1), 1299-1310.

Gaspar, P., F. Ogor, P.-Y. Le Traon, and O.-Z. Zanife (1994), Estimating the sea state bias of the TOPEX and POSEIDON altimeters from crossover differences, *J. Geophys. Res.*, *99*(C12), 24981-24994.

Hausman, J., and V. Zlotnicki (2010), Sea State Bias In Radar Altimetry Revisited, , *Marine Geodesy*, *33*(S1), 336---347.

Markus, T., et al. (2016), The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science 1 requirements concept, and implementation., *Remote Sensing of the Environment*, *in press*.

Menzies, R. T., D. M. Tratt, and W. H. Hunt (1998), Lidar In-Space Technology Experiment Measurements of Sea Surface Directional Reflectance and the Link to Surface Wind Speed, *Appl. Opt.*, *37*(24), 5550-5559.

Plant, W. J. (2015a), Short wind waves on the ocean: Wavenumber-frequency spectra, *J. Geophys. Res. Oceans*, *120*, 2147–2158.

Plant, W. J. (2015b), Short wind waves on the ocean: Long-wave and windspeed dependences, *J. Geophys. Res. Oceans*, *120*, 6436–6444.

Urban, T. J., and B. E. Schutz (2005), ICESat sea level comparisons, *Geophys. Res. Lett.*, *32*(23), L23S10.

Vandemark, D., B. Chapron, T. Elfouhaily, and J. W. Campbell (2005), Impact of high-frequency waves on the ocean altimeter range bias, *J. Geophys. Res.*, *110*(C11), C11006.

Wu, J. (1972), Sea-Surface Slope and Equilibrium Wind-Wave Spectra, *Physics of Fluids*, *15*(5), 741-747.