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

Algorithm Theoretical Basis Document (ATBD) for Ocean Surface Height

Prepared By:

ICESat-2 Science Definition Team Ocean Working Group

Contributors:

James Morison
David Hancock
Suzanne Dickinson
John Robbins
Leeanne Roberts
Ron Kwok
Steve Palm
Ben Smith
Mike Jasinski
Bill Plant
Tim Urban

This document may be cited as:

Morison, J., D. Hancock, S. Dickinson, J. Robbins, L. Roberts, R. Kwok, S. Palm, B. Smith, M. Jasinski, B. Plant, and T. Urban (2024). *Ice, Cloud, and Land Elevation Satellite (ICESat-2) Project Algorithm Theoretical Basis Document (ATBD) for Ocean Surface Height, Version 7*. ICESat-2 Project, DOI: 10.5067/4SZEQR2DHQOR

Goddard Space Flight Center
Greenbelt, Maryland

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 processing open ocean data 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:
Tom Neaumann, ICESat-2 Project Scientist
Mail Stop 615
Goddard Space Flight Center
Greenbelt, Maryland 207

Review/Approval Page

Prepared by:

Jamie Morison
Senior Principal Oceanographer
Affiliate Professor, Oceanography
University of Washington, Applied Physics Laboratory

Reviewed by:

Bea Csatho
Chair, Professor
University at Buffalo
Department of Geology

Tom Neumann
Project Scientist, ICESat-2
NASA Goddard Spaceflight Center, Code 615

Approved by:

Thorsten Markus
Program Manager, Cryospheric Sciences
NASA Headquarters

***** 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 7/27/2019 Section 5.3.4.1 code steps A: replace ***snoothrechist*** with ***smoothrechist*** 7/30/2019 – globally replace term Maximum Likelihood with Expectation Maximization to clarify that the Matlab developmental code and the ASAS FORTRAN code both use expectation maximization. 7/30/2019 Change the third paragraph of 5.3.4.2 (H) to: - To apply the expectation maximization approach, we first derive a sea surface height spatial series from the height distribution ***Y,*** which is defined over a range of heights the same as the received height histogram (i.e***., jlow*** to ***jhigh*** from surface finding). This is accomplished by first multiplying ***Y*** by 10,000, rounding and deleting any values less than or equal to zero (a few unrealistic weakly negative values to occur in ***Y*** associated with the Weiner deconvolution process) 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)***. The shift to heights including ***meanoffit2*** was accomplished in developmental code by adding ***meanoffit2*** to each value in ***XY***. (ASAS FORTRAN code does not add ***meanofffit2*** at this point but after the Gaussian Mixture calculation.) 7/30/2019 Add ***xbind*** as an output variable in **5.3.3** (14) and at the end of **5.3.3** To Table 6 add*:* *xbind* = 1 x 710 element array of 10-m bin averages of along-track distance And change the description of *xbin* to: "Center of 1 x 710 element array of 10-m bins. Note this may be included as a data description or other static array equal to [5, 15, 25, 35 ….. 7095 m] 7/30/2019 In Table 6 changed from 700 10-m bins to 710 10-m bins to account for ocean segments longer than 7,000 m 7/30/2019 Change 5.3.4.1 (G) points 1-4 to:	GG G	

	"The surface histogram should be the same length and have the sa horizontal (height) axis as the received histogram, ***rechist***. In this we compute the distribution of surface height as a probability den function (PDF), ***Y***, starting with the Fourier transform of that PDF output by the Weiner deconvolution. The steps are: 1. Compute ***Yfi***, as the inverse Fourier Transform of ***Yf*** divided by ***binsize. Yfi*** as it comes from the inverse Fourier transform is ***NFFT*** points long with the first half appearing displaced to the end of the array. We reorder this by first putting the last ***NFFT***/2 points ahead of the first ***NFFT***/2 points. To establish the surface height probability density function, ***Y***, we then remove the first (***rzbin***-1) points where ***rzbin*** is the index of the centroid (height equals zero) index of the received histogram. We keep as ***Y*** the remaining points equal to the length of the ***rechist***. In the case of the ASAS FORTRAN code, which pads each end of ***rechist*** to reach heights of ±15 m, ***rzbin*** is the center bin of ***rechist***. 2. Set the x-axis of ***Y***, ***derivedsshx*** equal to the x-axis, ***xrechist***, of the received histogram ***rechist***." 7/30/2019 Clarify bin centering in 5.3.1 (B) by changing (B) to: "Establish an initial coarse histogram array, $\boldsymbol{H_c}$, spanning ± 15 m with bin size $\boldsymbol{B_1}$ equal to 0.01 m with center bin centered at zero and bin centers at whole centimeters. Also establish a data array, $\boldsymbol{A_{coarse}}$, for up to 10,000 photon heights and associated information (index, geolocation, time) plus noise photon counts. This will be populated with data as we step through 14-geobin segments searching for an adequate number of surface photon candidates." and change 5.3.2 (A) to: "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***. In practice ***Bf*** equals 1 cm, ***Cntrpt*** values are whole centimeters, and ***Edges*** are at half centimeter values. Also establish a vector array, ***BinB***, as long as ***ht_initial*** for bin assignments."		

	8/6/2019 Change 5.3.4.1 (G) near the end from "***Y*** is defined over a vector of bin centers, ***sshx,*** but Y will be output as a 1001 element vector for every segment with the 501st center bin being for height zero, bin 501-***YlowX/binsize*** being the lowest bin with nonzero ***Y*** and bin 501+***YlowX/binsize*** being the highest bin with nonzero ***Y.***" To "***Y*** is defined over a vector of bin centers, ***sshx,*** but for ***binsize*** = 1 cm ***Y*** will be output as a 3001-element vector from -15 m to +15 m for every segment with the 1,501st center bin being for height zero. Values will be zero for bins outside the range of ***sshx.*** Owing to noise and the effect of the FFT-deconvolution-inverse-FFT process, small negative values occur in ***Y*** within the range of ***sshx***. These will be set to zero in the output ***Y*** vector." 8/28/2019 Replace section 5.3.3 to use simplified and more robust method of computing photon return rate in SSB calculation: ***xrbin*** = ***nbind***/10 **9/9/2019** in 5.3.3 7. - Add clause to end of sentence to clarify using only bins with photons for ssb calculation, i.e., 7. Compute, ***binAVG_Xr***, the average of binned data rate, ***xrbin,*** over the 10-m bins with photons. **11/5/2019** Modified 5.3.2 to change surface finding based on the distribution of photon heights to surface finding based on the photon height anomaly relative to a moving bin average of high confidence photon heights. This is done to exclude subsurface returns under the crests of surface waves that otherwise fall inside the histogram of true surface heights. Added to Table 5 ***conf_lim,*** the limiting confidence level to be included in the moving bin average, ***nphoton***, the number of photons either side of a central photon to be included in the moving bin average, e.g., for ***nphoton***=10, a 21 point average is used **11/18/2019** Added section 5.3.3.2 to compute ***nharms***+1 harmonic coefficients, ***a***, of surface height for surface waves and possible relation to degrees of freedom and uncertainty. Added to Table 5 nharms and added to Table 6 ***wn*** and ***a***. ***wn*** is a vector of ***nharms*** wavenumbers. ***a***(1) is the coefficient for the zero wavenumber component of the fit to test if this is a more stable estimate of average heights. ***a***(3,5,7,..2****nharms***+1) are sine		

	coefficients for ***wn***(1,2,3,..***nharms***). ***a***(2,4,6,..2****nharms***+1) are cosine coefficients for ***wn***(1,2,3,..***nharms***). **11/27/2019** Added a section to the end of 5.3.6.1 to estimate the effective number of degrees of freedom, ***NP_effect***, and associated uncertainty, ***h_uncrtn***, in sea surface height over an ocean segment. The estimates are based on the integral of autocorrelation of 10-m bin average surface height. Added ***h_uncrtn*** and ***NP_effect*** to Table 6. 12/19/2019 Removed (or corresponding ***ht_initial***, ***really?***) from 5.3.2(D) 12/19/2019 Added the sentence: "The received histogram, *Nrs*, is then the histogram of ***ht_initial2_surf*** only including from the lowest height with count>0 to the highest height with count>0." to 5.3.2 (P) to clarify that though the histogram of height anomalies, ***N***, is used for surface finding, we look to the corresponding histogram of heights, ***Nrs***, for the rest of the analysis. In a later section we form the probability density function, ***rechist***, from ***Nrs*** and pass it on to have the instrument impulse response removed and statistics of sea surface height computed. 12/19/2019 Pursuant to the change immediately above we changed the first sentence of 5.3.4.1 (A) "Start with the histogram of surface reflected photons heights, ***N***, from section (5.3.2 (M)) versus the …" to, " Start with the histogram of surface reflected photons heights, ***Nrs***, from section (5.3.2 (P)) versus the …" 02/06/2020: Corrected 5.3.3.2 paragraph 4) to properly include ***ht_initial2_surf*** in the derivation of harmonic coefficients: 4) For ***y*** equal to the column vector of ***ht_initial2_surf*** with ***meanoffit2*** added to it, compute the vector of harmonic coefficients, ***a***, by the Vericek least squared error or left pseudo inverse method, where ***a=F'F\F'y*** . The ***m*** x 1 vector of harmonic coefficients, ***a***, for the optimum harmonic fit to ***ht_initial2_surf*** plus ***meanoffit2*** versus ***trackdist_intial2_surf*** :		

	a(1)+ ***a*** (2)*sin(***wn***(1)* 2*π*x***) + ***a*** (3)*cos(***wn***(1)* 2*π*x***)+ ***a*** (4)*sin(***wn***(2)* 2*π*x***) + ***a*** (5)*cos(***wn***(2)* 2*π*x***)+ ***a*** (6)*sin(***wn***(3)* 2*π*x***) + ***a*** (7)*cos(***wn***(3)* 2*π*x***)+ : : + : : + ***a***(2*nharms)*sin(***wn***(***nharms***)* 2*π*x***) + ***a***(2*nharms +1)*cos(***wn***(***nharms***)* 2*π*x***) Backslash or left matrix divide A\B is the matrix division of A into B, which is essentially the same as $A^{-1}B$. 02/24/2020 Revised Section 5.3.3.2 to clear up various ambiguities and included for the harmonic fit only, the technique to fill gaps in the data greater than ***gaplimit***=3.2 m with white noise about ***meanoffit2*** 02/24/2020 Added ***gaplimit*** to control parameters Table 5 and added the harmonic fit signal to noise ratio ***SNR_harm*** and gap statistics, ***DXbar, Xvar***, and ***DXskew*** to outputs Table 6. 02/25/2020 Corrected 5.3.3.2 paragraph 7 by changing "***trackdist_initial2_surf"*** to***"*** the gap-filled ***trackdist_initial2_surf"*** in two places 3/19/2020 Correct 5.3.3.2 (1) to divide the third moment of data spacing by the 2nd moment to the 3/2 : skewness (***DXskew*** = Sum (***DX-DXbar***)^3/ ***DXvar³/²*** from 1 to ***N***-1 divided by ***N***-2)). 3/23/2020 Correct 5.3.3.2 formula for SNR-harm to remove mean fit in the numerator of the formula. It now reads: ***SNR_harm = E[(yfit-a(1))*(yfit-a(1))']/E[(y-yfit)*(y-yfit)']*** where ***E[]*** denotes expected value***.*** 4/7/2020 Modify 5.3.6.1 to clarify calculation of ***Lscale*** and ***NP_effect*** for uncertainty determination. This includes adding equations (41), (42), (43), and (44), identifying equation (40) and modifying the pseudo code section b) to include corrected formulations of ***Lscale*** : "Start with ***Lscale*** = 0 and at *i* = 1 where the corresponding lag, ***l***, is zero, and ***R_L***(*i*)=1. Increment i and the corresponding ***i*** by 1		

	and keep adding to ***Lscale*** while ***R_L(***i+1***)*** is greater than zero, i.e., ***Lscale*=*Lscale*** +((1-***l*/*Nbin10)****R_L*(i)+ (1-(***l+1)*/*Nbin10)****R_L*(i+1))/2 When ***R_L(***i+1***)*** becomes less than zero, terminate the integration by essentially assuming ***R_L(***i+1***)*** is zero, i.e., ***Lscale*=*Lscale*** + (1-***l*/*Nbin10)****R_L*(i))/2 " Also added ***Lscale*** and ***Nbin10*** to Output Table 6. 4/17/2020 – Modified 3.2 Grif=ded Sea Surface Height (ATL19) to allow for merging ATL10 with ATL12 data in the polar oceans. 4/17/2020 – Modified 4.5 and Figure 13 to add computing averages of SSH moments weighted by degrees of freedom. 4/20/2020 – In Table 6 ATL12 Outputs, change the name of photon noise rate, *photonns_rate,* to *photon_noise_rate* to avoid confusion with photon rate, *Photon_rate.* 4/20/2020 – Inserted as section 5.5, a description the gridding process including 5.6.1 Grids 5.6.2 Inputs to Gridding, and Table 7 of inputs from ATL12. 5/2/2020 Edited 5.3.6.1 paragraphs a) and b) to make it clear that we use the lagged covariance ***COV_L***, and lagged autocorrelation, ***R_L***, of binned surface height are used from lag, ***l***, equal to 0 to the length of the record. This included standardizing the spelling the lag=dependent variables as ***COV_L*** and ***R_L***. 5/7/2020 Changed Section 5.3.2 (B) ***conf_lim*** from 4 to 3. Now it reads "For example, in testing we used ***nphotons*** = 5 and ***conf_lim*** = 3 to make an 11-point running average centered on each photon height, and included in each average, only heights with confidence level 3 and 4. 5/7/2020 and 5/8/2020 Edited section 5.6 Gridding to include missing variables, regularize variable naming for averaged (_avg) and degree-of-freedom weighted averaged (_dfw)		

	variables. Added Appendix, a hierarchy of ATL12 and ATL19 variables. 5/19/2020 In the ATL19 section changed ***dot_sig_albm*** to ***dot_sigma_dfwalbm*** as a better descriptor. 5/19/2020 Substituted Dickenson's Table 6 edited to have correct folder labels and to be consistent with ATL12 output names and folders. 5/19/2020 Added folder headings to Table 7 to be consistent with edited Table 6 and made small edits to ATL19 to emphasize we would perform analyses er each strong beam until the all-beam combination. 6/6/2020 added ***n_ttl_photon*** to Table 7 inputs to ATL-19 gridding. 6/6/2020, added a paragraph to 5.19.4.1 to explain computation of grid cell-total surface photons, total photons, surface photon rate and noise photon rate. Also added a paragraph to 5.19.5 to explain computation of all-beam-total surface photons, total photons, surface photon rate and noise photon rate. 6/6/2020, added the grid cell total and all-beam grid cell total surface photon and all photon counts and photon rate and noise photon rate to Table 8 ATL19 Outputs. 6/6/2020 In response to Hancock "Photo select" email June 3, 2020, edited 4.2.1.1 on page 57 to clear up ambiguities in the use of the downlink band and confidence level from ATL03 to select candidate photons for the ATL12 surface finding routine. 6/29/2020 Edit 3.1.3.4 to distinguish between tide corrections already in ATL03 photon heights and those to be made in ATL12, Also emphasized the ATL03 solid earth tides are in a tide-free reference system as opposed to a mean-tide system, and for ATL12 this is appropriate when the ocean tides are added. 6/29/2020 Added to 4.2.1.1 and 5.2.1 sentences to include exclusion of data for which the pointing and orbit determination are suspect: "As of Release 3, the ATL03 includes a flag for each ground tract indicating the quality of the precision orbit and		

	pointing determination: ***gtx/geolocation/podppd_flag,*** and we have found unrealistic values for ocean segments when this flag is non-zero indicating degraded orbit and/or pointing data. Consequently ATL12 processing should only be done with data for which ***gtx/geolocation/podppd_flag***=0." Also added which ***podppd_flag*** to Table 2. 6/29/2020 Added to 5.3.3.1 (5) averaged variables - Set ***latbind*** to the mean of photon latitude in each 10-m bin. - Set ***lonbind*** to the mean of photon longitude in each 10-m bin Also added ***latbind*** and ***lonbind*** to Table 6, ATL12 Outputs 6/29/2020 Added a data editing paragraph to 5.19.2 "We find that for reasons that we are investigating, ATL12 produces some ocean segment heights that are unrealistic compared to the geoid. Consequently, prior to inclusion in the gridding process, each ATL12 ocean granule is to be edited of ocean segment heights, *h*, that do not survive a 2-pass 3-sigma filter on dynamic ocean topography equal to *h-geoid_seg*. For each granule compute the standard deviation and mean of *h-geoid_seg.* On the first pass any ocean segment for which the magnitude of *h-geoid_seg* – mean(*h-geoid_seg*)> 3 x std(*h-geoid_seg*) should be eliminated from the input data and the standard deviation of the remain *h-geoid_seg* re-computed from the first-pass edited data. The editing process is then repeated with any ocean segment from the first for which the magnitude of *h-geoid_seg* minus the re-computed mean(*h-geoid_seg*)> 3 x re-computed std(*h-geoid_seg*) should be eliminated from the data to be input to the gridding process. " 6/30/2020 - Section 3.1.3.1 Added explanation of ATL03 Release 4 changing the geoid provided to a tide-free EGM2008 in place of the mean-tide EGM2008 provided in earlier releases. Also noted ATL03 all will now include the conversion from tide-free to mean-tide geoid. 6/30/2020 Table 2, Inputs from ATL03, Added *geoid* and *geoid_free2mean* conversion factor to table 6/30/2020 Table 6 Outputs and Table 9, added specification that *earth_tide_seg* would be in the tide-free system		

	6/30/2020 Section 4.2.1.2 – Added specification that ATL03 geoid would now be in the tide-free system and require conversion to mean-tide with the provided conversion factor. 6/30/2020 Section 5.3.1 (A) 4)– Added specification that ATL03 geoid would now be in the tide-free system and require conversion to mean-tide with the provided conversion factor. 6/30/2020 Section 5.4.2– Added ***geoid_free2mean*** and ***tide_earth_free2mean*** to list of ancillary variables to be averaged over and ocean segment. 6/30/2020 Table 6 of ATL12 Outputs– Added ***geoid_free2mean_seg*** and ***tide_earth_free2mean_seg*** as segment averages of ***geoid_free2mean*** and ***tide_earth_free2mean*** to list of outputs. Noted that ***tide_earth_free2mean*** was not used in ATL12. 7/14/*2020 (with change of layer_switch default value to 0 on 7/15/2020)* Changed 4.2.1.2 as recommended by David Hancock 7/8/2020 concerning editing for *layer_flag* =1, i.e., 4.2.1.2 1) is now 1. Examine ocean depth and cloud parameters against control parameters to test for suitability for ocean processing. - If control parameter *layer_swtch* is equal to 1, then remove the ATL03 segment ids from ocean processing that are associated with the ATL09 segment ids where *layer_flag* is equal to 1. When *layer_swtch* equals 0 (the default value), the software will ignore the cloud cover in accepting data during coarse surface finding. - Ignore data collected over land or too close to land. If ocean depth, *depth_ocn*, is less than a depth, *depth_shore*, specifying the effective shoreline of water for which ocean processing is appropriate, then proceed to next 14 geo-bin (400-pulse) segment. The default value for control parameter *depth_shore* will be 10 m. 7/14/2020 Changed formatting of lines in Table 7, Input to gridding to “normal” in order for auto formatting of headings of 5.19 not to skip 5.19.3 and jump from 5.1.2 to 5.19.4 7/15/2020, Table 3 inputs from ATL09, added *layer_flag* as Representing presence (1) or absence (0) of significant cloud layers.		

	7/15/2020, Section 5.4.2 Added description of ***layer_flag_seg*** and ***cloudcover_percent_seg***: For each ocean segment, we output ***layer_flag_seg*** as equal to 1 for ocean segments with more than 50% of geosegs having ***layer_flag*** =1, indicating significant cloud layers were present. In addition, we compute ***cloudcover_percent_seg*** for each ocean segment as the percentage of geosegs in the ocean segment with ***layer_flag*** = 1, i.e., ***cloudcover_percent_seg*** = 100*(total number of geosegs used with ***layer_flag***=1)/(total number of geosegs used). 7/24/2020 Made numerous small edits and updates suggested by Patricia Vornberger, but with no change in the processing code 8/4/2020 Made numerous name updates, corrections, and deletions to variables in Tables 2 and 3 in accordance to Patricia's comments and John Robbins' ATBD_Table_2&3_Notes_JM Also substituted *asr_cloud_probability* for *cloud_flag_asr* in text in description of ATL09 processing for the *layer_flag* 8/11/2020 Corrected error that Leeanne spotted in equation 49. Corrected ***dot_m*** to ***dot_avg*** i.e. $c = dot\_avg - (a * grid\_lon + b * grid\_lat)$ 8/18/2020 Modified equation above 5.6.5 to correct placement of parentheses to read: $\boldsymbol{dot\_sigma\_dfwalbm = ((\text{Sum } [dof\_grid * (dot\_sigma\_dfw)^2]_{beams1,3,5}) / dof\_grid\_albm)^{1/2}}$ 8/28/2020 Sect 4.2.1.2 Paragraph 4: Add editing for *quality_ph*=0 i.e., "Select only photon heights for which the signal confidence (*gtxx/heights/signal_conf_ph*) is greater than or equal to 1 and indicator for saturation *gtxx/heights/quality_ph* is zero." - Similarly, added *quality_ph* to input Table 2 and - Sec 5.3.1 A) 3. Now includes "Select only photon heights for which the signal confidence (*gtx/heights/signal_conf_ph*) is greater than or equal to 1 and photon quality (*gxtx/heights/quality_ph*) is equal to zero."		

	8/28/2020 Sect 5.4.2, Added *full_sat_fract* and *near_sat_fract* to list of variables to be simply averaged over an ocean segment. And added the corresponding outputs of averaging, *full_sat_fract_prcnt* and *near_sat_fract_prcnt* to Table 6 9/8/2020 In section 5.6.2 we gave the second paragraph on pre-filtering its own subsection: 5.6.2.1 Pre-grid Filtering – Along-Track 9/8/2020 In Table 2 Inputs from ATL03, eliminated reference to *signal_conf_ph* = -1, -3, -4 and substituted, "With release 4 ***signal_conf_ph*** equal = -2 indicates possible TEP photon" 9/10/2020 In Table 6, ATL12 Outputs, changed *full_sat_fract_prcnt* tp *full_sat_fract_seg* and *near_sat_fract_prcnt* to *near_sat_fract_seg.* In the variable descriptions, "percentage" is changed to "fraction". 9/15/2020 In 4.2.1.2 5) corrected typo EGM208 to EGM2008 and Happy 2nd Birthday ICESat-2 and Happy 100th Birthday W. H. Morison 10/14/2020 - In Table 9 ATL12 outputs per orbit change the name of photon noise rate, *photonns_rate,* to *photon_noise_rate* to avoid confusion with photon rate, *Photon_rate.* 10/14/2020 – In Section 5.3.2 (Q) change the last sentence to: "The surface reflected photon rate per meter, ***photon_rate*** = $\boldsymbol{r_{surf}}$, is equal to ***n_photons*** divided by ***length_seg,*** and the noise photon rate per meter, ***photon_noise_rate*** = $\boldsymbol{r_{noise}}$, is equal to the difference of ***n_ttl_photon*** minus ***n_photons*** divided by ***length_seg***." 11/9/2020 – Section 5.6.3.2.1 regarding equation 48 description, changed "cross product" to "product sum" 1/4/2021 –In Table 6 Changed *Binsize* to *binsize* Deleted mention of *slope_seg* Added ***nharms*** (= 32) to description of output variable *ds_wn*, the vector of wavenumbers Added *last_geoseg* Deleted *segment_id*		

	1/4/2021 – Section 5.4.2 Ancillary Data – added ***last-geoseg*** 1/4/2021 – Deleted nearly all of Section 5.6 ATL 19 Gridded Product in favor of a short paragraph referring the reader to the ATL19 ATBD 1/4/2021 – In Table 6, changed *ds_wn* to *wn* and *ds_a* to *a* and moved to segments group. 1/5/2021- In Table 6 put *da_a* and *ds_wn* back in with corrected descriptions 1/5/2021 Table 6 – Put im alphabetical order wim groups and applied a consistent border scheme. 1/5/2021 – Deleted nearly all of Section 3.6 about plans for ATL 19 Gridded Product in favor of a short paragraph referring the reader to the ATL19 ATBD. 1/15/2021 Section 2.1, paragraph 3 – Updated and corrected the discussion of Tides, Medium-Frequency Changes in Circulation and Aliasing and the fourth paragraph on he Geoid, Narrowing the Window on Photon Returns, and Measuring Mean DOT. 1/15/2021 Section 2.2.3 - Updated previous sentence and the paragraph under ***A Priori*** **Estimation of SSB Using Only Altimeter Data** 1/18/2021 Section 3.1.1.6 – Annotated with output directories and variables e.g., *gtx/ssh_segments/heights/bimd, latbin, lonbin,htybin, xrbin, bin_ssbia, swh.* 1/18/2021 Section 3.1.15 – Annotated with output directories and variables, e.g., (*gtx/ssh_segments/stats/ n_photons;* (*gtx/ssh_segments/heights/y, ymean, yvar, yskew, ykurt, meanoffit2.* 1/18/2021 Added Section 3.1.1.9 **Ocean Segment Statistics** 1/18/2021 Section 4.5 on ATL19 gridded product: Added a reference to ATL19 ATBD and removed a small amount of out-of-date text leftover from previous edits.		

	1/18/2021 Section 5.2.3 Control parameters added: "These control parameters are output in *ancillary_data/ocean* for each granule/file ATL12" 1/18/2021 Section 3.1 near the end of the last paragraph, added: "where a geo-bin is the ATL03 geolocation segment and amounts to about 20-m along-track distance" 1/18/20 Section 5.4.2 at end of the first paragraph, added: "variables are located in */gtx/stats* group and include " 1/18/2021 Section 5.6 at the end corrected ATL-19 to ATL19 1/20/2021 Section 3.1.1.4 – Added an explanation of editing out photons in afterpulses due to saturation as indicate by the quality_ph flag being nonzero. Also added to Section 5.3.6 a reference to 3.1.1.4. Also in Table 2, noted in description for *fpb_param(n)* that it was not used at present and referred to Section 3.1.1.4. In Table 6 descriptions of *fpb_corr* and *fpb_corr_stdev* noted that they are "currently set to INVALID. See 3.1.1.4." 1/21/2021 Section 4.2.1.3 2): Added text describing change to 5.3.2 to change surface finding based on the distribution of photon heights to surface finding based on the photon height anomaly relative to a moving bin average of high confidence photon heights. 1/21/2021 Section 3.1.1.3: Added text describing the latest surface finding including "starting with Release 4, the surface finding, instead of being based on the distribution of photon heights, is based on the distribution of the photon height anomalies relative to a moving 11-point bin average of high-confidence photon heights. This excludes subsurface returns under the crests of surface waves obviating the immediate need for a subsurface return correction. 1/21/2021 Also to 5.6.3 added: "Also, for Release 4, the surface finding, instead of being based on the distribution of photon heights, is based on the distribution of the photon height anomalies relative to a moving 11-point bin average of high-confidence photon heights. This excludes subsurface returns under the crests of surface waves obviating the immediate need		

	for a subsurface return correction." 1/21/2021 Table 6: Specified the place-holder variables *ss_corr* and *ss_corr_stdev* would have values equal to INVALID pending further developments 1/21/2021 Updated Section 5.2.1 summary discussion of the use of quality variables and flags from ATL03, ***quality_ph, gtx/geolocation/podppd_flag***, ***signal_conf_ph,*** and ***layer_flag*** in editing and surface finding 1/22/2021 Changed previous Table 9 to 7 and Table 10 to 8 because previous Tables 7 and 8 were removed with ATL19 to their own ATBD. Also updaste the listd of Figures and Tables 2/23/2021 – In Section 5.2.2 added a paragraph for deriving ice concentration, ***ice_conc***, from NSIDC daily data. Also added ice concentration, ***gtx/ssh_segments/stats/ice_conc.***, to Table 6 ATL12 Outputs. 3/30/2021 – Section 4.2.1.1 Added specification of the data for which podppd_flag value would be processed in Release 5: "As of Release 5, the ***podppd_flag*** will have seven possible values: normal operations vales: 0=NOMINAL; 1=POD_DEGRADE; 2=PPD_DEGRADE; 3=PODPPD_DEGRADE; plus possible calibration maneuver related "CAL" values: 4=CAL_NOMINAL; 5=CAL_POD_DEGRADE; 6=CAL_PPD_DEGRADE; 7=CAL_PODPPD_DEGRADE. For ATL12, we will only use data when the POP/PPD flag indicates nominal normal operations or nominal CAL maneuvers, ***podppd_flag*** equal to 0 or 4. For each ocean segment processed, we will report the the higher of these two ***podppd_flag*** values, 0 or 4, of data used in the ocean segment." 3/30/2021 – Similar to above, for Section 5.2.1 added: "Per Section 4.2.1, for Release 5 with seven possible values of podppd_flag, ATL12 processing should only be done with data for which ***gtx/geolocation/podppd_flag***=0 indicating nominal operations or ***gtx/geolocation/podppd_flag***=4 indicating nominal calibration maneuvers" 3/30/2021 – Table 2 changed the description of ***podppd_flag*** to seven values and added ***podppd_flag_seg*** to Table 6 Outputs.		

	8/25/2021 – Table 7 change *slope_seg* to *p1*. 8/25/2021 -4.3.1 – Remove "However, pending further study, it is TBD whether they will be included in the ATL12 data output" 11/2/2021 Section 4.2.1 – Added text explaining the rationale for expanding the podppd_flag and changes to ocean scan and RTW scan operations to alleviate the Western Pacific ocean scan ATL12 data gap. 12/1/2021 Starting with the ATBD 11/02/2021 version with prior changes accepted we implemented slight changes to Table 6, mostly units, e.g., m in place of meters, for consistency. 12/3/2021 Change ". 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_initial2(ii)*** - the first value of ***trackdist_initial2(ii)***." To ". 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 maximum value of ***trackdist_initial2(ii)*** - the minimum value of ***trackdist_initial2(ii)***." 12/3/2021 Correcting ***trackdist_initial_surf*** to ***trackdist_initial2_surf*** as follows: In 5.3.2 (P) we define ***trackdist_initial2_surf***, as the along-track distances of the final selected surface photons: Identify the corresponding track distances, ***trackdist_initial2_surf***, in ***trackdist2_initial*** at the indices in ***ii***, [i.e., in Matlab code ***trackdist_initial2_surf*** = ***trackdist2_initial(ii)***]. But as of 12/3/20212 David Hancock and Suzanne Dickinson discovered that the 2 had been dropped from subsequent usage of the along-track distance, i.e., ***trackdist_initial_surf*** was used instead.		

	Corrections are made at the following subsequent places in the ATBD: In 5.3.3.1 : 1. Add ***meanoffit2*** to ***ht_initial2_surf***. 2. Convert ***trackdist_initial2_surf*** to meters if necessary and subtract the minimum of ***trackdist_initial2_surf*** to obtain distance from start of the ocean segment. 3. Sort height data in ascending order of along-track distance, ***trackdist_initial2_surf,*** to allow calculation of photon return rate per meter 4. Define ***x*** as the sorted ***trackdist_initial2_surf*** and define ***hty*** the associated heights from ***ht_initial2_surf***. and Estimating sea state bias over ocean 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 ***ht_initial2_surf*** and ***trackdist_initial2_surf*** along each ocean segment found during the coarse signal finding step. Section **5.3.1** Add mean from detrending used after first cut surface finding The means from the first cut of surface finding are computed from photon heights over surface waves and are therefore contaminated by the correlation of photon return rate with height of the surface over surface waves. If these are not added back into the height record, that contamination by SSB will be lost. Therefore, we add ***meanoffit2***, the mean of ***P0*** + ***P1*** x ***trackdist_initial2_surf*** (the along-track position of 2nd iteration surface photons), back into the photon height record, ***hty*** =***ht_initial2_surf*** +***meanoffit2*** In 5.4.3.2 (I) - The first moment or mean of the mixture, $\mu_{mix} = m_1\mu_1 + m_2\mu_2$ is the mean of the detrended surface photon heights relative to the geoid with the mean of the detrending process, ***meanoffit2*** (equal the mean of ***P0***+***P1*********trackdist_initial2_surf***), added back in prior t Gaussian mixture determination. Thus, it represents the mean DO uncorrected for sea state bias, along the ocean segment. Therefore mean sea surface height, ***SSH***, for the segment to be output in ATL		

	equal to the mean of the mixture, μ_{mix} , plus ***meangdht*** equal to the mean of geoid height, ***geoid***, over the segment: 10/4/2022 In section 5.3.3.1 added computation of standard deviation of photon heights in 10-m along-track bins by inserting a new line f) Set *htybin_std* to standard deviation of ***hty*** in each 10-m bin, If the number of photons in each 10-m, ***nbind*** <2, set *htybin_std* to NaN And modified line 14 to include the standard deviation as an output 14) Save vector arrays ***xbin***, ***xbind***, ***htybin***, ***stdhtybin***, and ***xrbin*** for segment output. And added *stdhybin* to Table 6 10/11/2022 Added section 4.2.3 "First Photon Bias Estimation" with introductory sentence, more to be added. 11/07/2022 Inserted a new 5.3.3.2 First Photon Bias Determination taken from ATL03 and interpreted by John Robbins. Previous 5.3.3.2 Harmonic Analysis is now 5.3.3.3. 11/10/2022 Added first photon bias parameters *fpb_10m, fpb_corr, and fpb_corr_std* to Table 6 Outputs under */gtx/ssh_segments/heights* 12/13/2022 – corrected 3 small typos in 3.1.1.6 per David Hancock email 12/12/2022 1/22/2023 Removed from Section 5.3.2 (G): "The ASAS 5.1 finds the limits of the histogram above (***jlow***) and below (***jhigh***) for which, searching from the index of the maximum in the 20-cm smoothed histogram ***smoothN***, the value in ***smoothN*** is greater than median of ***smoothN***. 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 the median of ***smoothN***.. At this point ***jlow*** = ***i***+1. Increment the index, ***i***, upward from ***imax*** until ***smoothN(i)*** is greater than the median of ***smoothN***. At this point ***jhigh*** = ***i***-1. As of 6/4/2019, the developmental code"		

	And removed from Section 5.3.2 (O): "The ASAS 5.1 finds the limits of the histogram above (***jlow***) and below (***jhigh***) for which, searching from the index of the maximum in the 20-cm smoothed histogram ***smoothN***, the value in ***smoothN*** is greater than median of ***smoothN***. 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 the median of ***smoothN***.. At this point ***jlow*** = ***i***+1. Increment the index, ***i***, upnward from ***imax*** until ***smoothN(i)*** is greater than the median of ***smoothN***. At this point ***jhigh*** = ***i***-1. As of 6/4/2019, the developmental code" 1/22/2023 Changed Section 5.3.3., step 2 to add saving and outputting the ocean segment start position, ***xbind_first_dist_x :*** "2. Convert ***trackdist_initial2_surf*** to meters if necessary and subtract the minimum of ***trackdist_initial2_surf*** to obtain distance from start of the ocean segment. Save and output as ***xbind_first_dist_x*** this start position equal to the minimum of ***trackdist_initial2_surf*** " And added ***xbind_first_dist_x*** to Output Table 6. 1/22/2023 Changed the name to in Table 6 and added its description and the end of section 5.3.2: "Compute ***seg_mean_dist_x*** as the ocean segment average of the along-track distance from the equator crossing to the start of the 20-m geolocation segments included in the ocean segment" 1/24/2023 In Section 5.3.2 (H) corrected: ***ht_initial2=ht_initial - P1*** x ***trackdist_initial+ P0***. To: ***ht_initial2=ht_initial*** - (***P1*** x ***trackdist_initial+ P0***). 1/24/2023 In Section 3.1.1.4 Corrected end of the pargraph to say: photons in these afterpulses are edited out by considering only photons with the index of photon height quality from ATL03 (*quality_ph*) equal to 0. Starting with Release 5 we deleted all photons coming from laser pulses that included any photons		

	with a quality index indicative of full saturation. For Release 7, we derive a first-photon-bias correction for 10-m bins within ocean segments and an ocean segment average first-photon-bias correction. 1/24/2023 – In Section 4.2.1.2, 4, update the use of *quality_ph_* flag to : Select only photon heights for which the signal confidence (*gtxx/heights/signal_conf_ph*) is greater than or equal to 1. For Release 7 onward, use only photon heights for which the indicator for saturation *gtxx/heights/quality_ph* is 0 indicating no saturation, or 10 indicating partial saturation for which a first-photon-bias correction will be computed. 1/24/2023 - In Section 4.2.1.3, 2. And in the Table of TBs we removed the TBD regarding for *Th_Nc_f* = 1.5: We have found a good value for *Th_Nc_f* is 1.5 in practice with Release 001 ATLAS data and results in greatly improved rejection of subsurface returns. 2/7/2023 - To 4.2.1.2 Added "For Release 7 onward, use only photon heights for which the indicator for saturation *gtxx/heights/quality_ph* is 0 indicating no saturation, or 10 indicating partial saturation for which a first-photon-bias correction will be computed" 2/7/2023 – To 5.2.1 added: "For Release 7, we will use only photons with *quality _ph* equal to 0 indicating no saturation, or 10 indicating partial saturation for which a first-photon-bias correction will be computed." 2/7/2023 - To Table 2 added "*quality_ph*=10 = partial saturation for which a first-photon-bias correction will be computed" 2/7/2023 – To Section 5.3.1 (A, 3) added: "Starting with Release 7 we will also accept photons for which *quality_ph* =10 indicating partial saturation and for which a first-photon-bias correction will be derived." 2/14/2023 – Modified section 4.2.5 to include using ATL07 to identify ATL12 10-m bins as bright leads characterized by the ATL07 *SSH_flag=1.*		

	3/6/2023 –Modified Section 5.2.3 to change to using AMSR2 for ice concentration consistent with ATL07/10 And In Section 5.2.4 Control Parameters added: Sea ice concentration- Two control parameters ***sic_filter_ssmi*** (default 0, not used) ***sic_filter_amsr2*** (default 1, use) Table 6 added ***ice_conc_SSMI*** and explanation of switch to AMSR2 for ***ice_conc***. 3/7/2023 - In Section 4.2.1.2, Paragraph 3 added: "As of this writing, March2023, that there are a few ocean locations for which there are no modeled tides (*tide_ocean* = INVALID). Select only photons which have a valid *tide_ocean*." And similarly In Section 5.3.1 Paragraph (A) 2, added "As of this writing, March2023, that there are a few ocean locations for which there are no modeled tides (*tide_ocean* = INVALID). Select only photons which have a valid *tide_ocean".* 3/7/2023 – Modified Section 4.2.5 SSH Determination in Ice-Covered Waters consistent with email exchanges with David Hancock. Added needed ATL07 variables to Input Table 3 Added Section 5.3.3.4 to compute DOT for 10-m bins and averaged in ocean segments in patches of bright leads identified in ATL07. The new variables are: the 10-m bin, ATL07 qualified bright lead flag, *xbin_atl07_dot*, equal to the DOT from the corresponding bright lead sea ice segments. the ocean segment average, *h_ATL07_ice_free*, of *xbin_atl07_dot* the ocean segment average, *h_ice_free*, of the ATL07 qualified 10-m bin DOT heights. Added *xbin_atl07_dot*, *h_ATL07_ice_free*, and *h_ice_free* to Output Table 6. 3/10/2023 – In Section 5.3.3.1 added saving ***bin_ssbias*** and ***bin_slopebias*** to paragraph 10 and end of section. 3/10/2023 – In Section 5.3.3.2 added saving and outputting ***fpb_corr*** and ***fpb_corr_stdev*** and the vector array ***fpb_10m*** for each ocean segment.		

	3/13/2023 In Section 5.3.4 (4) Changed INVALID_14B to INVALID and 999 to INVALID 3/12/2023 Added ***bin_slopebias*** to Output Table 6 and made inconsequential edits. 6/20/2023 Made numerous inconsequential edits for spelling and punctuation. 6/20/2023 Added to 4.2.5:"We will also identify all the ATL07 sea ice segments with *ssh_flag*= 1 (bright leads) in an ATL12 ocean segment and compute *h_atl07_inatl12oc*, the average DOT from all these ATL07 *ssh_flag* =1 qualified sea ice segments regardless of whether they have a corresponding ATL12 10-m bin. This will be important if there are only leads smaller than 10-m in the span of the ocean segment. " 6/20/2023 Added to 5.3.3.4 : " 6) We also identify all the ATL07 sea ice segments with *ssh_flag*= 1 (bright leads) in an ATL12 ocean segment and compute average, *h_atl07_inatl12oc* , of DOT (=*heights/height_segment_height* plus *geophysical/height_segment_mss* minus *height_segment_geoid* from ATL07) from all these ATL07 *ssh_flag* =1 qualified sea ice segments regardless of whether they have a corresponding ATL12 10-m bin. This will be important if there are only leads smaller than 10-m in the span of the ocean segment. " 6/20/2023 Added to Table 6 Outputs *h_atl07_inatl12oc* , the averaged DOT from all ATL07 *ssh_flag* =1 qualified sea ice segments in an ATL12 ocean segment with or without corresponding ATL12 10-m bins. 8/17/2023 Added to 5.3.3.1 computation of the correlation, ***magslopeXreturn***, of photon return rate with the absolute value of surface slope: d) Comput***magslopeXreturn*** as the product of the absolute value of the detrended slope, ***slopebinfit***, and the detrended binned data rate, ***xrbin***. e) Compute the covariance of the photon rate and the slope, ***binCOVmagslope_Xr***, as the sum of the ***magslopeXreturn*** divided by the number of 10-m bins with a valid ***slopebinfit***.		

	f) Compute the sea state slope magnitude bias, ***bin_magslopebias***, as the covariance of the data rate and the slope, ***binCOVmagslope_Xr***, divided by the mean of the data rate. 11. Save and output ***bin_ssbias, bin_slopebias, and bin_magslopebias,*** and vector arrays ***xbin***, ***xbind***, ***htybin***, ***htybin_std***, and ***xrbin*** for segment output. And at the end 53.3.1 augment the description bin_ssbias according to: Compute photon rate slope correlation. 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. If there are more than one 10-m bin in an ocean segment with a valid ***slopebinfit***, compute ***slopeXreturn*** as the product of the detrended slope, ***slopebinfit***, and the detrended binned data rate, ***xrbin***. Similarly compute ***magslopeXreturn*** as the product of the magnitude of the detrended slope, ***slopebinfit***, and the detrended binned data rate, ***xrbin***. Then compute the covariance of the photon rate and the slope, ***binCOVslope_Xr***, as the sum of the ***slopeXreturn*** divided by the number of 10-m bins with a valid ***slopebinfit*** and the covariance of the photon rate and the magnitude of slope, ***binCOVmagslope_Xr***, as the sum of the ***magslopeXreturn*** divided by the number of 10-m bins with a valid ***slopebinfit***. Compute the sea slope bias, ***bin_slopebias***, as the covariance of the data rate and the slope, ***binCOVslope_Xr***, divided by the mean of the data rate. Similarly compute the sea slope magnitude bias, ***bin_magslopebias***, as the covariance of the data rate and magnitude of the slope, ***binCOVmagslope_Xr***, divided by the mean of the data rate. Save variables ***bin_ssbias***, ***bin_slopebias,*** and ***bin_magslopebias*** vector arrays ***xbin, xbind***, ***htybin***, ***htybin_std,*** and ***xrbin*** for segment output. 8/17/2023 Added ***bin_magslopebias*** to Output Table 6. 8/21/2023 Corrected typo …maglope…corrected to magslope everywhere. Also change headers to read Rel 007 instead of Rel 006. Non-substantive changes		

	8/22/2023 In response to a problem Leeanne found in Table 6 and to be in a style consistent with slope bias, binCOVslope_Xr, changed binCOV_magslope_Xr to binCOVmagslope_Xr globally. 8/22/2023 Changes to allow the minimum number of weak beam photons defining an ocean segment to be ¼ the minimum number of strong beam photons. In 4.2.1.2 Paragraph 9 now reads: "If the total number of candidate signal photons for a strong beam is greater than or equal to a minimum number, *Th_Ps* (e.g., 8,000), of photons or S_{egmax} (e.g., 25) 14 geo-bin (400-pulse) segments have been collected this defines an ocean segment. For weak beams, and ocean segment is defined when the total number of candidate signal photons is greater than or equal to ¼ *Th_Ps* (e.g., 2,000) photons or S_{egmax} (e.g., 25) 14 geo-bin (400-pulse) segments have been collected. Once the ocean segment is defined, we 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* for a strong beam (1/4 *Th_Ps* for a weak beam), repeat steps 4.2.1.2-2 to 4.2.1.2-5 above" And Section 5.3.1, part G now reads: (G) 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*** (8,000) photons for a strong beam or ***¼ Th_Ps*** (2,000) photons for a weak beam, or if the number, ***Nseg400***, of 14 geo-bin (400-pulse) segments collected reaches S_{egmax} (e.g., 25 equivalent to ~7 km)), go on to 5.3.2 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*** for a strong beam or ***¼ Th_Ps*** for a weak beam, intake another 14 geo-bin (400-pulse) segment and repeat steps 5.3.1 A-G above. 8/29/2023 In table 6 clarified the description of *bin_slopebias* as: Indication of the dependence of reflectance on 10-m scale sea surface slope equal to covariance of the data rate and the surface slope, ***binCOVslope_Xr***, divided by the mean of the data rate. (5.3.3.1) And bin_magslopebias as: Indication of the dependence of reflectance on the magnitude of 10-n scale sea surface slope equal to covariance of the data rate		

	and magnitude of surface slope, ***binCOVmagslope_Xr***, divided by the mean of the data rate. (5.3.3.1 In Section 5.3.3.1 9c) and 9f) removed "state" from "seat state slope" for consistency with other text. 8/30/2023 To section 4.2.5 added correction of ATL07 DOT values to be relative to the geophysical/*ht_seg_dac* correction instead of the *geophysical/ht_seg_dynib* correction and add the check of requirimg ATL12 10-m bin photon rate to be between 16 and 20 photons per meter (strong beam) and between 4 and 5 (weak beam). The text now reads: Heights in the 10-m bins represent DOT, so for the sake of comparison, we flag the ATL12 10-m bins so qualified by ATL07 as open water with flag *xbin_atl07_dot* = ATL07 data *heights/height_segment_height* + *geophysical/height_segment_mss* - *geophysical/height_segment_geoid* (assuming this geoid is a mean tide geoid) + *geophysical/ht_seg_dynib* - *geophysical/ht_seg_dac*. Thus, we output these 10-m bin flags as ATL07 DOT values. Note that we will limit candidate bins for sea surface height determination to those 10-m bins for which the photon return rate, *xrbin*, is between 16 and 20 photons per meter (strong beams) or 4 to 5 photons per meter (weak beams) corresponding to quasi-specular but not fully saturated returns. This corresponds to one of the criteria for ATL07 segments to have *ssh_flag*=1, indicating bright leads. 8/30/2023 Similarly modify Section 5.3.3.2 to include converting the ATL07 IB correction to the DAC correction as used in ATL12 and imposing the photon rate criteria for selection of 10-m bins for corresponding ATL07 segments with ssh_flag: "3) By comparing the maps of 10-m bin boundaries and bright-lead sea ice segment boundaries, determine which 10-m bins also represent bright lead surfaces. Both boundaries of the 10-m bin should be in bright-lead segments. Note that we will limit candidate10-m bins for sea surface height determination to those bins for which the photon return rate, *xrbin*, is between 16 and 20 photons per meter (strong beams) or 4 to 5 photons per meter (weak beams) corresponding to quasi-specular but not fully saturated returns. This corresponds to one of the criteria for ATL07 segments to have *ssh_flag*=1, indicating bright leads. 4) Flag the so identified bright-lead 10-m bins with *xbin_atl07_dot* equal to the average of the DOT in the sea ice		

	segments bounding the 10-m bin. If the 10-m bin is in common with only one, the average DOT will simply be the DOT of that sea ice segment. The 10-m bin *xbin_atl07_dot* is set to INVALID where the ATL12 ice_conc < 15%. Where the ATL12 ice_conc >= 15%, if the beginning and end of a 10-m bin are within one ATL07 segment with *height_segment_ssh* =1, then *xbin_atl07_dot* = *heights/height_segment_height* plus *geophysical/height_segment_mss* minus *height_segment_geoid*) + *geophysical/ht_seg_dynib - geophysical/ht_seg_dac* from ATL07. If the beginning and end of the 10-m bin are in two different qualified ATL07 segments *xbin_atl07_dot* will equal the average of *heights/height_segment_height* plus *geophysical/height_segment_mss* minus *height_segment_geoid* from the two segments. If both ends of the 10-m segment are not in bright lead ATL07 segments, xbin_*atl07_dot* = INVALID." 9/5/2023 Added to section 4.2.5 in the first paragraph the requirement the ATL07 *height_segment/fit_quality_flag* <5 be added to the criteria for a valid bright open water surface height: "Furthermore, we can align ATL12 and simultaneous ATL07 and determine which ATL12 10-m bins are judged to be open water bins according to the ATL07 *height _segment_ssh_flag* identifying ATL07 height segments as bright, open-water surfaces. ATL10 uses an additional parameter to characterize the quality of fit of photon heights to the idealized ATLO7 photon height distribution, *height_segment_fit_quality_flag,* and requires *fit_quality_flag* to be 1, 2, 3, or 4 for use of the segment height for sea surface height. For ATL12 we will use the same criteria, ATL07 *ssh_flag*=1 and *fit_quality* = 1, 2, 3, or 4 to consider the ATL07 segment height suitable for SSH or DOT determination." 9/5/2023 Added to Section 3.2.4.4.2 Table 3, parameters input from other high level products, *fit_quality_flag* from ATL07. 9/5/2023 Changed Section 5.3.3.4 (2) to include consideration of *fit_quality_flag*: "2) For ***ice_conc*** > 15%, from ATL07, use ***seg_dist_x*** and ***height_segment_length_seg*** to establish the along-track distance boundaries of the ATL07 sea ice segments and use the ATL07 *ssh_flag* = 1 and the *fit_quality_flag* = 1, 2, 3, or 4 to identify which sea ice segments are the sea ice segments are candidate, bright-lead surfaces."		

	9/25/2023 Changed Sections 4.2.5 and 5.3.3.4 3) to allow a wider range of photon return rates thus: "Note that we will limit candidate bins for sea surface height determination to those 10-m bins for which the photon return rate, *xrbin*, is between 4 and 25 photons per meter (strong beams) or 1 to 7 photons per meter (weak beams) corresponding to quasi-specular but not fully saturated returns. This is a wider range than the criteria for ATL07 segments to have *ssh_flag*=1, indicating bright leads" 9/25/2023 Added to Section 5.3.3.4 paragraphs 7 and 8 defining ***ice_conc_12icefreebins*** and ***ice_conc_07icefreesegs***: 7) To aid in the interpretation of ***h_ice_free*** and ***h_atl07_ice_free***, we will also calculate and output an ice concentration of bright leads for the 10-m bin data as ***ice_conc_12icefreebins*** equal to the square of the complement of the fraction of 10-m bins classified as covering bright leads by ATL07, ***ssh_flag***=1. 8) Similarly, To aid in the interpretation of ***h_ice_free*** and ***h_atl07_ice_free***, we will also calculate and output an ice concentration of bright leads for the ATL07sea ice segment as ***ice_conc_07icefreesegs*** equal to the square of the complement of the fraction of ATL07 sea ice segments in the AL12 ocean segment classified as covering bright leads by ATL07, ***ssh_flag***=1 and Added ***ice_conc_12icefreebins*** and ***ice_conc_07icefreesegs*** to Table 6 Outputs 9/28/2023 Added new Section 5.3.3.4 (6) for calculation of *h_ice_free_uncrtn*. (on 10/1/2023 added *h_ice_free_uncrtn* to Table 6) 10/13/2023 Corrected typo in equation for ***h_ice_free_uncrtn*** to $h\_ice\_free\_uncrtn = \left((hicefree\_std\_inst)^2 + (hicefree\_std\_nat)^2 \right)^{1/2}$ 10/16/2023 Added to ***hicefree_std_inst*** and ***hicefree_std_nat*** to Table 6 10/24/2023 Deleted computation of ***ice_conc_07icefreesegs*** in Section 5.3.3.4 (9) and in Table 6. This computation would not be possible as the ATL07 ice segments overlap. 10/24/2023 Clarified 5.3.3.4 (8):		

	To aid in the interpretation of ***h_ice_free*** and ***h_atl07_ice_free***, we will also calculate and output an ice concentration based on the fractional area of bright-lead 10-m bins in an ocean segment: ***ice_conc_12icefreebins*** will be equal to complement (difference from 1) of the square of the fraction relative to the total number of 10-m bins in the ocean segment of 10-m bins classified as covering bright leads by ATL07, ***ssh_flag.*** Express as a percentage. Also modified description in Table 6 to be consistent with 5.3.3.4 (8). 10/30/2023 Corrected and clarified 5.3.3.4 (6) with: "The first, ***hicefree_std_inst,*** which over the smooth surface of a lead is likely instrumental and is manifest in the standard deviations of surface photon heights bright lead 10-m bin in the ocean segments. Let σ_i =***htybin_std*** in the i^{th} bright lead 10-m bin of thje ocean segment and let there be K bright lead-qualified 10-m bins in the ocean segment with n_i =10****xrbin*** photons in the i^{th} bin. Then the average standard deviation in the K bins, $\bar{\sigma}$, is 12/17/2023 Add 5.3.3.4 New paragraph 8 for computation of ***h_atl07_ice_free_uncrtn*** and ***h_atl07_inatl12oc_uncrtn*** and related natural and instrumental uncertainty component 12/17/2023 Add ***height_segment_surface_error_est*** to Table 3 inputs from ATL07. 12/17/2023 – Added to Table 6 Outputs: Uncertainties for ATL07 derived heights: ***h_atl07_ice_free_uncrtn, h_atl07_inatl12oc_uncrtn, h07icefree_std_inst , h07icefree_std_nat, h7in12icefree_std_inst, h7in12icefree_std_nat*** 12/18/2023 - Changed 5.3.3.4 (8) to using n_i = ***n_photons_used*** in each ATL07 sea ice segment in the computation of ***h07icefree_std_inst*** and ***h7in12icefree_std_inst***. Also added ***n_photons_used*** to Input Table 3. 12/19/2023 – In Table 6 deleted *fpb_corr* and *fpb_corr_std* from the stats group. They are in the heights group. Also deleted unused subsurface scattering correction, *ss_corr* and *ss_corr_stdev.* Our surface finding for the deep ocean does a good job of eliminating subsurface scattering. 12/30/2023 Corrected the limit of summation for ***h7in12icefree_std_inst,*** i.e., to K_{all} not K_{10}		

	h7in12icefree_std_inst = $\bar{\sigma}_{all}\left(\sum_{i=1}^{K_{all}} n_i/_2\right)^{-1/2}$ 1/5/2024 - Expanded Section 4.2.5 to better describe adjustments to ATL07 heights to convert them to DOT. Also added required variables from ATL07 to Table 3, Inputs. 1/6/2024 – In Section 5.3.3.4 for *h_atl07_ice_free_uncrtn* and *h_atl07_inatl12oc_uncrtn*, per email from Jesse Wimert, changed definition of σ_i to σ_i = ***height_segment_surface_error_est*** x (***n_photons_used***)$^{1/2}$ 1/8/2024 – Changed Section 5.3.3.4 for *h_atl07_ice_free_uncrtn* and *h_atl07_inatl12oc_uncrtn*, per email from Jesse Wimert, changed definition of σ_i to σ_i = ***hist_w*** and added ***hist_w*** to Table 3. 2/12-15/2024 Corrected 5.3.3.4 equations for ***hicefree_std_nat***, ***h07icefree_std_nat***, and ***h07in12icefree_std_nat***, to be standard deviation divided by square root of degrees of freedom. Also made editorial changes to text for clarity 2/15/2024 Change ***hicefree_std_nat***, ***h07icefree_std_nat***, and ***h07in12icefree_std_nat, hicefree_std_inst***, ***h07icefree_std_inst***, and ***h07in12icefree_std_inst*** to ***hicefree_uncrtn_nat***, ***h07icefree_uncrtn_nat***, and ***h07in12icefree_uncrtn_nat, hicefree_uncrtn_inst***, ***h07icefree_uncrtn_inst***, and ***h07in12icefree_uncrtn_inst*** 2/16/2024 In Table 6, corrected and improved descriptions of ***hicefree_uncrtn_nat***, ***h07icefree_uncrtn_nat***, and ***h07in12icefree_uncrtn_nat, hicefree_uncrtn_inst***, ***h07icefree_uncrtn_inst***, and ***h07in12icefree_uncrtn_inst*** 2/26/2024 Corrected Section 5.3.3.4 to apply correction for first photon bias (-***fpb_10m***) in the definition of: paragraph (4) for ***h_ice_free*** to average of ***htybin-fpb_10m*** for bright lead 10-m bins, and paragraph (6) for h_i to be equal to ***htybin-fpb_10m*** for bright lead 10-m bins. Also mention correction for first photon bias in Table 6 definition for ***hicefree_uncrtn_nat***		

	3/13/2024 - Added note about instrumental uncertainties for ***h07icefree_uncrtn_inst*** and ***h7in12icefree_uncrtn_inst*** "IMPORTANT NOTE: If and only if in the future, ATL07 sea ice segments are changed to non-overlapping, do not derate the each photon height by 2 so these uncertainties would be *h07icefree_uncrtn_inst* = $\bar{\bar{\sigma}}_{10}\left(\sum_{i=1}^{K_{10}} n_i\right)^{-1/2}$ and *h7in12icefree_uncrtn_inst* = $\bar{\sigma}_{all}\left(\sum_{i=1}^{K_{all}} n_i\right)^{-1/2}$." 3/18/2024 Modified Section 5.3.3.4 Paragraph 4 to add ocean segment slop back into htybin: "…. For each ocean segment, ***h_ice_free*** will be set to the average of the ***htybin-fpb_10m-meanoffit2+ (p0 +p1 x xbind)*** of ***ssh_flag****=1,* bright lead qualified ATL12 10-m bins and output. Similarly for each ocean segment, ***h_ATL07_ice_free*** will be set to the average of all the ***atl07_dot*** for all the sea ice segments coincident with bright lead qualified 10-m bins will be saved and output. Note 1: For the open ocean we added *meanoffit2* to the 10-m bin photon heights to allow computation of SSB without considering the trend in reflectance over a typical 7 km ocean segment. Typically, in open water the average slope is small ($<10^{-5}$), but with sea ice with 40-cm ice freeboard at one end of a 400-m ocean segment the slope, ***p1***, can be ten times as great (10^{-4}) and without the linear fit correction, ***(p0 +p1 x xbind)***, added back in and ***meanoffit2*** correction removed, the bias in 10-m bin htybin can be 40 cm. Note 2: Since the sea ice segments are short (extending about 150 photons) and often overlap by 50%, ***atl07_dot*** values will be found applicable for two or more 10-m bins…. " 8/12/2024 Applied changes as below as recommending by John Robbins in anticipation of the switch of tide model from GOT4.8 to FES2014b: • Section 3.1.3.4, under Ocean load tides: change GOT4.8 to FES2014b. And change -6 to 0 cm to -6 to +6 cm. • Section 3.1.3.4, under Ocean tides: replace whole paragraph with: Ocean tides (*tide_ocean*): Short period ocean tides (including diurnal and semi-diurnal, ±4 m) are computed from the		

	FES2014b extrapolated tide model. Details in ATL03 ATBD, section 6.3.1. • Section 3.1.4.3, under Equilibrium tides: change whole paragraph with [Richard Ray indicates that this set of tides is no longer in perfect equilibrium, but very close. 1 cm at the poles.]: Long-period tides (*tide_equilibrium*): Long-period tide computed from the FES2014b extrapolated tide model The S_a (solar annual) constituent was omitted due to it being deemed erroneous. Details in ATL03 ATBD, section 6.3.1. • Section 4.2.1.2, bullet 2: strike the whole sentence that starts with "As of this writing…" • Table 2: In the entry for tide_equlibrium, delete the word "Equilibrium" and make the description be: "Long-period tide, no S_a constituent" • Table 2: In the entry for tide_load, change text to: Crustal deformation due to Ocean Loading (-6 to +6 cm) • Table 2: In the entry for tide_ocean, change text to: Short-period ocean tides (diurnal and semi-diurnal, ±4m) • Section 5.3.1, (A), paragraph 2: strike the whole sentence that starts with "As of this writing…" • In Table 6 (Output Parameters): under the tide_equilibrium_seg entry, change text to: Ocean segment average of long-period ocean tides • In Table 6 (Output Parameters): under the tide_load_seg entry, change -6 to 0 cm to -6 to +6 cm. • In Table 6 (Output Parameters): under the tide_ocean_seg entry, change text to: Ocean segment average of short-period ocean tides (diurnal and semi-diurnal) • Please apply the same instructions for Table 6 (above) to Table 7 10/15/2024 Changed Sections 5.2.1 and 5.2.2 bathymetry data from GEBCO 2014 to GEBCO 2022		

List of TBDs/TBRs

Item No.	Location	Summary	Ind./Org.	Due Date
1	3.1, p52	The possible methods of distinguishing open ocean from sea ice and land and setting the ocean segment lengths in these mixed regions are TBD.		
2	3.2, p54	Applying a TBD first-photon-bias correction where photon rates are high		
3	4.2.1.3, p65	*Th_Nc_f* is TBD, but we have used *Th_Nc_f* = 1.5		
4	5.3.5 P107	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	Appendix A	Various parameters in other ICESat-2 data products		

Scope

The scope of the delivery of this edition of the **ICESat-2 Algorithm Theoretical Basis Document for Ocean Surface Height** points to the following ICESat-2 Project element targets:

L3A OC ATBD Version Identification:	Version 5.1
L3A OC ATBD Version Date:	*July 2, 2024*
NSIDC ATL12 Data Archival Version:	V7 (V07)

Table of Contents

Abstract ii
CM Foreword iii
Preface iv
Review/Approval Page v
Change History Log vi
List of TBDs/TBRs xxxvi
Table of Contents 1
List of Figures 4
List of Tables 6
1.0 INTRODUCTION 8
2.0 OVERVIEW AND BACKGROUND INFORMATION 11
2.1 Open Ocean Background 12
2.2 The Importance of Waves 14
2.2.1 Waves and Reflectance 14
2.2.2 Waves and Sea State Bias 16
2.2.3 ICESat-2 Height Statistics and Sea State Bias 21
3.0 Open Ocean Products 26
3.1 Open Ocean Surface Height (ATL12/L3A) 26
3.1.1 Height Segment Parameters (output group *gtx/ssh_segments/heights/*) 27
3.1.2 Input from IS-2 Products 29
3.1.3 Corrections to height (based on external inputs) 29
3.2 Gridded Sea Surface Height (ATL19/ L3B) 31
4.0 ALGORITHM THEORY 32
4.1 Introduction 32
4.2 ATL12: Finding the surface-reflected photon heights in the photon cloud 32
4.2.1 Selection of Signal Photon Heights 34
4.2.2 A Priori Sea State Bias Estimation and Significant Wave Height 40
4.2.3 First Photon Bias Estimation 40
4.2.4 Harmonics, Correlations Length, and Uncertainty 40
4.2.5 SSH Determination in Ice-Covered Waters 41

4.3 ATL12: Correction and Interpretation of the Surface Height Distribution 43
4.3.1 Separating the Uncertainty due to Instrument impulse response 43
4.3.2 Characterizing the Random Sea Surface 45
4.3.3 Expectation Maximization 47
4.4 Calculation of Uncertainty 47
4.5 ATL19: Gridding the DOT and SSH 48
5.0 Algorithm Implementation 49
5.1 Outline of Procedure 49
5.2 Input Parameters 49
5.2.1 Parameters Required from ICESat-2 Data 49
5.2.2 Parameters Required from Other Sources 53
5.2.3 Control Parameters for ATL12 Processing 53
5.3 Processing Procedure 54
5.3.1 Coarse Surface Finding and Setting of Segment Length 54
5.3.2 Processing Procedure for Classifying Ocean Surface Photons, Detrending, and Generating a Refined Histogram of Sea Surface Heights 57
5.3.3 Processing to Characterize Long Wavelength Waves, Dependence of Sample Rate on Long Wave Displacement, *A Priori* Sea State Bias Estimate, First Photon Bias, and Sea Surface Height in Ice-Covered Waters 63
5.3.4 Processing Procedure for Correction and Interpretation of the Surface Height Distribution 74
5.3.5 Applying *a priori* SSB Estimate 89
5.3.6 Expected Uncertainties in Means of Sea Surface Height 90
5.4 Ancillary Information 94
5.4.1 Solar Background Photon Rate and Apparent Surface Reflectance (ASR) 94
5.4.2 Additional Ancillary Data 94
5.5 Output Parameters 96
5.6 Gridding DOT for ATL19 104
5.7 Synthetic Test Data 105
5.8 Numerical Computation Considerations 111
5.9 Programmer/Procedural Considerations 111
5.10 Calibration and Validation 111

6.0 Browse Products 112
6.1 Data Quality Monitoring 112
6.1.1 Line plots (each strong beam) 112
6.1.2 Histograms (each strong beam) 112
7.0 Data Quality 113
7.1 Statistics 113
7.1.1 Per orbit statistics 113
8.0 TEST DATA 116
8.1 In Situ Data Sets 116
8.2 Simulated Test Data 116
9.0 CONSTRAINTS, LIMITATIONS, AND ASSUMPTIONS 117
9.1 Constraints 117
9.2 Limitations 117
9.3 Assumptions 118
10.0 References 119
ACRONYMS 120
APPENDIX A: ICESat-2 Data Products 121
APPENDIX B: Sea State Bias Computations for a Photon Counting Lidar 125
APPENDIX C: Expectation-Maximization (EM) Procedure 130
APPENDIX D: Fitting a Plane to Spatially Distributed Data 131
APPENDIX E: Hierarchy of ATL12 and ATL10 Variables 133

List of Figures

Figure Page
Figure 1 - Characteristics of an idealized distribution ……11
Figure 2. Geodetic height……13
Figure 3. Reflectance as modeled.……15
Figure 4. Typical Gram-Charlier distribution …… 16
Figure 5. Trochoidal shape of surface waves …… 17
Figure 6. Short wave MSS & standard deviation MSS……18
Figure 7. Simulated surface displacement and MSS……19
Figure 8. Height due to long-waves & MSS……19
Figure 9. Height due to long-waves & probability ……20
Figure 10. Surface return heights from MABEL ……24
Figure 11. ATL 12 processing block diagram……33
Figure 12 Coarse, smoothed, and fine histograms ……37
Figure 13. Block diagram for the ATL19 gridding…… 47
Figure 14. Coarse, smoothed, final Mabel histograms ……60
Figure 15. Instrument impulse response histogram ……74
Figure 16. Probability density functions of the synthetic SSH ……76
Figure 17. Synthesized received probability density functions ……77
Figure 18. Single-sided spectra of the received PDFs ……80
Figure 19. Single-sided spectrum of impulse response ……81
Figure 20. Single-sided amplitude spectra of the Wiener filters ……81
Figure 21. Single-sided amplitude spectra of the surface PDF ……82
Figure 22. PDF of SSH and calculated PDF for 8000 photons……84
Figure 23. PDF of SSH and calculated PDF for 800 photons……86

List of Tables

Table Page
Table 1 Reflectance……... 15
Table 2 Input parameters (Source: ATL03)…….. 50
Table 3 Input parameters (Source: ATL09) …… 52
Table 4 Control Parameters - Surface Finding…….. 54
Table 5 Control Parameters for Refined Surface Finding and Analysis……...…….56
Table 6 Output (See Appendix A for full product specifications)…….....96
Table 7 ATL12 Output Variables for per Orbit Statistics……..113
Table 8 ATL12 Test Data …..……...….116

1.0 INTRODUCTION

This ATBD will cover the retrieval of Sea Surface Height (SSH) from ICESat-2 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 ICESat-2 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.

This page intentionally left blank

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 ICESat-2, 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 ICESat-2 data.

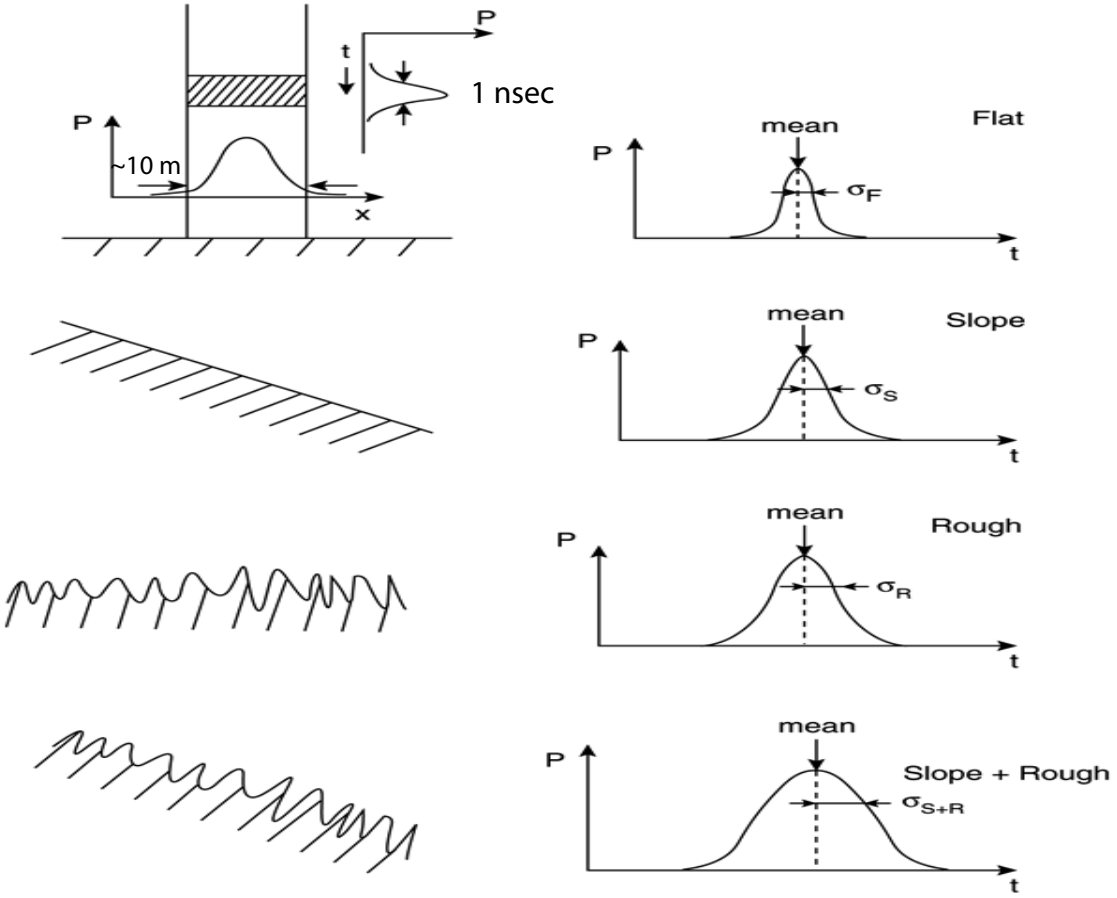

Figure 1 - Characteristics of an idealized distribution of ATLAS return photon arrival times assuming a Gaussian instrument impulse response and a reflective surface with Gaussian roughness. In general surface slope and roughness both broaden the return distribution. For the open ocean the slope effect 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 ICESat-2 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 ICESat-2 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 ICESat-2 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 outgoing pulse distribution and these will be part of the ICESat-2 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. ICESat-2 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 pose challenges with the interpretation and aggregation of meaningful SSH from ICESat-2 photon heights. It is useful to outline these issues here, keeping in mind that the magnitude of the most sought-after components of the SSH signal, changes in DOT and mean SSH, are significantly smaller than most of the other components of SSH.

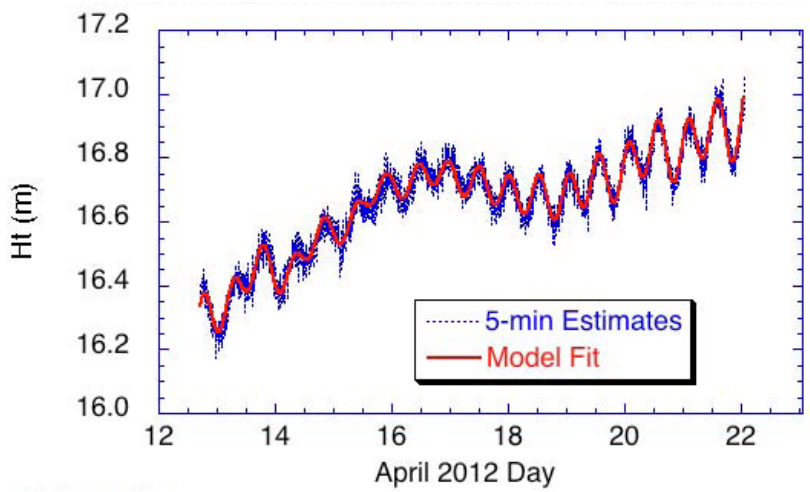

Figure 2. Geodetic height measured with a geodetic grad dual-frequency GPS with PPP processing at a drifting ice camp near the North Pole plotted with an *ad hoc* model of ocean tides and longer variations assumed due to the geoid and DOT variations. The geoid variations over the 60 km drift are estimated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

Tides, Medium-Frequency Changes in Circulation and Aliasing: ICESat-2 will complete one orbit of the earth in about 1.5 hours. The repeat period is 91 days, and commonly, especially at low latitudes, a 25-km grid cell in a grid-average of ICESat-2 might only a few satellite passes per month. Consequently, tides and medium to high-frequency changes in circulation will seriously alias into slowly sampled regions, making aggregation into valid long-term means difficult. This is addressed by correcting the photon heights to the SSH relative to the WGS84 ellipsoid using models of tides and the response of the ocean to atmospheric forcing including the inverse barometer effect., With the predicted tidal and circulation contributions to SSH, the residual can be averaged, and only the error in the modeled response is aliased. As we will show, for ATL12 processing we reduce the size of the signal we process by subtracting the EGM2008 mean tide geoid to process for dynamic ocean topography.

The Geoid, Narrowing the Window on Photon Returns, and Measuring Mean DOT: ICESat-2

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. Globally, variations in SSH and the geoid are hundreds of meters, while variations in DOT, equal to SSH minus the geoid, are only a couple of meters. Given the potentially sparse character of ICESat-2 ocean returns, the variations in the geoid from EGM2008 will be extracted from the photon heights to remove the inherent variability not due to ocean processes and reduce the range of variability we have to deal with in our signal processing. As a final step in processing, the EGM2008 geoid will be added back to DOT to yield SSH relative to the WGS84 ellipsoid as the primary AT L12 output.

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 ICESat-2 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 ICESat-2 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.

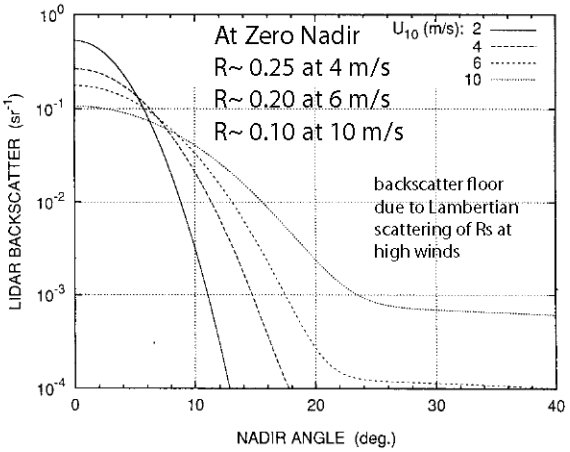

Figure 3. Reflectance as modeled by *Menzies et al.* [1998] as a function of wind speed and nadir angle.

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, $R_s \sim 1/ \langle S^2 \rangle$, due to wind generated surface waves. They use a relation between wind speed, U_{10}, and $\langle S^2 \rangle$ by Wu [*Wu*, 1972] based on results of Cox and Munk [*Cox and Munk*, 1954]; $\langle S^2 \rangle$ 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 ICESat-2 measurements of sea surface backscatter. Working solely with ICESat-2 observations, we can conceivably estimate U_{10} from reflectance. As we discuss below this estimate of U_{10} may be used with

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
10c	Wind = 10 m/s	0.1

SWH, estimated from the variance of the photon height distributions, to calculate the Sea State Bias in ICESat-2 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 ICESat-2.

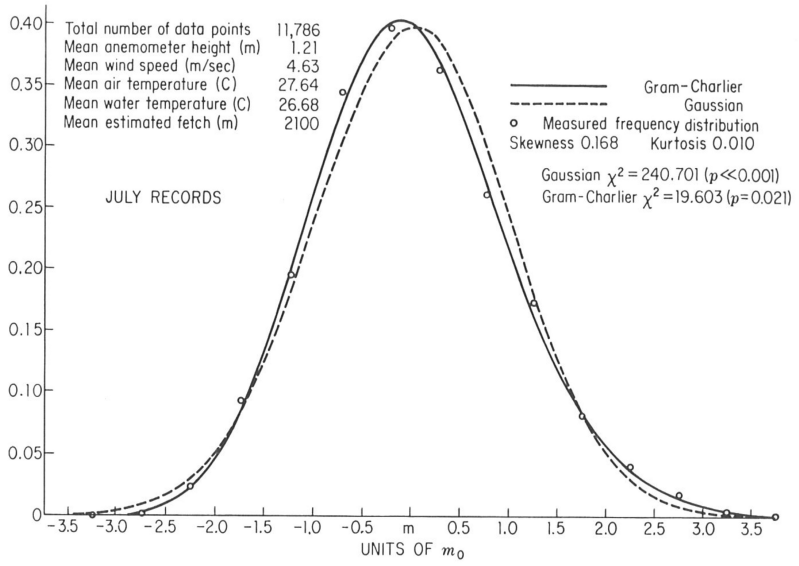

Figure 4. Typical Gram-Charlier distribution sea surface height in a wave environment. The positive skewness and excess kurtosis (Gaussian skewness and excess kurtosis = 0) is due to truncation at low heights and a long tail at high heights. Fig. 7.4-1 of *Kinsman* [1965].

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 ICESat-2 than photons striking the lower portions of wave surfaces; peaks are under sampled relative to troughs resulting in a sea state bias (SSB) in estimates of SSH based on the mean of an ICESat-2 photon height distribution or, in the case of ICESat, mean arrival time of a return pulse.

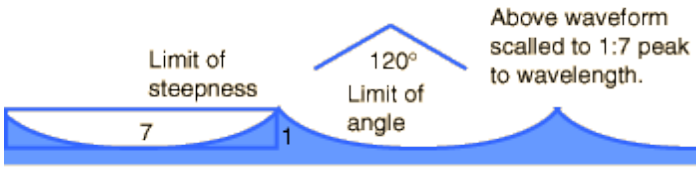

Figure 5. Trochoidal shape of surface waves at the limit of steepness, wavelength equal to seven times height. The sharply peaked trochoidal shape is common to all surface waves longer than capillary waves, which have an inverted trochoidal shape. At the steepness limit the waves begin to break. (Figure 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 despite 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 ICESat-2 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).

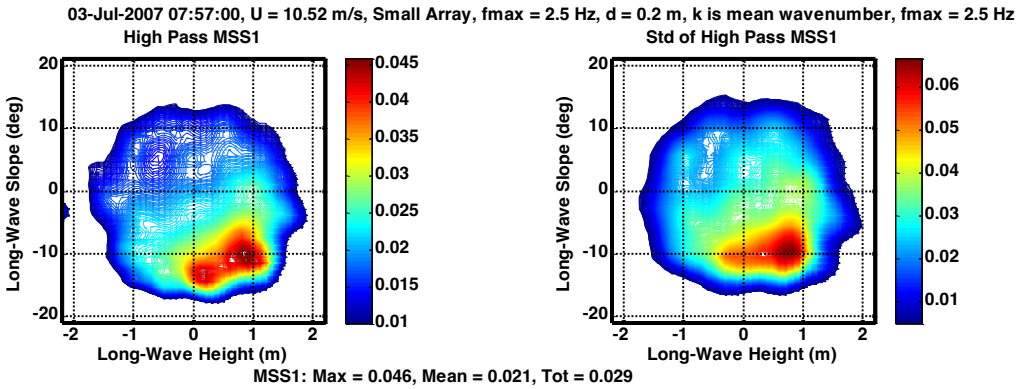

Figure 6. Left – Short wave MSS as a function of long-wave height and slope. Right – the standard deviation of short wave MMS as a function of the same variables.

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.

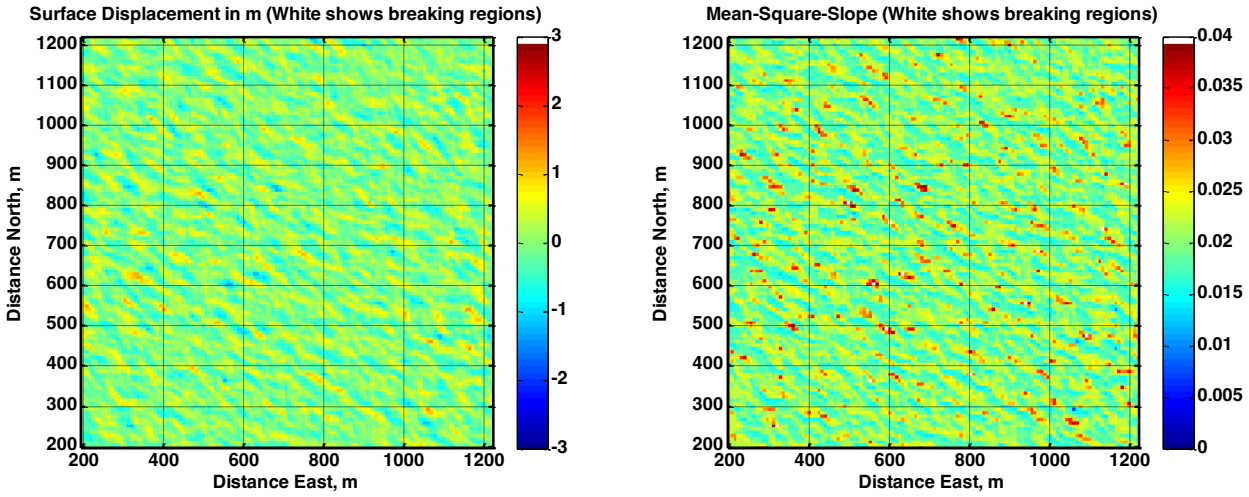

Figure 7. From *Plant* [2015b]. Left – Simulated 2D surface displacement, Right – 2D variation of short-wave mean MSS values that were calculated up to 2 Hz. Downwind direction is from upper right to lower left.

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].

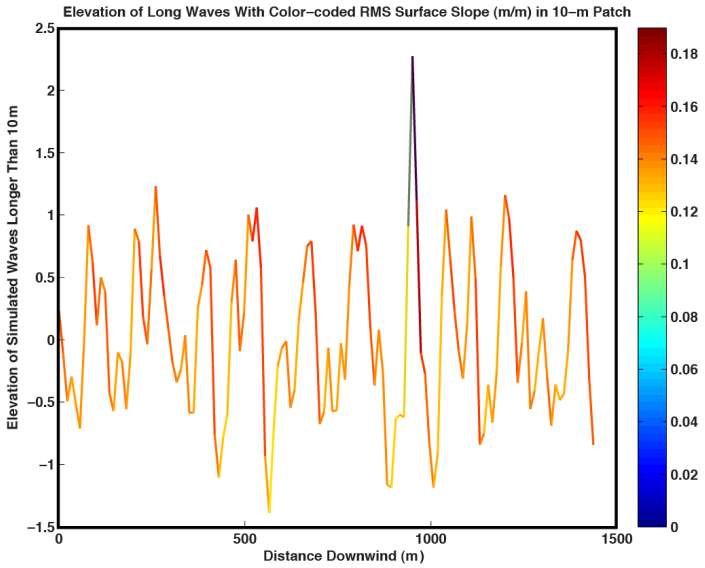

Figure 8. Height due to long-waves versus distance downstream (from upper right corner to lower left in Figure A-left) with color-coded corresponding MSS from Figure A-right.

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).

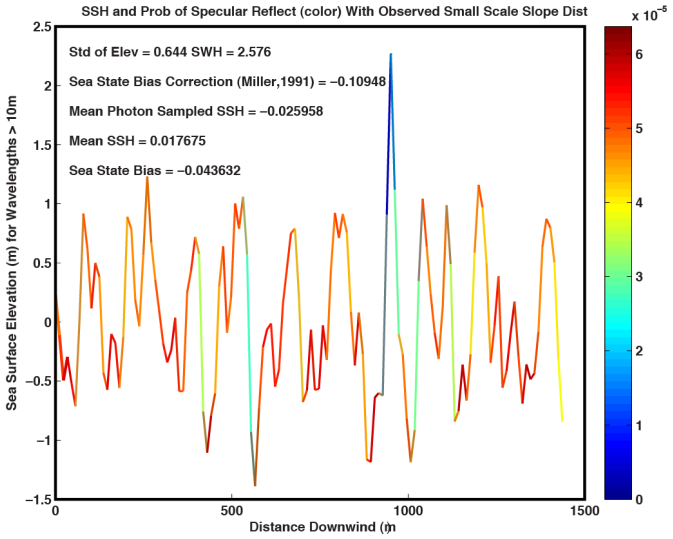

Figure 9. Height due to long-waves versus distance downstream (from upper right corner to lower left in Figure 6-left) with color-coded probability of specular reflection.

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 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. Consequently, lacking further data about the dependence of short wave characteristics on long wave position, we have derived a method whereby the SSB can be determined from ICESat-2 data themselves.

2.2.3 ICESat-2 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 ICESat-2.

A Priori Estimation of SSB Using Only Altimeter Data

The small footprint of ICESat-2 (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, presents a unique opportunity to capture the shape of the long wavelength, energy containing surface waves. This makes it possible to estimate ICESat-2 sea state bias solely based on contemporaneous ICESat-2 returns without resorting to outside data or even comparison with past ICESat-2 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:

$$Y = H_{SS} + A\sin\frac{2\pi}{L}x + N(0,\sigma) \tag{2}$$

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:

$$r = \bar{r} + \alpha A\sin\frac{2\pi}{L}x \ , \tag{3}$$

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

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

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

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

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

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

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

$$y_e = \frac{1}{X\bar{r}}\left[H_{SS}X\bar{r} + \frac{X}{2}\alpha A^2\right] = H_{SS} + \frac{\alpha A^2}{2\bar{r}} \tag{7}$$

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

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

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

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

where η_i is the measured surfaced displacement corresponding $A\sin(2\pi x/L)$ in our idealized example and σ_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:

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

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

$$\varepsilon = \frac{\sum_{i=1}^{N} b_i^0 \eta_i}{\sum_{i=1}^{N} b_i^0} = \frac{\frac{\alpha}{N}\sum_{i=1`}^{N} \eta_i^2}{\frac{1}{N}\sum_{i=1}^{N} b_i^0} = \frac{\alpha A^2}{2\overline{r}} = SSB \tag{11}$$

Appendix B gives a more detailed derivation of (9) in terms of rate of photon return for a photon counting lidar in place of cross-section. It also describes the corresponding EM or sea state bias in the higher moments of the surface distribution (Equation B20 in Appendix B).

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 ICESat-2; 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 wavenumbers for the spectral decomposition of the ocean surface. A more robust method is to average both the raw photon heights from the surface finding step and the rate of photon returns in evenly space bins of along-track distance. The covariance of these been averages is then used to evaluate equation 7 for the sea state bias. This method is easier to

automate and accounts for more of the energy in the wave field than does the harmonic method.

MABEL data from the Greenland Sea on April 19, 2012 illustrate 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

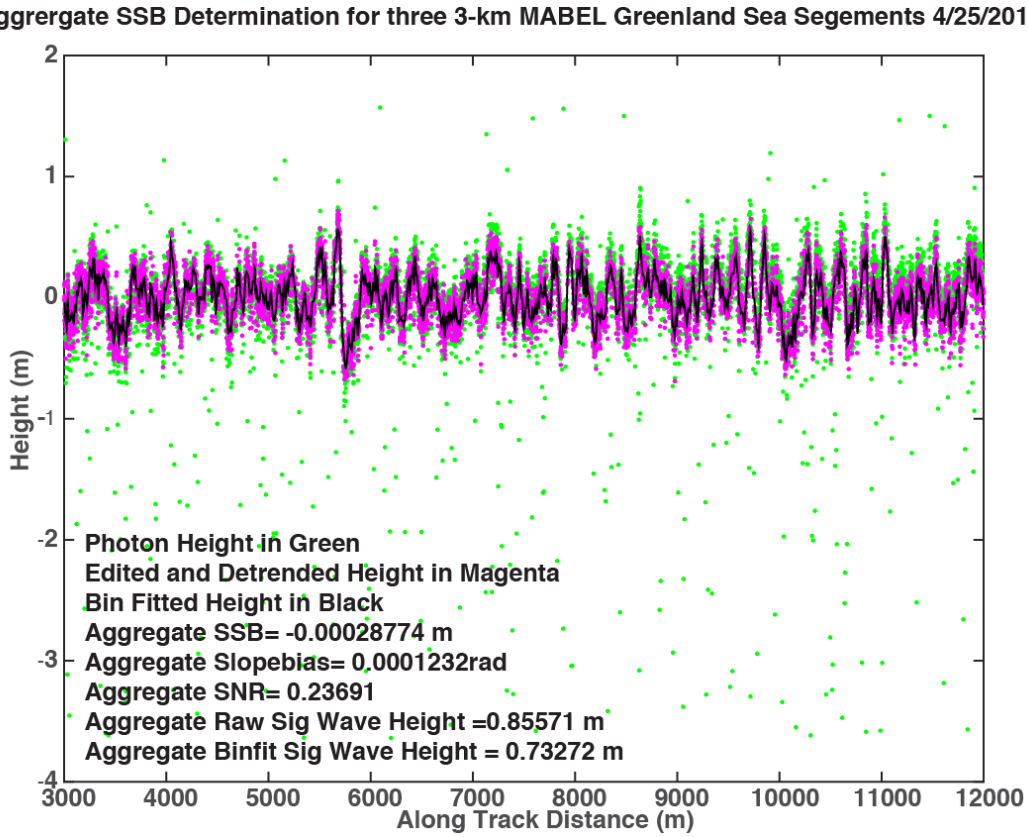

Figure 10. Surface return heights in green from MABEL transit over the Greenland Sea versus distance along-track. Three-sigma edited and detrended heights are plotted in green and a10-m bin averages are plotted in black. The bin averages and a priori SSB estimates were computed in three 3000- m segments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

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 ocean 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). Based on initial data ATLAS gets on the order of one surface reflected photon per laser pulse or 0.7 m of track, so 100 photon retrievals in the adaptive surface detection algorithm would result in segment lengths of about 70 m. Given the 17-m footprint of the ICESat-2 beams, this is equivalent to 4 independent (non-overlapping) spatial samples of the sea surface.

In future developments, heights in marginal ice zones or coastal zones may be defined for shorter variable length segments sampled at variable intervals along the ground track to adapt to the patches of water between sea ice in the first instance or near land in the second instance. These overlap regions are those that are classified as both sea ice and ocean (marginal ice zone) or land and ocean (coastal zone). The possible methods of distinguishing open ocean from sea ice and land and setting the ocean segment lengths in these mixed regions 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 (ATL13) 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 ICESat-2 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 necessary to accumulate large number of surface reflected photons to make meaningful estimates of the higher moments of the height distribution. Given the low

number of returns from open water (e.g., O[1 photon / pulse]), and considering that ocean waves can have wavelengths of hundreds of meters, to get meaningful statistics over open water with significant wave activity requires accumulating data over several kilometers. The atmospheric cloud flag (***layer_flag*** in ATL03) is used in part in setting the ocean segment length required to achieve a minimum number of photons. It and other geophysical data such as the background photon rate from ATL09 and the geoid are reported every 400 pulses (25 Hz) or 14 geo-bins where a geo-bin is the ATL03 geolocation segment and amounts to about 20-m along-track distance. Consequently, over the open ocean, we will seek a minimum photon count of 8,000 photons equivalent to about 5.6 km or 280 20-m geo-bins for meaningful higher order statistics.

3.1.1 Height Segment Parameters (output group *gtx/ssh_segments/heights/)*

3.1.1.1 Geolocation/Time

The location is the mean location of designated signal photons used as input to the surface finding procedure (***gtx/ssh_segments/ longitude & latitude,*** see Table 6***)***. The time (***gtx/ssh_segments***/*delta_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 (***gtx/ssh_segments/heights/h)***. The primary quality metric is the uncertainty in the average sea surface height, (***gtx/ssh_segments/heights*** */h_uncrtn)* determined as a function of height variance and length of the ocean segment divided by the height correlation length scale.

3.1.1.3 Subsurface-scattering corrections

Subsurface-scattering corrections are yet to be determined. We find subsurface returns in ICESat-2 ocean photon returns. They tend to be minimal for the deep open ocean where the water is relatively clear. Our surface finding is designed to include few subsurface returns in the first place by rejecting photon heights at greater noise levels below the surface than above the surface. More recently, starting with Release 4, the surface finding, instead of being based on the distribution of photon heights, is based on the distribution of the photon height anomalies relative to a moving 11-point bin average of high-confidence photon heights. This excludes subsurface returns under the crests of surface waves obviating the immediate need for a subsurface return correction.

Subsurface scattering is more 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 we may be able to develop relations between photon height distribution skewness and other evidence of scattering, and if so, 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

Generally, given the normally low reflectance of the ocean surface first-photon-bias corrections are not needed for the ocean products. However, in looking at on-orbit data, we find rare instances of quasi-specular returns that we think occur when the sea surface is smooth but has very short wavelength ripples. For these cases photon detector saturation can occur leading to first-photon bias and the appearance of subsurface afterpulses. In the current ATL12 processing (Release 4), photons in these afterpulses are edited out by considering only photons with the index of photon height quality from ATL03 (*quality_ph*) equal to 0. Starting with Release 5 we deleted all photons coming from laser pulses that included any photons with a quality index indicative of full saturation. For Release 7, we will use only photons with *quality _ph* equal to 0 indicating no saturation, or 10 indicating partial saturation for which a first-photon-bias correction will be computed. We derive a first-photon-bias correction for 10-m bins within ocean segments and an ocean segment average first-photon-bias correction.

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 (*gtx/ssh_segments/stats/n_photons)*, histogram of the population (*gtx/ssh_segments/heights/y)*, and, at minimum the mean, variance, and skewness, and kurtosis (*gtx/ssh_segments/heights/ymean, yvar, yskew, ykurt).* Note that *ymean* should be nearly zero and differs from the average sea surface height relative to the geoid by *meanoffit2* (*gtx/ssh_segments/heights/meanoffit2)*

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 (*gtx/ssh_segments/heights/ htybin)* and the rate of photon returns (*gtx/ssh_segments/heights/ xrbin)* in 10-m along-track segments (see *gtx/ssh_segments/heights/xbind, latbin, lonbin)*. The correlation of height and return rate normalized by the average photon return rate is the electromagnetic (EM) sea state bias (*gtx/ssh_segments/heights/bin_ssbias)*. 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 (*gtx/ssh_segments/heights/swh)* (=4xStd.) that can be compared with reanalysis products.

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).

3.1.1.9 Ocean Segment Statistics

Important statistics for each ocean segment will also be output (in the *gtx/ssh_segments/starts/* group). These include: the number of surface photons found for the segment, *n_photons;* the number of photons in the ±15-m ocean downlink band, *n_ttl_photon*; solar background rate from ATL03 averaged over the segment, *backgr_seg*; the the percentage of geo-segs in the ocean segment with *layer_flag* = 1, *cloudcover_percent_seg*; the segment average dynamic atmospheric correction to SSH, *dac_seg*; segment average ocean depth, *depth_ocn_seg*; *the o*cean segment average of the conversion factor added to ATL03 *geoid* to convert from the tide-free to the mean-tide system;, *geoid_free2mean_se* and the ocean segment average of EGM2008 geoid in the mean-tide system height above the WGS – 84 reference ellipsoid, *geoid_seg*.

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

ICESat-2 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. Prior to ATL03 Release 4, the EGM2008 was reported by ATL03 as taken from the mean-tide EGM2008 model. Starting with Release 4 ATL03 reports EGM2008 in the tide-free system and provides a conversion (geoid_free2mean) that can be added to the tide-free geoid to yield the mean-tide EGM2008 for use in the ATL12 ocean product.

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)

The tide related geophysical corrections applied to ATL03 ellipsoidal photon heights input to ATL12 include solid earth tides (in tide free system), ocean loading, and pole tides (both ocean and solid earth):
Solid earth tides (*tide_earth*): ATL03 photon heights input to ATL12 are relative to solid earth tides in the tide-free system derived using IERS (2010) conventions are ±30 cm max (details in ATL03 ATBD section 6.3.3). The ATL03 solid earth tides are appropriate to all higher level products, but they are appropriate to the ocean and sea ice products only insofar as they are the reference to which ocean tides are added for a full accounting of tides in ocean surface heights.

Ocean load tides (*tide_load*): The local displacement due to ocean loading (-6 to +6 cm) derived from ocean tide model FES2014b. 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)

Tidal corrections to be applied in ATL12 include short period ocean tides and longer period equilibrium tides:

Ocean tides (*tide_ocean*): Short period ocean tides (including diurnal and semi-diurnal, ±4 m) are computed from the FES2014b extrapolated tide model. Details in ATL03 ATBD, section 6.3.1..

Long-period tides (*tide_equilibrium*): Long-period tide computed from the FES2014b extrapolated tide model The S_a (solar annual) constituent was omitted due to it being deemed erroneous. Details in ATL03 ATBD, section 6.3.1.

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 with ICESat-2 derivations of SWH as part of the sea state bias calculation.

3.2 Gridded Sea Surface Height (ATL19/ L3B)

The ATL19 product includes gridded monthly estimates of dynamic ocean topography (DOT) taken from ATL12 ocean segment data. Ocean segments range in length roughly from 3 to 7km. The ocean segment data are averaged in ¼° or 25km grid cells. Data from all six beams (gtxr and gtxl) are used from the beginning to the end of each month, although commonly over most of the ocean only data from the strong beams will be available. The IS-2 data are averaged onto three grids, called mid-latitude, north-polar and south-polar. The ATL19 product has groups with the same names containing the gridded data for each of those regions. The grids, gridding schemes, input variables and output variables are described in detail in the ATBD for ATL19.

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 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.

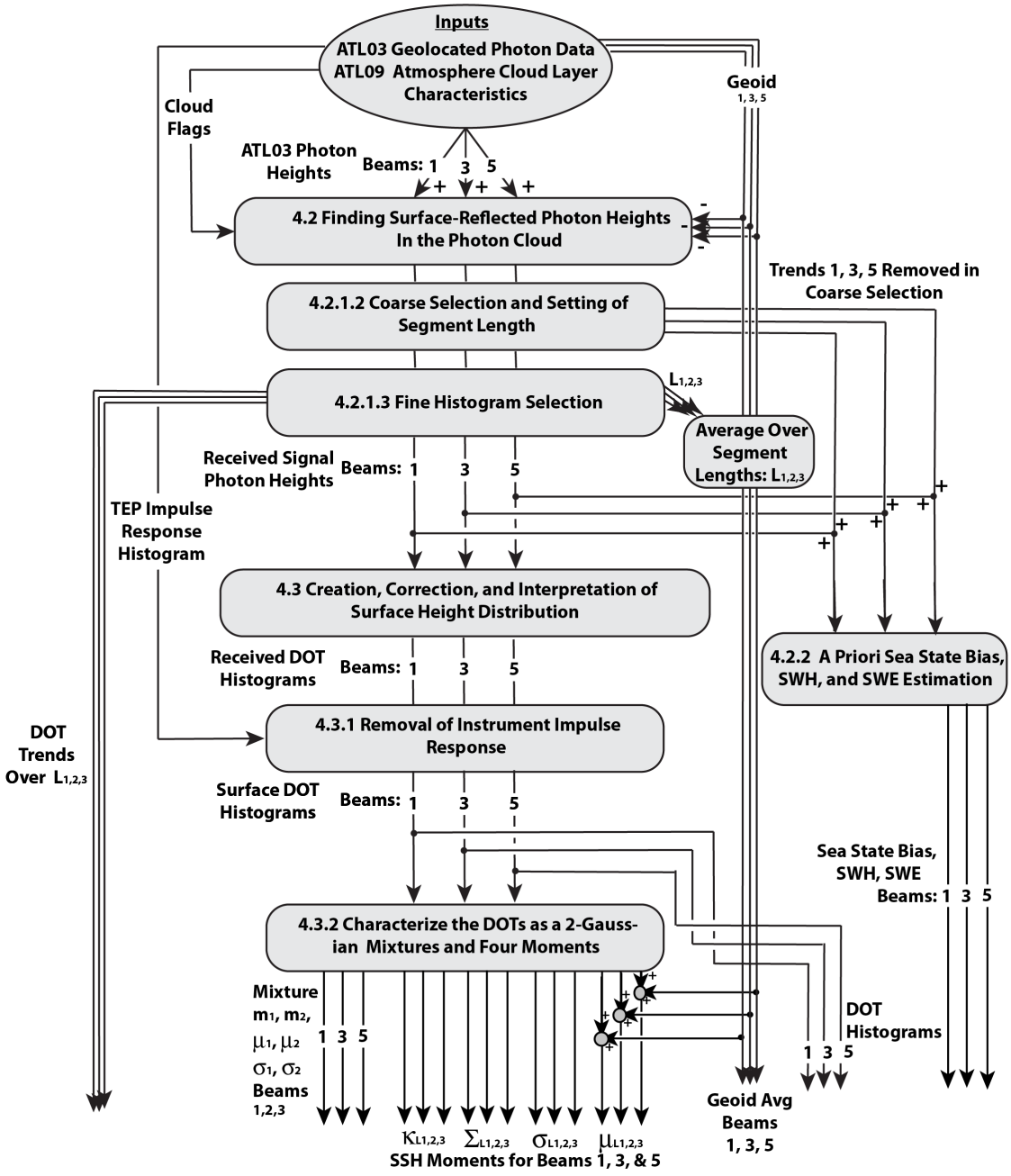

Figure 11. ATL 12 processing block diagram as discussed in Section 3 and 4. μ, σ, Σ, and K denote mean, standard deviation, skewness, and kurtosis respectively.

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:

ATL12 will require ATL03 photon heights for each pulse in the ocean downlink band and associated time geolocation, geophysical corrections (geoid, tides, dynamic atmospheric correction), and confidence level information, plus the cloud flag***, layer_flag***, every 400 pulses from ATL09.

As of Release 3, the ATL03 includes a flag for each ground track indicating the quality of the precision orbit and pointing determination: ***gtx/geolocation/podppd_flag,*** and we found unrealistic values for ocean segments when this flag was non-zero indicating degraded orbit and/or pointing data. Consequently, ATL12 processing for Release 4 is done with data for which ***gtx/geolocation/podppd_flag***=0.

It is important to note that non-zero values of the ***popppd_flag*** have been commonly associated with ocean scans over the Western Pacific Ocean or round-the-world (RTW) scans, which in either case require ICESat-2 to execute conical scans with 5° off-pointing. In the past, ***podppd_flag*** was set to zero in ATL03 throughout these maneuvers. As a consequence of the podppd-flag editing, ATL12 Release 4 had a significant data gap in the Western Pacific. However, with the present processing, the science data is reasonable in most parts of ocean scans and RTW scans other than transitions into them. Consequently, the podppd_flag has been refined.

As of Release 5, the ***podppd_flag*** will have seven possible values: normal operations vales: 0=NOMINAL; 1=POD_DEGRADE; 2=PPD_DEGRADE; 3=PODPPD_DEGRADE; plus possible calibration maneuver related "CAL" values: 4=CAL_NOMINAL; 5=CAL_POD_DEGRADE; 6=CAL_PPD_DEGRADE; 7=CAL_PODPPD_DEGRADE. For ATL12, we will only use data when the POP/PPD flag indicates nominal normal operations or nominal CAL maneuvers, ***podppd_flag*** equal to 0 or 4. For each ocean segment processed, we will report the higher of these two ***podppd_flag*** values, 0 or 4, of data used in the ocean segment.

Further, as of April 17, 2020, ICESat-2 started doing segmented RTW scans in three parts to cover a reference ground track excluding the ice sheets. More importantly, to reduce the Western Pacific data gap, on April 8, 2021, instead of two ocean scans per day ICESat-2 will only do two ocean scans on the same day as an RTW scan, and with two

(sometimes three) RTW scans per week, this will result in only four (sometimes five or six) ocean scans per week.

Based on early discussions, this ATBD and the developmental software developed from it was assumed that ICESat-2 over the ocean would downlink photon heights within ±15 m of the EGM2008 geoid. Aside from the ± 15-m band about the geoid, no classification as to surface/non-surface photons was to be made.

However, as configured in December 2018, the Flight Science Receiver Algorithms (FSRA) downlinks photon heights in bands about heights chosen by signal finding procedure described in text in "ATLAS Flight Science Receiver Algorithms Version 3.7c". This identifies center heights over 200-shot major frames as those where the most photon heights relative to noise are found in adjacent bins of a coarse altimetry band. When operating in local daylight and in the presence of even thin clouds, and perhaps over parts of the ocean with reduced reflectance, for example wave slopes, the most popular height bin can shift to chance concentrations of noise photons up to 200 to 300 m above or below the geoid. The resulting 30-m erroneous downlink bands of photon heights appear in 200-pulse major frames and one or two height bands supposedly bracketed by heights topping at (using the example of ground track 2 right) *gt2r/bckgrd_atlas/tlm_top_band1* and *gt2r/bckgrd_atlas/tlm_top_band2*, with heights given as *gt2r/bckgrd_atlas/tlm_height_band1* and *gt2r/bckgrd_atlas/tlm_height_band2*.

The most concerning problem is that when the downlink band is pulled away from the true surface, we can arguably lose true surface reflected photons, biasing sea surface heights in some unknown direction and violating the assumptions of the ATL12 sea state bias determination. In subsequent revisions of ATL03, the photons from major frames for which the center of the downlink band is far removed from the geoid are eliminated.

In practice we have found that the erroneous downlink bands can be edited out as those containing no ATL03 high confidence photons. Therefore, we use the signal confidence product in ATL03 to edit out the erroneous downlink bands. This grades each photon height return as to likelihood that it is a surface return by examining height bins at 50-pulse rate for signal to noise ratio in the context of noise level. Early ICESat-2 ocean data suggests that photon heights with high, medium, and low (4, 3, 2) confidence levels very seldom appear in the erroneous downlink bands. Further, the confidence level software assigns a fill-in confidence level of 1 to what would be zero confidence heights in order to fill a ±15 bin about the high confidence heights. Therefore, in practice confidence level 1 or greater essentially covers the geoid ±15-m band and can be used to edit out the erroneous downlink bands. In processing, ATL12 considers photons with confidence level equal to 1 or greater as the possible surface photons subject to the surface finding routine.

Because we are using this ground-based editing, in the future we will have to investigate the effect on the sea surface height and SSB calculation of missing true surface photons not downlinked in the presence of erroneous downlink bands, probably by comparing daytime and nighttime data over the same region.

4.2.1.2 Coarse Selection and Setting of Segment Length

The process is basically to first establish a variable length (e.g., up to 7 km) ocean segment with enough (e.g.) photon heights to yield a histogram with reasonable moments up to 4th order.

1. Examine ocean depth and cloud parameters against control parameters to test for suitability for ocean processing.
   - If control parameter *layer_swtch* is equal to 1, then remove the ATL03 segment ids from ocean processing that are associated with the ATL09 segment ids where *layer_flag* is equal to 1. When *layer_swtch* equals 0 (the default value), the software will ignore the cloud cover in accepting data during coarse surface finding.
   - Ignore data collected over land or too close to land. If ocean depth, *depth_ocn*, is less than a depth, *depth_shore*, specifying the effective shoreline of water for which ocean processing is appropriate, then proceed to next 14 geo-bin (400-pulse) segment. The default value for control parameter *depth_shore* will be -10 m for depths given by GEBCO as negative elevation values.
2. Correct photon heights for short period and long period ocean tides (*gtxx/geophys_corr/tide_ocean and tide_equilibrium*), the inverse barometer effect and the modeled dynamic response to the atmosphere (dynamic atmospheric correction, *gtxx/geophys_corr/dac*). Select only photons which have a valid *tide_ocean*.
3. Select only photon heights for which the signal confidence (*gtxx/heights/signal_conf_ph*) is greater than or equal to 1. For Release 7 onward, use only photon heights for which the indicator for saturation *gtxx/heights/quality_ph* is 0 indicating no saturation, or 10 indicating partial saturation for which a first-photon-bias correction will be computed. This will select photons in the likely surface ± 15-m buffer window. Also delete any photon heights outside the band defined by geoid plus or minus half the height of the ocean downlink window (presently geoid ±15-m, to become geoid ± 20-m). (**Is the height of the downlink band reported in ATL03 perhaps at the geoseg rate and if so, what is the variable name? The height is listed at the major frame rate. But it would be handier if it was at the geoseg rate**)
4. Subtract the EGM2008 geoid in the mean-tide system from photon heights to reference all heights to the geoid. Note that as of Release 4, EGM2008 provided by ATL03 is in the tide-free system. However, ATL03 also provides the conversion that should be added to the tide-free geoid to yield EGM2008 in the mean-tide system.
5. 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.
6. 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 (e.g., 0.01 m). This is used to estimate a running total number of signal and noise photons. We suggest a bin size of 0.01 for consistency with later steps in the processing. The 14 geo-bin (400-

pulse) segment should be aligned so that the once per 14 geo-bin (400-pulse, approximately 280 m along-track distance) cloudiness flag, ***layer_flag*** in ATL09, represents the aggregate effect of cloudiness and background radiation derived from solar zenith, ***asr_cloud_probability***, ***cloud_flag_atm***, and ***bsnow_con***, during the 280-m segment.

7. Photons in the 14 geo-bin (400-pulse) histogram bins with greater than the *Th_Nc_c* 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. We have been using *Th_Nc_c* = 1 in our tests with early ICESat-2 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.
8. 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.
9. If the total number of candidate signal photons for a strong beam is greater than or equal to a minimum number, *Th_Ps* (e.g., 8,000), of photons or S_{egmax} (e.g., 25) 14 geo-bin (400-pulse) segments have been collected this defines an ocean segment. For weak beams, and ocean segment is defined when the total number of candidate signal photons is greater than or equal to ¼ *Th_Ps* (e.g., 2,000) photons or S_{egmax} (e.g., 25) 14 geo-bin (400-pulse) segments have been collected. Once the ocean segment is defined, we 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* for a strong beam (1/4 *Th_Ps* for a weak beam), repeat steps 4.2.1.2-2 to 4.2.1.2-5 above and intake another 14 geo-bin (400-pulse) segment until the ocean segment reaches a length of 7 km maximum. If the 7-km length is reached and there are less

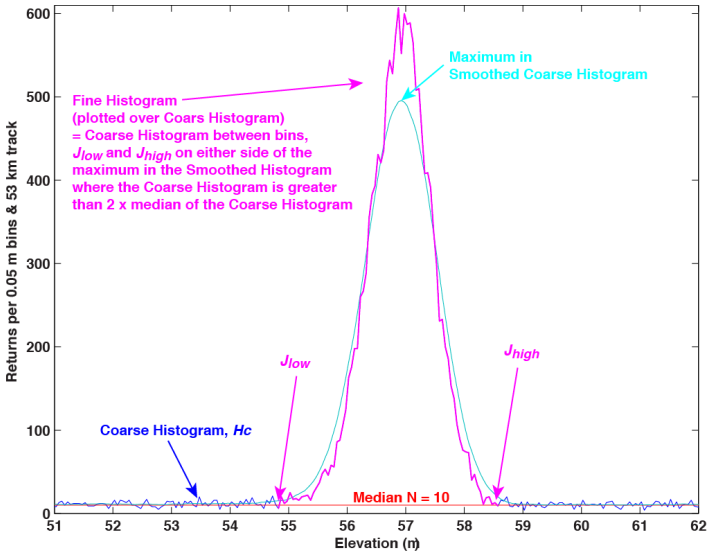

Figure 12. Illustrative coarse, smoothed, and fine histograms from steps 1 and 2 of 4.2.1.3 as applied to 10 minutes of April 2012 MABEL data over Chesapeake Bay. The corresponding figure for detrended data in the second iteration (steps 4-6) are similar.

than a threshold number of photons (e.g., *photon_min* = 4000), ignore that ocean segment and repeat 4.2.1.2-2 to 4.2.1.2-6. For photon counts above this in the 7-km segment, proceed to 4.2.1.3 and fine histogram-based signal finding.

4.2.1.3 Fine Histogram Selection

1. Considering the coarse histogram array, H_c, perform a N_{bmv} = 20-cm/B1 bin (e.g., 21 bins over 20-cm) 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.
2. The ASAS 5.1 finds the bin limits of the fine histogram as all the coarse histogram bins on either side of the maximum in the 20-cm smoothed histogram, from the J_{low} bin position to the J_{high} bin, where the smoothed histogram bin photon count is greater than the median of the smoothed histogram photon count. In practice with preliminary ATLAS data the median has usually been zero. Essentially, we have moved out from the middle of the histogram until the histogram levels fall to zero.

   As of 6/4/2019, the developmental code finds preliminary bin limits of the fine histogram as all the coarse histogram bins on either side of the maximum in the 20-cm smoothed histogram, from the preliminary J_{low} bin position to the preliminary J_{high} bin, where the coarse bin photon count is greater than the coarse histogram median. Then the average counts in the bins above the preliminary J_{high} are calculated and set equal to *tailnoisehigh* and the average counts in the bins below the preliminary J_{low} are calculated and set equal to *tailnoiselow*. Then we find the bin limits of the fine histogram as all the coarse histogram bins on either side of the maximum in the 20-cm smoothed histogram, from the final J_{low} bin position to the final J_{high} bin, where the final J_{low} bin position is that below which the smoothed histogram bin photon count is less than *Th_Nc_f* times *tailnoiselow*, and the final J_{high} bin position is that above which the smoothed histogram bin photon count is less than *Th_Nc_f* times *tailnoisehigh.* This procedure is better able to differentially remove subsurface noise returns and above surface returns. We have found a good value for *Th_Nc_f* is 1.5 in practice with Release 001 ATLAS data and results in greatly improved rejection of subsurface returns.

   As of 11/5/2019 for Release 4, we modified the surface finding procedure as described in section 5.3.2. Instead of basing surface finding on the distribution of photon heights, it is now based on the distribution of the photon height anomalies relative to a moving 11-point bin average of high-confidence photon heights. This excludes subsurface returns under the crests of surface waves that otherwise fall inside the histogram of true surface heights. This further reduces any error due to subsurface returns, obviating the immediate need for a subsurface return correction.
3. Consider the time series of all the heights stored in A_{coarse} with heights anomalies that lie between anomalies corresponding the J_{low} and the J_{high} bin positions in the distribution of height anomalies. 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 constant and first order

coefficients, *P0* and *P1*, of the linear fit used for detrending 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 including creating the 11-point moving average of high-confidence photons the height anomalies about this moving average, and the histogram of height anomalies.
5. Considering the detrended coarse histogram array, H_d, perform a N_{bmv} = 20-cm/B1 bin (e.g., 21 bins) 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. The ASAS 5.1 finds the bin limits of the fine histogram as all the coarse histogram bins on either side of the maximum in the 20-cm smoothed histogram, from the J_{low} bin position to the J_{high} bin, where the smoothed histogram bin photon count is greater than the median of the smoothed histogram photon count. In practice with preliminary ATLAS data the median has usually been zero. Essentially, we have moved out from the middle of the histogram until the histogram levels fall to zero.

   As of 6/4/2019, the developmental code finds preliminary the bin limits of the fine histogram as all the coarse histogram bins on either side of the maximum in the 20-cm smoothed histogram, from the preliminary J_{low} bin position to the preliminary J_{high} bin, where the coarse bin photon count is greater than the coarse histogram median. Then the average counts in the bins above the preliminary J_{high} are calculated and set equal to *tailnoisehigh* and the average counts in the bins below the preliminary J_{low} are calculated and set equal to *tailnoiselow*. Then we find the bin limits of the fine histogram as all the coarse histogram bins on either side of the maximum in the 20-cm smoothed histogram, from the final J_{low} bin position to the final J_{high} bin, where the final J_{low} bin position is that below which the smoothed histogram bin photon count is less than *Th_Nc_f* times *tailnoiselow*, and the final J_{high} bin position is that above which the smoothed histogram bin photon count is less than *Th_Nc_f* times *tailnoisehigh.* This procedure is better able to differentially remove subsurface noise returns and above surface returns. We have used *Th_Nc_f* = 1.5 in practice with Release 001 ATLAS data and this has resulted in greatly improved rejection of subsurface returns.

   As of 11/5/2019 for Release 4, as for the first pass, for the second pass we modified the surface finding procedure as described in section 5.3.2. Instead of basing surface finding on the distribution of photon heights, it is now based on the distribution of the photon height anomalies relative to a moving 11-point bin average of high-confidence photon heights. This excludes subsurface returns under the crests of surface waves that otherwise fall inside the histogram of true surface heights. This further reduces any error due to subsurface returns, obviating the immediate need for a subsurface return correction. Note that although surface finding is based on the height anomalies, the final product of these steps is the histogram of the actual heights of the selected surface photons, H_R.

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 wave and sea state bias analysis (Section 4.2.2).

4.2.2 A Priori Sea State Bias Estimation and Significant Wave Height

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. This height record has been detrended during surface finding and the mean of this fit (linear fit coefficients *P0* and *P1* in 4.2.1.3) over the positions of the surface detected points must be added back to the surface heights to include all SSB in the record. With this height record, we 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 the ocean segments (equivalent to 8,000 candidate signal photons or 7 km whichever is less), we will determine surface height variation and photon return rate in 10-m along-track bins. To calculate the photon return rate per meter it will be necessary to sort the photon height record in ascending order of along-track distance. From the height and photon rate records 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.

4.2.3 First Photon Bias Estimation

We will also use the spatial record of signal photon heights accumulated in the 10-m along-track bins (Section 4.2.2) to estimate the average first-photon bias (FPB) arising from saturation of the ATLAS photon detectors in instances of quasi-specular reflectance from the surface. In addition to knowledge of the photon detectors' aggregate deadtime obtained from ATL03, to calculate FPB, we need photon rate and width of the photon height histogram over an appropriate duration. At the photon rates likely to cause FPB, e.g., greater than 8 photons per meter, 10-m along-track provides and appropriate number of photon heights to calculate FPB. Photon rate in the 10-m bins is calculated for SSB determination (4.2.2). For FPB, we only need to additionally compute the standard deviation in photon heights over each 10-m bin to represent the distribution width. Thus, we can obtain the FPB for each 10-m bin, and average these to obtain the ocean segment-average FPB.

4.2.4 Harmonics, Correlations Length, and Uncertainty

The uncertainty in the ocean segment average SSH is equal to the standard deviation of the height observations divided by the square root of the degrees-of-freedom, which is the number of independent samples. If each photon height were independent, degrees of

freedom over an ocean segment would be enormous, but due to surface waves photon heights are not independent and degrees-of freedom tend to be the number of dominant wave lengths in the ocean segment. Consequently, we use the 10-m bin averages of 4.2.2 to estimate the correlation length scale, significant wave height equal to four times the standard deviation of height, along-track wave harmonics, and uncertainty in ocean segment-average SSH.

4.2.5 SSH Determination in Ice-Covered Waters

We would like our sea surface height and dynamic ocean topography determinations in ATL12 and ATL19 to be valid across the whole North Polar and South Polar areas from open water to ice-covered regions, i.e., ice concentration being greater than 15%, without being biased by sea ice freeboard. To do this we need measurements of height at a fine enough horizontal scale to resolve leads in sea ice. This we have in the ATL12 10-m averages if we are content with considering leads greater than 10-m wide. Furthermore, we can align ATL12 and simultaneous ATL07 and determine which ATL12 10-m bins are judged to be open water bins according to the ATL07 *height _segment_ssh_flag* identifying ATL07 height segments as bright, open-water surfaces. ATL10 uses an additional parameter to characterize the quality of fit of photon heights to the idealized ATLO7 photon height distribution, *height_segment_fit_quality_flag,* and requires *fit_quality_flag* to be 1, 2, 3, or 4 for use of the segment height for sea surface height. For ATL12 we will use the same criteria, ATL07 *ssh_flag*=1 and *fit_quality_flag* = 1, 2, 3, or 4 to consider the ATL07 segment height suitable for SSH or DOT determination. So, we flag the ATL12 10-m bins so qualified by ATL07 and ATL10 as valid bright open water with flag consisting of the corresponding ATL07 sea ice segment height, *xbin_atl07_dot*.

While heights in the ATL12 10-m bins represent DOT relative to the EGM2008 mean-tide geoid, ATL07 sea ice segment heights are given as the departure from a mean sea surface (MSS) derived from other altimetry data, primarily CryoSat-2. Consequently, we must adjust the ATL07 heights to convert them to DOT relative to EGM2008. We also need to remove the dynamic inverted barometer correction used in ATL07 Release 006 heights and apply the dynamic atmospheric correction used in ATL12. (Note: we believe the DAC is a more appropriate correction even in ice covered waters. If in future releases of ATL07, the DAC is used, this conversion won't be necessary.) The pertinent ATL07 variables and their descriptions are as of Release 006:

In the *gtxx/sea_ice_segments/geophysical/* group:

height_segment_dac: Dynamic atmospheric correction DAC includes an inverted barometer IB effect.

height_segment_dynib: Inverted barometer effect calculated from dynamic mean ocean sea level pressure (computed using ANC10 and ANC48, */ancillary_data/sea_ice/ mean_ocean_slp*)

height_segment_geoid: Geoid height above the WGS 84 reference ellipsoid (range -107 to 86m) based on the EGM 2008 model. The geoid height is in the tide free system.

height_segment_geoid_free2mean: Additive value to convert geoid heights from the tide-free system to the mean-tide system. Add to height_segment_geoid to get the geoid heights in the mean-tide system.

height_segment_mss: Mean sea surface height above WGS84 reference ellipsoid (range: -105 to 87 m) based on the DTU 13 model. The MSS height (from ANC15) is adjusted to be relative to the tide free system. (Note: A mean sea surface is conventionally defined as in the first sentence, relative to a reference ellipsoid e.g. WGS84. We ignore the second sentence in the description).

And in the *gtxx/sea_ice_segments/heights/* group

height_segment_height: Mean height from along-track segment fit determined by the sea ice algorithm. The sea ice height is relative to the tide-free MSS.

With these variables

xbin_atl07_dot = ATL07 data *heights/height_segment_height*
+ (*height_segment_dynib* - *height_segment_dac)* to correct to DAC
+ height_segment_mss to convert to SSH
– (*height_segment_geoid* + *height_segment_geoid_free2mean*) to convert to DOT

Note that we will limit candidate bins for sea surface height determination to those 10-m bins for which the photon return rate, *xrbin*, is between 4 and 25 photons per meter (strong beams) or 1 to 7 photons per meter (weak beams) corresponding to quasi-specular but not fully saturated returns. This is a wider range than the criteria for ATL07 segments to have *ssh_flag*=1, indicating bright leads.

When we account for FPB as in Section 4.2.3 above, ATL12 heights in the 10-m bins so flagged can also be considered ice-free sea surface heights. The average (termed *h_ice_free*) of these qualified ATL12 sea surface DOT heights and the average (termed *h_ATL07_ice_free*) of the corresponding ATL07 *xbin_atl07_dot* should be added to the ATL12 product to provide ATL12 and ATL07 renditions of ATL12 ocean segment ice-covered DOT. These two measures can be compared as an indication of how accurate it is to assume surface finding over open water produces the same result in ATL12 and ATL07.

We will also identify all the ATL07 sea ice segments with *ssh_flag*= 1 (bright leads) in an ATL12 ocean segment and compute *h_atl07dot_inatl12oc*, the average DOT from all these ATL07 *ssh_flag* =1 qualified sea ice segments regardless of whether they have a corresponding ATL12 10-m bin. This will be important if there are only leads smaller than 10-m in the span of the ocean segment.

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:

$$M_n(Y) = \frac{1}{m}\sum_{i=1}^{m}\left(Y_i - \overline{Y}\right)^n \tag{13}$$

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

$\int_{-\infty}^{\infty}\left(Y - M_1\right)^n H(Y)dY$, 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 to calculate sample moments for Y from the sample moments of T and R:

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

$$M_2(Y) = M_2(R) - M_2(T) \approx \frac{1}{m}\sum_{i=1}^{m}\left(R_i - \overline{R}\right)^2 - \frac{1}{m}\sum_{i=1}^{m}\left(T_i - \overline{T}\right)^2 \quad (15)$$

$$M_3(Y) = M_3(R) - M_3(T) \approx \frac{1}{m}\sum_{i=1}^{m}\left(R_i - \overline{R}\right)^3 - \frac{1}{m}\sum_{i=1}^{m}\left(T_i - \overline{T}\right)^3 \quad (16)$$

$$\begin{aligned} M_4(Y) &= M_4(R) - M_4(T) - M_2(Y)M_2(T) \\ &\approx \frac{1}{m}\sum_{i=1}^{m}\left(R_i - \overline{R}\right)^4 - \frac{1}{m}\sum_{i=1}^{m}\left(T_i - \overline{T}\right)^4 - M_2(Y)M_2(T) \end{aligned} \quad (17)$$

These sample moments will be calculated for the ICESat-2 heights as a check on more exact methods. 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, $M_4/M_2^2 - 3$. 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, $\mathcal{F}(\ )$, of the convolution of two variables, $H_R(k)=\mathcal{F}(H_R)= \mathcal{F}(H_T*H_Y)$ is equal to the product of the Fourier transform of the two variables,

$$H_R(k)=\mathcal{F}(H_R)= \mathcal{F}(H_T*H_Y) = \mathcal{F}(H_T)\ \mathcal{F}(H_Y) \quad (18)$$

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, $\mathcal{F}^{-1}(\ )$ of the ratio of $\mathcal{F}(H_R)$ and $\mathcal{F}(H_T)$,

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

The problem with this approach is that there is invariably noise in the received signal histogram, H_R, that produces significant values in $\mathcal{F}(H_R)$ at wavenumbers where $\mathcal{F}(H_T)$ 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 $H_T(k) = \mathcal{F}(H_T)$, $H_Y(k) = \mathcal{F}(H_Y)$, $H_R(k) = \mathcal{F}(H_R)$, and $N(k) = \mathcal{F}(N)$,

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

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 $(H_Y(\mathrm{k})\text{-}H_{Yest}(\mathrm{k}))^2$ is a minimum. This is found for

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

and

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

For wave numbers where the inverse signal to noise ratio, $\mathcal{F}(Noise)^2/\mathcal{F}(H_R)^2$, 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 expectation maximization (ML). For a distribution of known type $f(x \setminus \mu_{1,2,3,4})$, the optimum choice of parameters, $\mu_{1,2,3,4}$, for data values $x_1, x_2, x_3 \ldots\ldots x_m$, are those which maximize the likelihood, L,

$$L\left(\mu_{1,2,3,4} \setminus x_1, x_2, x_3 \ldots\ldots x_m\right) = \prod_{i=1}^{m} f(x_i \setminus \mu_{1,2,3,4}) \tag{25}$$

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

$$\begin{aligned} f(x_i \setminus \mu_{1,2,3,4}) &= m_a N_a(x_i \setminus \mu_{a1,2}) + m_b N_b(x_i \setminus \mu_{b1,2}) \\ m_a + m_b &= 1 \end{aligned} \tag{26}$$

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 expectation maximization.

The aggregate moments of the Gaussian mixture *X* with *n* component Gaussian distributions are functions of the component means, $\mu_{i,}$ and variances, σ_i^2, and the mixing ratio, $m_{i,}$

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

For example, the aggregate mixture mean is:

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

and the aggregate variance is:

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

4.3.3 Expectation Maximization

In the Matlab code used for development of the ATL12 software, an expectation maximization code (gmdistribution.fit termed maximum likelihood by Matlab)) was used to derive a 2-component Gaussian Mixture fit to the derived surface distribution. The mean and variance of the two Gaussian distributions along with the mixing ratio of the two Gaussians allows the accurate determination of the most likely first four moments of the distribution of surface height. The kernel of gmdistribution.fit is essentially the same as used in derivation of ATL007 (ATL007, Appendix E and Appendix C in this ATBD). The ASAS Fortran code developed for ATL12 uses the same as that for ATL07 Appendix E to do the 2-Gaussian Mixture fit. For synthetic test data the Fortran expectation maximization produced equivalent result to the Matlab gmdistribution.fit routine approach.

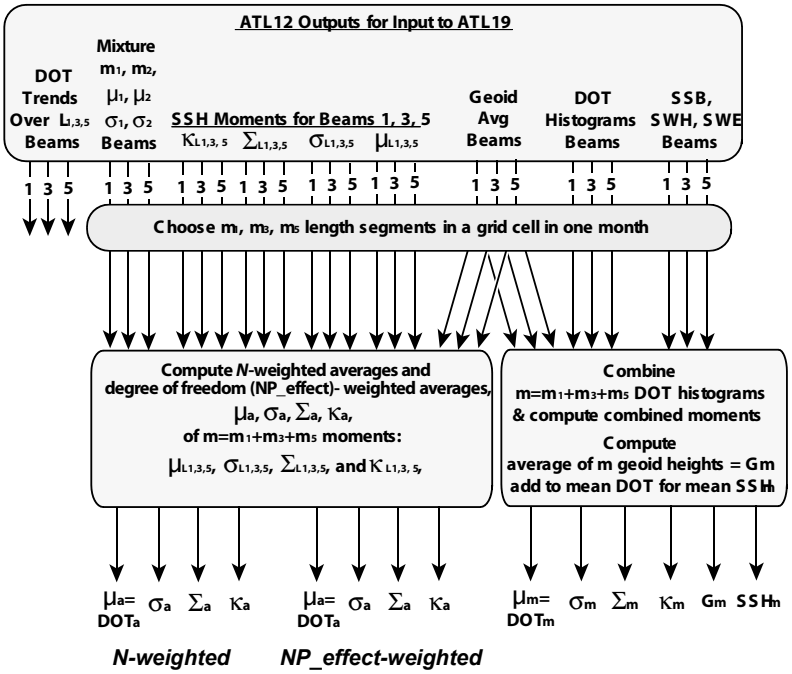

Figure 13. Block diagram for the ATL19 gridding procedure taking ATL12 ocean products as input. μ, σ, Σ, and K denote mean, standard deviation, skewness, and kurtosis respectively.

4.4 Calculation of Uncertainty

We have found the wave environment, rather than instrumental factors, is probably the biggest contributor to uncertainty in estimates of the mean SSH. This is illustrated by our analysis of multiple ocean segments in the southern part of the central North Pacific. These segments show variability in their mean sea surface heights of approximately ± 0.12, much greater than we'd expect from instrumental factors. The standard error, σ_L, in the estimates of the mean are given by

$$\sigma_L = \sigma_h \left(N_{df} \right)^{-1/2}$$

where σ_h is the standard deviation of the population and N_{df} is the number of degrees of freedom. If every one of the ~ 8,000 photon heights were independent, even with a SWH = 2.5 m (σ_h = .625 m), the uncertainty in estimates of the mean would be less than 1 cm. The problem is that owing to the presence of long large waves, successive photon heights are far from independent. The fact that the underlying wave signal is periodic makes the estimate of the effective degrees of freedom in terms of correlation length scale more problematic. As a worst case if we set N_{df} equal to the number of wave periods in the ocean segment, 15 in the case of long waves crossed at an acute angle, the uncertainty, σ_L, in the estimate of the mean equals 0.16 m, close to the mean SSH we see over many similar ocean segments. This problem is not instrumental; it is just made apparent by the ability of ICESat-2 to resolve waves. To lower this uncertainty, we are exploring using harmonic analysis of surface height over each ocean segment and use of the zero wave-number amplitude to represent SSH (D. Percival, personal communication, 2019) as a way of removing large periodic variations in height as a cause of uncertainty. A further advantage of this approach will be that it will add a measure of wave spectral properties to ATL 12. In any event, we will have to derive measures of the effective number of degrees of freedom as an ATL12 output for each ocean segment to properly account for uncertainties in the gridded SSH product.

4.5 ATL19: Gridding the DOT and SSH

This product, based on Product ATL12/L3A, contains gridded monthly estimates of sea surface height 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. ATL12 provides the histograms and first four moments of dynamic ocean topography over ocean segments ~ 3-7-km long (means of DOT are converted to mean sea surface heights by adding ocean segment mean geoid height, which is also output by ATL12). It also provides the number of photon heights, *n_photons*, going into the moments and an effective degrees-of-freedom, *NP_effect*, based on the correlation length scale of surface heights. Using these, ATL19 will produce monthly aggregate histograms of surface heights and averages of the ocean segment moments weighted by both *n_photons* and *NP_effect* (Fig. 13). See the ATL19 ATBD for a complete description.

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 2 and 3. For ATL12, the primary inputs will be photon heights in the downlink band and associated time, geolocation, the geoid (***geoid***), ocean tides (***tide_ocean***, ***tide_equilibrium***), and dynamic atmospheric correction (***dac***) information from ATL03, plus the cloud flag (***layer_flag*** in ATL09) every 400 laser pulses from ATL09 and ocean depth, ***depth_ocn***, from interpolated from GEBCO-2022 gridded ocean bathymetry. 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/geoid*** 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. In addition to the ± 15-m band about the geoid, we use several classifications and flags from ATL03 to edit photon heights.

The photon quality flag, ***quality_ph***, detects various conditions such as saturated laser pulses for which we want to discard photon height, so for Release 6 and earlier we only process heights with this flag equal to 0. For Release 7, we will use only photons with *quality _ph* equal to 0 indicating no saturation, or 10 indicating partial saturation for which a first-photon-bias correction will be computed.

We have also found unrealistic values for ocean segments when ATL03 ***gtx/geolocation/podppd_flag*** is non-zero indicating degraded orbit and/or pointing data. Per Section 4.2.1, for Release 5 with seven possible values of podppd_flag, ATL12 processing should only be done with data for which ***gtx/geolocation/podppd_flag***=0 indicating nominal operations or ***gtx/geolocation/podppd_flag***=4 indicating nominal calibration maneuvers.

Perhaps most importantly, we use the ATL03 confidence that the photon is surface reflected, ***signal_conf_ph*** at high levels (3 or 4) to determine if we have a viable surface in a downlink band, and we use only photons with ***signal_conf_ph*** greater than or equal to 1 in our surface finding procedure.

Also, the surface finding now (as of Release 4) uses the distribution of the anomaly of photon heights relative to a 11-point moving average of photons with ***signal_conf_ph*** reater than or equal to 2 to segregate surface photons from noise photons. This method is more sensitive than editing the height histogram and shows better rejection of subsurface returns under ocean waves.

The 14 geo-bin (400-pulse) ocean segment are aligned so that the once per 14 geo-bin (400-pulse) cloud flag (***layer_flag*** in ATL09) represents the cloud conditions during the ocean segment. We have found that we can get good surface heights even when the ***layer_flag*** indicates the presence of waves, but we do compute the average ocean segment cloud cover as the fraction of layer_flag =1 over the ocean segment.

Table 2 Input parameters (Source: ATL03)

Label	Description	Symbol
delta_time	Elapsed seconds since first data point in granule	***time_initial***
lat_ph	latitude of each received photon	***lat_initial***
lon_ph	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	Along-track segment ID number	***segment_id***

signal_conf_ph	Confidence level associated with each photon event selected as signal. Zero equals noise. One means added to allow for ± 15-m buffer around high confidence, but algorithm classifies as background. Two equals low confidence. Three equals medium. Four equals high. With release 4 ***signal_conf_ph*** equal = -2 indicates possible TEP photon	***signal_conf_ph***
quality_ph	Photon quality as to saturation and after-pulse • 0 = nominal, no saturation • 1 = possible after-pulse • 2 = possible long-tail impulse response • 3 = possible TEP (also indicated by *signal_conf_ph* = -2) • 10 = partial saturation for which a first-photon-bias correction will be computed	***quality_ph***
ref_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.	
ref_elev	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.	
bckgrd_rate	Background count rate (*backgrd_atlas/bckgrd_rate*), simply averaged over the segment	
fpb_parm(n)	Parameter needed to compute first-photon bias correction to a mean height. This is not used as of Release 4. See Section 3.1.1.4.	TBD
dac	Dynamic atmospheric correction (DAC) includes inverted barometer (IB) effect (± 20 cm)	
tide_earth	Solid Earth Tides in the tide-free system (±30 cm, max)	
tide_equilibrium	Long-period tide, no S_a constituent	
tide_load	Crustal deformation due to Ocean Loading (-6 to +6 cm)	
tide_oc_pole	Ocean Pole Tide -Loading due to centrifugal effect of polar motion on the oceans (2 mm, max)	
tide_ocean	Short-period ocean tides (diurnal and semi-diurnal, ±4m)	
tide_pole	Pole Tide -Rotational deformation due to polar motion (-1.5 to 1.5 cm)	
neutat_delay_total	Total neutral atmospheric delay correction	***neutat_delay_total***
surf_type	Five-element array in the geolocation group indicating the surface masks that apply to that segment	
geoid	Geoid height, EGM2008 in the tide-free system (mean-tide in Rel. 3 and earlier)	***geoid***
geoid_free2mean	Conversion factor to be added to *geoid* to convert to a mean-tide system	

podppd_flag	Indication of degraded pointing or orbit determination. Possible values are: : 0=NOMINAL; 1=POD_DEGRADE; 2=PPD_DEGRADE; 3=PODPPD_DEGRADE; possible CAL values are: 4=CAL_NOMINAL; 5=CAL_POD_DEGRADE; 6=CAL_PPD_DEGRADE; 7=CAL_PODPPD_DEGRADE.	***podppd_flag***

Table 3 Input parameters from Other High Level ICESat-2 Products (Source: ATL09 & ATL07)

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	
asr_cloud_probability	Cloud probability based on Apparent Surface Reflectivity p=(1-asr/t)*100	
apparent_surf_reflec	Apparent Surface Reflectivity (25Hz)	
Layer_flag	Represents presence (1) or absence (0) of significant cloud layers.	
ATL07 Label gtxx/sea_ice_segments/heights, geophysical, and stats	Description	Symbol
delta_time	Elapsed GPS seconds since the ATLAS SDP epoch	
latitude	Latitude, WGS84, North=+, Lat of segment center	
longitude	Longitude, WGS84, East=+, Lon of segment center	
seg_dist_x	Along-track distance from the equator crossing to the segment center	
height_segment_length_seg	Along-track length of segment	
height_segment_height	Mean height of surface in the segment	
height_segment_ssh_flag	Identifies whether segment is a candidate bright lead sea surface (=1) or ice (=0)	
height_segment_fit_quality_flag	With value = 5 identifies ATL07 photon heights that are not fit well by the assumed distribution shape and are unsuitable for SSH or DOT determination and ATL10 and ATL12 will only consider heights for DOT for which *fit_quality_flag* = 1, 2, 3, or 4.	
height_segment_surface_error_est	Error estimate for *height_seg_height* equal to standard deviation of heights/(sqrt(n_photons_used))	
n_photons_used	Number of photons used in the trimmed histogram fit used to determine ATL07 sea ice segment average	

height_segment_dac	dynamic atmospheric correction DAC includes an inverted barometer IB effect	
height_segment_dynib	Inverted barometer effect calculated from dynamic mean ocean sea level pressure (computed using ANC10 and ANC48, */ancillary_data/sea_ice/mean_ocean_slp*)	
height_segment_geoid	Geoid height above the WGS 84 reference ellipsoid (range -107 to 86m) based on the EGM 2008 model. The geoid height is in the tide free system.	
height_segment_geoid_free2mean	Additive value to convert geoid heights from the tide-free system to the mean-tide system. Add to height_segment_geoid to get the geoid heights in the mean-tide system	
height_segment_mss	Mean sea surface height above WGS 4 reference ellipsoid (range: -105 to 87 m) based on the DTU 13 model	
height_segment_height	Mean height from along-track segment fit determined by the sea ice algorithm. The sea ice height is relative to the tide-free MSS.	
hist_w	Standard deviation of the trimmed observed photon height histogram (from J. Wimert, 1/4/2024)	

5.2.2 Parameters Required from Other Sources

Bathymetry – Because the ocean mask extends into non-ocean areas we need a way to determine which parts of the ocean data are over water for which our analysis procedure is appropriate. For this we will take the ocean depth, ***depth_ocn***, from the General Bathymetric Chart of the Oceans (GEBCO) 2022 gridded data set (https://www.gebco.net/data_and_products/gridded_bathymetry_data/) evaluated at the geo-segment ***latitude*** and ***longitude***.
Ice Concentration – Because the ocean mask extends into sea ice covered oceans (the province of ATL07 and ATL10), in gridding and other analysis we seek ways of reconciling ATL12 ocean surface heights with DOT with ATL10 freeboard and ocean surface heights as measured in open water regions amongst the sea ice. To do this we need to output ice concentration, ***ice_conc***, in ATL12 at the average latitude and longitude of each ocean segment. These values of sea ice concentration are to be from the same sources used in the derivations of the ATL07 and ATL10 products found as daily values of ice concentration from the high resolution AMSR2 passive microwave SIC dataset for filtering, but maintain both AMSR2 to now be called ***ice_conc*** in ATL12 and the National Snow and Ice Data Center: Near-Real-Time NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 1 to be called *ice_conc_SSMI*, available at https://nsidc.org/data/g10016. Ocean segment mean latitude and longitude, ***gtx/ssh_segments/latitude*** and ***longitude,*** should be converted to the NSIDC Northern and Southern Hemisphere polar stereographic grids and these grid locations used to interpolate the ice concentration, ***gtx/ssh_segments/stats/ice_conc*** and ***gtx/ssh_segments/stats/ice_conc_SSMI.***

5.2.3 Control Parameters for ATL12 Processing

Bathymetry – The course surface finding described in 5.3.1 will ignore data collected in ocean waters with a depth, ***depth_ocn***, less than a depth, ***depth_shore***, defining the

effective shoreline of water for which our analysis procedure is appropriate. The default value for ***depth_shore***, will be 10 m.

Cloud cover – In the course surface finding described in 5.3.1 we may need to ignore data collected in ocean waters where and when the cloud cover is heavy as indicated by ***layer_flag***, equal to 1. The control parameter ***layer_swtch*** will specify if the cloudiness test is to be applied. If control parameter ***layer_swtch*** is equal to 1 and ***layer_flag*** is equal to 1, photon heights will not be included in the ATL12 processing. When ***layer_swtch*** equals the default value, 0, the software will ignore the cloud cover in accepting data during coarse surface finding.

Ice concentration - Sea ice concentration- Two control parameters ***sic_filter_ssmi*** (default 0, not used) ***sic_filter_amsr2*** (default 1, use).

These control parameters are output in *ancillary_data/ocean* for each granule/file ATL12.

5.3 Processing Procedure

5.3.1 Coarse Surface Finding and Setting of Segment Length

Table 4 Control parameters – Coarse surface finding

Parameter	Description	Value
B_c	Bin size of coarse histogram	5 cm
Th_Nc_c	Number of photons times median required to classify as surface bin	1 typical
Th_Ps	Minimum number of candidate surface photons per segment	8,000
Segmax	Max number of 400-pulse segments in a surface finding segment	25
layer_swtch	Flag determining if layer_flag will (*layer_swtch*=1) or won't (*layer_swtch* =0) be editing photon heights.	Default = 0
depth_shore	Prescribed depth defining (*depth_ocn>depth_shore*) boundary of ocean processing	Default= 10 m

The process is basically to first establish a variable length (e.g., up to 7 km) ocean segment with enough (e.g.) photon heights that will yield a histogram with reasonable moments up to 4th order. We start with 14 geo-bin (400-pulse) segments of ATL03 photon heights (see 5.2.1)

(A) Ocean Data Selection and Basic Geophysical Corrections

---

1) Stepping through 14-geobin segments, examine ocean depth and cloud parameters against control parameters to test for suitability for ocean processing. If ocean depth, *depth_ocn*, is less than a depth, *depth_shore*, specifying the effective shoreline of water for which ocean processing is appropriate, then proceed to next 14 geo-bin (400-pulse)

segment. The default value for control parameter *depth_shore* will be 10 m.

Similarly, if control parameter *layer_swtch* is equal to 1 and *layer_flag* is equal to 1, then proceed to next 14 geo-bin (400-pulse) segment. When *layer_swtch* equals the default value, 0, the software will ignore the cloud cover in accepting data during coarse surface finding.

2) Correct for photon heights for short period and long period ocean tides (*gtxx/geophys_corr/tide_ocean and tide_equilibrium*), the inverse barometer effect and the modeled dynamic response to the atmosphere (dynamic atmospheric correction, *gtxx/geophys_corr/dac*).
3) Select only photon heights for which the signal confidence (*gtx/heights/signal_conf_ph*) is greater than or equal to 1 and photon quality (*gxtx/heights/quality_ph*) is equal to zero. Starting with Release 7 we will also accept photons for which *quality_ph* =10 indicating ppartial saturation and for which a first-photon-bias correction will be derived. This will select photons in the likely surface ± 15-m buffer window. Also delete any photon heights outside the band defined by geoid plus or minus half the height of the ocean downlink window (presently geoid ±15-m, to become geoid ± 20-m). (**Is the height of the downlink band reported in ATL03 and if so, what is the variable name?**)
4) To narrow the range of variability, subtract the EGM2008 geoid height (***gtxx/geophys_corr/geoid***), varying over meters to tens of meters, from the photon heights so that we are dealing dynamic ocean topography varying over centimeters to 10s of centimeters. from photon heights to reference all heights to the geoid. Note that as of Release 4, EGM2008 provided by ATL03 is in the tide-free system, but ATL03 also includes a correction mean-tide system, which should be added to the tide-free EGM2008 to obtain the mean-tide EGM2008.

(B) Establish a Working Histogram Array

Establish an initial coarse histogram array, $\boldsymbol{H_c}$, spanning ± 15 m with bin size $\boldsymbol{B_1}$ equal to 0.01 m with center bin centered at zero and bin centers at whole centimeters. Also establish a data array, $\boldsymbol{A_{coarse}}$, for up to 10,000 photon heights and associated information (index, geolocation, time) plus noise photon counts. This will be populated with data as we step through 14-geobin segments searching for an adequate number of surface photon candidates.

(C) Step Through 14-Geo-bin Segments

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 $\boldsymbol{B_1}$ (e.g., 0.01 m). This is used to estimate a running total number of signal and noise photons. We suggest a bin size of 0.01 for consistency with later steps in the processing. The 14 geo-bin (400-pulse) segment should be aligned so that the once per 14 geo-bin (400-pulse, approximately 280 m along-track distance) cloudiness flag, ***layer_flag*** in ATL09, represents the aggregate effect of cloudiness and background radiation derived from solar zenith, ***asr_cloud_probability***, ***cloud_flag_atm***, and ***bsnow_con***, during the 280-m segment.

(D) Test for Surface Photons

Photons in the 14 geo-bin (400-pulse) histogram bins with greater than the ***Th_Nc_c*** 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_c*** is TBD, but we have been using ***Th_Nc_c*** = 1 in our tests with early ICESat-2 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.

Table 5 Control parameters – Fine surface finding and analysis

B_f	Bin size of fine histogram	5 cm, MABEL 1 cm, ICESat-2
Th_Nc_f	Surface qualified= smooth hist > *Th_Nc_f* * tail noise	1.5, ICESat-2
pts2bin	Bins in boxcar smoother	21, ICESat-2
conf_lim	Minimum confidence level to be included in moving average	Typically, 3 or 4 for ICESat-2
nphoton	Number of photons either side of central photon to consider averaging	5 for an 11 point average
nharms	Number of harmonics to fit to selected surface photons	32 default
gaplimit	Gap between photons above which gaps are filled with white noise for harmonic fit	3.2 m default

(E) Increase the Count of Surface Photons

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

(F) Add All Photons to Data Array

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

(G) 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*** (8,000) photons for a strong beam or ***¼ Th_Ps*** (2,000) photons for a weak beam, or if the number, ***Nseg400***, of 14 geo-bin (400-pulse) segments collected reaches S_{egmax} (e.g., 25 equivalent to ~7 km)), go on to 5.3.2 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*** for a strong beam or ***¼ Th_Ps*** for a weak beam, intake another 14 geo-bin (400-pulse) segment and repeat steps 5.3.1 A-G 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.

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. 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***, ***lon_initial,*** a vector of along-track distance, ***trackdist_initial***., and the vector of confidence level, ***conf_initial***, for each photon height.

(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 in steps of 1cm. Also establish the vector, ***Cntrpt***, of bin center points with length equal to ***N***. In practice ***Bf*** equals 1 cm, ***Cntrpt*** values are whole centimeters, and ***Edges*** are at half centimeter values. Also establish a vector array, ***BinB***, as long as ***ht_initial*** for bin assignments.

(B) Compute Moving Average of High Confidence Surface Photon Heights

---

Here we want a moving average of high confidence surface heights, ***mvng_avg,*** with a value corresponding to each photon raw photon height in ***ht_initial***. For the value of ***mvng_avg,*** compute the average of the photon heights in ***ht_initial*** within ***nphoton*** either side of the central photon, including in the average only heights with ***conf_initial*** greater than or equal to ***conf_lim***. For the first ***nphoton*** photons, pad ***mvng_avg*** with the average for the ***nphoton*** + 1 photon in the sequence. For the last ***nphoton*** photons, pad the running average with the average for the last photon minus ***nphoton***. For example, in testing we used ***nphotons*** = 5 and ***conf_lim*** = 3 to make an 11-point running average centered on each photon height, and included in each average, only heights with confidence level 3 and 4.

C) Compute the Anomaly in Height About the Moving Average

For each height in ***ht_initial***, compute ***anom_initial*** equal to ***ht_initial*** minus ***mvng_avg.***

(D) 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 ***anom_initial*** that are in the corresponding bin boundaries of ***Edges***. Also, for each value in ***anom_initial***, populate the vector ***BinB*** with the bin number to which ***anom_initial*** is assigned from one to the length, ***LN***, of ***N*** (in Matlab, [N, BinB]=histcounts(anom_initial, Edges).

(E) Find Median of Initial Histogram

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

(F) 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*** where ***pts2bin*** is a control parameter in Table 5. It should equal ***nbin*** +1 where ***nbin*** is an even number. We have chosen ***nbins***=20 and ***pts2bin*** set at 21 based on success with the developmental code. In this ATBD with 1-cm bin size, we refer to this nominally as the "20-cm smoother" although it actually smooths over 21 bins. (In the ASAS 5.1 code used for ICESat-2 Rel. 001, ***nbin*** equals 100 corresponding to ***pts2bin*** = 101 and a 100-cm smoother).
- 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.

(G) Find Limits of Valid Histogram

Find the limits of the histogram above a preliminary ***jlow*** and below a preliminary ***jhigh*** for which, searching from the index of the maximum in the 20-cm smoothed histogram ***smoothN***, the value in ***N*** is greater than ***Th_Nc_f*** *median of ***N***.

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

- Increment the index, ***i***, upward from ***imax*** until ***N(i)*** is greater than the median of ***N***. At this point preliminary ***jhigh*** = ***i***-1.
- Then the average counts in all the bins above the preliminary J_{high} are calculated and set equal to ***tailnoisehigh*** and
- The average counts in all the bins below the preliminary J_{low} are calculated and set equal to ***tailnoiselow.***
- Increment the index, ***i***, downward from ***imax*** until ***smoothN(i)*** is less than the ***Th_Nc_f * tailnoiselow.*** At this point ***jlow*** = ***i***+1.
- Increment the index, ***i***, upward from ***imax*** until ***smoothN(i)*** is greater than ***Th_Nc_f * tailnoisehigh***. At this point ***jhigh*** = ***i***-1.

(H) 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***.

(I) 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.***

(J) 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.,

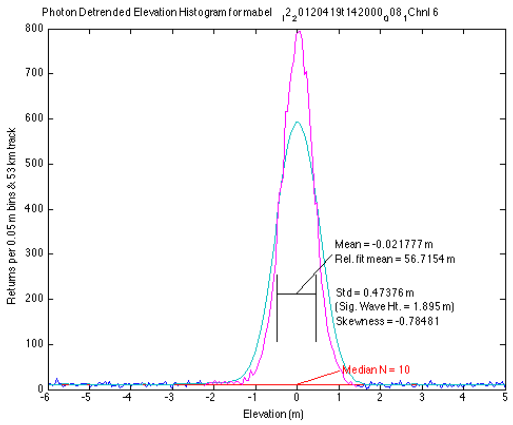

Figure 14. Coarse (blue), smoothed (cyan), and final product (magenta) histograms of detrended, 10 minutes of April 2012 MABEL data over Chesapeake Bay. The magenta histogram is ultimate output of the surface finding phase of the ocean analysis steps 3-6 of 4.2.1.3. The corresponding figure without trend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

ht_initial2=***ht_initial*** - (***P1*** x ***trackdist_initial***+ ***P0***).

REPEAT SURFACE FINDING AFTER REMOVING TREND

In steps (K) through (P) below we repeat Steps (B) Through (H) Acting on the Detrended Elevations Starting With: Compute Moving Average of High Confidence Surface Photon Heights

(K) Compute Moving Average of High Confidence Detrended Surface Photon Heights
---
Here we want a moving average of high confidence surface heights, ***mvng_avg2,*** with a value corresponding to each photon raw photon height in ***ht_initial2***. For the value of ***mvng_avg2,*** compute the average of the photon heights in ***ht_initial2*** within ***nphoton*** either side of the central photon, including in the average only heights with ***conf_initial*** greater than or equal to ***conf_lim***. For the first ***nphoton*** photons, pad ***mvng_avg2*** with the average for the ***nphoton*** + 1 photon in the sequence. For the last ***nphoton*** photons, pad the running average with the average for the last photon minus ***nphoton***. For example, in testing we used ***nphotons*** = 5 and ***conf_lim*** = 4, to make an 11-point running average centered on each photon height, but included in each average, only heights with confidence level 4.

L) Compute the Anomaly in Detrended Height About the Moving Average of Detrended Height

---

For each height in ***ht_initial2***, compute ***anom_initial2*** equal to ***ht_initial2*** minus ***mvng_avg2.***

Populate the histogram array ***N*** such that each element equals the number of values from ***anom_initial2*** that are in the corresponding bin boundaries of ***Edges***. Also, for each value in ***anom_initial2***, populate the vector BinB with the bin number to which ***anom_initial2*** is assigned from one to the length, ***LN***, of ***N*** [in Matlab, [N,BinB]=histcounts(anom_initial2,Edges)].

(M) Find Median of Initial Histogram

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

(N) 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*** where ***pts2bin*** is a control parameter in Table 5. It should equal ***nbin*** +1 where ***nbin*** is an even number. We have chosen ***nbins***=20 and ***pts2bin*** set at 21 based on success with the developmental code. In this ATBD with 1-cm bin size, we refer to this nominally as the "20-cm smoother" although it actually smooths over 21 bins. (In the ASAS 5.1 code used for ICESat-2 Rel. 001, ***nbin*** equals 100 corresponding to ***pts2bin*** = 101 and a 100-cm smoother).

- 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.

(O) Find Limits of Valid Histogram

Find the limits of the histogram above a preliminary ***jlow*** and below a preliminary ***jhigh*** for which, searching from the index of the maximum in the 20-cm smoothed histogram ***smoothN***, the value in ***N*** is greater than median of ***N***.

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

- Then the average counts in all the bins above the preliminary J_{high} are calculated and set equal to ***tailnoisehigh*** and
- The average counts in all the bins below the preliminary J_{low} are calculated and set equal to ***tailnoiselow*.**
- Increment the index, ***i***, downward from ***imax*** until ***smoothN(i)*** is less than the ***Th_Nc_f* tailnoiselow*.** At this point ***jlow*** = ***i***+1.
- Increment the index, ***i***, upward from ***imax*** until ***smoothN(i)*** is less than ***Th_Nc_f* tailnoisehigh***. At this point ***jhigh*** = ***i***-1.

(P) 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_initial2_surf***, in ***trackdist2_initial*** at the indices in ***ii***, [i.e., in Matlab code ***trackdist_initial2_surf*** = ***trackdist2_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_initial2 _surf***. The received histogram, *Nrs***,** is then the histogram of ***ht_initial2_surf*** only including from the lowest height with count>0 to the highest height with count>0.

An example of the second pass and the histogram of surface heights for MABEL data are shown in Figure 14.

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

The number of surface reflected photons in the ocean 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 maximum value of ***trackdist_initial2(ii)*** - the minimum value of ***trackdist_initial2(ii)***.

The total number of photons, ***n_ttl_photon***, equals the length of ***time_initial***, and as above, the number of photons reflected from the surface equals ***n_photons*** equal to the length of ***ii***. The surface reflected photon rate per meter, ***photon_rate*** = r_{surf}, is equal to ***n_photons*** divided by ***length_seg,*** and the noise photon rate per meter, ***photon_noise_rate*** = r_{noise}, is equal to the difference of ***n_ttl_photon*** minus ***n_photons*** divided by ***length_seg***.
Compute ***seg_mean_dist_x*** as the ocean segment average of the along-track distance from the equator crossing to the start of the 20-m geolocation segments included in the ocean segment.

5.3.3 Processing to Characterize Long Wavelength Waves, Dependence of Sample Rate on Long Wave Displacement, *A Priori* Sea State Bias Estimate, First Photon Bias, and Sea Surface Height in Ice-Covered Waters

5.3.3.1 Processing Steps to Estimate Sea State Bias

In this part we compute the Sea State Bias proportional to the covariance of photo return rate and surface height in 10-m along-track bins. This constitutes what is commonly referred to as the Electro-Magnetic or EM bias that is expected to dominate ICESat-2 SSB. The steps are:

1. Add ***meanoffit2*** to ***ht_initial2_surf***.
2. Convert ***trackdist_initial2_surf*** to meters if necessary and subtract the minimum of ***trackdist_initial2_surf*** to obtain distance from start of the ocean segment. Save and output as ***xbind_first_dist_x*** this start position equal to the minimum of ***trackdist_initial2_surf***.
3. Sort height data in ascending order of along-track distance, ***trackdist_initial2_surf,*** to allow calculation of photon return rate per meter
4. Define ***x*** as the sorted ***trackdist_initial2_surf*** and define ***hty*** the associated heights from ***ht_initial2_surf***.
5. Do the following analysis for each 10-m bin in the ocean segment. In practice 710 bins to cover the max 7-km ocean segment length. For bins with no photon returns, the bin-averaged variables (***xbind, htybin***, ***xrbin***, ***slopebinfit, latbind, lonbind***) will be set to NaNs (not a number)
   a) Set ***xbin*** to middle of 10-m bin.
   b) Set ***xbind*** to the mean of ***x*** in each 10-m bin.
   c) Set ***latbind*** to the mean of photon latitude in each 10-m bin.
   d) Set ***lonbind*** to the mean of photon longitude in each 10-m bin
   e) Set ***htybin*** to mean of ***hty*** in each 10-m bin.
   f) Set *htybin_std* to standard deviation of ***hty*** in each 10-m bin, If the number of photons in each 10-m, ***nbind*** <2, set ***htybin_std*** to NaN.
   g) Compute the photon data rate per meter, ***xrbin,*** by dividing the number of photons in each 10-m, ***nbind***, bin by 10.
   h) Establish ***slopebinfit***:
      I. = NaN if only one photon in a 10-m bin
      II. = change in ***hty*** over change in ***x*** if only two photons in a 10-m bin

III. = linear fit if three or more photons in a 10-m bin

6. Compute covariance of bin averaged height and data rate, ***binCOVHtXr***, as the average over all 10-m bins of the product of the detrended binned heights, ***htybin***, and the detrended binned data rate, ***xrbin***.
7. Compute, ***binAVG_Xr***, the average of binned data rate, ***xrbin,*** over the 10-m bins with photons.
8. Compute the sea state bias, ***bin_ssbias***, as ***binCOVHtXr*** divided by ***binAVG_Xr***.
9. If there are more than one 10-m bin in an ocean segment with a valid ***slopebinfit***:
   a) Compute ***slopeXreturn*** as the product of the detrended slope, ***slopebinfit***, and the detrended binned data rate, ***xrbin***.
   b) Compute the covariance of the photon rate and the slope, ***binCOVslope_Xr***, as the sum of the ***slopeXreturn*** divided by the number of 10-m bins with a valid ***slopebinfit***.
   c) Compute the sea slope bias, ***bin_slopebias***, as the covariance of the data rate and the slope, ***binCOVslope_Xr***, divided by the mean of the data rate.
   d) Compute ***magslopeXreturn*** as the product of the absolute value of the detrended slope, ***slopebinfit***, and the detrended binned data rate, ***xrbin***.
   e) Compute the covariance of the photon rate and the slope, ***binCOVmagslope_Xr***, as the sum of the ***magslopeXreturn*** divided by the number of 10-m bins with a valid ***slopebinfit***.
   f) Compute the sea slope magnitude bias, ***bin_magslopebias***, as the covariance of the data rate and the slope, ***binCOVmagslope_Xr***, divided by the mean of the data rate.
10. Save and output ***bin_ssbias, bin_slopebias, and bin_magslopebias,*** and vector arrays ***xbin***, ***xbind***, ***htybin***, ***htybin_std***, and ***xrbin*** for segment output.

Since the rationale for some of these steps may not be obvious, we provide the following explanations for the steps:

To determine the SSB using Eq. 11 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 each ocean segment found during the coarse signal finding step. Section **5.3.1**.

Estimating sea state bias over ocean 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 ***ht_initial2_surf*** and ***trackdist_initial2_surf*** along each ocean segment found during the coarse signal finding step. Section **5.3.1**

Add mean from detrending used after first cut surface finding

The means from the first cut of surface finding are computed from photon heights over surface waves and are therefore contaminated by the correlation of photon return rate with height of the surface over surface waves. If these are not added back into the height record, that contamination by SSB will be lost. Therefore, we add ***meanoffit2***, the mean of

P0 + ***P1*** x ***trackdist_initial2_surf*** (the along-track position of 2nd iteration surface photons), back into the photon height record, ***hty*** =***ht_initial2_surf*** +***meanoffit2***

Calculate the Photon Return Rate

After experimenting with more complicated schemes using photon spacing to determine photon return rate, we have concluded the return rate that most directly represents the EM sea state bias is to simply divide the number of photon returns in each 10-m bin by 10 to yield the photon return rate per meter in each bin.

Quantify surface height and slope and return rate in 10-m along-track bins.

Establish a vector ***xbin***, the middle of 10-m bins along a length of 7,100 m, i.e., 5:10:7095, where 7,100m is the understood maximum of our ocean segments. Assign each sample interval (*ii*) in the, $\boldsymbol{x}_i$, and $\boldsymbol{hty}_i$, to one of these bins with a bin index *ibin* where *ibin(ii)*=round-up (***x****(ii)*-min(***x***))/10) and cumulatively add the ***hty****(ii)* to ***htybin****(ibin(ii))* and increment a bin counter, *nbinc(ibin(ii))*, by one for each added sample.

Find averages, photon return rate and straight-line fits in each 10-m bin

For *ibin* equal to 1 through ***Nbin10***, the number of photons in each bin, ***nbind****(ibin)* = *nbinc(ibin).* Compute bin averages of surface height, ***htybin***, and data rate, ***xrbin***, as:
Compute the bin averages of height and data spacing:
htybin*(ibin)*= ***htybin****(ibin)/ nbinc(ibin) =* ***htybin****(ibin)/* ***nbind****(ibin)* and
xrbin*(ibin)*= ***nbind****(ibin) /* 10.
Set ***xbind*** to the mean of the along-track positions in each 10-m bin *=* ***x****(ibin)/* ***nbind****(ibin)*.

Also compute fits of surface height slope, ***slopebinfit***, in each bin
- In the cases with only 1 point in a bin, ***slopebinfit*** will be set to NaN (not a number).
- In the case of two points in a bin, ***slopebinfit*** will be equal to the change in ***hty*** over change in ***x***
- If the number of points in a bin is greater than or equal to 3, ***slopebinfit*** will equal the slope of a linear fit with distance over the points.

Compute Sea State Bias using all viable bins.

Compute covariance of bin averaged height and data rate, ***binCOVHtXr***, as the average over all 10-m bins of the product of the detrended binned heights, ***htybin***, and the detrended binned data rate, ***xrbin***, i.e., the sum over viable bins (bins with nonzero ***nbind***), *iii*, from 1 to ***Nbin10*** of (detrended (***htybin*** *(iii))* *detrended(***xrbin***(iii)))/***Nbin10***.

Similarly, compute ***binAVG_Xr***, the average of binned data rate, ***xrbin***, i.e., the average of ***xrbin****(iii)* from *iii*=1 to *Nbin10*.

Compute the bin average sea state bias estimate according to Equation X7:

bin_ssbias equals ***binCOVHtXr*** divided by ***binAVG_Xr***.

Compute Significant Wave Height, SWH
Compute significant wave height (***SWH***) equal to four times the standard deviation of the bin-averaged surface heights in all the10-m along-track bins in the segment.

Compute photon rate slope correlation.
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.

If there are more than one 10-m bin in an ocean segment with a valid ***slopebinfit***, compute ***slopeXreturn*** as the product of the detrended slope, ***slopebinfit***, and the detrended binned data rate, ***xrbin***. Similarly compute ***magslopeXreturn*** as the product of the magnitude of the detrended slope, ***slopebinfit***, and the detrended binned data rate, ***xrbin***.

Then compute the covariance of the photon rate and the slope, ***binCOVslope_Xr***, as the sum of the ***slopeXreturn*** divided by the number of 10-m bins with a valid ***slopebinfit*** and the covariance of the photon rate and the magnitude of slope, ***binCOVmagslope_Xr***, as the sum of the ***magslopeXreturn*** divided by the number of 10-m bins with a valid ***slopebinfit***.

Compute the sea slope bias, ***bin_slopebias***, as the covariance of the data rate and the slope, ***binCOVslope_Xr***, divided by the mean of the data rate. Similarly compute the sea slope magnitude bias, ***bin_magslopebias***, as the covariance of the data rate and the slope, ***binCOVmagslope_Xr***, divided by the mean of the data rate.

Save variables ***bin_ssbias***, ***bin_slopebias,*** and ***bin_magslopebias*** vector arrays ***xbin, xbind***, ***htybin***, *htybin_std,* and ***xrbin*** for segment output.

5.3.3.2 First Photon Bias Determination

In this part we compute the first photon bias using deadtime tables from ATL03, the photon rate in each 10-m bin, and the standard deviation of photon heights in each 10-m bin. Beyond calculating photon rate and width of the height distribution in 10-m bins, the method is taken from the ATL03 ATBD. The steps are:

1. For each strong beam of an ATL03 granule, establish the deadtime (X-index) for the ATL03 granule by computing a mean of the first 16 items in the table found in: *ancillary_data/calibrations/dead_time/gtx/dead_time.* Call this ***deadtm_mean*** and convert it into nanoseconds by multiplying by 10^9. Use this value to ascertain the corresponding row for the table: *ancillary_data/calibrations/first_photon_bias/gtx/dead_time*. Call the corresponding row number ***index_x***.
2. For each weak beam in an ATL03 granule, compute the dead time by computing a mean of the last 4 items in the table found in: ancillary_data/calibrations/dead_time/gtx/dead_time. Call this ***deadtm_mean*** and convert it into nanoseconds by multiplying by 10^9. Use this value to ascertain the corresponding row for the table:

ancillary_data/calibrations/first_photon_bias/gtx/dead_time. Call the corresponding row number index_x.

3. For each beam and for each 10-m bin, do the following:
   a) Establish the Y-index for return strength.
   Compute the return strength as 0.7 * xrbin. Call this return_strength. Use this value to ascertain the corresponding column in the ATL03 table: ancillary_data/calibrations/first_photon_bias/gtx/strength. Note that this table has six rows (one for each beam), thus index_x is also required for this table look up. Call the corresponding column number, index_y.
   b) Establish the Z-index for width.
   Compute the width as 2.355(htybin_std) / (c/2) where c is the speed of light (299,792,458 ms^{-1}) and convert to nanoseconds. Use this value to ascertain the corresponding column in the ATL03 table: ancillary_data/calibrations/first_photon_bias/gtx/width. Note that this table has six rows (one for each beam), thus index_x is also required for this table look up. Call the corresponding column number, index_z.
4. Given the indices, ***index_x***, ***index_y*** and ***index_z***, associated with each 10-m bin and each beam, use these indices to obtain (by table look-up) the first photon bias. The table is found in ancillary*_data/calibrations/first_photon_bias/gtx/ffb_corr*. The value corresponding to the indices is provided in picoseconds (1ps = 10^{-12} s) and needs to be converted into seconds, then scaled by the speed of light (in meters per second) divided by 2 (*c*/2) to give the ***fpb_10m*** value. [This is written to the product as a two-dimensional parameter, akin to htybin.]
5. For each ocean segment, compute, save, and output the mean of ***fpb_10m***. This will become the parameter ***fpb_corr*** for each ocean segment, for each beam, on ATL12. When it is non-zero, ***fpb_corr*** can be from the ocean segment average sea surface height to correct for first photon bias.
6. For each ocean segment, compute the standard deviation of ***fpb_10m***. This will become the parameter ***fpb_corr_stdev***, for each ocean segment, for each beam, on ATL12.
7. Output ***fpb_corr*** and ***fpb_corr_stdev*** and the vector array ***fpb_10m*** for each ocean segment.

5.3.3.3 Harmonic Analysis of Along-track Heights With Gap Determination and Filling

Starting with the ***n_photons*** points in the along-track sorted ***trackdist_initial2_surf***, and corresponding ocean heights in ***ht_initial2_surf*** with ***meanoffit2*** added back in, our aim is to compute the harmonic coefficients for ***nharm*** harmonics plus the zero-frequency coefficient best representing the sea surface.

The Vaníček method to be used is a least squared error technique that is not dependent on evenly space data. However, we have found that significant gaps in the data can result in wide and unrealistic variations in the resulting harmonic coefficients. This is because under the technique, the coefficients can vary unrestrained to fit the existing data, while producing wildly unrealistic Fourier series in the gaps where unconstrained by data. The

solution to this problem is to find and fill the data gaps with white noise from a Gaussian distribution with mean equal to meanoffit2 and standard deviation equal to that of the existing data. The first step is identifying significant data gaps in an ocean segment.

1) Letting ***X(j)*** be each value of ***trackdist_initial2_surf(j)***, if we take the difference, ***DX(1 to N-1)*** between ***X*** (2to***N***) and ***X*** (1 to ***N***-1), it will be the vector of spaces between points, and we will know that there is no data for ***DX(j)*** meters after ***X(j).*** We can think of the ***DX*** as a random variable so compute the sample median, mode, mean (***DXbar*** = Sum ***DX*** from 1 to ***N***-1 divided by ***N***-1), sample variance (***DXvar*** = Sum (***DX-DXbar***)^2 from 1 to ***N***-1 divided by ***N***-2), and sample skewness (***DXskew*** = Sum (***DX-DXbar***)^3/ ***DXvar***$^{3/2}$ from 1 to ***N***-1 divided by ***N***-2)). If there are no major gaps, the mean will be 0.4 to 0.7 m. If there are many small gaps, the variance and skewness will increase but the mean, mode, and median will probably stay close to 1 m. If there is one big gap, the skewness should become large. It will be helpful on many fronts to compute these statistics for ocean segments as indications of gaps in the ocean segment data. For example, in our tests with one ocean granule, we found unreasonable harmonic coefficients came from the Vaníček method only when ***DXskew*** >10. The user could ignore ocean segment harmonic coefficients and/or ocean heights altogether if these indicators suggested there were many gaps in the data.
2) Alternatively, for the purposes of computing harmonic coefficients we will fill any gaps greater than ***gaplimit*** (e.g., ***gaplimit*** =3.2 m) with random numbers every 0.7 m drawn from a Gaussian distribution with mean equal to ***meanoffit2*** and standard deviation equal to the standard deviation of ***ht_initial2_surf***. To do this:
3) Identify the ***DX(l)*** > ***gaplimit*** and for each create gap-filling vectors of corresponding distances vectors, ***xgap,*** every 0.7 m from ***X***(l) to **X**(l)+**DX**(l)/0.7 and of heights, ***hgap*** chosen as random numbers from a Gaussian distribution with mean equal to ***meanoffit2*** and standard deviation equal to the standard deviation of ***ht_initial2_surf***. For each the ***DX(l)*** > ***gaplimit*** concatenate the ***xgap*** and ***hgap*** onto the previous ***xgap*** and ***hgap*** vectors until the final ***DX(l)*** > ***gaplimit*** is reached.
4) When the final ***DX(l)*** > ***gaplimit*** is reached, concatenate ***xgap*** onto the end of ***trackdist_initial2_surf*** to make ***xtemp*** and ***hgap*** onto the end of ***ht_initial2_surf*** to make ***htemp.*** Sort both x***temp*** and ***htemp*** on ascending order in ***xtemp*** to insert the filled gaps in their proper order versus along-track distance. Set ***x***=***xtemp*** and ***y***=***htemp***, and the resulting gap-filled records will be the ***x*** and ***y*** in the harmonic coefficient determination below. The length of the gap-filled ***x*** is called ***n_photontemp***. Reference ***x*** to the start by subtracting ***x***(1) from each ***x*** value. Note that the gap-filled ***x*** and ***y*** will only be used for computing harmonics and will not be used for the other calculations of 10-m bin averages and sea surface height statistics.
5) For ***nharm*** harmonics, compute the wavelengths starting with the fundamental wavelength equal to the maximum of ***x*** minus the minimum of ***x***. The vector of harmonic wavelengths, ***lwave***, is given by ***lwave***(*i*)= ***length_seg*** /*i*, *i*= 1 to ***nharms***, and the vector of harmonic wavenumbers is given by inverse of the wavelengths ***wn(i)***=1/***lwave(i)*** , *i*= 1 to ***nharms.***

6) Create matrix ***F*** with ***n_photontemp*** rows and 1+2 * ***nharms*** columns. Populate ***F*** as follows:

   F(*j*=1 to ***n_photontemp***, *i*=1) = 1
   F (*j*, 2**i)* = sin(2*π***wn(i)*x(j)*) for *j*=1 to ***n_photontemp*** and *i*=1 to ***nharms***
   F (*j*, 2**i*+1) = cos(2*π***wn(i)*x(j)*) for *j*=1 to ***n_photontemp*** and *i*=1 to ***nharms***,
   And where ***x(j)*** is each value of gap-filled ***trackdist_initial2_surf(j)***.

7) Form left pseudo inverse and compute coefficients. For ***x*** = the gap-filled ***trackdist_initial2_surf,*** a ***n_photontemp*** x 1 column vector***, F*** is a ***n_photontemp*** rows by ***m***= 1+ 2**nharms* columns matrix. The transpose of ***F***, here called ***F'***, is a ***m*** by ***n_photontemp*** matrix. The matrix product, ***F'F*** is ***m*** by ***m.*** The product of ***F'*** and ***x***= the gap-filled ***trackdist_intial2_surf***, is a ***m*** by 1 vector.

8) For ***y*** equal to the column vector of gap-filled ***ht_initial2_surf*** with ***meanoffit2*** added to it, versus x, the gap-filled along-track distance, compute the vector of harmonic coefficients, ***a***, by the Vaníček least squared error or left pseudo inverse method, where

   a*=*F'F**F'y .

   The ***m*** x 1 vector of harmonic coefficients, ***a***, for the optimum harmonic fit of ***y*** to ***x***:

   yfit = ***a***(1)+
   a (2)*sin(***wn***(1)* 2*π**x***) + ***a*** (3)*cos(***wn***(1)* 2*π**x***)+
   a (4)*sin(***wn***(2)* 2*π**x***) + ***a*** (5)*cos(***wn***(2)* 2*π**x***)+
   a (6)*sin(***wn***(3)* 2*π**x***) + ***a*** (7)*cos(***wn***(3)* 2*π**x***)+
   : : +
   : : +
   a(2**nharms*)*sin(***wn***(***nharms***)* 2*π**x***) + ***a***(2**nharms* +1)*cos(***wn***(***nharms***)* 2*π**x***)

   Backslash or left matrix divide A\B is the matrix division of A into B, which is essentially the same as $A^{-1}B$.

9) Comparison tests between gap-free complete records and identical records to which gaps have been added and then filled with white noise about meanoffit2 as in (1)-(4) have shown that the vector harmonic coefficients, ***a***, are also the optimum harmonic fit of ***ht_initial2_surf*** plus ***meanoffit2*** versus ***trackdist_intial2_surf*** along-track distance, 0 to segment length (***length_seg***). In the gap regions, the harmonic fit makes small oscillations constrained to be near ***meanoffit2***. As a test, with ***y*** and ***yfit*** as row vectors, compute the signal to noise ratio ***SNR_harm = E[(yfit-a(1))*(yfit-a(1))']/E[(y-yfit)*(y-yfit)']*** where ***E[]*** denotes expected value***.*** Over many ocean segments we will compare ***SNR_harm*** with the measures the number and size of record gaps, ***DXbar, DXvar***, and ***DXskew***, to determine the point at which record gaps are so dominant that the harmonic fit is not useable.

5.3.3.4 Sea Surface Height Determination in Ice-Covered Waters

1) For ***ice_conc*** > 15%, use ***xbind_first_dist_x*** to establish the along-track distance boundaries of the 10-m bins.
2) For ***ice_conc*** > 15%, from ATL07, use ***seg_dist_x*** and ***height_segment_length_seg*** to establish the along-track distance boundaries of the ATL07 sea ice segments and use the ATL07 *ssh_flag* = 1 and the *fit_quality_flag* = 1, 2, 3, or 4 to identify which sea ice segments are the sea ice segments are candidate, bright-lead surfaces.

3) By comparing the maps of 10-m bin boundaries and bright-lead sea ice segment boundaries, determine which 10-m bins also represent bright lead surfaces. Both boundaries of the 10-m bin should be in bright-lead segments. Note that we will limit candidate10-m bins for sea surface height determination to those bins for which the photon return rate, *xrbin*, is between 4 and 25 photons per meter (strong beams) or 1 to 7 photons per meter (weak beams) corresponding to quasi-specular but not fully saturated returns. This a wider range than the criteria for ATL07 segments to have *ssh_flag*=1, indicating bright leads.

4) Flag the so identified bright-lead 10-m bins with ***xbin_atl07_dot*** equal to the average of the DOT in the sea ice segments bounding the 10-m bin, ***atl07_dot*** (defined below). The 10-m bin *xbin_atl07_dot* is always set to INVALID where the ATL12 *ice_conc* < 15%. For *ice_conc* greater than or equal to 15%, we do the following:

First, check to see if ATL07 is using the dynamic-IB correction as in Release 006. Then for ATL07 segments with *height_segment_ssh* =1, we define ***atl07_dot*** (using on-product variable names from ATL07) as:

atl07_dot = *heights/height_segment_height* + *geophysical/height_segment_mss* - (*geophysical/height_segment_geoid* +*geophysical/height_segment_geoid_free2mean*) + *geophysical/height_segment_dynib* - *geophysical/height_segment_dac*.

This places the ATL07 height into the same mean-tide geoid-based and DAC-based system as ATL12 variables. If ATL07 is using the DAC correction already:

atl07_dot = *heights/height_segment_height* + *geophysical/height_segment_mss* - (*geophysical/height_segment_geoid* +*geophysical/height_segment_geoid_free2mean*)

If a 10-m bin is coincident with only one ATL07 segment, then ***xbin_atl07_dot*** will simply be the single ***atl07_dot*** value for that ATL07 segment.

If two (or more) qualified ATL07 segments enter, are contained within, or exit a 10m bin, then ***xbin_atl07_dot*** equals the average of ***atl07_dot*** all the coincident sea ice segments.

If both ends of the 10-m segment are not in bright lead ATL07 segments, ***xbin_atl07_dot*** = INVALID.

For each ocean segment, ***h_ice_free*** will be set to the average of the ***htybin-fpb_10m-meanoffit2+ (p0 +p1 x xbind)*** of ***ssh_flag****=1,* bright lead qualified ATL12 10-m bins and output. Similarly for each ocean segment, ***h_ATL07_ice_free*** will be set to the

average of all the ***atl07_dot*** for all the sea ice segments coincident with bright lead qualified 10-m bins will be saved and output.

Note 1: For the open ocean we added *meanoffit2* to the 10-m bin photon heights to allow computation of SSB without considering the trend in reflectance over a typical 7 km ocean segment. Typically, in open water the average slope is small ($<10^{-5}$), but with sea ice with 40-cm ice freeboard at one end of a 400-m ocean segment the slope, ***p1***, can be ten times as great (10^{-4}) and without the linear fit correction, ***(p0 +p1 x xbind)***, added back in and ***meanoffit2*** correction removed, the bias in 10-m bin htybin can be 40 cm.

Note 2: Since the sea ice segments are short (extending about 150 photons) and often overlap by 50%, ***atl07_dot*** values will be found applicable for two or more 10-m bin, ***xbin_atl07_dot*** averages. However, even if a bright lead sea ice segment overlaps more than one ATL12 10-m bin, its value of ***atl07_dot*** should only be counted once in the average of ***h_ATL07_ice_free***. Otherwise, wide leads would be a unfairly overweighted.

Adding ***h_ice_free*** and ***h_ATL07_ice_free*** to the ATL12 product will provide ATL12 and ATL07 renditions of ATL12 ocean segment ice-covered DOT in ATL07 ***ssh_flag***=1 qualified bright leads captured in the ATL12 10-m bins. Agreement between ***h_ATL07_ice_free*** and ***h_ice_free*** will indicate the surface finding of ATL12 and ATL07 in leads give the same surface height.

5) We also identify all the ATL07 sea ice segments with *ssh_flag*= 1 (bright leads) in an ATL12 ocean segment and compute average, ***h_atl07_inatl12oc***, of ***atl07_dot*** from all these ATL07 *ssh_flag* =1 qualified sea ice segments regardless of whether they have a corresponding ATL12 10-m bin. This will be important if there are only leads smaller than 10-m in the span of the ocean segment.

6) We will quantify and output the uncertainty in ***h_ice_free*** from two sources. The first, ***hicefree_uncrtn_inst,*** which over the smooth surface of a lead is likely instrumental and is manifest in the standard deviations of surface photon heights bright lead 10-m bin in the ocean segments. Let σ_i =***htybin_std*** in the i^{th} bright lead 10-m bin of the ocean let there be K bright lead-qualified 10-m bins in the ocean segment with n_i =10****xrbin*** photons in the i^{th} bin. Then the average standard deviation in the K bins, $\bar{\sigma}$, is

$$\bar{\sigma} = \left(\sum_{i=1}^{K} n_i \sigma_i^2 \Big/ \sum_{i=1}^{K} n_i \right)^{1/2}$$

Then the uncertainty in the mean height or standard deviation in the estimate of the mean height, ***h_ice_free,*** due to instrumental noise is ***hicefree_uncrtn_inst***:

$$\textbf{\textit{hicefree\_uncrtn\_inst}} = \bar{\sigma} \left(\sum_{i=1}^{K} n_i \right)^{-1/2}$$

The second, ***hicefree_uncrtn_nat***, is due to the natural variability in the bright lead 10-m bin averages, h_i = ***htybin-fpb_10m*** in the i^{th} bright lead bin and $\bar{h}$ = ***h_ice_free***. It is zero if K <2 but for K greater than or equal to 2, the uncertainty in the estimate of the

mean height is equal to the standard deviation in the bright lead ***htybin*** divided by the square root of the number of bright lead 10-m bins, *K*:

$$\textbf{\textit{hicefree\_uncrtn\_nat}} = \left(\frac{\sum_{i=1}^{K}(h_i-\bar{h})^2}{K}\right)^{\frac{1}{2}} / K^{1/2} = \left(\sum_{i=1}^{K}\left(h_i-\bar{h}\right)^2\right)^{1/2} / K$$

We will output ***hicefree_uncrtn_inst*** and ***hicefree_uncrtn_nat.*** We will also compute and output the total uncertainty, ***h_ice_free_uncrtn***, equal to the square root of the sum of ***hicefree_uncrtn_inst*** squared and ***hicefree_uncrtn_nat*** squared,

$$\textbf{\textit{h\_ice\_free\_uncrtn}} = \left(\textbf{\textit{(hicefree\_uncrtn\_inst)}}^2 + (\textbf{\textit{hicefree\_uncrtn\_nat}})^2\right)^{1/2}$$

7) We also need uncertainties for ***h_atl07_ice_free*** to be called ***h_atl07_ice_free_uncrtn***, and for ***h_atl07_inatl12oc*** to be called ***h_atl07_inatl12oc_uncrtn***. They will be computed in a manner analogous to ***h_ice_free_uncrtn***. Considering instrumental and natural variability.

Instrumental Variability Component: ATL07 gives estimates of the uncertainty in The uncertainties due to instrumental variability in individual sea ice segment heights, ***height_segment_surface_error_est = hist_w /(n_photons_used***)$^{1/2}$, where ***hist_w*** equals the standard deviation of the trimmed observed photon height histogram (outliers are trimmed from the initial return). The ***hist_w*** in ATL07 segments is analogous to ***htybin_std*** in ATL10-m bins and will be used to expand the ATL07 uncertainties to multiple sea ice segments in and ATL12 ocean segment. The formulas are similar to those above for ***h_ice_free_uncrtn*** except ***hist_w*** will be used for standard deviation of photon heights, and the number of photons, n_i, for each ATL07 sea ice segment is given by n_i=***n_photons_used***.

a) For the computation of ***h_atl07_ice_free_uncrtn***, take σ_i = ***hist_w*** in the i^{th} bright lead (*ssh_flag*=1) sea ice segment of K_{10} bright lead sea ice segments contacting 10-m bins and contributing to ***h_atl07_ice_free*** (Note: Per paragraph 4 above, in some cases two or more bright lead sea ice segments might be used for computation of an individual 10-m bin value of ***xbin_atl07_dot***, so K_{10} will be greater than the number of bright-lead 10-m bins. On the other hand no sea ice segment will contribute more than once to the average ***h_atl07_ice_free***.).

b) For the computation of ***h_atl07_inatl12oc_uncrtn***, take σ_i = ***hist_w*** in the i^{th} bright lead (*ssh_flag*=1) sea ice segment of all, K_{all}, bright lead sea ice segments in the ocean segment and contributing to ***h_atl07_inatl12oc***. Then the average standard deviation in the *K* bins, $\bar{\sigma}$, are:

$\bar{\sigma}_{10} = \left((\sum_{i=1}^{K_{10}} n_i\,\sigma_i^2)/\sum_{i=1}^{K_{10}} n_i\right)^{1/2}$ for calculating *h_atl07_icefree_uncrtn*

$\bar{\bar{\sigma}}_{all} = \left(\left(\sum_{i=1}^{K_{all}} n_i \sigma_i^2\right) / \sum_{i=1}^{K_{all}} n_i\right)^{1/2}$ for calculating *h_atl07_inatl12oc_uncrtn*

The uncertainty in the estimates of the mean heights, ***h_atl07_ice_free*** and ***h_atl07_inatl12oc*** due to instrumental noise would be equal to the average standard deviation divided by the square root of the number of sample points. For example, in the case of K_{10} sea ice segments, the number of sample points would equal $\sum_{i=1}^{K_{10}} n_i$. However, sea ice segments overlap, and in cases where adjacent and overlapping sea ice segments are needed to qualify a 10-m bin as a bright lead, or anywhere a bright lead appears in consecutive sea ice segments, photons are used twice, and their heights are not independent. In an effort to account for this in computation of instrumental uncertainty we derate each photon height by a factor of 2.

$$\textbf{\textit{h07icefree\_uncrtn\_inst}} = \bar{\bar{\sigma}}_{10}\left(\sum_{i=1}^{K_{10}} n_i/_2\right)^{-1/2}$$

$$\textbf{\textit{h7in12icefree\_uncrtn\_inst}} = \bar{\sigma}_{all}\left(\sum_{i=1}^{K_{all}} n_i/_2\right)^{-1/2}$$

IMPORTANT NOTE: If and only if in the future, ATL07 sea ice segments are changed to non-overlapping, do not derate the each photon height by 2 so these uncertainties would be $h07icefree\_uncrtn\_inst = \bar{\bar{\sigma}}_{10}\left(\sum_{i=1}^{K_{10}} n_i\right)^{-1/2}$ and $h7in12icefree\_uncrtn\_inst = \bar{\sigma}_{all}\left(\sum_{i=1}^{K_{all}} n_i\right)^{-1/2}$.

Natural Variability Component: The uncertainties due to the natural variability in the bright lead sea ice segment averages are taken to be the standard deviation of the sea ice segment averages. These are important to consider if the surface is rougher than that indicated by the uncertainty due to instrumental noise.

For the computation of natural component of ***h_atl07_ice_free_uncrtn***, take $\bar{h}$= ***h_atl07_ice_free*** and considering the i^{th} of K_{10} bright lead sea ice segments with corresponding 10-m bins, take h_i equal to the corresponding values of ***ATL07_dot***,

and:

$$\textbf{\textit{h07icefree\_uncrtn\_nat}} = \left(\sum_{i=1}^{K_{10}}\left(h_i - \bar{h}\right)^2\right)^{1/2} / K_{10}$$

For computation of the natural component of ***h_atl07_inatl12oc_uncrtn***, take $\bar{h}$= ***h_atl07_inatl12oc,*** and considering the i^{th} of K_{all} bright lead sea ice segments in all of the ocean segment, take h_i equal to the contributing values of ***ATL07_dot***.

and:

$$\textbf{\textit{h7in12icefree\_std \_nat}} = \left(\sum_{i=1}^{K_{all}}\left(h_i - \bar{h}\right)^2\right)^{1/2} / K_{all}$$

We will output ***h07icefree_uncrtn_inst*** and ***h07icefree_uncrtn_nat*** and ***h7in12icefree_uncrtn_inst*** and ***h7in12icefree_uncrtn_nat.***

We will also compute and output the total uncertainty, ***h_atl07_ice_free_uncrtn*** equal to the square root of the sum of ***h07icefree_uncrtn_inst*** squared and ***h07icefree_uncrtn_nat*** squared,

$$h\_atl07\_ice\_free\_uncrtn = ((h07icefree\_uncrtn\_inst)^2 + (h07icefree\_uncrtn\_nat)^2)^{1/2}$$

and the total uncertainty, ***h_atl07_inatl12oc_uncrtn*** equal to the square root of the sum of ***h07icefree_uncrtn_inst*** squared and ***h07icefree_uncrtn_nat*** squared,

$$h\_atl07\_inatl12oc\_uncrtn = ((h7in12icefree\_uncrtn\_inst)^2 + (h7in12icefree\_uncrtn\_nat)^2)^{1/2}$$

(Note: These combinations will represent slight overestimates of uncertainty if the natural uncertainty is equal to the instrumental uncertainties)

8) To aid in the interpretation of ***h_ice_free*** and ***h_atl07_ice_free***, we will also calculate and output an ice concentration based on the fractional area of bright-lead 10-m bins in an ocean segment: ***ice_conc_12icefreebins*** will be equal to complement (difference from 1) of the square of the fraction relative to the total number of 10-m bins in the ocean segment of 10-m bins classified as covering bright leads by ATL07, ***ssh_flag***=1 . Express as a percentage.

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.

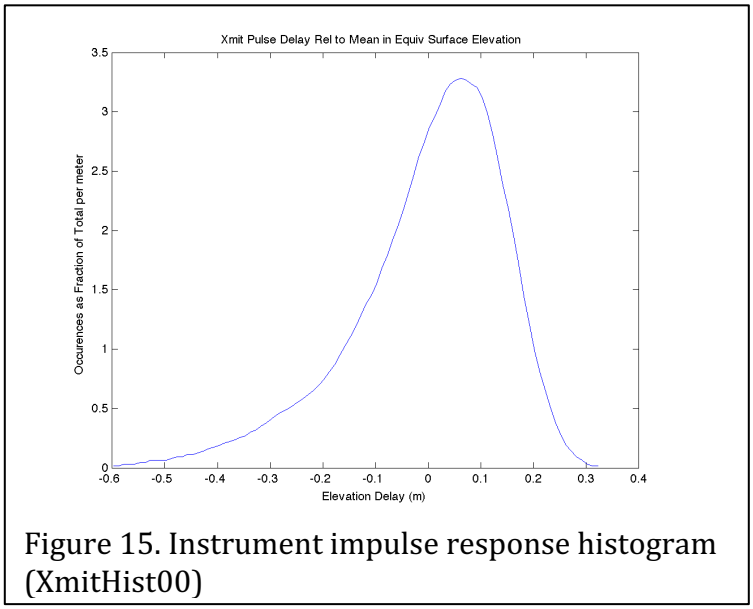

Figure 15. Instrument impulse response histogram (XmitHist00)

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. 15) with a synthetic sea surface height, 2-Gaussian mixture distribution with 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 the true surface distribution, namely Wiener deconvolution to remove the instrument impulse response distribution. This is where the processing of ATL03 histograms begins. Part (H) in 5.2.5 (Characterizing the Random Sea Surface) breaks the processed histogram into a Gaussian Mixture and computes the aggregate distribution statistics for ATL12 using equation 17.

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, ***Nrs***, from section (5.3.2 (P)) 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 12^{th} 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 ***smoothrechist*** 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***).

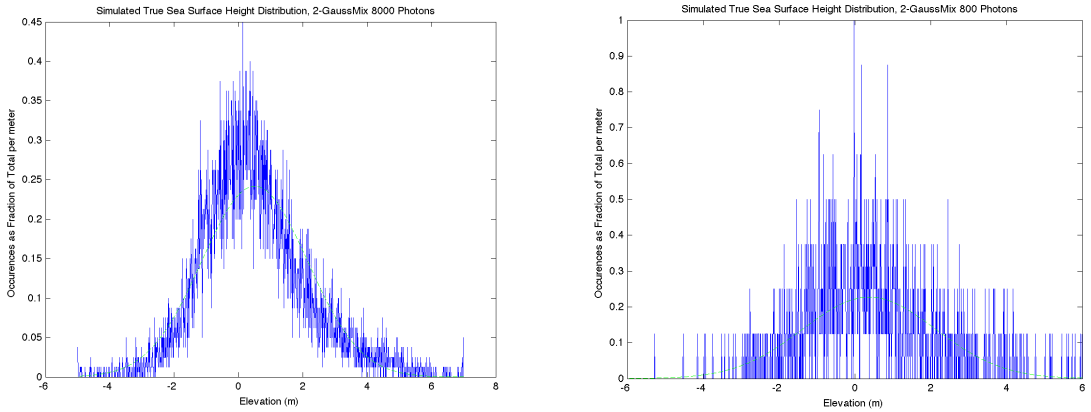

Figure 16. Left - Probability density function of the Synthetic SSH as a Gaussian mixture, 4000 samples from N(0,,1) and 4000 from N(1,2). Right – Same but with 800 photons for which the raw synthetic PDF is fairly noisy, but two peaks and skewness are faintly discernable.

(B) 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.

The instrument impulse response distribution is calculated from the most recent transmit echo pulse (TEP) passed from ATL03 *in /atlas_impulse_response/pce1_spot1 or /pce2_spot3*.

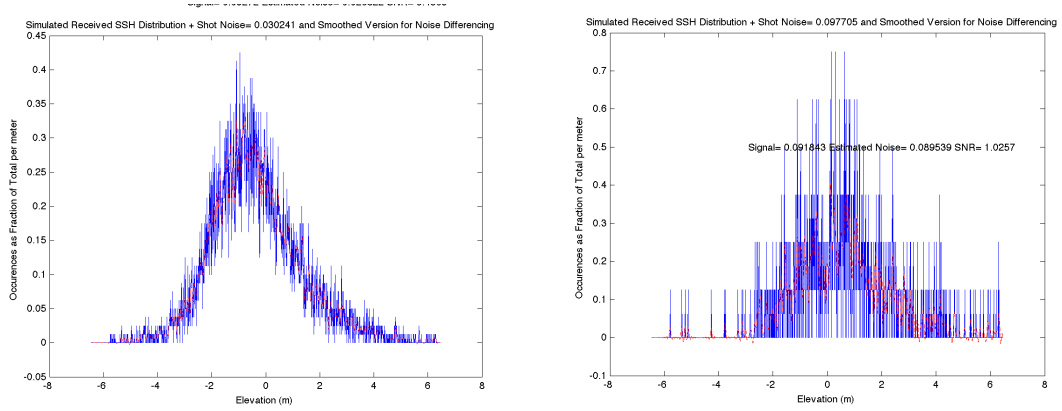

Figure 17. Synthesized received probability density functions (blue) as a convolution of the instrument impulse response PDF (Fig. 15) and the synthesized true SSH PDF (Fig. 16) plus pulse noise. The PDF for 8000 photons is at left and for 800 photons is on the right. These PDFs represent the data as 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.

The Transmit Echo Pulse

The TEP are derived from photons detected via the transmitter echo path (TEP, see Section 7.2.2 of the ATL03 ATBD). These photons are picked off from the transmitted laser pulse, and routed into the ATLAS receiver, for ATLAS strong beams 1 and 3. These data monitor the internal timing bias of ATLAS, as well as the instrument impulse response for the beam. We will assign TEP 1 or TEP 3 to each beam we process, mainly 1, 3, and 5, according to the recommendations in ***tep_valid_spot*** , a 6x1 array indicating which TEP to use for each spot. Order of the array is by ground track (gt1l; gt1r; gt2l; gt2r; gt3l; gt3r), from ATL03. (Note also, that when the spacecraft is flying forward beam 1 with a TEP is covering spot gt1r and beam 3 with a TEP is covering spot gt2r. When the spacecraft is flying backward beam 1 with a TEP is covering spot gt3l and beam 3 with a TEP is covering spot gt2l.) According to ATL03, the parameters are generated as often as sufficient TEP data exists in the granule, typically about every five granules (granules each span about $1/14^{th}$ of an orbit.). Otherwise data from the prior granule is retained in the current granule. The TEP photon event arrival times are provided in a histogram, ***tep_hist_times***.

The TEP photon events are downlinked when the TEP photon events happen to occur within the telemetry band approximately twice per orbit, for about three hundred seconds. The number of TEP photons per shot for a particular beam varies. However, the ratio of the number of TEP photons between beams 1 and 3 appears stable with a value of 0.7, with beam 1 being the stronger of the two. Averaging across a range of instrument configurations, beam 1 generates approximately 0.09 TEP photons per laser pulse, and beam 3 generates approximately 0.07 TEP photons per laser pulse. Assuming one TEP

photon event per twenty pulses as the lower limit for the TEP strength, downlinking TEP photons for approximately three hundred seconds should yield about 150,000 TEP photons. TEP photon events are captured any time there are more than thirty seconds of contiguous possible TEP photon events identified in ATL02 (e.g., 15,000 TEP photons), and there are at least 2000 possible TEP photons in the set. All available TEP photons within a given set are used; there is no upper limit on the duration or number of TEP photons. The set is ended when the TEP photons are no longer present in the telemetry band, or five seconds have passed with no additional TEP photons. With a sufficient number of possible TEP photon events identified, these will either be background photon events, misclassified signal photon events, or TEP photon events. The time of day of the start of the TEP data used in subsequent analyses is reported on ATL03 as ***tep_tod***, while the duration of TEP data is reported as ***tep_duration***.

The histogram of these possible TEP photon events is formed using 50 picoseconds bin widths. There is more than a single mean arrival time of TEP photons; with two bands with the primary band being at ~ 20 nanoseconds and the secondary band being at 46 nanoseconds. The actual time band limits for each TEP will be indicated by the ATL03 variable ***tep_range_prim***. Additional noise photons are also present. The mean background photon rate is determined by excluding the bins between 17 and 24 ns (the region of the primary return) and the region between 43 and 50 ns (the region of the secondary return) and calculating the mean of the remaining bins. This value is reported on the ATL03 output product as ***tep_bckgrd***, and is subtracted from all bin counts. The number of counts in the remaining histogram is reported as ***tep_hist_sum***. The number of counts in each bin is normalized by by ***tep_hist_sum*** and reported for each bin as ***tep_hist***. Time is measured relative to the centroid transmit pulse, and on average for the TEP represents the travel time back and forth through the ATLAS optical path. The time of the histogram bin centers (the mean of the times of the leading and trailing edges of the bins) is recorded as ***tep_hist_time***. The parameter ***tep_tod*** indicates when the TEP data on a given ATL03 granule was generated. The parameters derived from TEP data are part of the ATL03 data product in the groups /atlas_impulse_response/pce1_spot1/ and /atlas_impulse_response/pce2_spot3/.

Turning the TEP histogram into the Instrument Impulse Response

We convert from ***tep_hist*** expressed as normalized counts in *tepbinsize*=50 picosecond wide bins to an impulse response expressed as the probability density function with ***binsize*** wide bins of surface height. For the purpose of calculating the length of the derived surface distribution it is conceptually useful to interpolate to values in ***XmitHist000*** 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 process of deriving this impulse response from a TEP first requires identifying the appropriate TEP to use for a particular beam specified by ***tep_valid_spot*** , isolating the primary band of times specified by ***tep_range_prim***, trimming the TEP to primary histogram levels above the background noise rate, converting from travel time to surface photon height offset, converting to ***binsize*** wide bins starting from negative offsets, and normalizing to a probability density function by dividing by the sum of the histogram levels multiplied by ***binsize***. For the purpose of calculating the length of the derived surface distribution it is conceptually

useful to interpolate values in ***XmitHist000*** 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.
In part this pseudo code comes from the Matlab program BinconvertWF_JM3.m.

- The appropritate TEP histogram, ***tep_hist***, as specified by ***tep_valid_spot*** for a particular beam includes both the primary and secondary arrivals. Choose the primary TEP arrivals in bins between ***tep_range_prim***(1) and ***tep_range_prim***(2) (typically around 17 and 24 ns) for use in ATL12 processing. e.g., (iprimary=find(tepT>= ***tep_range_prim***(1) & tepT<= ***tep_range_prim***(2); tepT=tepT(jprimary); tepP=tepP(jprimary);
- Because ***tep_bckgrd*** is subtracted from the from the TEP histogram, the tails of the TEP will include bins with negative values. We trim off these tails where negative values appear indicating the TEP is below the noise floor. Working from time of TEP maximum out, set the TEP value to zero at the first bins where TEP is less than zero, and delete all bins beyond those points.
- Convert from travel time to surface height coordinate for the histogram by multiplying it times minus half the speed of light.
- Reverse the order of the histogram and its distance axis so that the array starts at negative height offsets.
- Center the TEP histogram so that it has height offset equals zero at the centroid of the distribution.
- Confirm the TEP bin size, ***bin_TEP***, as the distance between bin centers 50 picoseconds/ (2 x speed of light).
- Convert TEP bin centers to bin boundaries in meters
- Establish an array of impulse response histogram bin boundaries with an odd number of bins, a bin size equal to ***binsize***, interior bin centered on zero, and all within the range of the TEP bin boundaries.
- Compute the cumulative histogram of the TEP versus the TEP bin boundaries and interpolate to the impulse response histogram bin boundaries.
- Differentiate the resulting cumulative impulse response histogram and form a corresponding array of bin centers to produce the instrument impulse response histogram.
- Divide the impulse response histogram by the product of ***binsize*** and the sum of values in the histogram to yield the impulse response probability density function.

At this point, we have described the averaged pulse shape characteristics from the threshold crossing times posted at the nominal ~20 meters along-track segment rate. We also have described the TEP photon events and have formed a histogram and related parameters based on TEP events. Since the times of these respective groups are known, it is possible to examine how the SPD-based statistics vary during the period spanned by the TEP data. However, combining these two groups quantitatively is complicated by two aspects. First, the SPD-based data and the TEP-based data have unique paths through the instrument and so the reported times of these data will be substantially different. While these differences are characterized with pre-launch data, we have little insight into how the

temporal bias between these data will evolve through time. Secondly, the SPD-based data is derived from the reported times when a filtered version of the transmitted laser pulse crosses a particular amplitude threshold (measured in volts). The TEP-based data on the other hand is a histogram of counts of a massively attenuated version of the transmitted laser pulse. As such a quantitative correspondence between voltages and counts is problematic. A relationship could possibly be made between the standard deviation of the threshold crossing times and the apparent width of the TEP-based histogram at several points along the profile. This may allow a user to better predict how the transmitted pulse is changing between realizations of the TEP.

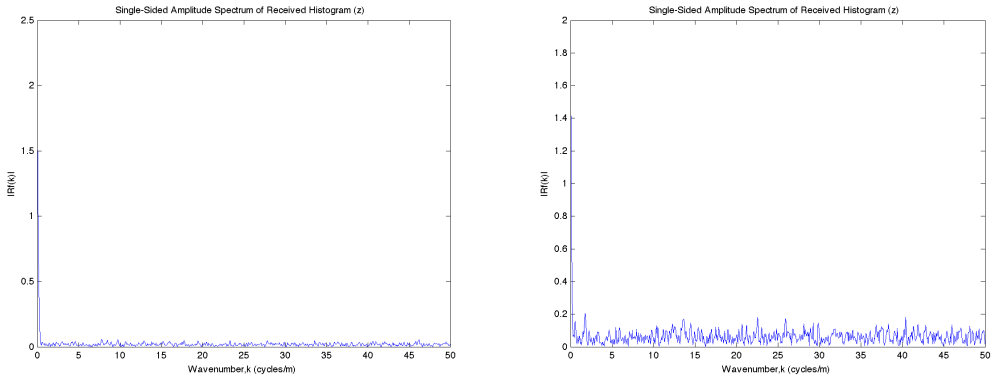

Figure 18. Single-sided amplitude spectra of the received PDFs (Fig. 17). The transform for 8000 photons is on the left and 800 photons is on the right. Note the higher wavenumber variability for 800 photons.

(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***

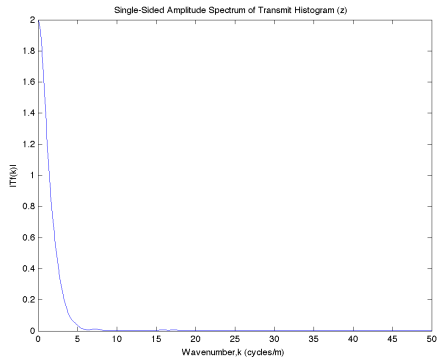

Figure 19. Single-sided amplitude spectrum of the instrument impulse response PDF (Fig. 15).

– 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
– 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 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***
– 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.

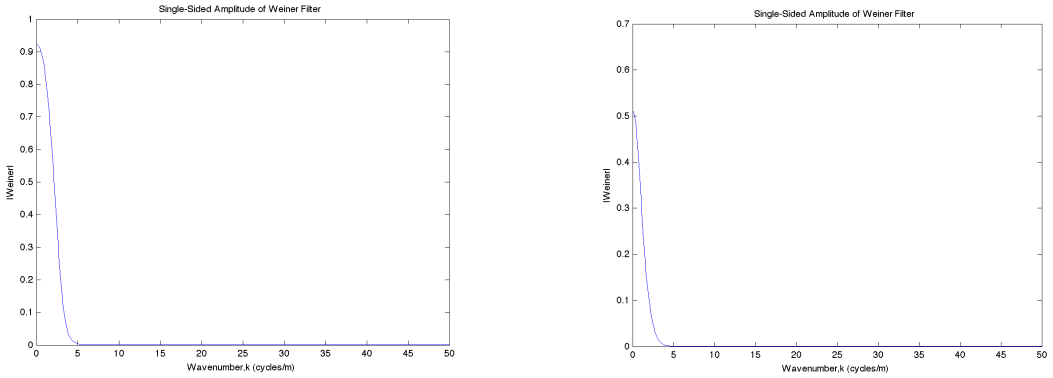

Figure 20. Single-sided amplitude spectra of the Wiener filters computed from the Fourier transform of the instrument impulse response PDFs (Fig. 19) and the signal to noise ratio estimated from the received height PDFs and smoothed versions of them (Fig. 17). Owing to a lower signal to noise ratio,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

(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.

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

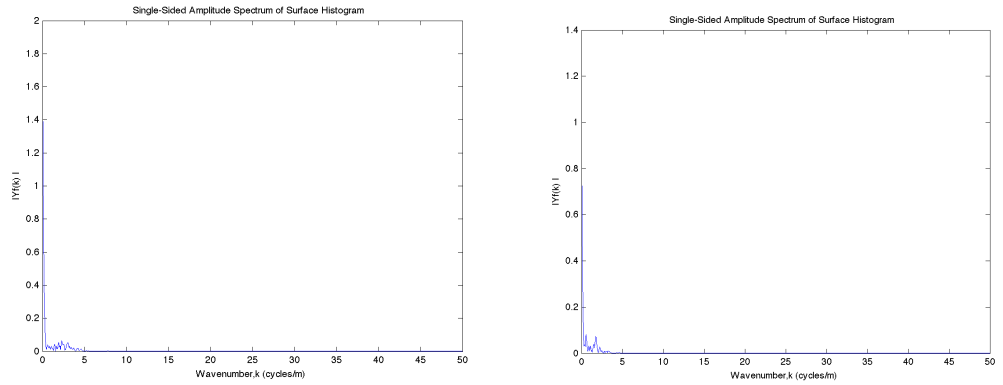

Figure 21. Single-sided amplitude spectra of the true surface PDF estimated as the Wiener filter (Fig. 20) times the Fourier transforms of the received PDFs (Fig. 18) divided by the Fourier transform of the instrument impulse response PDFs (Fig. 19). The result for 8000 photons is on the left and 800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

G) Compute the Surface Height Distribution

The surface histogram should be the same length and have the same horizontal (height) axis as the received histogram, ***rechist***. In this part, we compute the distribution of surface height as a probability density function (PDF), ***Y***, starting with the Fourier transform of that PDF output by the Weiner deconvolution. The steps are:

8. Compute ***Yfi***, as the inverse Fourier Transform of ***Yf*** divided by ***binsize. Yfi*** as it comes from the inverse Fourier transform is ***NFFT*** points along with the first half appearing displaced to the end of the array. We reorder this by first putting the last ***NFFT***/2 points ahead of the first ***NFFT***/2 points. To establish the surface height probability density function, ***Y***, we then remove the first (***rzbin***-1) points where ***rzbin*** is the index of the centroid (height equals zero) index of the received histogram. We keep as ***Y*** the remaining points equal to the length of the ***rechist***. In

the case of the ASAS FORTRAN code, which pads each end of ***rechist*** to reach heights of ±15 m, ***rzbin*** is the center bin of ***rechist***.

9. Set the x-axis of ***Y***, ***derivedsshx*** equal to the x-axis***, xrechist***, of the received histogram ***rechist***.
10. Normalize ***Y***, by dividing ***Y*** by the sum of the product of ***Yfi*** and ***binsize***.
11. Compute the mean or centroid of ***Y***, ***meanY***, by summing the product of ***Y*** and ***derivedsshx*** and dividing that sum by the sum of ***Y***.
12. Compute the standard deviation of ***Y***, ***stdY***:
    a) square ***derivedsshx*** minus ***meanY***
    b) Multiply a) by ***Y***
    c) Take sum of b)
    d) Divide c) by sum of ***Y***
    e) Take square root of d)
13. Compute the skewness of Y, ***skewnessY***:
    a) Cube ***derivedsshx*** minus ***meanY***
    b) Multiply a) by ***Y***
    c) Take sum of b)
    d) Divide c) by sum of ***Y***
    e) Divide d) by the cube ***stdY***.
14. Compute the kurtosis of ***Y***:
    a) Compute ***derivedsshx*** minus ***meanY*** to the fourth power
    b) Multiply a) by ***Y***
    c) Take sum of b)
    d) Divide c) by sum of ***Y***
    e) Divide d) by the ***stdY*** to the fourth power
    f) Subtract three from e)

Since the rationale for some of these steps may not be obvious, and the steps may have to be modified depending on the differences between the Matlab (used in the developmental code) and Fortran (used in the ASAS code), we provide the following explanations for the steps:

– 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 , 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 vector x-axis of the received height histogram, ***xrechist***, to the appropriate x-axis for the underlying surface height histogram, by deleting the first ***zbin*-1** points and the last ***lxmithist-zbin*** points from ***xrechist*** to yield the vector of x-axis of the surface histogram***, derivedsshx,*** with length, ***derivedlsshhist*** equal to length(***xrechist***)-(***lxmithist***-1).

Trimming the vector of surface histogram values is complicated by the necessity of padding received histogram with zeros to a length equal to a power of two for computation of its fast Fourier transform. As a result, at least in the prototype Matlab code, the inverse Fourier transform of the output of the Weiner deconvolution process, ***Yfi***, also has a length equal to this same power of two and should be truncated to the appropriate length.

– Consequently, truncate ***Yfi*** keeping only the first number of points in the original received histogram less (***lxmithist***-1), i.e., set ***derivedlsshhist*** equal to length(***xrechist***)-(***lxmithist***-1). This yields the surface histogram, ***Y,*** defined over ***derivedsshx,*** that will 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***).
– As a check on the deconvolution computation, the centroid of ***rechist*** and ***Y*** should both be near zero on account of the detrending of the surface heights and the assumed zero mean of the instrument impulse response computation
– Examples of ***Y*** and the difference between ***Y*** and the synthetic surface examples that we started

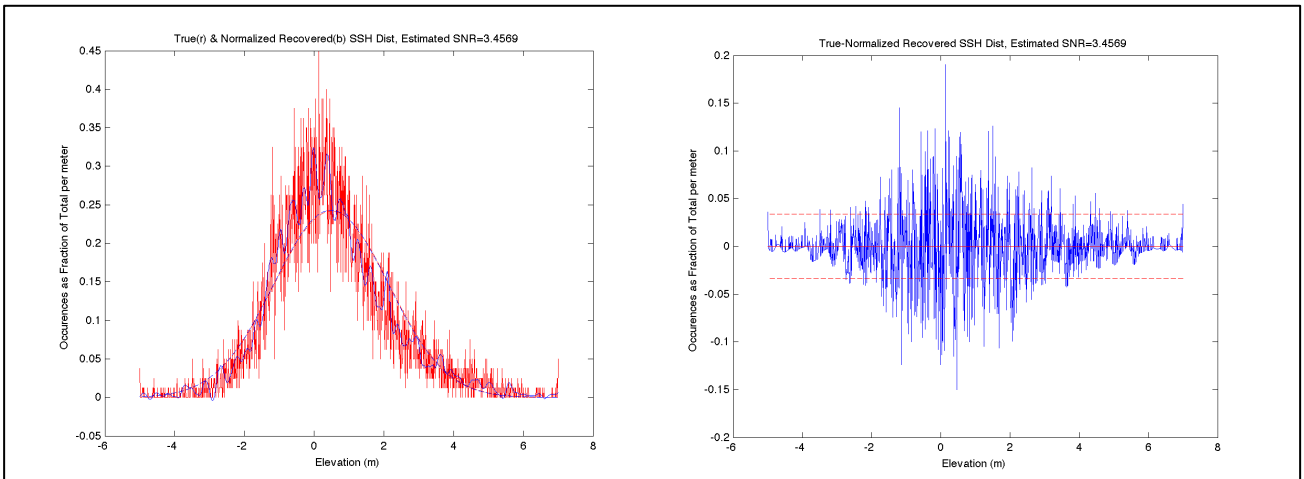

Figure 22. For the 8000 photon realization: On left in red, the PDF of true SSH (Fig. 16) with a Gaussian curve fit (dashed-red). Also in blue the PDF of SSH calculated by the Wiener deconvolution process (inverse transform of Fig. 21) with a Gaussian curve fit (dashed-blue). On the right is the difference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

with are given in Figure 22 for 8000 photons and Figure 23 for 800 photons.

- For subsequent use in the derivation of the ATL19 gridded product, the surface pdf, ***Y,*** will also be output for each segment along with the total number of surface photons, ***N***, and bin size, ***binsize.***
- ***Y*** is defined over a vector of bin centers, ***sshx,*** but for ***binsize*** = 1 cm ***Y*** will be output as a 3001-element vector from -15 m to +15 m for every segment with the 1,501st center bin being for height zero. Values will be zero for bins outside the range of ***sshx.*** Owing to noise and the effect of the FFT-deconvolution-inverse-FFT process, small negative values occur in ***Y*** within the range of ***sshx***. These will be set to zero in the output ***Y*** vector.
- As a simple check on the optimum Gaussian Mixture determination of the higher moment of the sea surface distribution we will compute and output the mean, variance, skewness, and excess kurtosis (***Ymean, Yvar, Yskew, Ykurt***) related to the higher moments of ***Y*** directly.
-

$$Ymean = \mu_Y = \left(\sum_{Y_i>0} X_i Y_i \right) \Big/ \left(\sum_{Y_i>0} Y_i \right) \tag{30}$$

$$Yvar = \left(\sum_{Y_i>0} (X_i - \mu_Y)^2 Y_i \right) \Big/ \left(\sum_{Y_i>0} Y_i \right) \tag{31}$$

$$Yskew = \left[\left(\sum_{Y_i>0} (X_i - \mu_Y)^3 Y_i \right) \Big/ \left(\sum_{Y_i>0} Y_i \right) \right] \Big/ Yvar^{3/2} \tag{32}$$

$$Ykurt = \left[\left(\sum_{Y_i>0} (X_i - \mu_Y)^4 Y_i \right) \Big/ \left(\sum_{Y_i>0} Y_i \right) \right] \Big/ Yvar^{2} - 3 \tag{33}$$

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: Expectation Maximization 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
---

- At this point we have the probability density function, ***Y,*** of detrended DOT over height bins with centers at ***sshx***. The detrending process used in surface finding removed an average, ***meanoffit2***, from the original heights. Once we have determined the mean and higher moments of ***Y***, we could add ***meanoffit2*** to the mean of ***Y*** to learn the mean DOT over the ocean segment. However, preliminary analysis suggests if we add ***meanoffit2*** back into the distribution before the optimal Gaussian Mixture determination, the resulting higher moments are less noisy over many ocean segments. This can be done by shifting ***sshx*** by adding ***meanoffit2*** to each value in ***sshx***. Then the

true average DOT will be a derived as the aggregate mean of the optimally derived Gaussian mixture. For the developmental Matlab code a slightly different but equivalent approach was used.

- The Matlab development code used gmdistribution.fit, an expectation maximization approach, to compute the 2-component Gaussian mixture corresponding to the observed surface height distribution, ***Y***. The ASAS Fortran code uses the expectation maximization routine from ATBD ATL07 Appendix E.
- To apply the expectation maximization approach, we first derive a sea surface height spatial series from the height distribution ***Y*,** which is defined over a range of heights the same as the received height histogram (i.e**.,** ***jlow*** to ***jhigh*** from surface finding). This is accomplished by first multiplying **Y** by 10,000, rounding and deleting any values less than or equal to zero (a few unrealistic weakly negative values to occur in ***Y*** associated with the Weiner deconvolution process) 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)***. The shift to heights including ***meanoffit2*** was accomplished in developmental code by adding ***meanoffit2*** to each value in ***XY***. (ASAS FORTRAN code does not add ***meanofffit2*** at this point but after the Gaussian Mixture calculation.) Although it is not strictly necessary, the values of ***XY*** are randomly permutated to produce a realistic distribution of heights with the probability density function ***Y*** on heights given in

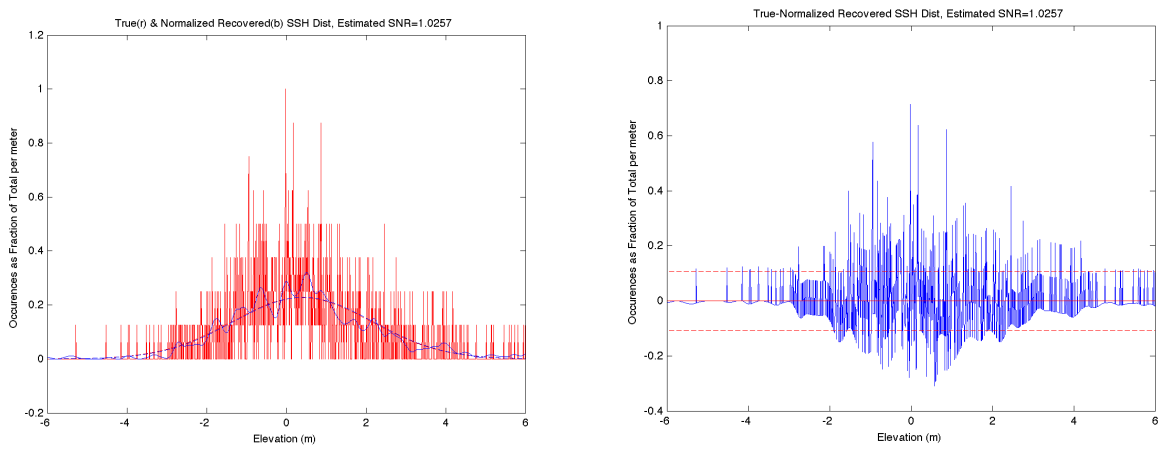

Figure 23. For the 800 photon realization: On left in red, the PDF of true SSH (Fig. 13) with a Gaussian curve fit (dashed-red). Also in blue the PDF of SSH calculated by the Wiener deconvolution process (inverse transform of Fig. 21) with a Gaussian curve fit (dashed-blue). On the right is the difference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

sshx shifted by adding ***meanoffit2***.

- We then compute the 2-compnent Gaussian mixture variables, of the two-Gaussian mixture using expectation maximization, Matlab 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 dimension:
<u>Component 1:</u>
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 dimension:
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:

$$\mathrm{E}\left[\left(X-\mu_{mix}\right)^{j}\right]=\sum_{i=1}^{2}\sum_{k=0}^{j}\frac{j!}{k!(j-k)!}\left(\mu_i-\mu_{mix}\right)^{j-k}m_i\mathrm{E}\left[\left(X_i-\mu_i\right)^{k}\right] \qquad (34)$$

$$E\left[\left(X-\mu_{mix}\right)^{j}\right]=\sum_{i=1}^{L}\sum_{k=0}^{j}\frac{j!}{k!(j-k)!}\left(\mu_i-\mu_{mix}\right)^{j-k}m_iE\left[\left(X_i-\mu_i\right)^{k}\right]$$

Note that in computing the moments (34), 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 (34) 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, $\mu_{mix}=m_1\mu_1+m_2\mu_2$ (28) is the mean of the detrended surface photon heights relative to the geoid with the mean of the detrending process, ***meanoffit2*** (equal to the mean of ***P0+P1*trackdist_initial2_surf***), added back in prior to the Gaussian mixture determination. Thus, it represents the mean DOT uncorrected for sea state bias, along the ocean segment. Therefore, the mean sea surface height, ***SSH***, for the segment to be

output in ATL12 is equal to the mean of the mixture, μ_{mix}, plus ***meangdht*** equal to the mean of geoid height, ***geoid***, over the segment:

$$SSH = \mu_{mix} + meangdht = m_1\mu_1 + m_2\mu_2 + meangdht \tag{35}$$

- 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. Here we don't include the variance due to the segment trend in sea surface height due to underlying trends in the DOT or geoid height. From (29), the second moment, or squared standard deviation, σ^2_{mix} , of the mixture is:

$$SSHvar = \sigma^2_{mix} = m_1\left(\left(\mu_1 - \mu_{mix}\right)^2 + \sigma_1^2\right) + m_2\left(\left(\mu_2 - \mu_{mix}\right)^2 + \sigma_2^2\right) \tag{36}$$

- Using (34), compute third and fourth moments, we obtain the sea surface height skewness, ***SSHskew***, and excess kurtosis ***SSHkurt***,

$$SSHskew = \mathrm{E}\left[\left(Y - \mu_{mix}\right)^3\right] \Big/ \sigma^3_{mix} \tag{37}$$

$$SSHkurt = \mathrm{E}\left[\left(Y - \mu_{mix}\right)^4\right] \Big/ \sigma^4_{mix} - 3 \tag{38}$$

and output these as part of ATL12 to better characterize the sea surface. Here we don't include the skewness and kurtosis due to the segment trend in surface height.

5.3.4.3 Conclusions for Section 5.3.4:

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 one 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 gmdistribution.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 gmdistribution.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 produce 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 Expected Uncertainties in Means of Sea Surface Height

The ocean products of ATL12 are the distributions of sea surface height over ocean segments typically 5 to 7 km long and including about 8,000 candidate surface reflected photons. The height distributions are characterized up through their 4th moments (mean, standard deviation, skewness, and kurtosis) using the 2-Gaussian mixture approach. In this subsection, we examine the magnitude of the expected uncertainties in the mean of the surface height distribution. For the open ocean, we consider uncertainty in the mean height due to:
1. Signal-to-noise considerations in a distribution of photon heights.
2. Uncertainties in the mean impulse response.
3. Subsurface scattering.
These are borrowed from and consistent with the ATL07 Sea Ice ATBD except due the low rate of photon returns (about 1 photon per pulse) over the ocean surface, for the most part we exclude first-photon bias (See Section 3.1.1.4). We also do not consider sea state bias discussed in previous chapters because we are calculating SSB directly and the errors in this calculation will have to be determined during the ICESat-2 cal/val process.

5.3.6.1 Signal-to-noise considerations and the dominant effect of surface wave roughness

We expect the largest contributions to uncertainty in the mean sea surface height to be due to surface roughness, and at least over short averaging lengths in daytime, background noise. The total number of photons
(NP_{tot}) in a height window, W, is the sum of the number of signal photons, the background rate (B_s) and the number of pulses (N_{pulses}) within that window:

$$NP_{tot} = NP_{signal} + NP_{bkg}$$

Background photons are uniformly distributed within the height window and thus the standard deviation, σ_{bkg} , can be written as,

$$\sigma_{bkg} = \mathrm{W}\ (12)^{-1/2} = 0.2887\ W,$$

where the factor .2887 is the standard deviation of a uniform random variable on a unit interval. The number of background photons NP_{bkg}) in the window (W) is a Poisson variable with a mean value given by:

$$NP_{bkg} = 2\ N_{pulses}\, Bs\ (\mathrm{W/c})$$

If the distribution of the received signal photons is approximately Gaussian, the variance depends on the transmit-pulse width (σ_{pulse}) and the surface roughness (σ_{rough}):

$$\sigma_{signal}^{\ 2} = \sigma_{rough}^{\ 2} + \sigma_{pulse}^{\ 2}$$

The expected composite variance of surface height can then be written as:

$$\sigma_{surf}^2 = \frac{NP_{signal}\sigma_{signal}^2 + NP_{bkg}\sigma_{bkg}^2}{NP_{signal} + NP_{bkg}}$$

Except for the roughness, all of the quantities in this equation can be estimated from the data: the background and signal photon counts can be estimated from the total number of photons and the background rate. For a relatively smooth flat surface, the expected variance in the estimation of surface height in a window of N_{pulses} can be calculated using:

$$\hat{\sigma}_{surf}^2 = \frac{\sigma_{surf}^2}{NP_{total}}$$

Ultimately ICESat-2 ocean data will likely extend into the marginal ice zone (MIZ) for where the data will be analyzed for short along-track distances and smooth water. From the ATBD for Sea Ice (ATL07), over a flat lead with mean returns of ~1 photon per pulse for 100 pulses (70-m of smooth open water) with a nominal reflectance of 0.2 and no background noise, the expected uncertainty is 1.0 cm due entirely to uncertainty due to pulse width. With a 3 MHz background rate the uncertainty increases to 1.1 cm. Even at short distances, the contribution of background noise is small.

In true open ocean conditions, the dominant source of uncertainty in the mean height of the sea surface over each ocean segment will be the roughness of the surface due to waves.

$$\hat{\sigma}_{surf} = \frac{\sigma_{surf}}{\sqrt{NP_{total}}} \approx \frac{\sigma_{waves}}{\sqrt{NP_{total}}} \qquad (39)$$

For a typical situation with a significant wave height (SWH=4 σ_{waves}) equal to 2 m, σ_{waves} equals 0.5 m, and over a typical ocean segment with 8,000 photons, the uncertainty in the ocean segment mean sea surface height, $\hat{\sigma}_{surf}$, would be 0.56 cm if the heights were uncorrelated, and over 8,000 pulses, $\hat{\sigma}_{surf}$ associated with surface roughness would remain less than 1 cm for SWH less than 3.6 m. Unfortunately, analysis of initial ICESat-2 data suggests that for large waves, the ocean segment to segment average heights vary much more than this. This is because successive height measurements are correlated due to the harmonic character of the surface. In fact, the uncertainty, $\hat{\sigma}_f$, of (39) appears to more applicable to the ICESat-2 segment-to-segment height variability if we take the degrees of freedom, NP_{total}, equal to the number of wavelengths of the dominant waves in the ocean segment. For long waves, especially at an angle to the ground track, this number can easily be 15 so that the uncertainty can be in the 20 to 30 cm range. This is not a characteristic of ICESat-2, but rather a characteristic of the ever-changing ocean surface.
To address this natural uncertainty, we will calculate the approximate degrees of freedom from the autocorrelation of 10-m bin averaged heights (5.3.3)

From Brockwell and Davis [*Brockwell and Davis*, 2016], considering *n* discrete samples of *h*, the mean squared error of the sample mean, μ_h, from the true mean, $\overline{H}$, is

$$E\left(\mu_h - \bar{H}\right)^2 = n^{-1}\sum_{i=-n}^{n}\left(1-\frac{|i|}{n}\right)\gamma(i)$$

where $\gamma(i)$ is the autocovariance of *h* at lag, *i*. In terms of distance, for a distance increment, δ_x, the length of record, *L*, equals $n\delta_x$ and taking $\lambda = i\delta_x$

$$E\left(\mu_h - \bar{H}\right)^2 = \left(n\delta_x\right)^{-1}\sum_{i\delta_x=-n\delta_x}^{n\delta_x}\left(1-\frac{|i\delta_x|}{n\delta_x}\right)\gamma(i\delta_x)$$
$$= L^{-1}\sum_{\lambda=-L}^{L}\left(1-\frac{|\lambda|}{L}\right)\gamma(\lambda)$$

The autocovariance is symmetric with lag, λ, and is equal to the autocorrelation of *h* times the variance in *h*. Therefore

$$E\left(\mu_h - \bar{H}\right)^2 = 2L^{-1}\sum_{\lambda=0}^{L}\left(1-\frac{|\lambda|}{L}\right)\gamma(\lambda) = \sigma_h^2 2L^{-1}\sum_{\lambda=0}^{L}\left(1-\frac{|\lambda|}{L}\right)R(\lambda)$$

This corresponds to (39) for:

$$\hat{\sigma}_{surf} = \frac{\sigma_{waves}}{\sqrt{NP_{effect}}} = \frac{\sigma_{waves}}{\sqrt{\dfrac{L}{2\sum_{\lambda=0}^{L}\left(1-\dfrac{|\lambda|}{L}\right)R(\lambda)}}} \qquad (40)$$

So the correlation length scale, L_{scale}, and effective degrees of freedom $\boldsymbol{NP_{effect}}$ become:

$$L_{scale} = \sum_{\lambda=0}^{L}\left(1-\frac{|\lambda|}{L}\right)R(\lambda) \quad \& \quad NP_{effect} = \frac{L}{2L_{scale}} = \frac{L}{2\sum_{\lambda=0}^{L}\left(1-\dfrac{|\lambda|}{L}\right)R(\lambda)}\,. \qquad (41)\ \&\ (42)$$

Equations (41) and (42) for ***Lscale*** and ***NP_effect*** are written in terms of a record length, ***L***. We can think of this a length ***L*** in meters. It is equivalent and more convenient to think of the record length in terms of the number of 10-m bins in the ocean segment, *N*=***Nbin10***, for which (41) and (42) become:

$$L_{scale} = \sum_{\lambda=0}^{N}\left(1-\frac{|\lambda|}{N}\right)R(\lambda) \quad \& \quad NP_{effect} = \frac{L}{2L_{scale}} = \frac{L}{2\sum_{\lambda=0}^{N}\left(1-\dfrac{|\lambda|}{N}\right)R(\lambda)}\,. \qquad (43)\ \&\ (44)$$

Pseudo Code for Determining ***NP_effect*** and ***h_uncrtn***
First, estimate the degrees of freedom using an approach inspired by the Matlab function degrees_freedom.m by Kathy Kelly (personal communication, 2020).

a) Compute as function of lag, ***l***, the lagged autocorrelation vector ***R_L(l)*** of ***htybin,*** which is defined for 10-m bins with centers ***xbin.*** Lags will be increments of 10-m bin indices and extend from ***l*** =0 to ***l*** equal the number of bins in the ocean segment, ***Nbin10*** =***L***/10m. (Note: To consider lags up to the length of the record, it may be advisable to double the length of ***htybin*** by zero padding each end ***Nbin10***/2. This would not usually be necessary because we should only need to use the covariance over the first 10 to 20 lags, but in cases where the ocean surface is not dominated by long waves, the autocorrelation may never be less than zero and it may be necessary to perform the integration of (41) out to a lag equal to the record length). Lagged covariance ***COV_L*** at lag ***l*** will be the sum over all *i*= 1 to length, ***Nbin10***, of the product ***htybin(i)* htybin(i+l).*** Where either ***htybin(i)***or ***htybin(i+l)*** does not exist, no addition is made to the sum. The autocorrelation, ***R_L*** as a function of lag, ***l***, equals ***COV_L(l)*** divided by the covariance, ***COV_L*** , at zero lag so the ***R_L*** at zero lag, ***l*** = 0, is equal to 1, ***R_L(***0***)*** =1.

b) Integrate (1-***l/L***)* ***R_L*** from zero lag to the first zero crossing to get decorrelation scale, ***Lscale*** .
Start with ***Lscale*** = 0 and at *i* = 1 where the corresponding lag, ***l***, is zero, and ***R_L***(*i*)=1. Increment *i* and the corresponding ***l*** by 1 and keep adding to ***Lscale*** while ***R_L(***i+1***)*** is greater than zero, i.e.,

$$Lscale = Lscale + ((1-l/Nbin10)*R\_L(i) + (1-(l+1)/Nbin10)*R\_L(i+1))/2$$

When ***R_L(***i+1***)*** becomes less than zero, terminate the integration by essentially assuming ***R_L(***i+1***)*** is zero, i.e.,

$$Lscale = Lscale + (1-l/Nbin10)*R\_L(i))/2$$

c) Degrees of freedom, ***NP_effect***, equals length of the ocean segement, ***Nbin10***, divided by two times the ***Lscale***. Note that Matlab code from Kathy Kelly (personal communication, 2020) that has informed this code development deletes the division by 2 and the (1-***l/Nbin10***) term in the integration and will give larger degrees of freedom than this routine. This code is more conservative.

d) ***h_uncrtn*** equals the standard deviation of sea surface height divided by the square root of ***NP_effect.***

5.3.6.2 Uncertainties in the mean impulse response

In our processing we, account for the impact of the instrument impulse response (IIP) on the distributions of surface heights by deconvolving from the received height

distributions the impulse response estimated from the transmit echo pulse (TEP). To the extent that the TEP represents the impulse response this should eliminate uncertainty due to the IIP. Uncertainties due to a mismatch between the TEP and true IIP will have to be inferred by examination of the TEP in the on-orbit data, which remains to be done.

5.3.6.3 Uncertainties due to subsurface scattering.

Because the blue-green laser of ATLAS penetrates water, true subsurface returns have always been a concern, and the higher subsurface density of photons may be due in part to true subsurface scattering in the ocean. However, we see similarly enhanced subsurface densities over, clear deep ocean waters and even over land where penetration and backscatter shouldn't occur. Consequently, present thinking expressed in the ATL03 Known Issues is that the subsurface noise level is due to forward scattering delays in the atmosphere of surface reflected photons.

Whatever their cause, some subsurface photon heights are being included in the surface height histogram, creating what we think may be an order 1-3 cm bias in average SSH. To reduce the sensitivity to subsurface returns, the processing code first makes a simple estimate of the high and low histogram limits and then uses these to determine separate above surface and subsurface noise levels, ***tailnoisehigh*** and ***tailnoiselow***. The ultimate high limit is then chosen where the smoothed histogram falls below a factor ***Th_Nc_f*** (e.g., ***Th_Nc_f*** = 1.5) times ***tailnoisehigh*** and low limit is chosen where the smoothed histogram falls below the same factor times ***tailnoiselow***. Also, for Release 4, the surface finding, instead of being based on the distribution of photon heights, is based on the distribution of the photon height anomalies relative to a moving 11-point bin average of high-confidence photon heights. This excludes subsurface returns under the crests of surface waves obviating the immediate need for a subsurface return correction dependent on sea state.

5.4 Ancillary Information

5.4.1 Solar Background Photon Rate and Apparent Surface Reflectance (ASR)

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 ***backgr_seg***.
The apparent surface reflectance should be taken from ATL09 (***/profile_x/high_rate/apparent_surf_reflec***) at the 25 Hz rate and simply averaged over the length of the height segment to produce ***asr_seg***.

5.4.2 Additional Ancillary Data

Variables from the ATL03 ***geolocation/*** and ***geophys_corr/*** groups will also simply be averaged over each ocean segment and saved. Segment averages of the following variables, corresponding to the heights stored in A_{fine} will be returned in ATL12. Those variables are locted in */gtx/stats* group and include

neutat_delay_total
ref_azimuth
ref_elev
seg_dist_x
solar_azimuth
solar_elevation
geoid
geoid_free2mean
dac
tide_earth
tide_earth_free2mean
tide_load
tide_oc_pole
tide_ocean
tide_equilibrium
tide_pole
full_sat_fract
near_sat_fract

Also, the first and last of the geosegment IDs (***segment_id***) for each ocean segment will be returned in ATL12 as ***first_geoseg*** and ***last _geoseg***.

The percentages of each ***surf_ty***pe of the photons from Fine Select in the ocean segment will also be returned as ***surf_type_prcnt.*** This will be a 5-element variable for each ocean segment with each element corresponding to the percentage of photons from each of the 5 surface types.

For each ocean segment, we output ***layer_flag_seg*** as equal to 1 for ocean segments with more than 50% of geosegs having ***layer_flag*** =1, indicating significant cloud layers were present. In addition, we compute ***cloudcover_percent_seg*** for each ocean segment as the percentage of geosegs in the ocean segment with ***layer_flag*** = 1, i.e., ***cloudcover_percent_seg*** = 100*(total number of geosegs used with ***layer_flag***=1)/(total number of geosegs used).

5.5 Output Parameters

Table 6 ATL12 Output (See Appendix A for full product specifications)

Product Label	Units	Description	Symbol
ancillary_data/			
ds_a		Dimension scale (1:65) for the harmonic coefficients (*a*) in the *gtx/ssh_segments/heights* group. *a*(1) is the coefficient for wavenumber equal zero (constant part), Above a(1), *a*(even index i) is the sine coefficient for *wn*(i/2), and *a*(odd index i) is the cosine coefficient for *wn*((i-1)/2)	*a*
ds_surf_type		land (1), ocean (2), seaice (3), landice (4), inland water (5)	
ds_wn		Dimension scale (1:32) for the wavenumbers (*wn*) in the *gtx/ssh_segments/heights* group.	*wn*
ds_xbin	m	Center of 1 x 710 element array of 10-m bins. Note this may be included as a data description or other static array equal to [5, 15, 25, 35 ….. 7095 m]	*xbin*
ds_y_bincenters	m	-15 to +15, by 1cm bins	
ancillary/ocean/			
Hist_bin_size	m	Bin size for Y and sshx	*binsize*
gtx/			
gtx/ssh_segments/			
delt_seg	seconds	Time duration segment	
delta_time	seconds	Mean time of surface photons in segment	*t_seg*
latitude	°N	Mean latitude of surface photons in segment	*lat_seg*

longitude	°E	Mean longitude of surface photons in segment	*lon_seg*
gtx/ssh_segments/heights			
A	m	Vector of 2 x ***nharms*** +1 coefficients for each harmonic component in harmonic analysis of heights (5.3.3.2). *a*(1) is the coefficient for wavenumber equal zero (constant part), Above a(1), *a*(even index i) is the sine coefficient for *wn*(i/2), and *a*(odd index i) is the cosine coefficient for *wn*((i-1)/2)	***a***
bin_slopebias	1	Indication of the dependence of reflectance on 10-m scale sea surface slope equal to covariance of the data rate and the surface slope, ***binCOVslope_Xr***, divided by the mean of the data rate. (5.3.3.1)	***bin_slopebias***
bin_ssbias	m	Sea state bias estimated from the correlation of photon return rate with along-track 10-m bin averaged surface height (4.3.1)	***bin_ssbias***
bin_magslopebias	1	Indication of the dependence of reflectance on the magnitude of 10-n scale sea surface slope equal to covariance of the data rate and magnitude of surface slope, ***binCOVmagslope_Xr***, divided by the mean of the data rate. (5.3.3.1)	***bin_magslopebias***
Dxbar	m	Mean distance between surface reflected photons	***DXbar***
Dxskew		Skewness of the distribution of distance between surface reflected photons	***DXskew***
Dxvar	m^2	Variance of the distribution of distance between surface reflected photons	***DXvar***
fpb_10m	m	Estimated first photon bias in each 10-m bin	*fpb_10m*
fpb_corr	m	Ocean segment average of 10-m first photon biases. When non-zero, *fpb_corr* should be subtracted from ocean	*fpb_corr*

		segment average height to correct for first photon bias.	
fpb_corr_stdev	m	Standard deviation of *fpb_10m* over the ocean segment.	*fpb_corr_stdev*
h	m	Mean sea surface height relative to the WGS84 ellipsoid	***SSH***
h_kurtosis		Excess kurtosis of sea surface height histogram	***SSHkurt***
h_skewness		Skewness of photon sea surface height histogram	***SSHskew***
h_uncrtn	m	Uncertainty in the mean sea surface height over an ocean segment	***h_uncrtn***
h_var	m^2	Variance of best fit probability density function ($meters^2$)	***SSHvar***
htybin	m	1 x 710 element array of 10-m bin average heights from SSB calculation	***htybin***
htybin_std	m	Standard deviation of photon heights in 10-m along-track bins.	***htybin_std***
h_atl07_ice_free	m	Average DOT from ATL07 *ssh_flag=1* qualified sea ice segments with corresponding ATL12 10-m bins	***h_atl07_ice_free***
h_atl07_ice_free_uncrtn	m	Uncertainty in *h_atl07_ice_free*	***h_atl07_ice_free_uncrtn***
h07icefree_uncrtn_inst	m	Instrumental uncertainty in *h_atl07_ice_free* related to average standard deviation of photon DOTs in ATL07 bright lead sea ice segments common with ATL12 10-m bins.	***h07icefree_uncrtn_inst***
h07icefree_uncrtn_nat	m	Natural uncertainty in h_atl07_ice_free related to standard deviation of average DOTs in ATL07 bright lead sea ice segments common with ATL12 10-m bins	***h07icefree_uncrtn_nat***
h_atl07_inatl12oc	m	Averaged DOT from all ATL07 *ssh_flag* =1 qualified sea ice segments in an ATL12 ocean segment with or without corresponding ATL12 10-m bins	***h_atl07_inatl12oc***

h_atl07_inatl12oc_uncrtn	m	Uncertainty in *h_atl07_inatl12oc*	***H_atl07_inatl12oc_uncrtn***
h7in12icefree_uncrtn_inst	m	Instrumental uncertainty in h_atl07_inatl12oc related to average standard deviation of photon DOTs in ATL07 bright lead sea ice segments in the ATL12 ocean segment	***h7in12icefree_uncrtn_inst***
h7in12icefree_uncrtn_nat	m	Natural uncertainty in h_atl07_inatl12oc related to standard deviation of average DOTs in ATL07 bright lead sea ice segments in the ATL12 ocean segment	***h7in12icefree_uncrtn_nat***
h_ice_free	m	Average DOT from ATL12 10-m bins in common with *ssh_flag* qualified ATL07 sea ice segments	*h_ice_free*
h_ice_free_uncrtn	m	Uncertainty in average DOT from ATL12 10-m bins in common with *ssh_flag=1* qualified ATL07 bright lead sea ice segments	***h_ice_free_uncrtn***
hicefree_uncrtn_inst	m	Instrumental uncertainty in *h_ice_free* related to average standard deviation of photon DOTs in ATL12 10-m bins in common with ssh_flag=1 qualified ATL07 bright lead sea ice segments.	***hicefree_uncrtn_inst***
hicefree_uncrtn_nat	m	Natural uncertainty in *h_ice_free* related to standard deviation of average photon DOTs corrected for first photon bias from ATL12 10-m bins in common with ssh_flag=1 qualified ATL07 bright lead sea ice segments.	***hicefree_uncrtn_nat***
l_scale	deci-meters	Correlation length scale expressed in units of number of 10-m bins (may be fractional, i.e., decimeters)	***Lscale***
latbind	°N	1 x 710 element array of 10-m bin averages of latitude	***latbind***

length_seg	m	Length of segment (m)	***length_seg***
lonbind	°E	1 x 710 element array of 10-m bin averages of longitude	***lonbind***
meanoffit2	m	Mean of linear fit removed from surface photon height	***meanoffit2***
mix_m1		Fraction of component 1 in 2-component Gaussian mixture	***m1***
mix_m2		Fraction of component 1 in 2-component Gaussian mixture	***m2***
mix_mu1	m	Mean of component 1 in 2-component Gaussian mixture	***mu1***
mix_mu2	m	Mean of component 1 in 2-component Gaussian mixture	***mu2***
mix_sig1	m	Standard deviation of component 1 in 2-component Gaussian mixture	***Sig1***
mix_sig2	m	Standard deviation of component 1 in 2-component Gaussian mixture	***Sig2***
n_pulse_seg	counts	Number of laser pulses in segment	***n_pls_seg***
nbin10	counts	Number of 10-m bins in an ocean segment	***Nbin10***
np_effect		Effective degrees of freedom of the average sea surface height for the ocean segment	***NP_effect***
p0	m	Constant of linear fit versus along-track distance to surface photon height over the ocean segment	***P0***
p1		Slope of linear fit versus along-track distance to surface photon height over the ocean segment	***P1***
snr_harm		Signal to noise ratio of harmonic fit with coefficients in a to the surface reflected photons including ***meanoffit2*** and with data gaps greater than ***gaplimit*** filled with Gaussian white noise about ***meanoffit2***	***SNR_harm***
swh	m	Significant wave height estimated as 4 times the standard deviation of along-track 10-m bin averaged surface height	***SWH***
wn	m^{-1}	nharms (= 32) wavenumbers equal to the inverse of wavelengths for each harmonic component in	*wn*

		harmonic analysis of heights (5.3.3.2)	
xbind	m	1 x 710 element array of 10-m bin averages of along-track distance	*xbind*
xbind_first_dist_x	m	The along-track distance from the equator crossing to the start of the 10m bins [xbin(1)] of the ocean segment	*xbind_first_dist_x*
xrbin	Photon m^{-1}	1 x 710 element array of 10-m bin average photon rate from SSB calculation	*xrbin*
xbin_atl07_dot	m	A 10-m bin qualified bright-lead flag: = INVALID_I4B where the ATL12 ice_conc < 15%. = ATL07 sea ice segment DOT where the ATL12 ice_conc >= 15%, and beginning and end of a 10-m bin are within ATL07 segments with *height_segment_ssh* =1, = 999 otherwise *xbin_atl07_dot* = 999	*xbin_atl07_dot*
Y	m^{-1}	Probability density function of photon surface height	*Y*
Ykurt		Excess Kurtosis = (fourth moment of *Y*)/ $Yvar^2$ -3	*Ykurt*
Ymean	m	Mean=first moment of Y, should be ~0 = *h* -meanoffit2	*Ymean*
Yskew		Skewness = (third moment of *Y*)/ $Yvar^{3/2}$	*Yskew*
Yvar	m^2	Variance = second moment (m^2) of *Y*	*Yvar*
gtx/ssh_segments/stats			
backgr_seg	m^{-1}	backgrd_atlas/bckgrd_rate from ATL03 averaged over the segment	***backgr_seg***
cloudcover_percent_seg	%	The percentage of geosegs in the ocean segment with **layer_flag** = 1.	***cloudcover_percent_seg***
dac _seg	m	Ocean segment average of dynamic atmospheric correction (DAC) includes inverted barometer (IB) affect	***dac _seg***
depth_ocn_seg	m	The average of depth ocean of geo-segments used in the ocean segment.	***depth_ocn_seg***

first_geoseg	counts	The first of the geosegment ids (**segment_id**) for each ocean segment	***first_geoseg***
first_pce_mframe_cnt	counts	First Major Frame ID in the SSH segment	***first_pce_mframe_cnt***
first_tx_pulse	counts	First Transmit pulse in along-track segment	***first_tx_pulse***
full_sat_fract_seg		fraction of all pulses in all geosegs used that were fully saturated	***full_sat_fract_seg***
geoid_free2mean_seg	m	Ocean segment average of the conversion factor added to ATL03 geoid to convert from the tide-free to the mean-tide system	***geoid_free2mean_seg***
geoid_seg	m	Ocean segment average of geoid in the mean-tide system height above the WGS – 84 reference ellipsoid (range -107 to 86 m)	***geoid_seg***
ice_conc	%	The percent ice concentration at the average ocean segment position from the high resolution AMSR2 passive microwave SIC dataset used for ATL07	***ice_conc***
ice_conc_12icefreebins	%	ATL12 ocean-segment ice concentration based on the complement of the square of the fraction of 10-m bins classified as bright leads per ATL12 ATBD Rel. 7	***ice_conc_12icefreebins***
ice_conc_SSMI	%	The percent ice concentration at the average ocean segment position from NSIDC passive microwave data formerly used for ATL12 and ATL07.	***ice_conc_SSMI***
last_geoseg	counts	The last of the geosegment ids (**segment_id**) for each ocean segment	***last_geoseg***
last_pce_mframe_cnt	counts	Last Major Frame ID in the SSH segment	***Last_pce_mframe_cnt***
last_tx_pulse	counts	Last Transmit pulse in along-track segment	***last_tx_pulse***
layer_flag_seg		The layer flag from ATL09 that is in effect over 50% of the ocean segment, 0 indicating absence of clouds and forward scattering, and 1 indicating possibility of forward scattering as in ATL09	***layer_flag_seg***
n_photons	counts	Number of surface photons found for the segment	***n_photon***

n_ttl_photon	counts	Number of photons in the ±15-m ocean downlink band	***n_ttl_photon***
near_sat_fract_seg	counts	fraction of all pulses in all geosegs used that were nearly saturated	***near_sat_fract_seg***
neutat_delay_total_seg	m	Ocean segment average of total neutral atmosphere delay correction (wet + dry)	***neutat_delay_total_seg***
orbit_number	counts	Unique identifying number for each planned ICESat-2 orbit	***orbit_number***
photon_noise_rate	m^{-1}	Noise photon count rate, averaged over the segment	r_{noise}
photon_rate	m^{-1}	Photon count rate, averaged over the segment	r_{surf}
podppd_flag_seg		Higher, 0 or 4, of ATL03 **gtx/geolocation/podppd_flag** of data used in the ocean segment	*podppd_flag*
ref_azimuth_seg	radians	Ocean segment average of azimuth of the unit pointing vector for the reference photon in the local ENU frame in radians. The angle is measured from North and positive towards East	***ref_azimuth_seg***
ref_elev_seg	radians	Ocean segment average of elevation of the unit pointing vector for the reference photon in the local ENU frame in radians. The angle is measured from the East-North plane and positive towards Up	***ref_elev_seg***
seg_mean_dist_x	m	Ocean segment average of the along-track distance from the equator crossing to the start of the 20-m geolocation segments included in the ocean segment	***seg_mean_dist_x***
solar_azimuth_seg	degrees	Ocean segment average of the azimuth of the sun position vector from the reference photon bounce point position in the local ENU frame. The angle is measured from North and is positive towards East. The average is provided in degrees.	***solar_azimuth_seg***
solar_elevation_seg	degrees	Ocean segment average of the elevation of the sun position vector from the reference photon bounce point position in the local ENU frame. The angle is measured from the	***solar_elevation_seg***

		East-North plane and is positive towards Up. The average is provided in degrees.	
surf_type_prcnt		The percentages of each **surf_ty**pe of the photons in the ocean segment as a 5-element variable with each element corresponding to the percentage of photons from each of the 5 surface types	***surf_type_prcnt***
tide_earth_free2mean_seg	m	Ocean segment average of the factor that needs to be added to tide_earth_seg to convert from the tide-free to the mean-tide system. This is not used in ATL12.	*tide_earth_free2mean_seg*
tide_earth_seg	m	Ocean segment average of solid earth tides in the tide-free systems	***tide_earth_seg***
tide_equilibrium_seg	m	Ocean segment average of long-period ocean tides	***tide_equilibrium_seg***
tide_load_seg	m	Ocean segment average of local displacement due to ocean loading (-6 to +6 cm)	***tide_load_seg***
tide_oc_pole_seg	m	Ocean segment average of oceanic surface rotational deformation due to polar motion (-2 to +2 mm)	***tide_oc_pole_seg***
tide_ocean_seg	m	Ocean segment average of short-period ocean tides (diurnal and semi-diurnal)	***tide_ocean_seg***
tide_pole_seg	m	Ocean segment average of solid pole tide - Rotational deformation due to polar motion (-1.5 to 1.5 cm)	***tide_pole_seg***

5.6 Gridding DOT for ATL19.

The ATL19 product includes gridded monthly estimates of dynamic ocean topography (DOT) taken from ATL12 ocean segment data. Ocean segments range in length roughly from 3 to 7km. The ocean segment data are averaged in ¼° or 25km grid cells. Data from all six beams (gtxr and gtxl) are used from the beginning to the end of each month, although commonly over most of the ocean only data from the strong beams will be available. The IS-2 data are averaged onto three grids, called mid-latitude, north-polar and south-polar. The ATL19 product has groups with the same names containing the gridded data for each of those regions. The grids, gridding schemes, input variables and output variables are described in detail in the ATBD for ATL19.

5.7 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 RMS difference 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/ICESat-2_SDT/ICESat-2_OceanATBD_Development/PhotonSourceDistribution
load XmitHist000
meanXmit00=sum(XmitHist000(:,1).*XmitHist000(:,2))/sum(XmitHist000(:,2))
stdXmit00=sqrt(sum(XmitHist000(:,1).*XmitHist000(:,1).*XmitHist000(:,2))/sum(XmitHist000(:,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(XmitHist000(:,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.


% Convert back to PDF
Noise=std(noisey/(NptsMix*binsize)-rcvhist);
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).

5.8 Numerical Computation Considerations

TBD – as needed specific considerations on method of code computation

5.9 Programmer/Procedural Considerations

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

5.10 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 the project personnel can do it, 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 perhaps funded outside the Cryosphere program.
3) Direct comparison to in situ precision GPS measurements. Buoys are being built with NOPP funding to do this as an add-on to similar measurements for NASA Oceanography's SWOT cal/val effort.

6.0 BROWSE PRODUCTS

These browse figures show images that allow user to easily evaluate if the data would be useful and of quality for their research and will aid in the quick approval or disapproval of products prior to public distribution.

6.1 Data Quality Monitoring

6.1.1 **Line plots** (each strong beam)

a) mean sea surface height for each segment
b) standard deviation of height distribution and associated SWH for each segment
c) skewness of height distribution for each segment
d) kurtosis of height distribution for each segment

6.1.2 **Histograms** (each strong beam)

a) Line plots of parameters of 2-Gaussian fit for each segment
b) TBD quality measure of each 2-Gaussian fit

7.0 DATA QUALITY

7.1 Statistics

Calculate the following statistics for each open ocean granule and broken down by strong beam ground track – so there is one set for each ground track.
• Aggregate histogram and mean/standard deviation of each of the following parameters

7.1.1 Per orbit statistics

Table 7 ATL12 Output Variables for Per Orbit Statistics

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_sig1	Standard deviation of component 1 in 2-component Gaussian mixture	***Sig1***
mix_m2	Fraction of component 1 in 2-component Gaussian mixture	***m2***
mix_mu2	Mean of component 1 in 2-component Gaussian mixture	***mu2***
mix_sig2	Standard deviation of component 1 in 2-component Gaussian mixture	***Sig2***
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 (m)	***length_seg***
P0	Constant of linear fit versus along-track distance to surface photon height over the ocean segment	***P0***
P1	Slope of linear fit versus along-track distance to surface photon height over the ocean segment	***P1***
n_pulse_seg	Number of laser pulses in segment	***n_pls_seg***
meanoffit2	Mean of linear fit removed from surface photon height	***meanoffit2***
Y	Probability density function of photon surface height	***Y***
Ymean	Mean=first moment of Y, should be ~0 = *h* -meanoffit2	***Ymean***
Yvar	Variance = second moment (m^2) of *Y*	***Yvar***
Yskew	Skewness = (third moment of *Y*)/ *Yvar*$^{3/2}$	***Yskew***
Ykurt	Excess Kurtosis = (fourth moment of *Y*)/ *Yvar*^{2} -3	***Ykurt***

bin_ssbias	Sea state bias estimated from the correlation of photon return rate with along-track 10-m bin averaged surface height (4.3.1)	***bin_ssbias***
SWH	Significant wave height estimated as 4 times the standard deviation of along-track 10-m bin averaged surface height	***SWH***
Height_segment_stat		
n_ttl_photon	Number of photons in the ±15-m ocean downlink band	***n_ttl_photon***
n_photons	Number of surface photons found for the segment	***n_photon***
Photon_rate	Photon count rate, averaged over the segment	r_{surf}
Photon_noise_rate	Noise photon count rate, averaged over the segment	r_{noise}
backgr_seg	*backgrd_atlas/bckgrd_rate* from ATL03 averaged over the segment	***backgr_seg***
first_pce_mframe_cnt	First Major Frame ID in the SSH segment	***first_pce_mframe_cnt***
first_tx_pulse	First Transmit pulse in along-track segment	***first_tx_pulse***
layer_flag_seg	The layer flag from ATL09 that is in effect over 50% of the ocean segment, 0 indicating absence of clouds and forward scattering, and 1 indicating possibility of forward scattering as in ATL09	***layer_flag_seg***
depth_ocn_seg	The average of depth_ocn of geo-segments used in the ocean segment.	***depth_ocn_seg***
orbit_number =	Unique identifying number for each planned ICESat-2 orbit	***orbit_number***
ss_corr	Subsurface scattering correction, placeholder = zero pending further findings to the contrary	***ss_corr***
ss_corr_stdev	Estimated error of subsurface scattering correction, placeholder = zero pending further findings to the contrary	***ss_corr_stdev***
neutat_delay_total_seg	Ocean segment average of total neutral atmosphere delay correction (wet + dry)	***neutat_delay_total_seg***
seg_dist_x _seg	Ocean segment average of the along-track distance from the equator crossing to the start of the 20-m geolocation segments included in the ocean segment	***seg_dist_x _seg***
solar_azimuth_seg	Ocean segment average of the azimuth of the sun position vector from the reference photon bounce point position in the local ENU frame. The angle is measured from North and is positive towards East. The average is provided in degrees.	***solar_azimuth_seg***
solar_elevation_seg	Ocean segment average of the elevation of the sun position vector from the reference photon bounce point position in the local ENU frame. The angle is measured from the East-North plane and is positive towards Up. The average is provided in degrees.	***solar_elevation_seg***
geoid_seg	Ocean segment average of geoid height above the WGS – 84 reference ellipsoid (range -107 to 86 m)	***geoid_seg***
dac _seg	Ocean segment average of dynamic atmospheric correction (DAC) includes inverted barometer (IB) affect	***dac _seg***

tide_earth_seg	Ocean segment average of solid earth tides in the tide-free system	***tide_earth_seg***
tide_load_seg	Ocean segment average of local displacement due to ocean loading (-6 to +6 cm)	***tide_load_seg***
tide_oc_pole_seg	Ocean segment average of oceanic surface rotational deformation due to polar motion (-2 to +2 mm)	***tide_oc_pole_seg***
tide_ocean_seg	Ocean segment average of short-period ocean tides (diurnal and semi-diurnal)	***tide_ocean_seg***
tide_equilibrium_seg	Ocean segment average of long-period ocean tides	***tide_equilibrium_seg***
tide_pole_seg	Ocean segment average of solid pole tide - Rotational deformation due to polar motion (-1.5 to 1.5 cm)	***tide_pole_seg***
surf_type_prcnt	The percentages of each ***surf_ty***pe of the photons in the ocean segment as a 5-element variable with each element corresponding to the percentage of photons from each of the 5 surface types	***surf_type_prcnt***

8.0 TEST DATA

This section describes the test data sets that have been used to derive and verify the performance of the ATL12 code for surface finding, deconvolution of the instrument impulse response, and Gaussian mixture determination.

8.1 In Situ Data Sets

Table 8 ATL12 Test Data

Instrument	Date	Location
MABEL	April 2012	Fram Strait and Greenland Sea
MABEL	July 2014	North Pacific

8.2 Simulated Test Data

Development of ATL12 used the simulated data set described in Section 5.6 Synthetic Data. For these synthetic data we started with a synthetic sea surface height record comprised of a known 2-Gaussian mixture that we convolved with a known instrument impulse response and added random measurement noise. We exercised the developmental Matlab code and the ASAS PGE code on this data to see that they output the same 2-Gaussian mixture that was the basis of the synthetic ocean surface

9.0 CONSTRAINTS, LIMITATIONS, AND ASSUMPTIONS

In this section, we list notable constraints, limitations, and assumptions important to ATL12 that affect the coverage, quality and interpretation of the sea surface height retrievals and our objectives for the ATL12. These topics have been discussed throughout the ATBD. In order to compare with other open ocean topography measures, the overarching objective for the ATL12 product is that it should be a solid description of sea surface height probability distribution over each strong beam ground track (and weak beam ground track where available). This objective and the constantly changing wave-covered character of the ocean surface impose the primary constraint on the ATL12 product. To achieve standard errors in mean sea surface height better than 1-cm under typical sea states, we accumulate the distributions over distances of 5 to 7 km. This is far longer than the ultimate spatial resolution of the ICESat-2 instrument, which we in fact use in the ATL12 analysis for sea state bias estimation. Users can aggregate the ATL 12 distributions to obtain greater statistical significance, and we do this to produce the ocean gridded product, ATL19. Where applicable we and other users can go back to ATL03 photon heights and derive specialized ocean products using ATL12 statistics as a guide and benchmark.

9.1 Constraints

The following are constraints imposed by the inherent capability of the instrument.
• At 532 nm, clouds will affect visibility of the open ocean. First impressions looking at preliminary ICESat-2 data the lidar appears able detect the ocean surface through at least thin clouds better than expected.
• The reflectance of the ocean surface is lower than the other surfaces ICESat-2 operates over with the result that only on the order of one photon per pulse is returned from the surface. Thus, longer path lengths are required to accumulate desired statistical significance of ocean height measurements than would be required for other surfaces. On the other hand, we don't have to worry about detector saturation and first photon bias over the ocean.

9.2 Limitations

These limitations stem from on our current understanding of the altimetric returns from the sea ice cover.
• Subsurface Scattering - Quantification of the impact of subsurface scattering on height retrievals due to multiple scattering from bubbles or particles in the water remains to be addressed. For the central parts of the deep ocean, these effects are likely reduced by the scarcity of scatterers in the water. We might also expect that just as surface waves reduce the reflectance of the surface, they may also reduce the transmittance of subsurface returns up through the ocean surface, a benefit offsetting the generation of bubbles by wind waves. Preliminary analysis of ICESat-2 deep ocean data suggests long, but low energy negative tails in the raw height distribution that are mainly removed in the trimming of the histogram in the fine surface finding step. Subsurface scattering will be most important in coastal regions where sediment and biological productivity increase the density of scatterers. Gaining a better understanding the subsurface return problem will be a high priority for future ICESat-2 oceanography.
• Sea State Bias – ICESat-2 will be unique in estimating what we think will be the main source of sea state bias by calculating the correlation of photon return rate and surface height at scales

resolving the energy containing surface waves. This constitutes what is termed EM-bias in other altimeters. Because we don't use retracking in the traditional sense or assume a Gaussian surface distribution, we anticipate ICESat-2 will not suffer from sea state bias due to these factors, and the ICESat-2 sea state bias is wholly due to EM-bias. We will assess this assumption as part of the validation process.

9.3 Assumptions

These are assumptions made by the retrieval routines that will have to be tested.
• Ocean Surface Height retrievals – As discussed above key assumptions are that our fine surface finding clips off the negative histogram tail due to subsurface return and that what is commonly termed EM-bias is the sole source of sea state bias in our analysis of ICESat-2 over the ocean.

10.0 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.

ACRONYMS

Acronym List

ASAS	ATLAS Science Algorithm Software
ATLAS	ATLAS Advance Topographic Laser Altimeter System
GSFC	Goddard Space Flight Center
ICESat-2 MIS	ICESat-2 Management Information System
IIP	Instrument Impulse Response
MIZ	Marginal Ice Zone
PSO	Project Science Office
PSO	ICESat-2 Project Support Office
SDMS	Scheduling and Data Management System
SIPS	Science Investigator-led Processing System
TEP	Transmit Echo Pulse

APPENDIX A: ICESat-2 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

APPENDIX B: Sea State Bias Computations for a Photon Counting Lidar

For ICESat-2, we are looking over all the received surface photon heights to get a received height distribution and then deconvolving the instrument impulse response to get the surface height distribution. If we are getting a disproportionate share of photons back from the wave crests or troughs because of a difference in reflectance between crests and troughs, the correlation of height and reflectance (or in this case rate of photon returns) and the sea state bias it causes are unknown to us. However, if as a separate step we take advantage of the fine spatial resolution of ICESat-2 to obtain along-track distance bin average estimates of height and photon return rate unbiased by the number of photons in the bins, and over bins that are short compared to the dominate surface wavelength, we can estimate the sea state bias in the surface height distribution using the relation of *Arnold et al.* [1995]. To do this we need to consider the *Arnold et al.* relation in terms of bin averages rather than aggregate photon averages.

First we establish the relation between average height computed over all surface photons and the average taken over all photons grouped in K along-track distance bins. The photon average sea surface height is the sum the heights of N surface reflected photons divided by N:

$$H_{photonavg} = \frac{1}{N}\sum_{i=1}^{N}\eta_i \tag{B1}$$

However, if we express N as the sum over all K along-track bins of the number of photons grouped into each bin, r_j, we see that:

$$N = \sum_{j=1}^{K} r_j \tag{B2}$$

Similarly, we can group the heights into the contributions from each bin

$$\sum_{i=1}^{N}\eta_i = \sum_{j=1}^{K}\sum_{l=1}^{r_j}\eta_l \tag{14- B3}$$

where η_l is the height of the l^{th} photon in a bin with a total of r_j photons in it. Therefore, the average height over all photon heights is equal to the bin average heights weighted by the relative share of photons in each bin, combining (12-14)

$$\frac{1}{N}\sum_{i=1}^{N}\eta_i = \frac{1}{N}\sum_{j=1}^{K}\sum_{l=1}^{r_j}\eta_l = \frac{\sum_{j=1}^{K}\sum_{l=1}^{r_j}\eta_l}{\sum_{j=1}^{K} r_j} \tag{15- B4}$$

In the relation of *Arnold et al.*, [1995] (9), we replace the backscatter coefficient of σ_i^0, by the number of photons per bin r_l as an expression of the energy returned from the surface in each bin.

SSB is the photon average or photon weighted bin average minus the true average of surface heights over all the distance bins, $\sum_{j=1}^{K}\eta_j \Big/ K$:

$$SSB = \frac{1}{N}\sum_{i=1}^{N}\eta_i - \frac{\sum_{j=1}^{K}\eta_j}{K} = \frac{\sum_{j=1}^{K}\sum_{l=1}^{r_j}\eta_l}{\sum_{j=1}^{K} r_j} - \frac{\sum_{j=1}^{K}\eta_j}{K} \tag{B5}$$

where η_j is the true average of height in the j^{th} bin not biased by the number photons, and so the second term on the right is μ_η, the true arithmetic mean of surface height averaged over K bins.

$$\mu_\eta = \frac{\sum_{j=1}^{K}\eta_j}{K} \tag{B6}$$

The number of samples in along-track bin j, is r_j, and though the sample rate with each sample, ρ_j, is potentially different for each sample, we take ρ_j as the average sample rate in each δ_x wide bin such that

$$\delta_x \rho_j = r_j \tag{B7}$$

The average of photon heights in each bin is

$$\bar{\eta}_j = \frac{1}{r_j}\sum_{l=1}^{r_j}\eta_l \tag{B8}$$

So

$$r_j \bar{\eta}_j = \sum_{l=1}^{r_j}\eta_l \tag{B9}$$

and (B5) becomes

$$SSB = \frac{\sum_{j=1}^{K} r_j \bar{\eta}_j}{\sum_{j=1}^{K} r_j} - \frac{\sum_{j=1}^{K} \eta_j}{K} \tag{B10}$$

and with (B6), (B10) becomes

$$SSB = \frac{\sum_{j=1}^{K} r_j \bar{\eta}_j}{\sum_{j=1}^{K} r_j} - \mu_\eta = \frac{\sum_{j=1}^{K} r_j \bar{\eta}_j - \sum_{j=1}^{K} r_j \mu_\eta}{\sum_{j=1}^{K} r_j} = \frac{\sum_{j=1}^{K} r_j \left(\bar{\eta}_j - \mu_\eta\right)}{\sum_{j=1}^{K} r_j} \tag{B11}$$

This looks like the same relation as (9) by *Arnold et al.* [1995] but it is a little unclear if Arnold et al meant to use the backscatter coefficient in his equation or the variation in backscatter coefficient about the mean. To check this for our relation we break height and return rate in $\sum_{j=1}^{K}\sum_{l=1}^{r_j}\eta_l \Big/ \sum_{j=1}^{K} r_j$ into mean, (μ) and time varying (primed) components. Speaking of the along-track bin quantities, $\bar{\eta}_j = \mu_\eta + \eta_j'$ and $r_i = \mu_r + r_j'$. From (B9):

$$\frac{1}{N}\sum_{i=1}^{N}\eta_i = \frac{1}{N}\sum_{j=1}^{K}\sum_{l=1}^{r_j}\eta_l = \frac{\sum_{j=1}^{K}\sum_{l=1}^{r_j}\eta_l}{\sum_{j=1}^{K} r_j} = \frac{\sum_{j=1}^{K} r_j \bar{\eta}_j}{\sum_{j=1}^{K} r_j} \tag{B12}$$

If the along-track bins are small enough so that the sea surface height variation within a bin is much smaller than the variation of height between bins (bins are small compared to the wavelength of the dominant waves), we can assume that the average of photon heights within bin, $\bar{\eta}_j$, equals the true average of height within a bin, η_j, and the true average height is equal to the average of $\bar{\eta}_j$ over all bins.

$$\mu_\eta = \frac{\sum_{j=1}^{K} \eta_j}{K} = \frac{\sum_{j=1}^{K} \bar{\eta}_j}{K} \tag{B13}$$

This assumption is equivalent to saying the bias due to variation of the number of photons within bins is small compared to the bias due to the differences in the number of photons among bins. Similarly, the average photon rate, μ_ρ, is

$$\mu_\rho = \frac{\sum_{j=1}^{K} r_j}{K} \tag{B14}$$

Applying (B9) to (B4) and breaking into mean and varying parts

$$\frac{1}{N}\sum_{i=1}^{N}\eta_i = \frac{\sum_{j=1}^{K}\sum_{l=1}^{r_j}\eta_l}{\sum_{j=1}^{K} r_j} = \frac{\sum_{j=1}^{K} r_j\,\bar{\eta}_j}{\sum_{j=1}^{K} r_j} = \frac{\sum_{j=1}^{K} r_j\,\eta_j}{\sum_{j=1}^{K} r_j} = \frac{\sum_{j=1}^{K}(r_j'+\mu_r)(\eta_j'+\mu_\eta)}{\sum_{K=1}^{K} r_j}$$

$$= \frac{\sum_{j=1}^{K}(r_j'\eta_j'+r_j'\mu_\eta+\mu_r\eta_i'+\mu_r\mu_\eta)}{\sum_{j=1}^{K} r_j} \tag{B15}$$

$$= \frac{\sum_{j=1}^{K}(r_j'\eta_j'+\mu_r\mu_\eta)}{\sum_{j=1}^{K} r_j} = \frac{\sum_{j=1}^{K}(r_j'\eta_j')}{\sum_{j=1}^{K} r_j} + \frac{K\mu_r\mu_\eta}{\sum_{j=1}^{K} r_j} = \frac{\sum_{j=1}^{K}(r_j'\eta_j')}{\sum_{j=1}^{K} r_j} + \mu_\eta$$

So in fact the product in the SSB expression of *Arnold et al.* (1995) is the product of the variations in distance bin photon rate and photon height:

$$SSB = \frac{\sum_{j=1}^{K}(r_j'\eta_i')}{\sum_{j=1}^{K} r_j} \tag{B16}$$

Alternatively, consider breaking down (B11)

$$SSB = \frac{\sum_{j=1}^{K} r_j\,\bar{\eta}_j}{\sum_{j=1}^{K} r_j} - \mu_\eta = \frac{\sum_{j=1}^{K} r_j\,\bar{\eta}_j - \sum_{j=1}^{K} r_j\,\mu_\eta}{\sum_{j=1}^{K} r_j} = \frac{\sum_{j=1}^{K} r_j\left(\bar{\eta}_j - \mu_\eta\right)}{\sum_{j=1}^{K} r_j} \tag{B17}$$

with, $\bar{\eta}_j = \mu_\eta + \eta_j'$ and $r_i = \mu_r + r_j'$.

$$SSB = \frac{\sum_{j=1}^{K} r_j \bar{\eta}_j}{\sum_{j=1}^{K} r_j} - \mu_\eta = \frac{\sum_{j=1}^{K} (r_j' + \mu_r)(\eta_j' + \mu_\eta)}{\sum_{j=1}^{K} r_j} - \mu_\eta = \frac{\sum_{j=1}^{K} (r_j' \eta_j')}{\sum_{j=1}^{K} r_j} + \mu_\eta - \mu_\eta = \frac{\sum_{j=1}^{K} (r_j' \eta_j')}{\sum_{j=1}^{K} r_j} \quad \text{(B18)}$$

consistent with (B16).

We propose that there is a sea state bias in the higher moments also. If we consider higher sample n^{th} moments of height

$$M(n) = \frac{\sum_{j=1}^{K} (\eta_j - \mu_\eta)^n}{K} = \frac{\sum_{j=1}^{K} g_j(n)}{K} \quad \text{(B19)}$$

As with (B15):

$$M(n) = \frac{1}{N} \sum_{i=1}^{N} g_i(n) = \frac{\sum_{j=1}^{K} \sum_{l=1}^{r_j} g_j(n)}{\sum_{j=1}^{K} r_j} = \frac{\sum_{j=1}^{K} r_j \bar{g}_j(n)}{\sum_{j=1}^{K} r_j} = \frac{\sum_{j=1}^{K} r_j g_j(n)}{\sum_{j=1}^{K} r_j} = \frac{\sum_{j=1}^{K} (r_j' + \mu_r)(g_j(n)' + \mu_g)}{\sum_{K=1}^{K} r_j}$$

$$= \frac{\sum_{j=1}^{K} (r_j' g_j(n)' + r_j' \mu_g + \mu_r g_j(n)' + \mu_r \mu_g)}{\sum_{j=1}^{K} r_j}$$

$$= \frac{\sum_{j=1}^{K} (r_j' g_j(n)' + \mu_r \mu_g)}{\sum_{j=1}^{K} r_j} = \frac{\sum_{j=1}^{K} (r_j' g_j(n)')}{\sum_{j=1}^{K} r_j} + \frac{K \mu_r \mu_g}{\sum_{j=1}^{K} r_j} = \frac{\sum_{j=1}^{K} (r_j' g_j(n)')}{\sum_{j=1}^{K} r_j} + \mu_g$$

Thus, the $(\eta_j - \mu_\eta)^n$, n$^{\text{th}}$ moment, would have a similar derivation for their sea state bias, *SSBM(n)*:

$$SSBM(n) = \frac{\sum_{j=1}^{K} r_j' ((\eta_j - \mu_\eta)^n - avg(\eta_j - \mu_\eta)^n)}{\sum_{j=1}^{K} r_j} \quad \text{(B20)}$$

APPENDIX C: Expectation-Maximization (EM) Procedure

From Appendix E of ATBD ATL07

EM algorithm for estimating the parameters of two-component normal mixtures
The EM algorithm estimates the parameters of a two-component normal mixture that best describe a distribution of random variables (*Dempster et al.,* 1977*; Bilmes*, 1998).
The two-component mixture model:

$$f(x_i) = \alpha\phi_i(x_i;\mu_1,\sigma_1) + (1-\alpha)\phi_2(x_i;\mu_2,\sigma_2)$$

$$\phi_i(x;\mu_i,\sigma_i) = \frac{1}{\sqrt{2\pi\sigma_i}} e^{\frac{-(x-\mu_i)^2}{2\sigma_i^2}} \quad \text{(Normal Distribution)}$$

The EM steps to compute the parameters of the mixture model with T random variates $x_t(\alpha,\mu_1,\sigma_1,\mu_2,\sigma_2)$ are:

1. Initialize with: $\alpha,\mu_1,\sigma_1,\mu_2,\sigma_2$
2. Expectation (E) step: compute:

$$\gamma_i = \frac{\alpha\phi_i(x_t;\mu_1,\sigma_1)}{\alpha\phi_i(x_t;\mu_1,\sigma_1) + (1-\alpha)\phi_2(x_t;\mu_2,\sigma_2)}$$

for t = 1,...,T.

3. Maximization (M) step. compute:

$$\mu_1 = \frac{\sum_{t=1}^{T}\gamma_t x_t}{\sum_{t=1}^{T}\gamma_t}, \quad \sigma_1^2 = \frac{\sum_{t=1}^{T}\gamma_t (x_t-\mu_1)^2}{\sum_{t=1}^{T}\gamma_t}$$

$$\mu_2 = \frac{\sum_{t=1}^{T}(1-\gamma_t) x_t}{\sum_{t=1}^{T}(1-\gamma_t)}, \quad \sigma_2^2 = \frac{\sum_{t=1}^{T}(1-\gamma_t)(x_t-\mu_2)^2}{\sum_{t=1}^{T}(1-\gamma_t)}, \text{ and}$$

$$\alpha = \sum_{t=1}^{T}\frac{\gamma_t}{T}$$

IIterate E and M steps until parameters converge.

APPENDIX D: Fitting a Plane to Spatially Distributed Data

To evaluate the average DOT at the center of a grid cell we fit a plane to all the samples within the grid cell and evaluate the height of the plane at the center of the grid cell.
To do this we first have to make a least squares fit of a plane to DOT at *N* locations within the grid cell. Following *Eberly* (2019), given data (DOT) as a function of *x* and *y*, $h_i = f(x_i, y_i)$ at *i*=1 to *N* locations, find a least squares fit of a plane, coefficients *a*, *b*, and *c*, to h_i with mean, $\overline{h} = \sum_{i=1}^{N} h_i$. (Also note $\overline{x} = \sum_{i=1}^{N} x_i$ and $\overline{y} = \sum_{i=1}^{N} y_i$)

$$h_i \approx a x_i + b y_i + c \tag{1}$$

And we want to choose *a*, *b*, and *c* such that the error, *E,*

$$E(a,b,c) = \sum_{i=1}^{N} ((a x_i + b y_i + c) - h_i)^2$$

is minimized. According to *Eberle* (2019) the solution is more robust and the equations are simpler if we initially eliminate the need to determine *c* by taking the average of (1) and subtracting it from (1), to get:

$$h_i - \overline{h} \approx a(x_i - \overline{x}) + b(y_i - \overline{y})$$

and so for the deviations $h_i' = h_i - \overline{h}$, $x_i' = x_i - \overline{x}$, and $y_i' = y_i - \overline{y}$ (2)

$$h_i' \approx a x_i' + b y_i' \tag{3}$$

and we will choose *a* and *b* to minimize:

$$E(a,b) = \sum_{i=1}^{N} ((a x_i' + b y_i') - h_i')^2 \tag{4}$$

Then c will be given by $c = \overline{h} - (a\overline{x} + b\overline{y})$. To minimize *E* with respect to *a* and *b* we find *a* and *b* for which:

$$\partial E(a,b) \Big/ \partial a = 2\sum_{i=1}^{N} x_i'((a x_i' + b y_i') - h_i') = 2\sum_{i=1}^{N} a x_i' x_i' + b x_i' y_i' - x_i' h_i' = 0$$

$$\partial E(a,b) \Big/ \partial b = 2\sum_{i=1}^{N} y_i'((a x_i' + b y_i') - h_i') = 2\sum_{i=1}^{N} a y_i' x_i' + b y_i' y_i' - y_i' h_i' = 0$$

or

$$a\sum_{i=1}^{N} x_i' x_i' + b\sum_{i=1}^{N} x_i' y_i' = \sum_{i=1}^{N} x_i' h_i'$$
$$a\sum_{i=1}^{N} y_i' x_i' + b\sum_{i=1}^{n} y_i' y_i' = \sum_{i=1}^{N} y_i' h_i'$$

Letting

$$L_{xx} = \sum_{i=1}^{N} x_i' x_i' \qquad R_{xh} = \sum_{i=1}^{N} x_i' h_i'$$
$$L_{yy} = \sum_{i=1}^{n} y_i' y_i' \quad \text{and} \quad R_{yh} = \sum_{i=1}^{N} y_i' h_i' \tag{5}$$
$$L_{xy} = \sum_{i=1}^{N} y_i' x_i'$$

a, b, and c are given by

$$a = \frac{R_{xh}L_{yy} - R_{yh}L_{xy}}{L_{xx}L_{yy} - L_{xy}^2}$$
$$b = \frac{R_{yh}L_{xx} - R_{xh}L_{xy}}{L_{xx}L_{yy} - L_{xy}^2} \tag{6}$$
$$c = \overline{h} - (a\overline{x} + b\overline{y})$$

Eberly, D., 2019, Least Squares Fitting of Data by Linear or Quadratic Structures
David Eberly, Geometric Tools, Redmond WA 98052
https://www.geometrictools.com/

Created: July 15, 1999
Last Modi_ed: February 14, 2019
https://www.geometrictools.com/Documentation/LeastSquaresFitting.pdf

APPENDIX E: Hierarchy of ATL12 and ATL10 Variables

ATL12 Inputs to ATL19

Segment Averages: ***SSH- geoid_seg, lat_seg***, ***lon_seg, bin_ssbias, geoid_seg, depth_ocn_seg, length_seg,*** and ***surf_type_prcnt***.
Segment Moments: ***SSHvar, SSHskew, SSHkurt, SWH***
Segment Histogram: ***Y, n_photons***
Segment Degrees-of_Freedom: ***NP_effect***

5.6.4.1.1 Output ATL19 Averaging Over ***n_segsinbin*** Bins

Grid Cell Averages: ***dot_avg***, ***grid_lat, grid_lon, ssb_avg, geoid_avg***, ***depth_avg***, ***length_avg, surf_avg***
Grid Cell Average Moments: ***dot_sigma_avg, dot_skew_avg, dot_kurt_avg, SWH_avg***
Grid Cell Total Histogram: ***dot_hist_grid***
Grid Cell Totals: ***n_segsinbin, n_photons_gridttl, length_sum***

5.6.4.1.2 Output ATL19 Averaging Weighted by Degrees-of-Freedom

Grid Cell Degree-of-Freedom Weighted Averages ***dot_dfw, grid_lat_dfw, grid_lon_dfw, , ssb_dfw, geoid_dfw***, ***depth_dfw***, ***length_dfw***, ***surf_dfw,***
Grid Cell Degree-of-Freedom Weighted Moments ***dot_sigma_dfw, dot_skew_dfw***, ***dot_kurt_dfw, SWH_dfw***
Grid Cell Degrees-of-Freedom and DOT Uncertainty: ***dof_grid , dot_dfw_uncertain***

5.6.4.2.1 Output ATL19 Interpolated to Bin Centers (***gridcntr_lon*** and ***gridcntr_lat***)

Average DOT at Grid Cell Center: ***dot_avgcntr***
5.6.4.2.2 Output ATL19 Interpolated to Bin Centers (***gridcntr*** and ***gridcntr_lat***)
DOF Weighted Average DOT at Grid Cell Center: ***dot_dfwcntr***

5.6.4.5.1 Output ATL19 Merging All 3 Strong Beams of DOT Averaged by ***n_segsinbin***

Average DOT at Grid Cell Center: ***dot_allbeam***

5.6.4.5.2 Output ATL19 Merging All 3 Strong Beams of DOT Averaged by ***dof_grid***

Degree-of_Freedom Weighted Average DOT at Grid Cell Center: ***dot_dfwallbeam***
Degree-of_Freedom Weighted Average DOT at Grid Cell Center: ***dot_sigma_dfwalbm***
Degree-of_Freedom Weighted Average DOF and Uncertainty at Grid Cell Center***:***
dof_grid_allbeam, , dot_dfwalbm_uncrtn