Ice, Cloud and Land Elevation Satellite 2 (ICESat-2) Algorithm Theoretical Basis Document (ATBD) For Coastal and Nearshore Along-Track Bathymetry Product (ATL24)

Lori Magruder[1], Christopher Parrish[2], Jeff Perry[1], J.P. Swinski[3], Matthew Holwill[1], Keana Kief[2], and Forrest Corcoran[2]

[1]University of Texas at Austin, Center for Space Research, Austin, Texas
[2]School of Civil and Construction Engineering, Oregon State University, Corvallis, Oregon
[3]NASA Goddard Space Flight Center, Greenbelt, Maryland

March 19, 2025

Key Points

- ATL24 is the Level 3 along-track bathymetry product using ICESat-2 measurements
- ATL24 provides the photon level classification of sea surface and seafloor using AI/ML methods for coastal and nearshore regions
- ATL24 provides both sea surface and refraction-corrected seafloor heights, with associated point classification confidence values and height uncertainty values

Version: 1.0
Release Date: April 1, 2025
This document should be cited as:
Magruder, L., Parrish, C., Perry, J., Swinski, J.P, Holwill, M., Kief, K., and Corcoran, F. (2025). *Ice, Cloud and Land Elevation Satellite (ICESat-2) Project Algorithm Theoretical Basis Document (ATBD) for Coastal and Nearshore Along-track Bathymetry Product (ATL24).* 1.0. NASA. DOI: 10.5067/PXJMCZD0MYLN

Change History Log

Table 1: *History of changes*

Change Log		
Date	Version	Change
July 8, 2024	0.5	Initial draft
March 17, 2025	1.0	ATL24 r001

Contents

List of Figures

List of Tables

List of Algorithms

1 Introduction

Along with solid earth, cryosphere, vegetation structure, and hydrology, coastal processes are recognized focus topics by the National Academy of Science as in need of better understanding regarding how they respond to climate variability and how those changes will subsequently affect societal resources (National Academies of Sciences and Medicine 2018). Improving our understanding of coastal processes such as wave structure, tides, currents, sediment transport (erosion and deposition), and benthic environments requires knowledge of shallow water bathymetry.

Unfortunately, there is currently a global lack of nearshore bathymetric data, which hinders a number of coastal and marine science, management and engineering applications. Important uses of nearshore bathymetry include analysis of nearshore hydrodynamics, morphology, and sediment transport (Plant et al. 2002), tsunami inundation modeling (Titov et al. 2005), and benthic habitat mapping (Costa et al. 2009), among others (Gao 2009). Due to the importance of shallow nearshore bathymetric data and the challenges in acquiring the data with conventional, ground-based surveying technologies, remote sensing methods have been of interest since early studies into the feasibility of bathymetric measurement from aerial photography in the 1940s (Moore 1947). This interest in remotely-sensed bathymetry began accelerating in the mid-1970s (Polcyn and Lyzenga 1975; Warne 1978), as satellite imagery began to become publicly available, and has only continued to grow in subsequent decades (Christopher E. Parrish, L. A. Magruder, et al. 2019; Gao 2009). Coastal communities, in particular, face critical issues associated with the influence of climate change on local sea level and natural resource availability. The coasts experience heightened risks continually, but recent studies indicate an increased timeline of negative impact making future studies with remote sensing critical to improving resiliency and deriving mitigation strategies to protect coastal environments.

Nearshore bathymetry is notoriously difficult to measure. Conventional, terrestrial surveying technologies, such as GNSS and total stations, are not designed for underwater data acquisition. Hence, these technologies are only viable in the shallowest waters, in which a person can stand with a survey rod, and, even then, such methods can be laborious, dangerous (especially, in high-energy nearshore environments), and expensive. Meanwhile, boat-based acoustic hydrographic surveying technologies, such as multibeam echosounders (MBES), enable efficient and highly-accurate data acquisition in deeper waters ($>$ 4-5 m). However, MBES is inefficient in very shallow areas, and ships and even small boats face safety challenges in the shallowest areas, especially in the presence of nearshore hazards to navigation. These hazards prompted the National Oceanic and Atmospheric Administration (NOAA) hydrographic surveying fleet to implement a policy of not surveying shoreward of a pre-defined boundary, termed the navigable area limit line (NALL). The NALL is generally the seaward-most of: a) the 3.5-m depth contour, or b) a distance of 0.8 mm at chart scale seaward of the shoreline (e.g., 64 m seaward of the shoreline at a chart scale of 1:80,000) (NOAA 2021). However, the NALL can be set even further offshore in the presence of hazards, such as kelp, rocks, or breaking waves.

Another local method for acquiring shallow water bathymetry along the NALL is airborne bathymetric lidar. These surveys are quite detailed and provide high resolution, highly accurate seafloor retrievals but are limited in terms of cost and are often infeasible in remote coastal locations. The motivation for airborne surveys is both in filling a measurement gap along the coastlines to characterize benthic environmental conditions and also to pursue a

seamless topographic/bathymetric surface in support of better modeling associated sediment transport, hydrodynamics, inundation and discharge that are improved in the absence of discontinuities. Specific examples of priority shallow-water benthic habitats include coral reefs, eelgrass beds, and kelp forests, all of which are important components of our nearshore communities in terms of resources and water quality. In the U.S., airborne bathymetric lidar has been collected over a number of habitats, including in Florida, Hawaii, and the U.S. Virgin Islands (Brock et al. 2004; Wedding et al. 2008; Wilson et al. 2019) with a focus on coral reef health and protection of marine ecosystems. In this example, the airborne lidar is critical, but there remain voids in a myriad of other locations where local airborne data is not available which hinders long-term monitoring efforts at the global scale. Beyond coral reefs, numerous other important and fragile benthic habitats exist worldwide, with each responding differently to climate variability and significant storm events. It is important that we determine a capability through direct measurement or data fusion that we can achieve the vertical accuracy and spatial extent required to understand drivers of nearshore change and inform policies relative to quantitative results of the structural and spectral variations that will influence our future communities.

An opportunity to capture shallow water bathymetry at the global scale was realized in 2018 after the launch of NASA's Earth observing laser altimeter mission, ICESat-2 (Ice, Cloud and Land Elevation Satellite-2). ICESat-2 was developed as a means to provide the science community significant insight into changes in the mass of ice sheets, sea ice, and vegetation heights and follows the predecessor mission, ICESat (2003-2009) (Schutz et al. 2005). Both missions were motivated by similar science objectives, primarily associated with monitoring the cryosphere for temporal and spatial variability, and secondary goals for vegetation structure. For ICESat-2 specifically, there were nine level 1 science requirements that included elevation change rates of our polar ice sheets and sea ice, in both the Southern Ocean and the Arctic sea (Markus et al. 2017. The requirements were increasingly more challenging than the predecessor mission which invoked development of new lidar technology focused primarily on increasing the possible spatial resolution both along-track (increased laser pulse repetition rates) and across-track (multiple beams). The resulting instrument, ATLAS (Advanced Topographic Laser Altimeter System), is a photon-counting, 532 nm wavelength, 6 beam, 10 kHz system that provides 0.7 m along-track resolution and six elevation profiles across 6.6 km of terrain swath. The orbit is in a 91-day repeat configuration to facilitate capture of seasonal trends of surface change and contribute to annual statistics of variability (L. Magruder et al. 2021, L. A. Magruder, T. Neumann, et al. 2024, L. A. Magruder, Farrell, et al. 2024).

Although it is well known that wavelengths of 532 nm can penetrate water and are often utilized for bathymetric mapping, due to the micro-pulse energy level and 500 km distance between the spacecraft and the Earth's surface, the bathymetric measurement capabilities of ICESat-2 were largely unknown pre-launch and only investigated in a few studies (e.g., N. Forfinski-Sarkozi and C. Parrish 2016; N. A. Forfinski-Sarkozi and Christopher E. Parrish 2019; Jasinski et al. 2016). Dedicated pre-launch preparation for orbit observations in support of bathymetry was not initiated, as was the case for other science disciplines. The development of the algorithms to produce the initial set of mission data products was initiated in parallel to the technology development during the pre-launch period. The along-track Level 3a (L3a) data products provided to the community by the ICESat-2 Project Science Office (PSO) are specific to surface type. These products include land ice, sea ice, inland water, ocean,

land/vegetation and atmospheric profiles. Each of these L3a data types use ATL03 as input. ATL03, the Level 2b geolocated photon cloud, contains both signal and noise attributed to atmospheric backscatter and/or system response (e.g. saturation effects) T. A. Neumann et al. 2019a. The production algorithms for the L3a products have customized signal finding, optimized length scales, performance metrics and environmental characterization parameters based on expected surface structure, reflectivity and potential science application needs.

Once on-orbit, the ICESat-2 bathymetric capability was revealed, and the community has embraced the contribution to coastal sciences. ATLAS has demonstrated the ability to reach depths up to 40 m in certain environments (Christopher E. Parrish, L. A. Magruder, et al. 2019) and shows great utility with the measurements themselves or in conflation with other data sources. Currently, ICESat-2 has been on-orbit for 6+ years and remains in nominal condition as it operates within its extended mission timeline. Recently, the ICESat-2 PSO has initiated the development of a dedicated L3a product for bathymetry in response to the community's need for data and the complexities and challenges related to signal finding and surface classification (water surface, seafloor). This new product, ATL24, will provide automated bathymetry extraction, sea surface/wave parameters, and water column statistics in all regions that provide adequate conditions for probable measurements. In contrast to other ATLAS data products that were designed pre-launch, a great benefit to the ATL24 development effort is the opportunity to learn from and leverage the great wealth of ICESat-2 bathymetry studies that have been published over the past six years.

2 Context/Background

2.1 Historical Perspective

The algorithms used in this ATBD draw on the long history of airborne lidar bathymetry (ALB) development. An excellent summary of ALB development history is provided in Guenther 2007, starting with the pioneering work of Hickman and Hogg 1969 in the late 1960s, and extending through the NASA-NOAA Airborne Oceanographic Lidar (AOL) in the 1970s (Goodman 1979; Hoge et al. 1980), the experimental advanced airborne research lidar (EAARL) and follow-on EAARL-B in, respectively, the early 2000s and early 2010s (Brock et al. 2004; Wright et al. 2016) to modern commercial ALB systems (Kastdalen et al. 2024. In turn, ALB theory and development extensively leveraged theory and tools from the field of ocean optics (Jerlov 1976; Curtis D Mobley 2001; C. Mobley et al. 2022). Additionally, ATL24 development leverages the significant work that has been published on ICESat-2 bathymetry since the 2018 ICESat-2 launch, including, but not limited to, work done by the ATL24 team and the Bathymetry Working Group of the ICESat-2 Science Team (e.g., Babbel et al. 2021; Christopher E. Parrish, L. Magruder, et al. 2022; T. C. Thomas et al. 2021; Ranndal et al. 2021b; Albright and Glennie 2020).

Beyond processing algorithms, ATL24 also leverages recent developments in open science and data dissemination. NASA has embraced open access and open science in providing important Earth data to the world without constraint. The primary distribution architecture is the Earth Observing system Data and Information system (EOSDIS) with major facilities at the Distributed Active Archive Centers (DAAC) located across the United States. The DAACs act as stewards of the Earth observing satellites and field measurement programs and are responsible for processing, archiving, documenting and distributing data. For the original ICESat mission and for the ICESat-2 mission the designated DAAC is the National Snow and Ice Data Center (NSIDC) in Boulder Colorado. NSIDC is the primary DAAC for snow and ice processes with a particular focus in snow, ice, atmosphere and ocean interactions and has been operational since 1976.

The partnership between the NSIDC team and the ICESat-2 mission has been extremely successful in data product development and distribution via Earthdata Search. The mechanisms of searching and selecting data repositories in NSIDC are well established and provide useful tools for data filtering and visualization.

3 ATL24 Overview

3.1 ATL24 Data Workflow

Similar to other ICESat-2 Level 3a products, the input to the ATL24 pipeline is Level 2a, Global geolocated point cloud data; ATL03 (T. A. Neumann et al. 2019b). ATL03 provides every detected photon (signal and noise), with the calculated geolocation (geodetic latitude and longitude) and associated parameters (operational) for each of ATLAS's six beams. The ATL03 product also provides signal confidence flags and estimated uncertainties at the photon level. Gridded surface masks for land ice, sea ice, land, ocean and inland water products are used within the L3b processing workflows to reduce the volume of data processed and guide the production of these surface-specific, higher-level ICESat-2 data products. The ATL24 workflow requires a similar search approach to limit data processing to coastal and nearshore environments that present a reasonable opportunity for capturing bathymetry. As such, a gridded bathymetry mask based on possible retrievability was created to guide the processing extents and is discussed in detail in section subsection 6.4.

Figure 1 provides the overarching processing pipeline for ATL24, including the search mask step to identify relevant ATL03 granules. The ATL24 algorithm's main goal is to provide a solution for robust, global bathymetric and sea surface signal extraction and classification. The specific classifications in ATL24 are are sea surface (41), and bathymetry (40), all other photons, not classified as 40 or 41 are considered unclassified (0). These point class designations are from the American Society for Photogrammetry and Remote Sensing (ASPRS) LAS Domain Profile for Topo-Bathy Lidar (ASPRS 2013).

In many regards, the classification step in the workflow is the most important and challenging aspect in producing the product. In airborne bathymetric lidar, significant human analyst time is typically spent manually providing the classifications of the data. This human-in-the-loop is primarily due to the lack of any individual algorithm capable of handling the full extent of seafloor morphologies, substrates, depth ranges and cover types (e.g. coral, seagrass, macroalgae) that exist worldwide. Certainly, customized classification algorithms do currently exist for ALB as well as ICESat-2 but are all optimized for location-specific environmental conditions which precludes the scalability of any single solution to a global level.

A novel aspect of the ATL24 workflow is to use an ensemble model for sea surface and seafloor point classification. The overarching idea is that the whole is greater than the sum of the parts: by combining the outputs of a number of base models or algorithms, it is possible to attain better classification results than can be achieved with any individual model or algorithm. For purposes of ICESat-2 seafloor point classification, use of an ensemble is particularly important, since: 1) as noted above, in contrast to the case with ALB, which is typically acquired over small project areas, ATL24 is a global product, and, hence, manual point editing is infeasible–the classification must be entirely automated; and 2) any individual algorithm or model is unlikely to be able to handle the extremely wide range of seafloor types, morphologies, depth ranges, water types, and cover types (e.g., coral, seagrass, macroalgae) that exist throughout coastal and nearshore areas worldwide. For example, one algorithm or model may work well in shallow estuaries, while another works well in shallow coral reefs, and another in deeper, open water. By combining the outputs of diverse models and algorithms, we can achieve good results across a broad range of conditions through a fully-automated approach. Furthermore, the ensemble outputs classification confidence scores, which can be

used alone or in combination with total propagated uncertainty (TPU) values to filter the data.

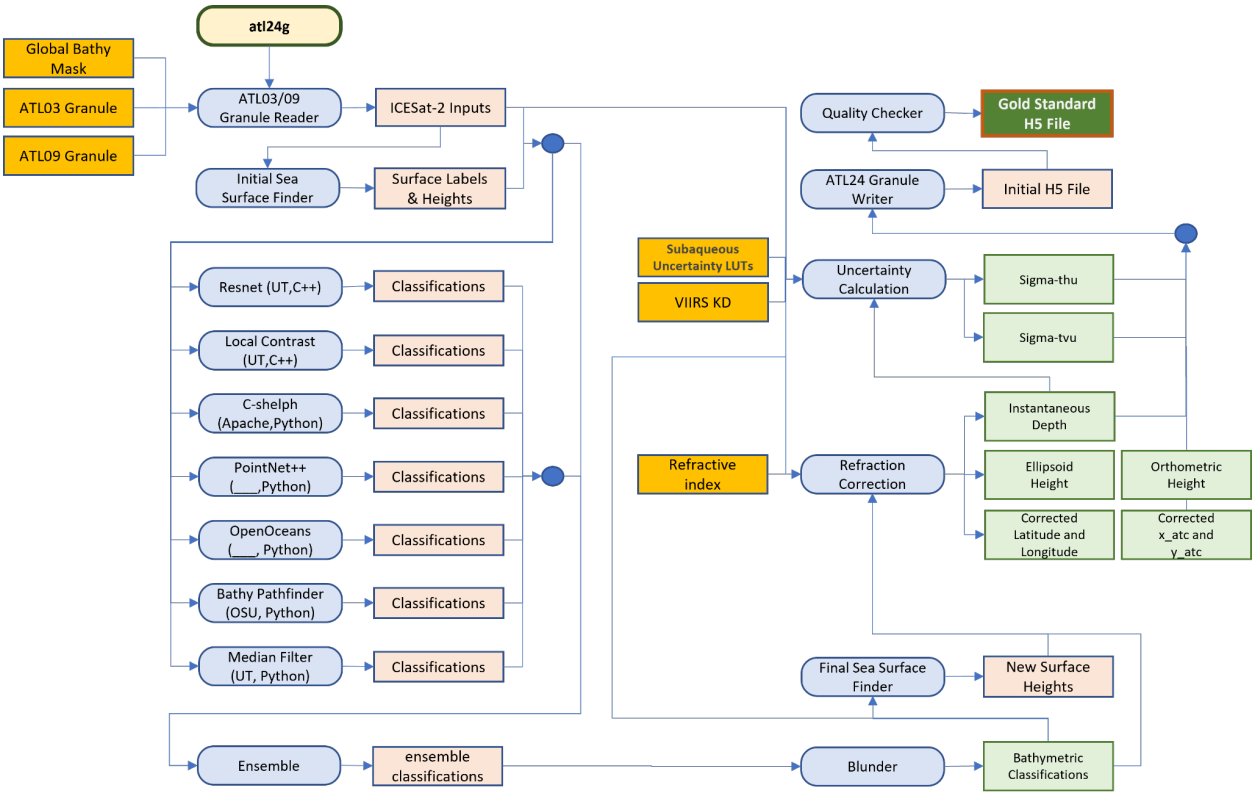

Figure 1: *ATL24 flow diagram of computational architecture*

3.2 Data Dissemination

The implementation architecture chosen for ATL24 r001 will follow the standard approach for the dissemination of the ICESat-2 mission products through the National Snow and Ice Data Center (NSIDC). However, in the last few years NSIDC has pushed the products to the cloud for modernized access to the data, which will be an additional pathway to ATL24 data access for the scientific community. Both means of access will provide capabilities for geographical and temporal sub-setting to the user. Additionally, for ATL24 specifically, there is also a planned, parallel capacity for allowing on-demand and customized, science-ready bathymetry product from ATL03 granules via SlideRule, a public web application programming interface (API) for processing of science data in the cloud (Shean et al. 2023). ATL24 will eventually present a family of data products, which collectively will be referred to as ATL24.x. This ATBD describes the version referred to as the "gold standard" version, ATL24.g, hosted by NSIDC and available in Earthdata cloud, with the metadata for the granules registered in NASA's Common Metadata Repository (CMR). Subsequent versions, ATL24.s and ATL24.p will leverage the full capabilities of SlideRule to provide a subsetting service and on-demand product generation service using a Python client, Javascript client, or web map GUI. This functionality will enable users to optimize the output data product for their particular science need, resulting in truly "science-ready" data. The descriptions of each planned ATL24.x product goals and client service plans are listed below:

ATL24.g The gold standard product will be generated by a private instantiation of SlideRule running in the AWS US-West-2 data center. The granules will initially exist in SlideRule's private S3 bucket prior to transfer to NSIDC. Moving forward, the ATL24.g option will exist in SlideRule as a client facing product with subsetting capabilities.

ATL24.s and ATL24.p Web-services will be provided by the public instantiation of SlideRule. Includes interfacing to the client, and reading the ATL24 granules from S3.

Graphical web interface The interface will be hosted in AWS S3 and served by Amazon's CloudFront at https://client.slieruleearth.io

The gold standard ATL24 product will be generated on a per-granule basis using SlideRule and following the prescribed nearshore/coastal bathymetry mask to coordinate and execute the full suite of contributing classification algorithms. This gold standard data product will be a global resource using the most current algorithmic workflow and will be available to users via sub-setting. Ultimately, the ATL24.g product provides the most robust algorithm parameterization for global applications but does not provide the option for users to adjust the input parameters. Figure 1 shows the execution flow from an incoming ATL24.g request all the way to the output of a gold standard h5 granule.

3.3 ATL24 ATBD Sections

ATL24 primary input is ATL03, using the geolocated photon point cloud to determine classifications of sea surface and seafloor. The many algorithms combined to provide a robust means to signal finding and photon labeling are presented in Section 4.1. This section describes 7 independent methods that each produce predictive classes within the granule and then describes an innovative ensemble machine learning model that provides the final label. Additional description in Section 4.1 is given for the specific input variables and output variables of the ATL24 workflow. After the photons are labeled, there is a correction applied to adjust the spatial coordinates for refraction at the air-water interface. The theory and application of this correction is also provided in Section 4.1. The performance of the algorithm/workflow output is discussed in Section 5. This section includes both the quality or accuracy of the photon labels but also the accuracy of the bathymetric heights relative to independent data sources. Section 5 also contains the approach to modeling the uncertainty of the ATL24 output, which is a parameter included on data product. Section 6 provides the approach to ATL24 production and dissemination, as this product will be the first provided as an on-demand, science ready product for the user as a web service. Section 6 will provide descriptions of SlideRule and both the deployment and development environments. The appendices contain the many acronyms used in the ATL24 ATBD lexicon.

4 Technical Considerations

4.1 Scientific Theory

The scientific theory underlying bathymetric lidar, including ICESat-2 bathymetry, is well-developed, based on over five decades of work, starting with the pioneering experiments of Hickman and Hogg 1969. While the extension from airborne to spaceborne bathymetric lidar is new to ICESat-2, most of the underlying theory and physics remain the same.

The operational principles of bathymetric lidar are largely equivalent to those of topographic lidar, with one major difference: namely, the laser light must penetrate the water column, reach the seafloor with sufficient pulse energy to register a detectable return signal (which is also a function of seafloor reflectance at the laser wavelength), and propagate back to the sensor, where it can be detected and used to compute spatial coordinates of the seafloor point.

Water clarity is consistently the limiting factor in bathymetric lidar. For a photon-counting system, such as ATLAS, an equation for the strength of the received signal, in terms of number of photoelectrons is given in N. Forfinski-Sarkozi and C. Parrish 2016 :

$$\eta_{bot} = \eta_q \eta_t \eta_r \frac{E_t}{h\nu} \rho \cos(\alpha_t) \frac{A_r}{\pi (R_a + R_w)^2} (1 - r_{int}(\alpha_s))^2 T_\lambda^2 e^{-2cR_w} \tag{1}$$

where η_q is the detector quantum efficiency, η_t is the transmitter optical efficiency, η_r is the receiver optical efficiency, E_t is beam transmit energy, h is Planck's constant, ν is the photon frequency, ρ is the reflectance of the seafloor and cover (e.g., coral, seagrass, macroalgae), α_t is the incidence angle on the seafloor, A_r is the receiver telescope lens' area, R_a is the range from the satellite to the water surface, R_w is the in-water range, r_{int} is the reflectance of the water surface, α_s is the incidence angle on the seafloor, T is the atmospheric transmittance (one way), and c is the effective attenuation coefficient. Most of the variables in Eq. 1 are specular (i.e., a function of wavelength), but since we are only interested in a single wavelength (532 nm, the wavelength of ATLAS's laser), the subscript λ is omitted from the variable names.

Importantly, the received signal strength depends strongly on the water clarity, which is contained within the effective attenuation coefficient, c, an exponential term in Eq. 1. Ignoring sensor-dependent factors, c can be reasonably approximated by the diffuse attenuation coefficient of downwelling irradiance, K_d, an apparent optical property:

$$K_d = -\frac{1}{E(\lambda)} \frac{\partial E(\lambda)}{\partial z} \tag{2}$$

where E is downwelling irradiance in watts per square meter, z is depth, and K_d has units of inverse meters.

Beyond a certain depth that is typically referred to as the extinction depth, which is a function of K_d, as well as system-dependent parameters contained in Eq. 1, the received optical power drops below the detection threshold, and bathymetric measurement is infeasible. For ICESat-2 bathymetry, the extinction depth can often be approximated visually as the point at which the seafloor returns peter out in a profile plot, such as Figure 2. In this particular example, the seafloor points begin to approach extinction at approximately 30 m (non-refraction-corrected).

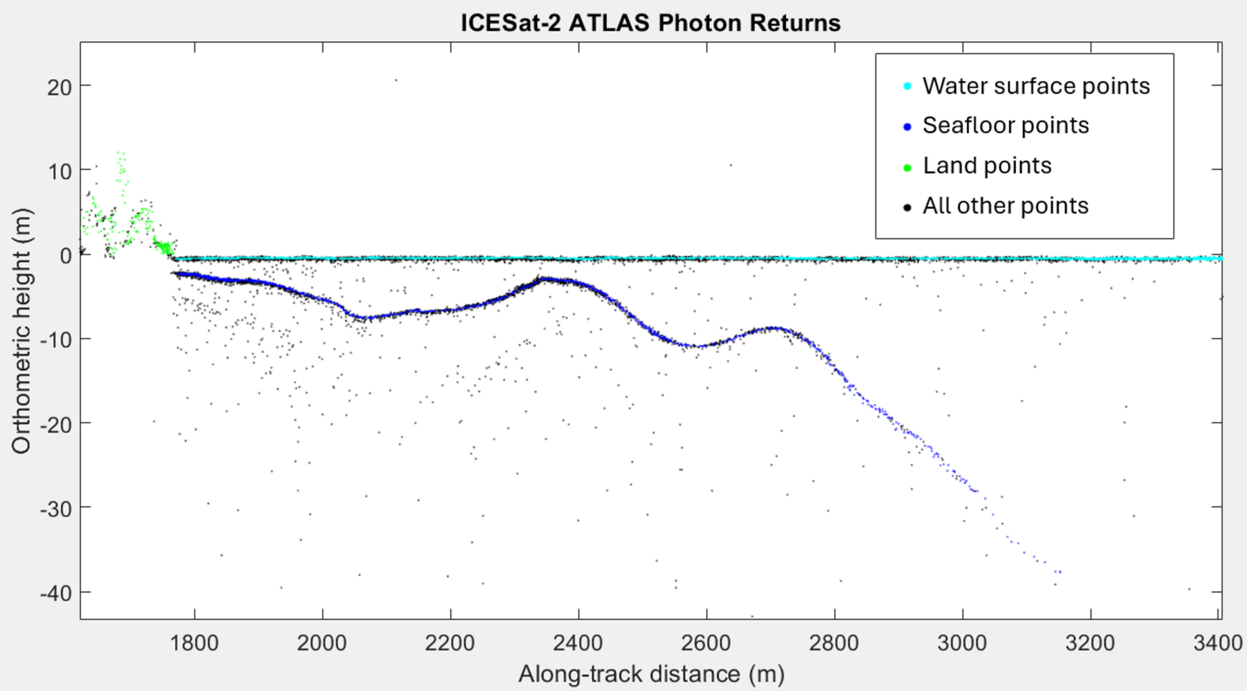

Figure 2: *Profile of ICESat-2 ATLAS along-track photon returns, with orthometric heights, based on the EGM08 geoid model. Here, water surface, seafloor, and land points have been manually labeled. Refraction correction has not yet been applied, so the seafloor points have a deep bias.*

With respect to water clarity, ICEsat-2 has a distinct advantage, compared to airborne bathymetric lidar. Water clarity in many regions is highly spatially and temporally variable, depending on factors including general geomorphic characteristics, wind and wave climate, salinity, precipitation and tide cycles, among many others.

With airborne bathymetric lidar, if some combination of the above factors are unfavorable for water clarity, the mission may result in no bathymetry being collected. This can also happen with an ICESat-2 overpass, but the revisit cycle provides many opportunities to re-observe at different times and under different environmental conditions, increasing the probability of obtaining usable bathymetry over the ICESat-2 data record.

Assuming the seafloor is detected, the geolocation of detected points can initially be the same as for topographic lidar, but coordinate corrections then need to be applied, based on the change in speed and direction of the laser light at the air water interface, a process known as refraction correction. Roughly following the procedures used in airborne bathymetric lidar, we start with return photon coordinates that assume topograpy, rather than bathymetry, and then apply refraction correction. The initial coordinates are from ATL03: specifically, the lat_ph, lon_ph, and h_ph in the /gtx/heights group.

4.2 ATL24 Input Variables

Table 2 captures each stage of the processing phase in the production of the ATL24 granule and explains the significance of each step. Table 2 also lists the required inputs for that stage with brief explanation. ATL03, the primary input, provides the heights above the WGS84 ellipsoid (ITRF2014 reference frame, through Release 06 of ATL03, after which ITRF2020

will be used), the latitude, longitude, and time tag for every ATLAS photon detection. These values are the primary input to ATL24 but many ATL03 specific parameters are passed through the processing workflow and attached to the ATL24 product for user utility and convenience. ATL03 parameters are all defined in the dedicated Algorithm Theoretical Basis Document and the Neumann et al., 2019 publication on the product (T. A. Neumann et al. 2019b). Table 2 provides a broad look at independent processing stages within the ATL24 workflow with the associated input data required to provide an understanding of the interdependencies of the algorithm on initial data products and intermediate products produced within the computational pipeline.

Table 2: *ATL24 input variables and processing details within each stage of the product production*

Stage	**Processing**	**Input Products**
Client Request	**ATL24.g** - accept an incoming HTTP request	**ATL03 granule name**
Input Data and Data Configuration	**ATL03/09 Granule Reader** - subset granules to bathymetry mask **Photon Filtering** - optionally filter photons based on certain fields (e.g. quality_ph)	**Global Bathymetry Mask** - a shapefile capturing all areas where there is bathymetry detectable by ICESat-2 **ATL03 Granule** - ICESat-2 standard data product h5 file **ATL09 Granule** - ICESat-2 standard data product h5 file
Photon Classifiers: Sea surface	**CoastNet** **Quantile Trees** **OpenOceans++**	**ICESat-2 inputs** **Trained ML models**
Photon Classifiers: Seafloor	**CoastNet** **Quantile Trees** **C-SHELPH** **OpenOceans++** **BathyPathFinder** **Median Filter**	**ICESat-2 inputs** **Trained ML models** **Sea Surface** - some classification algorithms require an initial estimate of sea surface

Ensemble Classification	**Ensemble**	**ICESat-2 Inputs** **Trained ML models** **Photon Classifications** - predicted sea surface (40), and seafloor (41) photon labels available from the classification algorithms
Error Correction	**Blunder Detection** - heuristically evaluates each photon classification against list of checks	**ICESat-2 input** **Photon Classifications**
Refraction Correction	**Refraction Correction** - calculates final depths of bathymetry photons using refraction index and surface heights	**Refractive index of water data layer**
Uncertainty Calculation	**Uncertainty Calculation** - calculates the vertical and horizontal uncertainty of each photon's location	**ICESat-2 Inputs** **VIIRS Kd490** **Subaqueous Uncertainty Look Up Tables** **Photon Classifications**
Final Outputs	**ATL24 Granule Writer** - writes h5 file **Quality Checker** - heuristically evaluates each h5 file and determines quality assessment	**All previous outputs**

Further description of the input components for the ATL24 processing workflow are listed in Table 3. These include the other ICESat-2 data used in production of the product and the independent data needed for processing implementation. An example is the refractive index of water data layer which is a global mask produced for the refractive correction to the subaqueous photon heights using temperature and salinity values from the E.U. Copernicus Marine Service Multi Observation Global Ocean 3D Temperature, Salinity, Height, Geostrophic Current, and Mixed Layer Depth (MLD) dataset. Another example is the bathymetry mask utilized as a processing guide for where bathymetry is both present and ATLAS is capable of observing it based on previous studies Christopher E. Parrish, L. A. Magruder, et al. 2019.

Table 3: *ATL24 component input for processing pipeline*

Item	Description
ATL03 Granule	ATLAS/ICESat-2 Level 2a global geolocated photon data
ATL09	ATLAS/ICESat-2 Level 3a calibrated backscatter profiles and atmopsheric layer characteristics
Global Bathymetry Mask	Used to spatially subset the ATL03/ATL09 granules prior to initiating the ATL24 workflow.(Note: this is also referred to internally as the "search mask")
VIIRS Kd490	NOAA product from the Visible Infrared Imaging Spectrometer to provide a diffuse attenuation coefficient at 490 nm to the uncertainty calculation.
Uncertainty Look up Table	Generated by an offline Monte Carlo simulation; used with turbidity, windspeed (ATL09), and photon depth to generate vertical and horizontal subaqueous geolocation uncertainty.
Refractive Index	Global data product using a seasonal value using temperature and salinity data from the E.U. Copernicus 2022 Multi Observation Global Ocean 3D Temperature Salinity Height Geostrophic Current and MLD dataset datasets

4.3 ATL24 Output Variables

Table 4: *ATL24 output variables for each stage of the product production*

Stage of Origin	Output Variables

ICESat-2 inputs	***Per Beam, Photon-level Variables (_ph)*** **lat_ph** - (deg) latitude computed from the ECEF Cartesian coordinates of the bounce point **lon_ph** - (deg) longitude computed from the ECEF Cartesian coordinates of the bounce point **sigma_h** - (m) ATL03 photon vertical aerial uncertainty **delta_time** - fractional seconds from ALTAS epoch to laser fire associated with photon measurement ***Per Granule Variables*** **wind_v** - (m/sec) the wind speed at the center photon of the subsetted granule; calculated from ATL09 *met_u10m* and *met_v10m*
Classifier input parametrization	**index_ph** - unique index of photon in the given granule **index_seg** - unique segment index
Ensemble	**class_ph** - photon classification: sea_surface (41), bathymetry (40), other (1) or unclassified (0) **confidence** - floating point value 0-1 based on ensemble score for accuracy of predicted photon classification **surface_h** - orthometric height of the sea surface at the location of the photon based on the EGM08 geoid model

Refraction Correction	**ellipse_h** - (m) height of photon as measured from ellipsoid surface of Earth (WGS84) **ortho_h** -(m) height of photon as measured from geoid of Earth (EGM08) **lat_ph** - ITRF2014 (until release 07, when it will be updated to IRTF2020) latitude of each received photon (degrees north); computed from the ECEF Cartesian coordinates of the bounce point **lon_ph** - ITRF2014 longitude of each received photon (degrees east); computed from the ECF Cartesian coordinates of the bounce point
Uncertainty Calculation	**sigma_thu** - (m) total horizontal uncertainty (1-sigma), which is the square root of the sum of the squares of the aerial and subaqueous horizontal uncertainty for each received photon **sigma_tvu** - (m) total vertical uncertainty (1-sigma), which is the square root of the sum of the squares of the aerial and subaqueous vertical uncertainty for each signal photon
Data Flags	**invalid_kd** - This binary flag indicates VIIRS Kd490 data is available (= 0) within the required timespan (=/- 1 day) from the ATL03 photon time-tag or when it is not available (= 1) **low_confidence_flag** - This binary flag indicates when the *confidence* parameter value is greater than 0.6 (= 0) or when it is lower than 0.6 (= 1) **night_flag** - This binary flag indicates when the data was acquired in the absence of sunlight (= 1) in in the presence of sunlight (= 0) **sensor_depth_exceeded** - This binary flag indicates when the ATL24 photon depth is a reasonable depth (= 0) or outside of the sensor capability (= 1) **invalid_wind_speed** - This binary flag indicates the absence (= 1) of ATL09 wind speed or when a wind speed estimation is avilable (= 0). Wind speed is used in the uncertainty estimation.

4.3.1 ATL24 Data Structure and Naming Conventions

The naming and internal organization of the final HDF5 file for each ATL24 granule will match the conventions used on the rest of the project. Each granule will also have the naming convention similar to the other ICESat-2 data products. Each variable within the name (place holders) are described in Table 5. Please note that there are two sets of version and revision numbers that represent both ATL03 input version and revision and then the version and revision of ATL24.

ATL24_[yyyymmdd][hhmmss]_[ttttccss]_[vvv_rr]_[vvv_rr].h5

Table 5: *ATL24 naming convention*

Variable	Description
ATL24	ICESat-2 Level 3 Global nearshore and coastal bathymetry
ATL09	ATLAS/ICESat-2 Level 3a calibrated backscatter profiles and atmopsheric layer characteristics
yyyymmdd	Year, month and day of data acquisition
hhmmss	Data acquisition start time, hour, minute, and second (UTC).
tttt	Four digit Reference Ground Track number. The ICESat-2 mission has 1,387 RGTs, numbered from 0001 to 1387.
cc	Cycle Number. Each of the 1387 RGTs is targeted in the polar regions once every 91 days. The cycle number tracks the number of 91-day periods that have elapsed since ICESat-2 entered the science orbit.
ss	Segment number. ATL03 data files are segmented into approximately 1/14th of an orbit. Segment numbers range from 01-14. Note that some segments may not be available.
vvv_rr	ATL03 Version and revision number
vvv_rr	ATL24 Version and revision number

4.4 Classification Algorithms

In many regards, the most important and challenging step in the ATL24 workflow is the classification of sea surface and seafloor (bathymetric bottom) returns in input ATL03 point clouds. In airborne bathymetric lidar, significant human analyst time is typically spent on this step. For ATL24, this step is even more challenging, because: a) it must be fully automated, b) ICESat-2 ATL03 granules are generally noisier than airborne bathymetric lidar, due to ATLAS's photon-counting design, and c) while airborne bathymetric lidar projects typically cover a fairly limited spatial extent, the ATL24 sea surface and seafloor classification must perform well for coastal and nearshore regions around the world, encompassing a wide range of seafloor morphologies, substrates, depth ranges, and cover types (e.g., coral, seagrass, macroalgae).

Of the considerations noted above, it is the exceptionally wide range of subaqueous environments that poses the most serious challenges for entirely automated seafloor classification. Any individual algorithm or model is unlikely to be able to handle the full extent of seafloor

types, morphologies, depth ranges, water types, and cover types that exist throughout coastal and nearshore areas worldwide. The ATL24 workflow addresses this hurdle by utilizing an ensemble of classification models. The overarching idea is that the whole is greater than the sum of the parts: by combining the outputs of a number of base models or algorithms, it is possible to attain better classification results than can be achieved with any individual model or algorithm. By combining the outputs of diverse models and algorithms, we can achieve good results across a broad range of conditions through a fully-automated approach.

It is important that the predictions of the individual models are distinct in their methods. This aspect is crucial as models with error modalities that are strongly correlated are more difficult to detect and correct. We have thus chosen an ensemble of models, each described below, that satisfy these requirements.

4.4.1 CoastNet Classification

Each ground track was arranged as a 2D raster representing the along-track profile. Fixed sized samples were then selected from this raster. Each pixel in the raster was 4 meters wide and 1 meter tall, and samples were 65×65 pixels (an odd numbered size was chosen so that a pixel was centered within the sample), resulting in 260×65 meter samples (fig. 3). Training samples pixels were colored black if the pixel contained a photon and were colored white if the pixel did not contain a photon. Each sample was labeled according to the sample's central photon label (shown in red). The task of the network was then to determine if each sample contained a noise photon at its center (colored gray in the figure), a bathymetry photon (magenta), or sea surface photon (cyan).

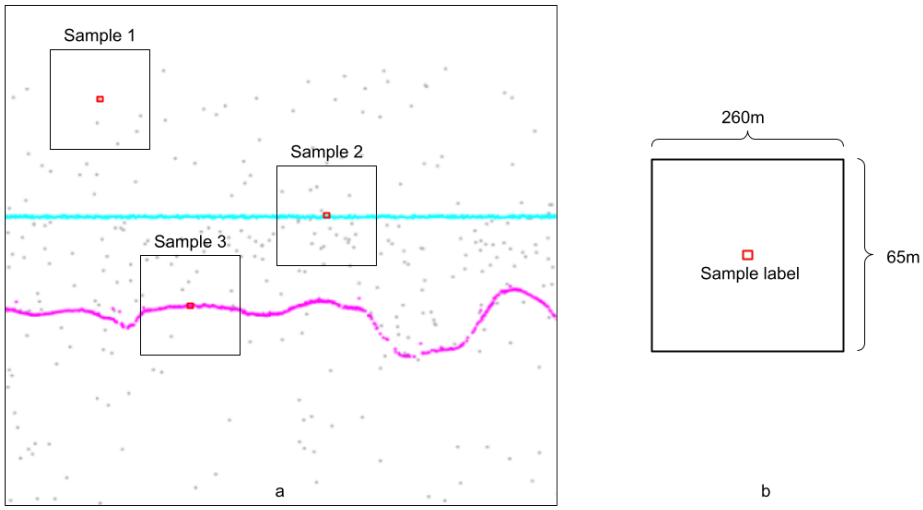

Figure 3: *a) Samples are centered on photons. b) Each sample is a 65×65 pixel image raster, where each pixel is 4 meters wide and 1 meter tall for a total coverage area of 260×65m. The sample's ground truth label is the label of the photon in the center of the sample: sea surface, bathymetry, or other.*

The machine learning network employed for the model was a 'deep-residual' network, a type of convolutional neural network (CNN) designed for network architectures with significant depth, often comprising tens or hundreds of layers.

CNNs are widely used in image recognition tasks, as they can effectively extract spatial features from input images. The deep-residual network architecture addresses a common problem in deep CNNs known as the 'vanishing gradient problem,' where gradients become exceedingly small in deeper layers, making it difficult to train the network effectively.

To overcome this, the deep-residual network introduces 'skip connections' or 'residual connections,' which directly connect layers in the network with layers further down the architecture. These skip connections enable gradients to flow more directly from the output layers to the earlier layers, facilitating efficient training and alleviating the vanishing gradient problem.

The deep-residual network has demonstrated superior performance on various image recognition tasks, including image classification and object detection. Its ability to handle complex and high-dimensional data makes it a powerful tool for various computer vision

applications.

In our experiments, we opted to build separate sea surface and bathymetry models instead of a single model that classified both. This helped us mitigate class imbalance and allowed for separate analysis of classification performance for each class. Bathymetry photon counts are considerably lower compared to noise and sea surface, which can significantly impact the scores depending on the chosen metric. This aspect will be discussed in more detail later.

Algorithm 1 Coastnet Classification Algorithm

Input: ICESat-2 granule
Output: Photon label predictions
  $G \leftarrow$ granule photons
  $M \leftarrow$ pretrained model
  for $g_i \subseteq G$, $c_i \subseteq$ context photons centered at g_i **do**
    $R_i \leftarrow 63 \times 63$ raster formed from c_i
    $prediction_i \leftarrow \mathrm{M}(R_i)$
  end for

4.4.2 BathyPathFinder Classification

BathyPathfinder, first described in Corcoran et al. 2024, is a bathymetric surface extraction algorithm that leverages techniques from network analysis in order to extract a representative sample of seafloor photon returns from an ATLAS profile containing bathymetry. Unlike other common methods, BathyPathfinder does not rely on the spatial density of neighboring photons in order to perform classification. Instead, following sea surface removal, all subaqueous photon returns are organized into a weighted, undirected graph $G = (V, E)$, where V represents the set of photons (vertices) and E represents the graph connections (edges). The edges are weighted proportional to the distance between the pairs of photons they connect. This process is done using the KD Tree algorithm (Bentley 1990), with the value k representing the number of neighbors connected to each subaqueous photon.

Following the graph generation step, edges with outlier weights (i.e., those exceeding a threshold percentile) are pruned from the tree. The default value for this threshold, n, is set to the 99th percentile, based on previous empirical results. The algorithm then selects pairs of "target" and "source" photons, which represent photons at either end of a contiguous bathymetry surface. Many techniques for target and source selection can be employed, however, for the sake of simplicity, the algorithm uses the photons with the median elevations within 5m of the bathymetric surface edges.

Finally, using the target and sources pairs, the algorithm performs an optimal "least-cost" network traversal using the A^* algorithm (Hart et al. 1968). The photons used by A^* to connect the target and source photons are then extracted and classified as bathymetry. In the case of discontinuous bathymetry, due to occlusion of the sensor or sensor saturation, the KD Tree may generate two or more unconnected subgraphs, denoted $S_1, S_2, ..., S_m$. In this situation, Bathy Pathfinder will automatically separate these subgraphs and apply the remaining steps of the algorithm to each S_i $(i = 1, 2, ..., m)$ individually. Once the bathymetric surface photons have been extracted, they are combined to recreate the original, segmented bathymetric surface. A segment of the BathyPathFinder pseudo code is found in (Algorithm 2).

Algorithm 2 Bathy PathFinder Algorithm

Input: Point Cloud, Sea Surface n, k
Output: Bathymetry

```
G ← KD Tree(Point Cloud \ Sea Surface)
G ← Prune(G, percentile=n)
Bathymetry ← [ ]
for S_i ⊆ G do
    t_i, s_i ← TargetSourceSelection(S_i)
    Path ← A*(S_i, s_i, t_i)
    Bathymetry ← Append(Bathymetry, Path)
end for
```

4.4.3 OpenOceans++ Classification

OpenOceans++ (OO++) is loosely based on an adaptable, opensource approach developed at University of Texas at Austin called OpenOceans (https://github.com/jonm3D/OpenOceans). The overarching inspiration of OO++ is from historical full-waveform lidar signal processing techniques (Leigh and L. A. Magruder 2016, L. A. Magruder 2010) but implemented in C++. OO++ is designed to leverage a vertical histogram approach for classification of sea surface and sea floor photons. At the OO++ basic level the algorithm is calculating 'pseudo-waveforms' from the histograms using a Gaussian fit from the identified peak magnitudes, modes and full-width-half-maximum values in the ATL03 geolocated photon input. These Gaussian distributions are subsequent to an along-track binning process to partition the photons. Below is the description of the general OO++ algorithm that details how the histogramming and waveform fitting is performed, in addition to several filters and quality checks implemented. There are separate algorithm descriptions provided to address both the sea surface and seafloor classifications that use probability mass functions from the input photon elevations but differ in statistical analyses and data filtering steps. After algorithm completion, all photons will be labeled as 'sea surface' or 'bathymetry', and all photons will contain both a surface and a bathymetry elevation estimate, even in places where no nearby sea or seafloor surface was detected. This allows the user to obtain estimates of bathymetry and sea surface without having to specify how far away (horizontally) any particular photon is from a labeled bathymetry or surface photon.

Algorithm 3 OO++ General Algorithm

Input: ATL03 track consisting of a list of photons. Each photon contains an along-track distance, x, and an elevation, z.

Output: ATL24 track, where each photon has an associated classification (unclassified, sea surface, or bathymetry), an estimate of the sea surface elevation at that photon's along-track distance, and an estimate of the bathymetry elevation at that photon's along-track distance.

1. Compute a surface elevation estimate for the entire collection of photons. This estimate is used as an input to the surface and bathymetry detection algorithms in steps 3b and 3d.
   (a) Filter photons that are beyond a threshold distance (e.g. ± 20 meters) from zero mean sea level elevation
   (b) Calculate the median elevation of the filtered photons
   (c) Filter the photons again that are beyond a threshold distance (e.g. ± 1 meter) from the median elevation
   (d) Compute the median elevation of the second set of filtered photons. This is the initial surface estimate for the entire set of photons.
2. Partition the photons into along-track bins whose boundaries are spaced at, for example, 10 meter intervals.
3. For each along-track bin
   (a) Identify sea surface photons in the bin. See Algorithm 4 below for a description of the sea surface identification algorithm.
   (b) Compute the surface estimate for this bin by taking the average of all photon elevations identified as sea surface. If no surface photons are detected in the bin, the surface elevation estimate for the bin is 0.0 meters.
   (c) Identify bathymetry photons in the bin. See Algorithm 5 below for a description of the bathymetry identification algorithm.
   (d) Compute the bathy estimate for this bin by taking the average of all photon elevations identified as bathymetry. If no bathymetry photons are detected in the bin, the bathy elevation estimate for the bin is 0.0 meters.
4. Apply a Gaussian smoothing filter based upon the sea surface estimates associated with each bin. The standard deviation of the sea surface smoothing function is large (for example, 200 meters) because the ocean surface typically does not change drastically over short along-track distances.
5. Apply a Gaussian smoothing filter to the bathymetry estimates associated with each bin. The standard deviation of the bathymetry smoothing function is smaller than the sea surface standard deviation, for example, 100 meters.
6. Assign surface and bathymetry estimates at each photon according to the smooth estimates in steps 4 and 5.

Algorithm 4 OO++ sea surface identification algorithm

Input: List of photons all corresponding to a single along-track bin. Sea surface estimate for the entire track.

Output: List of photon indexes identified as sea surface.

1. Construct a probability mass function (PMF) from the input photon elevations.
2. Smooth the PMF with a Gaussian filter, (e.g. stddev = 0.5 meters)
3. Get a list of PMF peaks
4. Eliminate peaks from the list that are below a prominence threshold (e.g. 0.01) and are too close to adjacent peaks (e.g. less than 2 vertical bins).
5. Using the shape of the input sea surface estimate (mean and variance) determine the valid range of sea surface estimates. E.g. ±3 standard deviations from the sea surface mean.
6. Eliminate peaks from the list that fall outside of the valid range of sea surface estimates computed in the preceding step.
7. Of the remaining peaks in the list, choose the two that contain the most mass. If no peaks remain, return an empty list of photon indexes.
8. If the two peaks are close to one another in mass (e.g. within 30corresponding to the sea surface. Otherwise, choose the peak with the most mass.
9. Compute the average and variance of the photons that lie within a certain elevation distance (e.g. ±1 meter) from the surface peak.
10. Use this average and variance to select the indexes of all photons that lie within a threshold number of standard deviations (e.g.: 2) from the average. These indexes are the return value of the algorithm.

Algorithm 5 OO++ bathymetry identification algorithm

Input: List of photons all corresponding to a single along-track bin. Sea surface estimate for the entire track.

Output: List of photon indexes identified as bathymetry photons.

1. Eliminate photons from the list that lie above the sea surface estimate for the track.
2. Construct a probability mass function (PMF) from the input photon elevations.
3. Smooth the PMF with a Gaussian filter, (e.g. stddev = 0.5 meters)
4. Get a list of PMF peaks
5. Eliminate peaks from the list that are below a prominence threshold (e.g., 0.01) and are too close to adjacent peaks (e.g., < 2 vertical bins).
6. If no peaks remain, return an empty list of photon indexes.
7. Select the peak with the highest elevation.
8. Compute the average and variance of the photons that lie within a certain elevation distance (e.g., ±1 meter) from the bathy peak.
9. Use this average and variance to select the indexes of all photons that lie within a threshold number of standard deviations (e.g., 2) from the average. These indexes are the return value of the algorithm.

4.4.4 Median Filter Classification

The median filter method is described in the paper Ranndal et al. 2021a as simple empirical method for extracting bathymetry profiles from ICESat-2. The technique was tested on four tracks over the Heron Reef, Australia and a region north-west of Sisimiut, Greenland. This method shows good success in extraction of the bathymetric signal in good conditions (e.g. relatively low turbidity) and is computationally inexpensive. This approach, along with the OpenOceans algorithm (section 4.4.3 and C-SHELPh (section 4.4.5), are the three classification algorithms that do not utilize machine learning. The ATL24 implementation of the median filter approach s as follows:

1. All ATL03 photon ellipsoid heights are converted to orthometric heights using the EGM08 geoid model.

2. Using the input sea surface classifications provided by QTrees (section 4.4.6), all photons more than 1.5 m below the sea surface median are kept for subsequent processing.

3. Windows for every 50 photons are created and the corresponding median elevation for each window is calculated. Any photon in a window with an elevation that is more than 4m different from the median is removed.

4. Another moving median is calculated but this time with a window size of 30 photons. This time, a moving standard deviation of elevations is calculated and any photons with an elevation greater than 0.6m from the median AND a standard deviation greater than 1.2 m are removed.

5. Now the retained photons are separated into 0.001 degrees latitude (100m) segments. For each segment, if there are more than 14 photons, the data are assumed to be bathymetry.

4.4.5 C-SHELPh Classification

The C-SHELPh (Classification of Sub-aquatic Height Extracted Photons) algorithm was developed by Thomas et al. in the effort to provide an open source tool for producing bathymetric maps (N. Thomas et al. 2022). The work, published in 2022, was one of the first techniques for automating the extraction of ICESat-2 shallow-water, regional bathymetry. The initial success of the approach was demonstrated at the Great Bahamas Bank around the island of Andros, Bahamas and included 224 ICESat-2 tracks.

At its core, C-SHELPh detects the dense clustering of photons as typically these clusters are indicative of surface returns. The density values are determined across a user specified grid with the default values being 0.5 m in the vertical direction and 10 m in the horizontal direction. This gridding convention provide surface heights and along-track latitudes. The photon clusters around a height of 0 m are labeled to be the ocean surface photons and surface height is estimated to be the median value of cluster. Photon clusters that occur below the ocean surface value are identified via evaluation per-grid-cell basis relative to user defined thresholds of the signal to noise. C-SHELPh proved very successful in the Caribbean environment as the water has fairly low turbidity, the depths range from 0-10 m and the seafloor is highly reflective; all of which provide an ideal scenario for density driven signal finding.

The ATL24 implementation of C-SHELPh bins photons into 0.001 degree along track blocks and 0.5 meter vertical blocks. The number of photons in each block is counted and then using these bin counts, a threshold is calculated based on the number of photons in all the blocks. The threshold is defined as the nth percentile of photon counts per block where n=0.5. Although CSHELPh has been used in the past to determine the sea surface, the ATL24 usage takes the surface input from the Quantile Trees output (section 4.4.6).

The threshold is calculated for the 85% and 65% and if these are equal, or the 65% threshold number of photons in a bin is less than five, the algorithm sets the required number of photons in a bin for bathymetry prediction to occur equal to five.

Algorithm 6 C-SHELPh Algorithm

Input: Point Cloud, Sea Surface n, k
Output: Bathymetry

```
binnedphotons ← bin photons into 0.001 degrees along track, 0.5m vertical (Point Cloud \ Sea Surface)
if thresh_85 == thresh_n then
    if thresh_85 < 5 then
        mincounts = 5
    else
        mincounts = thresh_n
    end if
    if thresh_n < 5 then
        mincounts = 5
    else
        mincounts = thresh_n
    end if
end if
Bathymetry ← [ ]
for bin_i ∈ binnedphotons do
    if bin_i - photon count > mincounts then
        Bathymetry ← Append(Bathymetry, bin_i photons)
    end if
end for
```

4.4.6 Quantile Trees Classification

Similar to the CoastNet algorithm, the Quantile Trees (or QTrees) algorithm arranges ground track profiles as two dimensional profile images. At each photon in the track, the algorithm will contstruct an image centered at the target photon. However, unlike the CoastNet algorithm, the QTrees algorithm places the distribution of photons within each image column into quantiles. N-quantiles are non-parametric descriptions of distributions that partitions a set of observations into N bins. In our algorithm, N is typically set to 32 bins.

The quantile bins in each column of the image are then used as features to a supervised learning algorithm. In our implementation, we use the extreme gradient boosted algorithm XGBoost. So, for example, for a given photon, the image will typically be divided into nine columns (an odd number so that the target photon lies within the center column). If we are using N-quantiles with N=32, then there will be 9 x 32 = 288 features, plus the elevation of the target photon, for a total of 289 input features. The algorithm will then train on ground tracks from hundreds of labeled datasets to produce a model that can predict *noise*, *sea surface*, and *bathymetry* labels for any given input ground track.

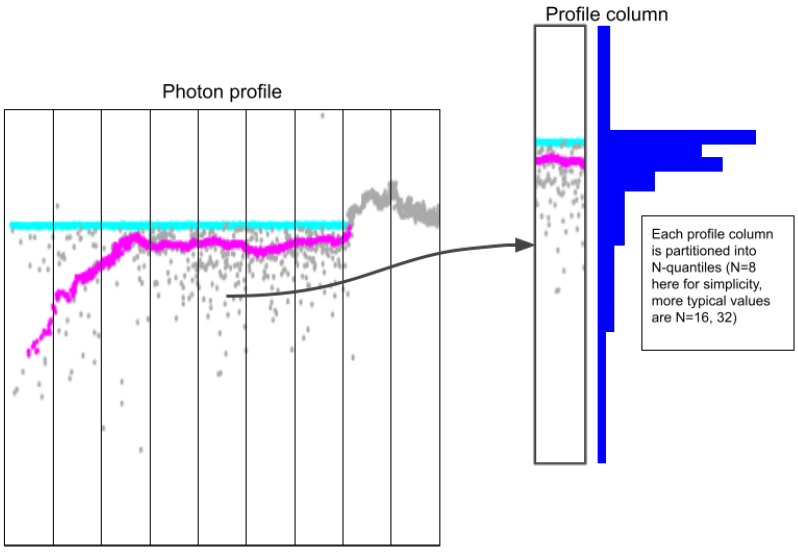

Figure 4: *Example of the QTrees algorithm. Each vertical column of the image is partitioned into N-quantiles. The quantile boundaries are then used as features, along with the target photon's elevation, which are used as inputs to an XGBoost classifier.*

Algorithm 7 QTrees Classification Algorithm

Input: ICESat-2 granule
Output: Photon label predictions
  $G \leftarrow$ granule photons
  $M \leftarrow$ pretrained model
  for $g_i \subseteq G$, $c_i \subseteq$ context photons centered at g_i **do**
    $Q_i \leftarrow$ 288 quantile bin locations formed from c_i
    $e_i \leftarrow$ elevation of photon g_i
    $prediction_i \leftarrow \mathrm{M}(e_i, Q_i)$
  end for

4.4.7 Ensemble Classification

An ensemble is a meta-machine learning technique used to combine predictions from multiple algorithms. Ensembles can generate predictions that are more accurate than the predictions of each of the combined algorithms individually. The particular type of ensemble used by ATL24 is a *stacking ensemble*. This method is different from a voting scheme because it allows the model to learn which combinations of predictions produce the best results. For example, one algorithm may be good at predicting bathymetry in certain environmental conditions, while the others might be statistically divided between choosing bathymetry versus sea surface labels for the given photons.

The ATL24 consists of a gradient-boosted decision tree (XGBoost) within a supervised machine learning framework. XGBoost is also known as "Extreme Gradient Boosting". The framework takes as input the photon classification predictions output from each of the classification algorithms (Median Filter, C-Shelph, etc.) and incorporates them into the decision tree model, which is then trained using hand-labeled water surface, bathymetry, and noise photons as the ground-truth reference. Other features like photon depth and signal confidence may also be used as model inputs. This technique allows the model to learn which prediction to select from the inputs for any given context.

For example, the model may learn that certain combinations of input predictions produce more accurate results than others. The model could also learn, for example, which algorithms work better in deep versus shallow water, or when there is more or less background noise, and then make its final prediction accordingly. These capabilities allow the model to make better predictions than many other techniques, such as linear regression or voting. Ultimately, the ensemble framework is configured to add any number of relevant input features in addition to an unlimited number of base classification predictions as new algorithms become available. However, as additional algorithm predictions are included, the ensemble model must be retrained on the comprehensive feature space before implementation.

Algorithm 8 Ensemble Classification Algorithm

Input: ICESat-2 granule, Photon label predictions from each algorithm
Output: Revised photon label predictions

```
G ← granule photons
P ← photon predictions from each algorithm
M ← pretrained model
for g_i ⊆ G do
    p_i ← P's length 7 vector of predictions for photon i
    e_i ← elevation of photon g_i
    prediction_i ← M(p_i, e_i)
end for
```

4.4.8 Blunder Detection

Blunder detection is a series of simple post-processing checks. These checks ensure that the photon classification predictions and elevation estimates make logical sense.

Prior to blunder detection, each photon has already been labeled as one of *unclassified*, *sea surface*, or *bathymetry*. In addition, each photon has an associated sea surface estimate and bathymetry depth estimate at that photon.

The checks that are performed are:

1. Surface elevation check: Ensure that photons labeled as *sea surface* are close to the sea level estimate. Photons that are not within, for example, ±30 meters of the sea level estimate at that photon are relabeled as *unclassified*.

2. Bathymetry elevation check: Photons labeled as *bathymetry* are relabeled as *unclassified* if they are deeper than, for example, 100 meters below the sea level estimate at that photon.

3. Bathymetry relative elevation check: Photons labeled as *bathymetry* are relabeled as *unclassified* if the photon's elevation is greater than the sea surface estimate at that photon.

4. Surface range check: Photons that are labeled as *sea surface* must be within some nominal range (e.g. ±5 meters) of the sea surface elevation estimate at that photon.

5. Bathymetry range check: Photons labeled as *bathymetry* must be within some nominal range (e.g. ±5 meters) of the bathymetry elevation estimate at that photon.

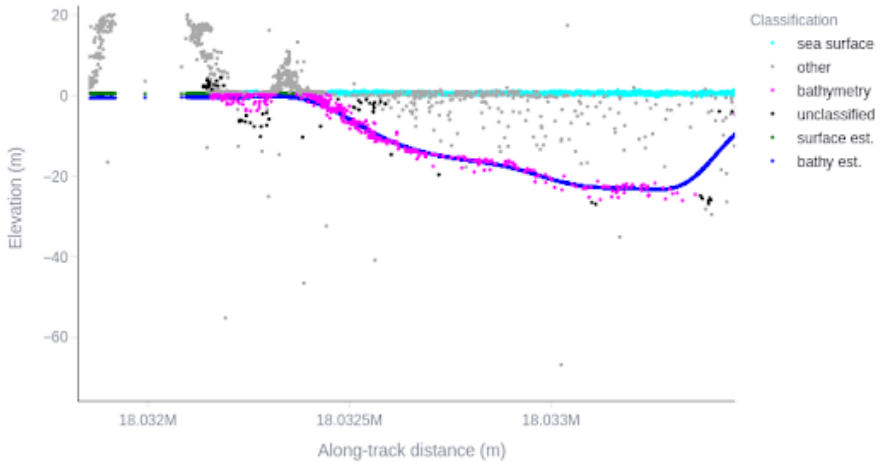

Figure 5: *Blunder detection example. Some bathymetry predictions (magenta) are reclassified (black) if they are not close to the bathymetry elevation estimate (blue). Likewise, some sea surface predictions (cyan) are reclassified (black) if they are not close to the sea surface elevation estimate (green, mostly occluded by cyan photons).*

Algorithm 9 Blunder Detection Algorithm

Input: ICESat-2 granule, ensemble predictions
Output: Modified predictions, sea level estimates, bathymetry elevation estimates

```
G ← granule photons
P ← ensemble predictions
for iteration ⊆ {1, 2, 3} do
    S ← sea level elevation estimates
    B ← bathymetry elevation estimates
    for g_i ⊆ G, p_i ⊆ P, s_i ⊆ S, b_i ⊆ B do
        if p_i == surface AND g_i elevation is not in ±30.m range then
            c_i ← unclassified
        else if p_i == bathymetry AND g_i elevation is not in 0.0m, -100.0m range then
            c_i ← unclassified
        else if p_i == bathymetry AND g_i elevation > s_i then
            c_i ← unclassified
        else if p_i == surface AND g_i elevation is not in range ±5.0m of > s_i then
            c_i ← unclassified
        else if p_i == bathymetry AND g_i elevation is not in range ±5.0m of > b_i then
            c_i ← unclassified
        end if
    end for
end for
```

4.5 Refraction Correction

Refraction correction is a critical step in which the change in direction and speed of laser light at the air-water interface are accounted for in computing corrected coordinates of seafloor points. If refraction correction is not applied, large errors will exist in the vertical and horizontal coordinates of bathymetric points. The vertical error is the larger of the two, being approximately 25% of the uncorrected depth (Christopher E. Parrish, L. A. Magruder, et al. 2019; Lambert and Christopher E Parrish 2023). The direction of the vertical component of the refraction correction is up (i.e., corrected bathymetric points will be at a higher elevation, or, equivalently, a shallower depth). Meanwhile, the magnitude of the horizontal component of the refraction correction is approximately $0.44D\tan(\theta_i)$, where θ_i is the off-nadir angle of the beam and D is the uncorrected depth. The direction of the horizontal correction is in towards the satellite's nadir point. The refraction correction algorithm used in ATL24 is the one introduced in Christopher E. Parrish, L. A. Magruder, et al. 2019 with a few minor modifications.

An important variable in the refraction correction is the index of refraction of water. In general, the index of refraction of seawater is a function of salinity, temperature, wavelength, and pressure (Austin and Halikas 1976; Quan and Fry 1995). For ATL24 bathymetry, the wavelength is fixed, based on ATLAS's wavelength of 532 nm, and, for the depth ranges in which ICESat-2 bathymetric measurement is feasible, it is reasonable to assume atmospheric pressure (Quan and Fry 1995). However, because of the large variation in salinity and temperature throughout the global ocean, large errors would result from using a single value of $n_w(532)$ everywhere. Hence, we developed a global refractive index layer using temperature and salinity data from E.U. Copernicus 2022 Multi Observation Global Ocean 3D Temperature Salinity Height Geostrophic Current and MLD dataset (Stephanie Guinehut et al. 2012, Guinehut et al. 2021) and the empirical equation of Quan and Fry 1995, simplified to consider only the wavelength of 532 nm. This resulted in a 1/4° resolution grid of the refractive index for use in the ATL24 refraction correction (Figure 6). These procedures are described in J. T. Dietrich and Christopher E Parrish 2025.

Many other versions of ICESat-2 refraction correction algorithms have been presented and discussed in the published literature. In fact, interestingly, after classification of seafloor points, refraction correction has been perhaps the most widely-researched aspect of ICESat-2 bathymetry processing, based on the number of published papers. Some of the proposal enhancements have include accounting for water surface slope and fluctuation of the water surface (Ma et al. 2020), wave effects (Xu et al. 2020), and modeled water surfaces and ray tracing (Zhang et al. 2022), among others. These refraction corrections are substantially more computationally complex than the one implemented in the ATL24 workflow, and it is not clear whether they can be sufficiently automated and robust to provide good results in all coastal locations around the world. However, for users who wish to customize the refraction correction, ATL24.p will provide the ability to do so, including through use of a custom water surface model.

Pseudocode for the refraction correction is given below in Algorithm 10.

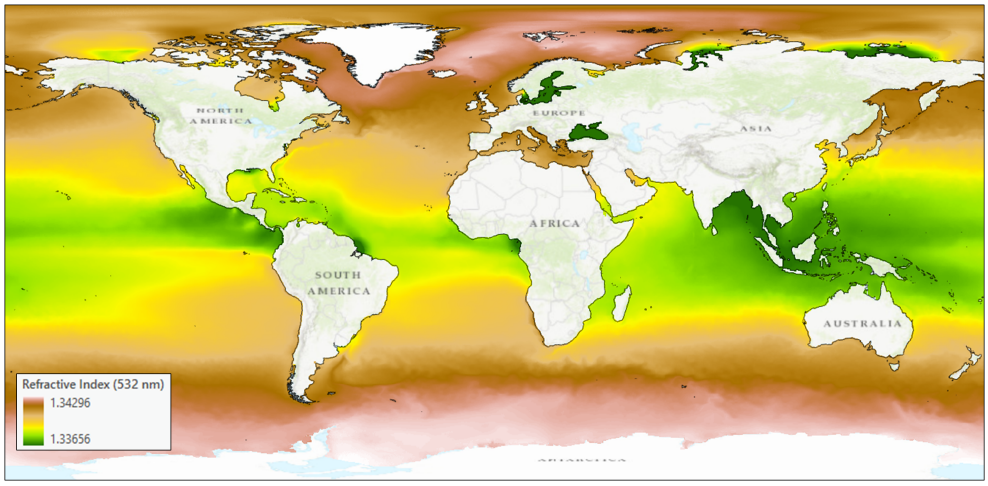

Figure 6: *Index of refraction layer computed from the E.U. Copernicus 2022 Multi Observation Global Ocean 3D Temperature Salinity Height Geostrophic Current and MLD dataset.*

Algorithm 10 Refraction Correction Algorithm

Input: bathymetric point latitude (lat_ph), longitude (lon_ph), and height (h_ph) from ATL03, water surface model (W), refractive index of air (n_a), refractive index of water (n_w), reference elevation from ATL03 (ψ), reference azimuth from ATL03 (κ)

Output: Coordinate corrections in easting, northing and height (ΔE, ΔN, ΔZ)

for bathymetric point in granule segment **do**

Angle of incidence (θ_1) $\leftarrow \pi/2 - \psi$

Angle of refraction (θ_2) $\leftarrow \sin^{-1}(n_a sin(\theta_1)/n_w)$

Depth (D) $\leftarrow W$ - h_ph

Uncorrected slant range (S) $\leftarrow D/\cos(\theta_1)$

Corrected slant range (R) $\leftarrow S n_a/n_w$

Distance between corrected and uncorrected points (P) $\leftarrow \sqrt{R^2 + S^2 - 2RScos(\theta_1 - \theta_2)}$

Bathymetric point altitude (β) $\leftarrow \pi/2 - \theta_1 - \sin^{-1}(R\sin(\theta_1 - \theta_2)/P)$

Correction in cross-track direction (ΔY) $\leftarrow P\cos(\beta)$

Correction in vertical direction (ΔZ) $\leftarrow P\sin(\beta)$

Easting correction (ΔE) $\leftarrow \Delta Y \sin(\kappa)$

Northing correction (ΔN) $\leftarrow \Delta Y \cos(\kappa)$

end for

The parameters in the refraction correction algorithm are depicted graphically in Figure 7. Here, the red point is the uncorrected point, while the green point is the refraction-corrected point.

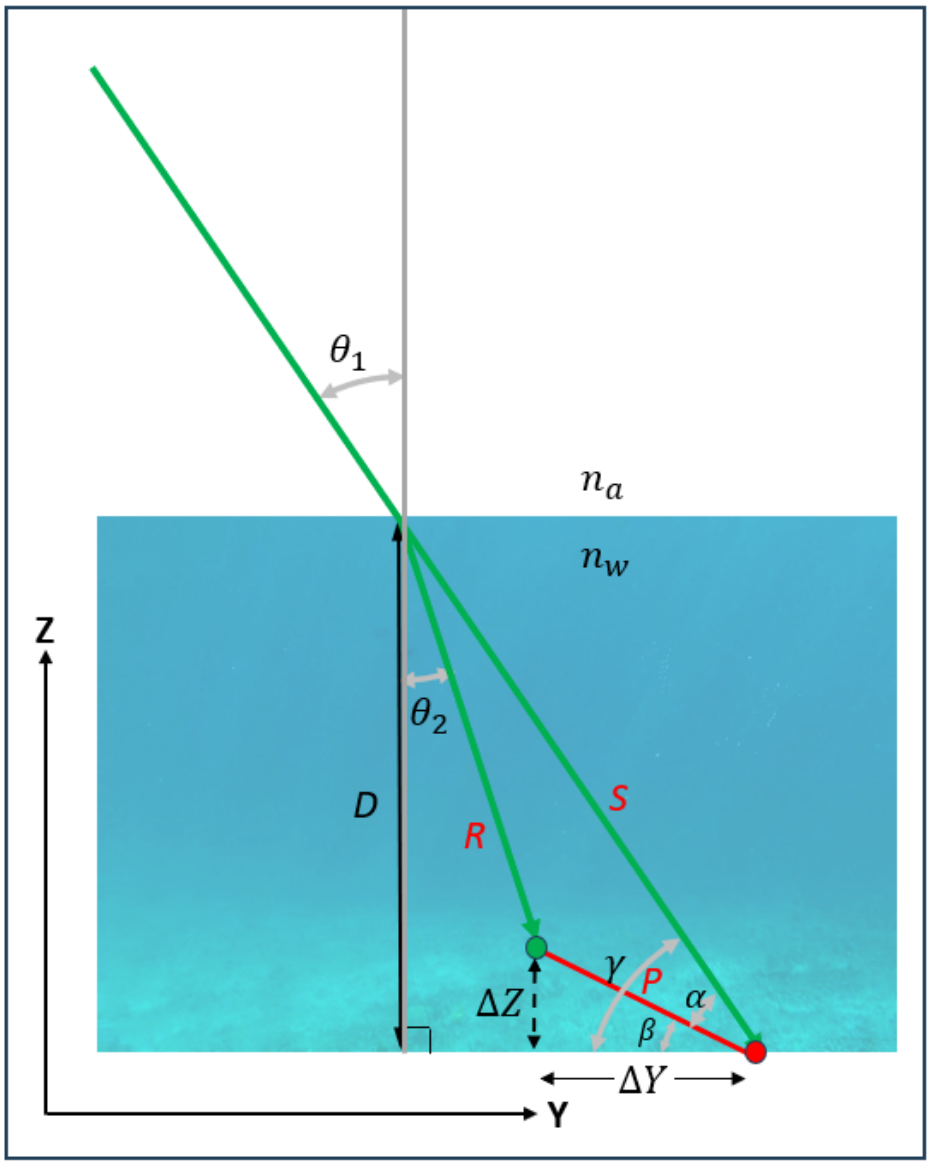

Figure 7: *Graphical illustration of variables in refraction correction algorithm. The red point is the uncorrected bathymetric point, and the green point is the refraction corrected point. Adapted from Christopher E. Parrish, L. A. Magruder, et al. 2019*

.

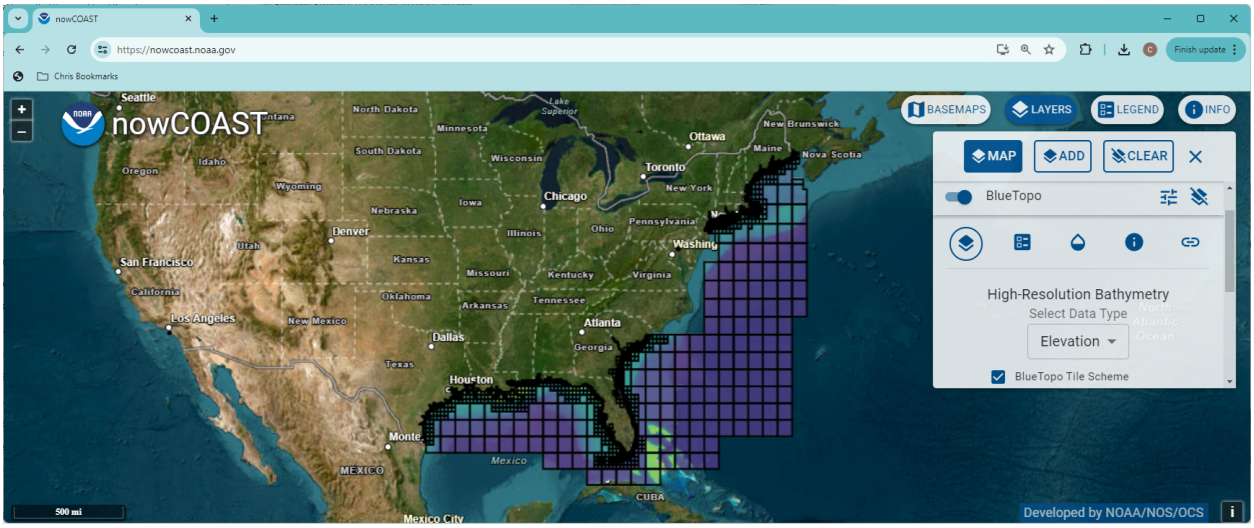

Figure 8: *NOAA BlueTopo data set, which will be used as the reference data set in testing the accuracy of ATL24.*

5 Performance Assessment and Validation

The accuracy of ATL24 will be tested using the NOAA BlueTopo data set, which is a compilation of best-available bathymetric data for U.S. waters, maintained and distributed by NOAA's Office of Coast Survey (OCS) through the National Bathymetric Source program (Rice et al. 2023). The reasons for using BlueTopo as the reference for the accuracy test are: 1) it is referenced consistently to a known vertical datum; 2) it has better spatial resolution than many other publicly-available data sets; 3) being produced and maintained by NOAA OCS for supporting NOAA's nautical charting mission, it was created with a focus on accuracy, datum consistency, and reliability; and 4) it is available via an AWS S3 bucket with code for access provided on a public GitHub repository, facilitating use in SlideRule ATL24. A disadvantage is that, while ATL24 has global coverage, BlueTopo currently only covers the U.S. East Coast and Gulf Coast. However, despite the limited geographic extent, BlueTopo was determined to be the best publicly-avaiable bathymetric data set for use as a reference data set in testing the accuracy of ATL24. Other publicly-available bathymetric data sets, such as ETOPO2022, CUDEM, and GEBCO 2023, have highly variable (and, in many cases, unknown) accuracy, which, in some places, is worse than the accuracy of ATL24, making them unsuitable as reference data sets for testing the accuracy of ATL24. Furthermore, the vertical datum of other data sets is not always known. Statistics to be computed in the accuracy test include: root mean square error (RMSE), R^2 of a linear regression of BlueTopo on ATL24, mean error (i.e., bias), standard deviation, and minimum and maximum (signed) errors. To the extent possible, the accuracy test will follow the procedures in the American Society for Photogrammetry and Remote Sensing (ASPRS) Positional Accuracy Standards for Digital Geospatial Data, 2nd Edition (Abdullah 2023).

5.1 Uncertainty Modeling

The ICESat-2 total propagated uncertainty (TPU) model leverages the project team's work on the cBLUE (comprehensive bathymetric lidar uncertainty estimator) software tool and subaqueous TPU model, which was initially developed for airborne bathymetric lidar (Eren et al. 2019). The TPU model underlying cBLUE breaks the uncertainty into two components: subaerial (comprising everything above the water surface) and subaqueous (water surface to seafloor). As applied to ICESat-2 bathymetry, this approach is illustrated in Fig. 9.

The subaerial vertical uncertainty for ATL24 is obtained from the sigma_h parameter in the /gtx/geolocation group in ATL03, which is an estimate of the total photon height uncertainty. While sigma_h is currently a static parameter, based on information supplied by the the ICESat-2 Project Science Office (PSO), in the future, it will vary on a point-by-point basis. The horizontal (or "planimetric") component of the subaerial model uses the *sigma_along* and *sigma_across* variables from the ATL03 /gtx/geolocation group.

The subaqueous component of the TPU model utilizes Monte Carlo ray tracing. Thousands of simulated rays are generated and refracted at a modeled sea surface. These simulated rays then undergo scattering within the water column, leading to an ensemble of plausible 3D spatial coordinates of detected bathymetric points, roughly within an "uncertainty ellipsoid" as depicted in Fig. 9. The planimetric and vertical spread of these points is then used to quantify the positional uncertainty. Because the Monte Carlo ray tracing requires long computation times, it is not run in real time at the point of generating a bathymetric data set. Instead, it is run in advance for many different combinations of wind speed and turbidity, with the latter represented by the diffuse attenuation coefficient of downwelling irradiance, K_d. The outputs of these runs are used to generate a set of look up tables (LUTs), each storing the parameters of a quadratic fit of depth to depth uncertainty for a given wind speed and K_d. The Monte Carlo ray tracing is based on the work of Curtis D. Mobley 1999, and full details are given in Eren et al. 2019.

The subaqueous uncertainty component is applied only to photons that are below the estimated sea surface, and not labeled by the ATL24 classifier as sea surface photons. The uncertainty model has a tendency to underestimate vertical uncertainty in very shallow waters. Under the assumption that there is a minimum level of uncertainty associated with the laser pulse entering the water, we establish a minimum subaqueous vertical uncertainty of 0.10m. If a given photon has no valid wind speed or K_d value, the maximum value for that parameter is used to retrieve LUT coefficients and offsets. When this happens, the invalid_kd or invalid_wind_speed flags are set to 1 to indicate invalid K_d or wind speed respectively.

As the last step in the uncertainty modeling process, the horizontal and vertical subaerial and subaqueous uncertainties are added in quadrature to compute the horizontal and vertical seafloor TPU.

Pseudocode for the uncertainty algorithm is given below:

Algorithm 11 Uncertainty Algorithm

Input: K_d, wind speed, orthometric height (OH), pointing angle, sea surface height (SH), along track and across track uncertainty coefficients and offsets, ATL24 classification, and along-track (sigma_along), across-track (sigma_across), and vertical uncertainties (sigma_h) from ATL03

Output: THU, TVU

```
Depth = SH - OH
for D_i ∈ Depth and C_i ∈ Classification do
    if D_i < 0 and C_i ! = 40 then
        Vertical uncertainty coefficients and offsets (A_v, B_v) ← K_d, wind speed, pointing angle
        Horizontal uncertainty coefficients and offsets (A_h, B_h) ← K_d, wind speed, pointing angle
        Subaqueous vertical uncertainty (SVU) ← A_v D_i + B_v
        Subaqueous horizontal uncertainty (SHU) ← A_h D_i + B_h
        if SVU < 0.10 then
            (SVU) ← 0.10
        end if
    else
        (SVU) ← 0
        (SHU) ← 0
    end if
end for
THU ← sqrt(sigma_along^2 + sigma_across^2 + SHU^2)
TVU ← sqrt(sigma_h^2 + SVU^2)
```

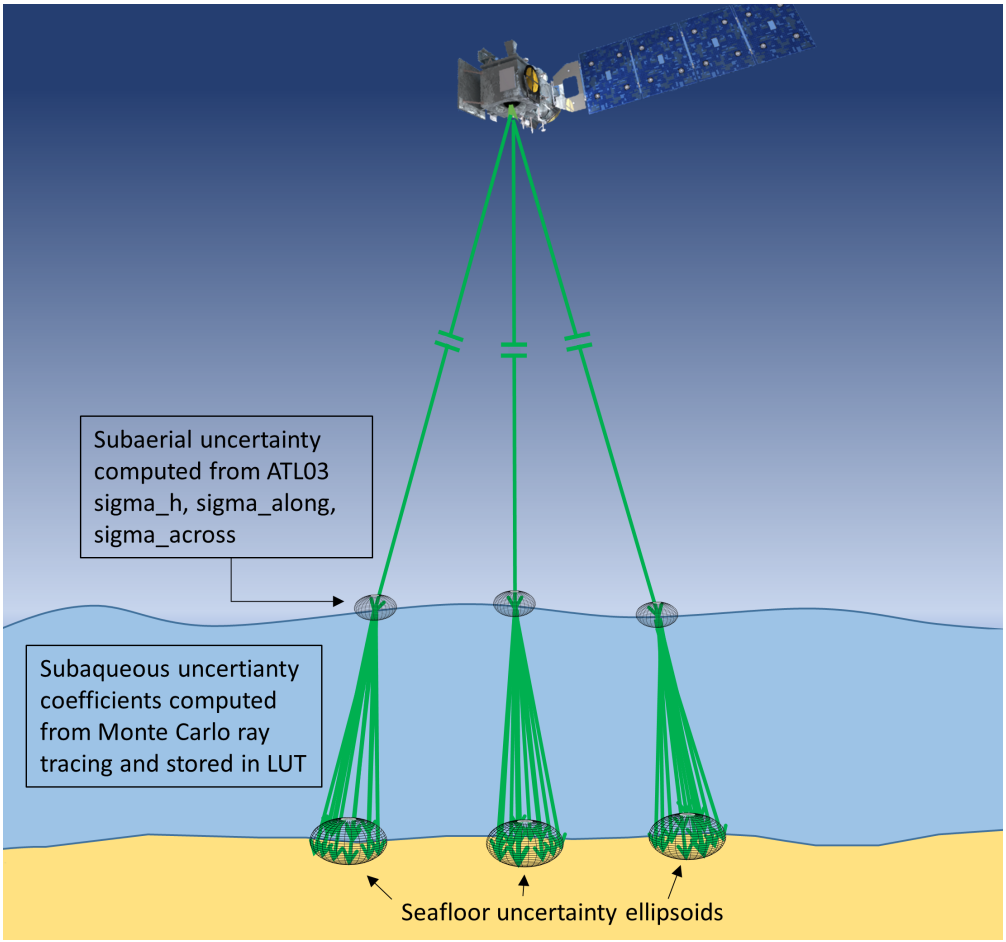

Figure 9: *Approach to ATL24 TPU modeling.*

5.2 Computational Performance

5.2.1 Model Training Data

A key step in the ATL24 processing workflow is training each of the machine learning classification algorithms to extract signal and accurate label the data to the best of each algorithm's ability given variations in environmental conditions. As such, a dataset of manually labeled ATL03 was created (Ohlwiler et al. 2023).

While the database was developed specifically for ATL24, it has been made available through a ScholarsArchive publication in hopes that the datasets will benefit other researchers in the scientific community, including those interested in developing and testing their own algorithms for auto-segmentation of seafloor and sea surface returns. These data sets are provided 'as is' and no claims are made about their suitability for any particular purpose. Please note that refraction correction has not been performed on these datasets. The reason is that these data are intended for use in training and testing algorithms for automatically detecting seafloor and sea surface returns in ATL03 datasets—steps which are performed before refraction correction. For users who would like to apply refraction correction, Python and MATLAB scripts are available on this GitHub repository: https://github.com/ICESat2-Bathymetry/Information

5.2.2 Individual algorithm performance

There are many performance metrics to choose from when evaluating classification algorithms. As such, it is difficult to select the appropriate metrics or combinations of metrics that suit a specific application. For example, some accuracy measures could be ill-suited for characterizing datasets that have imbalanced classification distributions. Additionally, the metrics can be selected from those that indicate how well the algorithm classifies the photons or they could address how well the identified sea floor photons match the elevations from independent reference data.

Furthermore, we could choose to make these measurements over individual ground tracks and then aggregate these measurements, or we could instead combine all the ground tracks into a single dataset and evaluate a single metric on this single dataset. These two methods might be similar or they may produce very different results depending on how the errors are distributed within each dataset and also on the sizes of the individual datasets.

When quantifying the ATL24 algorithms' classification performances, F1-scores, balanced accuracy and accuracy are used. Other scores like Matthews correlation coefficient (MCC, or phi coefficient), calibrated F1, and area under precision-recall curve (AUC-PR) were also investigated. However, these values were strongly correlated with F1 and BA, and ultimately did not offer any additional information on algorithm performance.

F1-scores characterize the overall (multi-class) performance of an algorithm and assist when measuring the error for individual (binary) classifications, for example, bathymetry versus non-bathymetry. F1 scores are more sensitive to errors in the minority class, especially when the dataset is unbalanced. However, a drawback of using F1-scores is that they do not include true negative predictions, whereas balanced accuracy is a measure that does include true negatives. In addition, balanced accuracy places more emphasis on overall performance across all data classes. ATL24, or rather, ATL03 input datasets, have a high-class imbalance. The imbalance is because there are a very high number of sea surface points relative to

both the background noise photons and the often, sparse amount of bathymetry due to signal attenuation (turbidity) or low density seafloor returns in deeper depths. In order to summarize scores across all classes, Macro and Micro F1 scores were used. Macro F1 scores are computed by averaging binary F1 scores across all classes without weighting. Similar to binary F1 scores, Macro F1 scores emphasize performance of the minority class. Micro F1 scores are computed by tallying true positives, false positives, and false negatives, regardless of photon class, and then computing an F1 score from the combined tallies. Similar to binary balanced accuracy, Micro F1 scores emphasize overall performance across all classes. By including the three types of accuracy metrics, the outcomes from each algorithm can be compared to determine the performance differences quantitatively, proving insight into why there are performance differences in specific locations and conditions.

Table 6: *Sea Surface cross validation results from 180 labeled datasets.*

Algorithm	Accuracy	F1-Score	Balanced Accuracy
CoastNet	0.959	0.973	0.945
Quantile Trees	0.947	0.965	0.928
OpenOceans++	0.946	0.965	0.910
Ensemble	0.970	0.981	0.946

Table 7: *Bathymetry cross validation results from 180 labeled datasets.*

Algorithm	Accuracy	F1-Score	Balanced Accuracy
Bathy Pathfinder	0.974	0.198	0.567
CoastNet	0.991	0.794	0.901
CShelph	0.982	0.406	0.635
Median Filter	0.987	0.671	0.798
OpenOceans	0.985	0.686	0.851
Quantile Trees	0.986	0.605	0.740
Ensemble	0.992	0.816	0.892

Table 8: *Cross validation results from 180 labeled datasets.*

Algorithm	Macro F1-Score	Micro F1-Score
Bathy Pathfinder	0.672	0.933
CoastNet	0.882	0.949
CShelph	0.748	0.941
Median Filter	0.839	0.946
OpenOceans	0.836	0.938
Quantile Trees	0.808	0.935
Ensemble	0.889	0.950

Table 6 provides the accuracies, F1-scores and balanced accuracy for the sea surface points when using a cross-validation strategy from 180 manually labeled datasets for the 4 algorithms that identify sea surface. Cross-validation is an important part of assessing an ML algorithm's performance by separating the data set into multiple training and testing sets and ensures that the model works well on data not previously seen. Data separation reduces overfitting by testing multiple data splits and provides a better estimate of model accuracy than a simple train-test split. A few observations from Table 6 is that CoastNet out performs the other two base algorithms for correctly identifying sea surface photons, although all three accuracy metrics between the 3 individual algorithm inputs are fairly close. However, the ensemble scores for sea surface photon classification are superior in all performance metric categories. Table 7 provides the cross-validation results for each of the individual algorithms and the ensemble with regard to identifying bathymetry photons. These results show fairly close accuracy numbers which is primarily attributed to the imbalance of the dataset, as most photons are sea surface. However, the F1 scores and the balanced accuracies are quite different across the algorithms. As with the sea surface results, the ensemble provides the best quality output in comparison to the other, individual approaches.

5.3 Data Quality and Filtering Flags

For the initial release of ATL24 the product provides some parameter values and binary flags to assist users with filtering data that meets certain criteria or requirements. One of the most useful parameters for filtering on the general quality of the classification accuracy is the *confidence* variable. The confidence value is a per photon value determined by the ensemble as a prediction probability. The prediction probability is determined similar to a softmax function but has a few differences. Ultimately, the prediction probability is the confidence that the classification is correct for each of the three classes (surface, bathymetry and other). These values are combined as log-odds scores across all decision trees in the model. By passing the combined score through a logistic function (the inverse of the logit function) the ensemble transforms the log-odds scores into a range of 0 to 1 as the prediction probability for that classification. The classification with the maximum prediction probability is used as the photon's classification (the argmax). The numerical value of the maximum prediction probability is used as the photon's confidence score.

Related to this *confidence* parameter is the *low_confidence_flag* parameter. This value is currently set to 0.6 as a best-estimate for removing those photon classification that have a lower probability of being correct. Removing these lower confidence photons (=1) will certainly eliminate many bathymetry false positives but could be at the cost of removing true positives in the case where fewer of the base algorithm inputs to the ensemble agreed. The threshold is determined by comparing the photon elevations and geolocations to a reference data set. A direct comparison between the provided the vertical disparity. For those comparisons that were within 5 m of separation, the corresponding F1 score was computed. This process was repeated as a function of the ensemble's *confidence* score. This functional relationship peak value occurred at 0.6. However, it should be noted that the 0.6 value is based on evaluation of the ensemble performance metrics given the initial repository of training data, the specific reference data and the collective contribution of the six base classification algorithms. This value is expected to change with future iterations of the ATL24 product.

In terms of ATL24 flags on the product, there are two that provide insight into situations where some of the ancillary data were not available for a given photon. The *invalid_kd* flag is a binary flag that indicates the absence (=1) of a VIIRS Kd490 value within ±1 day of the time tag on the photon. Having the VIIRS Kd490 value is an important component in the total propagated uncertainty calculation. An invalid flag indicates that the uncertainty estimate may be compromised. The invalid_wind_speed is also a binary flag to indicate the absence (=1) of a corresponding ATL09 wind speed. Absense of wind speed also impacts the total propagated uncertainty estimation. Other useful flag of interest when considering filtering for higher quality data is the *sensor_depth_exceeded*. This flag (=1) is for scenarios when the photon depth exceeds any plausible ICESat-2 depth but remains (=0) when the depth is reasonable. Finally, the *night_flag* is a binary flag to indicate when the photon was detected in the absence of sunlight (=1). This was added with the knowledge that often the presence of solar background noise degrades the classification accuracy. The combination of uncertainty values, ensemble confidence scores, *invalid_kd flags*, and *invalid_wind_speed* flags, and *night_flags* enables robust filtering of the data for specific use cases.

6 ATL24 Implementation Architecture and Product Accessibility

6.1 SlideRule Overview

The current user experience with ICESat-2 data is associated with downloading large volumes of standard data products from NSIDC and then developing independent routines if the goal is to explore new parameterizations or data resolutions for their own research. This paradigm is the case for ICESat-2 but it is also the scenario for many NASA satellite missions and the supporting NASA DAACs. Often the only data tools made available to users are those for geographical and/or temporal subsetting and although extremely useful in data downloads the requests can take hours depending on the size of the area. State-of-the-art solutions to length and voluminous data downloads seem to be leveraging on-demand, cloud-based processing. One example of this is the Alaska Satellite Facility's Hybrid Pluggable Processing Pipeline (ASFHyP3) for customized processing of SAR images across multiple missions. The OpenTopography Project provides another example through its support of web-based services for scalable capabilities in processing and analysis of Earth science-oriented topography data (Shean et al. 2023). These more modern approaches to data production and dissemination inspired SlideRule, with specific applications for ICESat-2.

SlideRule is an on-demand data processing system for rapid, scalable, open science, which is open to the public and accessible at https://slideruleearth.io. SlideRule runs in Amazon's cloud under GSFC Code 606's Science Managed Cloud Environment (SMCE) and has access to NASA's Cumulus data archives. SlideRule provides web-services for researchers and other data systems to generate custom data products in real-time using processing parameters supplied at the time of the request.

Scientists access SlideRule directly from any Python environment using a provided client; a Javascript client is also provided for integrating SlideRule into other web-based systems. SlideRule is currently being used by glacier, snow, and bathymetry researchers to process tens of thousands of ICESat-2 granules each month.

SlideRule also supports private instantiations of its infrastructure that require authenticated access. These instantiations, called private clusters, are managed by the SlideRule Provisioning System at https://ps.slideruleearth.io. Private clusters are used for executing large processing runs, providing dedicated compute resources, and running proprietary algorithms.

6.2 Deployment Environment

ATL24 will use SlideRule to provide the compute infrastructure for all four project objectives:

- The atl24g gold standard product will be generated by a private instantiation of SlideRule running in the AWS us-west-2 data center. The granules will initially exist in SlideRule's private S3 bucket prior to being transferred to the NSIDC.
- The atl24s and atl24p web services will be provided by the public instantiation of SlideRule that runs in the AWS us-west-2 data center.
- The graphical web interface will be hosted in AWS S3 and served by Amazon's CloudFront at https://client.slideruleearth.io.

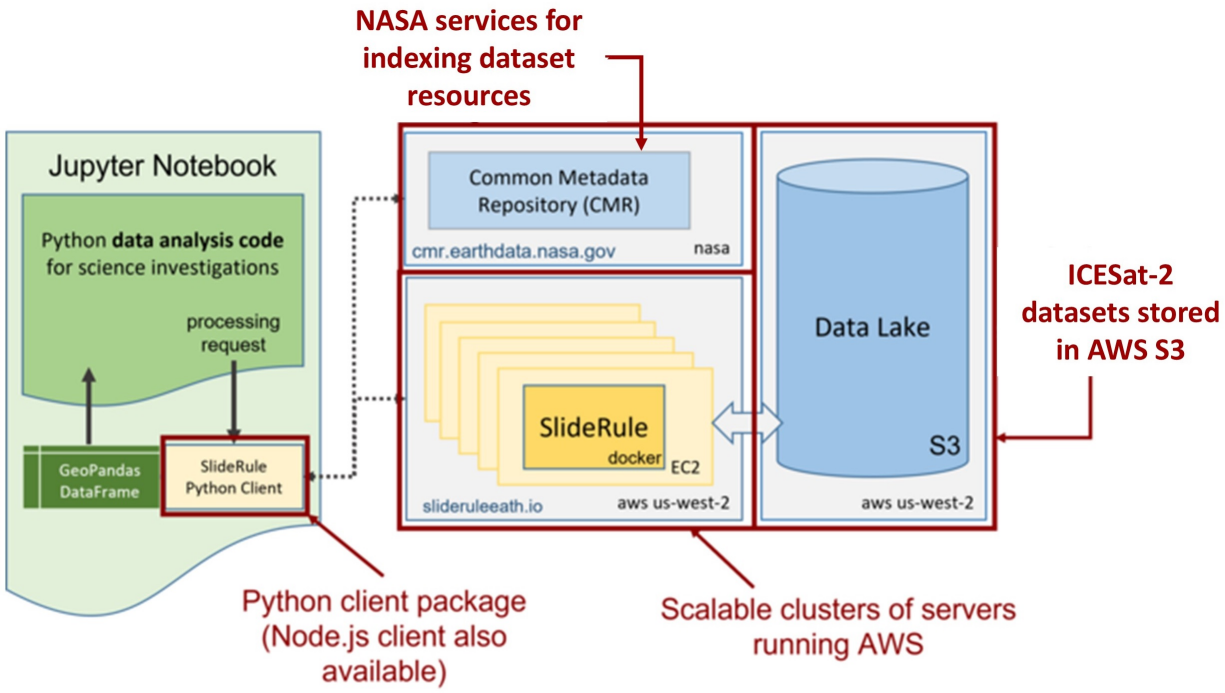

Figure 10: *Top Level SlideRule Architecture*

SlideRule Native Runtime

The native runtime environment for SlideRule services is an extended Lua interpreter where each request maps to a Lua script that instantiates custom classes written in C++ to perform the processing needed to fulfill the request.

The runtime is designed to quickly complete requests and return results back to users in near real-time. To that end, all requests are expected to complete within 10 minutes, and results are streamed back to the user as soon as they are available, over a TCP/IP connection that remains open for the entire time of the request. (It is typical for the users that request many granules to be processed at once to start receiving results for parts of their request that have finished before other parts of their request have even begun to be processed).

The native runtime environment will be used for the atl24s and atl24p endpoints, but will not be used exclusively for the atl24g endpoint, as the processing needed for that endpoint

includes long-running Python scripts. In addition to using the native runtime, the atl24g endpoint will also use the SlideRule Container Runtime.

SlideRule Container Runtime

The container runtime environment for SlideRule services is a new runtime implemented specifically to meet the needs of the atl24g endpoint. It uses the cluster management, intelligent load balancing, and job orchestration components of SlideRule to kick-off and communicate with Docker containers that are u

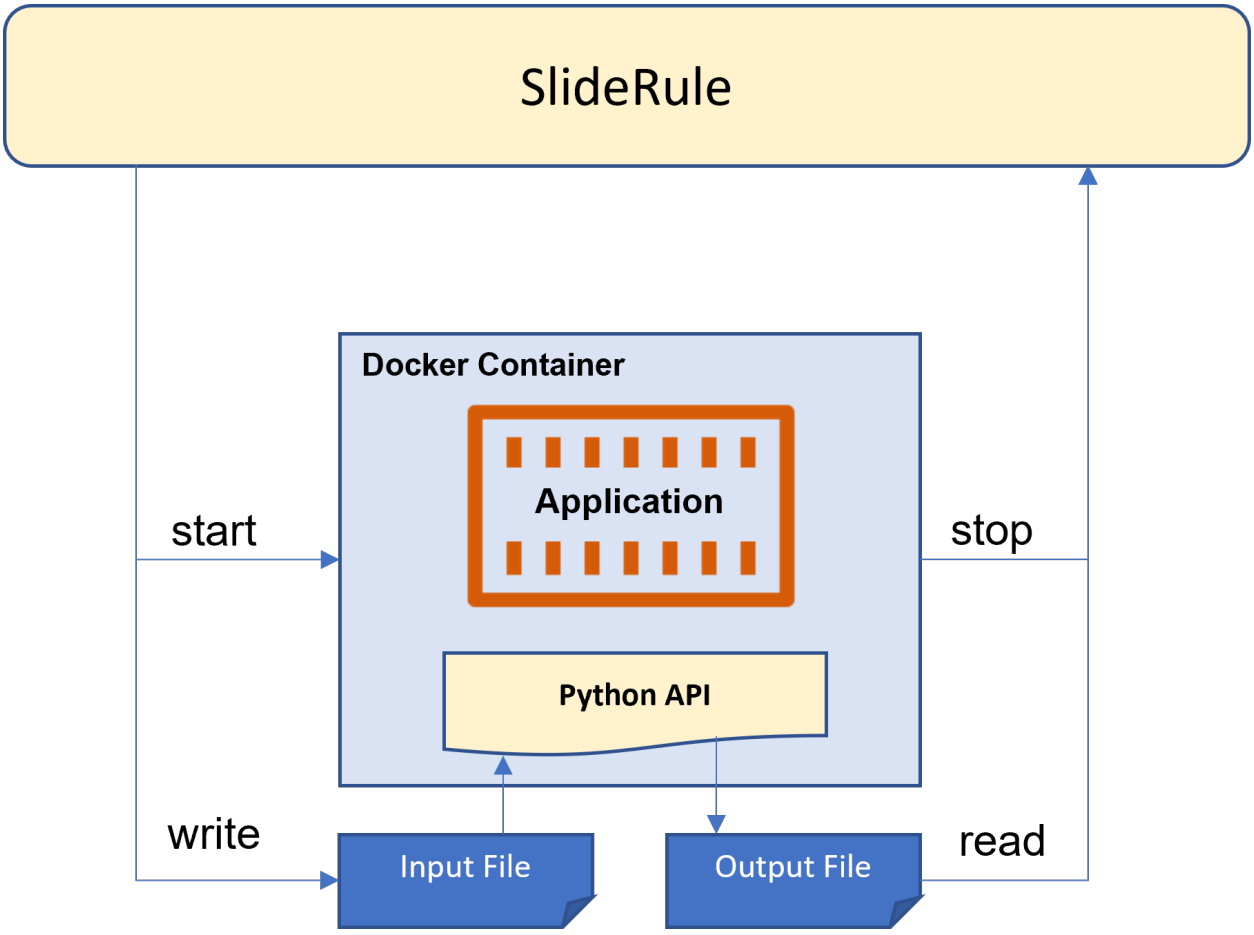

Figure 11: *Top Container Schematic of SlideRule runtime environment*

Applications written in Python will execute inside a Docker container running a Python environment, and will use a provided Python API to retrieve a list of input files and return a list of output files. Additional Docker container environments will be made available for programs not written in Python.

6.3 Development Environment

The development of the atl24g, atl24s, and graphical web page will be done on local development machines and coordinated through the GitHub ICESat-2 organization. This includes training models, writing source code, compiling code, and deploying to both test and production environments.

Both compute and storage services in AWS are available through the SlideRule SMCE account and will be used on an as-needed basis.

The following data resources will be stored in the SlideRule SMCE account S3 bucket:

- Labeled photon data
- Global bathymetry mask
- Refractive index
- Uncertainty lookup table

The following Docker images will be stored in the SlideRule SMCE account container registry:

- SlideRule server, intelligent load balancer, and monitor
- Python runtime environment

The following applications will be hosted in the SlideRule SMCE S3 bucket:

- Graphic web interface
- Documentation webpage

6.4 Bathymetric Processing Masks

Gridded surface masks for land ice, sea ice, land, ocean and inland water are used within the processing chain of ATL03 to reduce the volume of data processed and guide the production of surface-specific higher-level ICESat-2 data products.

A key step in the ATL24 workflow is applying a bathymetric search mask (Figure 12), which determines where to search for bathymetry. ATL03 granules within the mask are input to the algorithm to search for bathymetry, while those outside the mask are ignored in subsequent ATL24 processing steps. Use of the mask drastically reduces computation time by ignoring data that are on land or far too deep and/or too turbid for ATLAS bathymetric measurement. However, there is an important tradeoff in establishing the mask: including too much area unnecessarily increases processing time, while being too restrictive could lead to bathymetry being missed. Our guiding philosophy in establishing the mask was to err on the side of including too much area, to minimize the probability of missing the discovery of new bathymetric features, such as offshore sandbars, reefs, seamounts, or other submerged features, including those far from shore. Since discovery of such features could lead to major scientific advances, some increase in processing time was determined to be an acceptable tradeoff.

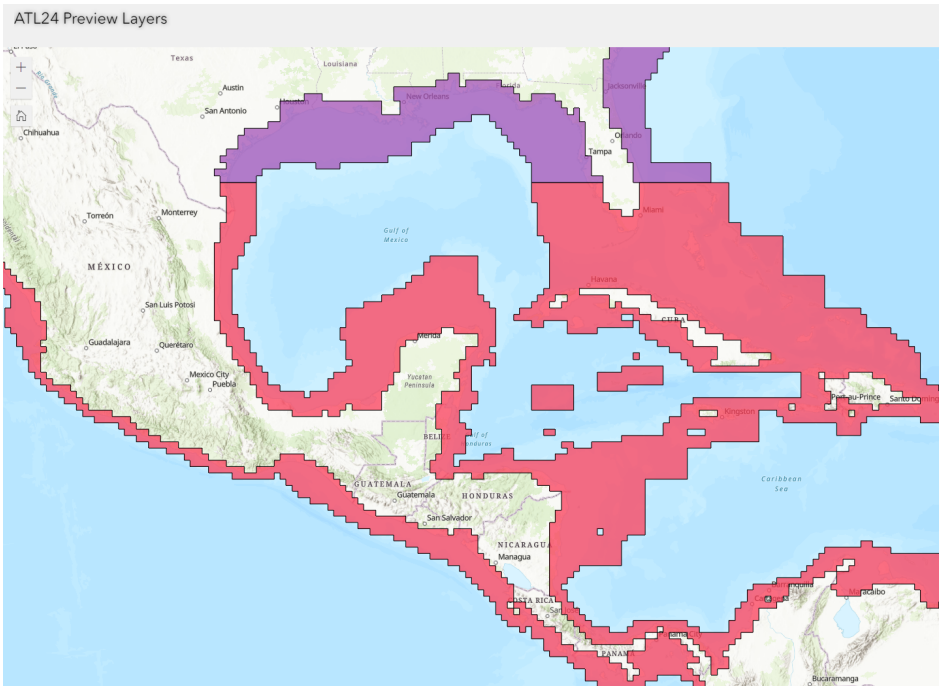

Figure 12: *Example ATL24 bathymetric search mask. The colors in the figure correspond to ATL03 granule regions.*

The general process used to create the mask was to use a coarse, global bathymetric data set to segment the areas in which ICESat-2 bathymetric measurement might be feasible, applying generous buffers, both vertically and horizontally, to extend the coverage extents and avoid missing "unexpected" bathymetry. The input consisted of the NOAA ETOPO 2022 30 arc-second data set (NOAA National Centers for Environmental Information, 2022). ETOPO 2022 is a seamless, global topographic-bathymetric dataset compiled from best-available sources with EGM 2008 orthometric heights. Offshore bathymetric data within ETOPO

2022 largely make use of General Bathymetric Chart of the Oceans (GEBCO) 2022 data, which in turn, use bathymetry prediction from satellite gravimetry for large areas of the global ocean. However, the nearshore data, which better correspond with the areas in which ICESat-2 bathymetric observations are feasible, often incorporate more accurate source data, including, in some cases airborne bathymetric lidar data and single beam and multibeam echosounder (MBES) data. Because of the unknown accuracy of the ETOPO 2022 dataset in many nearshore areas, as well as its relatively coarse resolution and the goal of not missing ICESat-2 bathymetry detections, both horizontal and vertical buffers were applied to generate the search mask. First, although the limit of ICESat-2 depth measurement in even the very clearest areas is 50 m (in part, due to the telemetry window settings, as described in J. Dietrich et al. 2023), an additional 50 m depth buffer was applied. Hence, the initial mask boundary was set at the 100-m depth contour generated from the ETOPO 2022 dataset. The 100-m depth contour was intersected with a 0.25° grid and the grid cells that were selected by the intersection are the basis for the masks. Next, a horizontal buffer of -m was applied. Then, manual editing was performed to fill gaps (i.e., join disconnected, but proximate areas) and improve the mask in challenging areas, such as around small islands, archipelagos, and complex coastal geometries. Throughout the process, the guiding philosophy was to err on the side of including too much area. An additional data layer used in this analysis was the ICESat-2 Bathymetric Retrievability Index, developed by the ATL24 team and described in J. Dietrich et al. 2023). The inputs to this layer consisted of Kd490, diffuse attenuation coefficient of downwelling irradiance at 490 nm, data from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensors on the on the Suomi NPP and NOAA-20 satellite missions, and ETOPO2022. Kd490 was converted to Kd532 (corresponding to ATLAS's wavelength), and then converted to Secchi depth, Zsd, using an empirical relationship (Guenther 1985). Monthly Zsd averages were overlaid on the ETOPO2022 layer. Masks were then generated for all areas for which Zsd $> Z_{ETOPO2022}$, representing the areas in which ICESat-2 is predicted to be able to measure bathymetry. A final raster layer was generated representing the number of months per year that ICESat-2 bathymetry retrieval is anticipated to be possible at each grid cell. This layer was disseminated via an ArcGIS Online webGIS: https://experience.arcgis.com/experience/474b9d16f9da4ca4b2830fcfa92852d9

6.5 Known Issues

ATL24 product known issues can be categorized as data quality issues, user utility issues, or product parameterization. Each of these categories has several different facets that range in complexity of the problems and the possibility of a solution.

- **Data Quality**

  1. **Classification performance:** The most notable issue with data quality will likely be the photon-classification accuracy, namely Class 40 false positives. Although the innovation of using an ensemble improves the classification over the individual algorithm approaches, there are still many locations where ATL24 is reporting bathymetry incorrectly. Certainly, the ensemble confidence score helps limit some of these false positives but there can be improvements. Table 9 provides a summary of the classification issues on version 1.0 of ATL24 and mentions possible solutions for future data releases.

Table 9: *Known issues, reasons and possible solutions to ATL24 classification accuracy*

Classification Issue	Reason	Possible Solution
Class 40 false negatives in deep water	Classification algorithms have been developed on shallow depth environments	Create new training data in locations with increase depth opportunities
Class 40 false positives for daylight granules	Solar background creates a different scenario of signal finding	create a separate model pipeline for day and night granules
Class 40 false positives in water column	Turbidity, instrument response, fish and mermaids	Create training data with higher levels of noise (turbity andor solar background)
Class 40 false positives in depths past extinction depth	High levels of noise within the entire vertical range window	Filter those class 40 points that have high vertical uncertainty values
Class 40 false positives in close proximity of sea surface	The ATL24 algorithms often mistake dynamic or highly reflective sea surface as bathymetry. The false positives are also often related to the instrument response signal (echo)	Filter by the ATL03 flag *return_source* where if the flag is (=3) it is predicted to be an echo

  2. **TPU Estimation:** The uncertainty estimates from the ATL24 TPU model are currently optimistic when compared with outputs of empirical accuracy assessments. Improvements to the TPU model for uncertainty determination are in the works for a future ATL24 release.
  3. **EGM08 Geoid:** The EGM08 geoid model used for computing ATL24 orthometric heights is outdated and contains errors.

4. **Refraction correction:** The refractive index of water layer usid in the ATL24 refraction correction is based on global, 0.25 degree resolution temperature and salinity datasets processed using the Quan-Fry equation Quan and Fry 1995. The current version of the refractive index layer uses only annual averages of salinity and temperature at each geographical location and does not currently account for temporal variability.

- **User Utility**

1. **Uncertainty value:** For the subaqueous photons the uncertainty value on each photon is a combination of the uncertainty in ATL03 and the TPU model uncertainty. If the user would like to separate these values the original ATL03 *sigma_h, sigma_lat,* the *index_ph* value can provide the link back to the original ATL03 photon's uncertainties.
2. **Refraction correction:** For the subaqueous photons the correction value applied from the index of refraction data layer can be removed by using the *index_ph* to link back to the original ATL03 photon's position.

- **Product Parameterization**

1. **Waves:** Currently there are no parameters on ATL24 related to wave characteristics derived from the sea surface photons
2. **Classification confidence:** When the value of the *confidence* is less than 0.6 the *low_confidence* = 1. This threshold value will change with future iterations of ATL24.

7 Appendices

7.1 Appendix A: Acronyms

Abbreviation	Full Form
ATLAS	Advanced Topographic Laser Altimetry System
AWS	Amazon Web Services
CMR	Common Metadata Repository
CSV	Comma Separated Vector
CUDEM	NOAA NCEI Continuously Updated Digital Elevation Model
DAAC	Distributed Active Archive Center
ECEF	Earth Centered Earth Fixed
ETOPO	Earth Topography Global relief model (NOAA)
GEBCO	General Bathymetric Chart of the Oceans
GSFC	Goddard Space Flight Center
ICESat-2	Ice, Cloud, and land Elevation Satellite
LUT	Lookup table
JSON	JavaScript Object Notation
MBES	Multibeam Echosounder
NCEI	National Centers of Earth Information
NDWI	Normalized Difference Water Index
NOAA	National Oceanic and Atmospheric Administration
NSIDC	National Snow and Ice Data Center
OSU	Oregon State University
S3	Amazon Simple Storage Solution
SMCE	Science Managed Cloud Environment
TEP	Transmitter Echo Path
TPU	Total propagated uncertainty
UTexas	University of Texas at Austin
UTM	Universal Transverse Mercator
VIIRS	Visible Infrared Imaging Radiometry Suite
ZETOPO	

7.2 Appendix B: User Notes

8 References

References

Abdullah, Qassim (2023). "The ASPRS Positional Accuracy Standards, Edition 2: The Geospatial Mapping Industry Guide to Best Practices." *Photogrammetric Engineering & Remote Sensing* 89.10.

Albright, Andrea and Craig Glennie (May 2020). "Nearshore Bathymetry From Fusion of Sentinel-2 and ICESat-2 Observations". *IEEE Geoscience and Remote Sensing Letters* 18.5, pp. 900–904. DOI: 10.1109/LGRS.2020.2987778.

ASPRS (2013). *LAS Domain Profile Description: Topo-Bathy Lidar*. URL: https://www.asprs.org/a/society/divisions/lidardivision/LAS_Domain_Profile_Description_Topo-Bathy_Lidar.pdf (visited on 06/30/2024).

Austin, Roswell W and George Halikas (1976). "The index of refraction of seawater".

Babbel, Benjamin J., Christopher E. Parrish, and Lori A. Magruder (Feb. 2021). "ICESat-2 Elevation Retrievals in Support of Satellite-Derived Bathymetry for Global Science Applications". *Geophysical Research Letters* 48.5. DOI: 10.1029/2020GL090629. URL: https://onlinelibrary.wiley.com/doi/10.1029/2020GL090629 (visited on 12/14/2021).

Bentley, Jon Louis (1990). "*K* -d trees for semidynamic point sets". en. *Proceedings of the sixth annual symposium on Computational geometry - SCG '90*. Berkley, California, United States: ACM Press, pp. 187–197. DOI: 10.1145/98524.98564. URL: http://portal.acm.org/citation.cfm?doid=98524.98564 (visited on 04/09/2024).

Brock, John C, C Wayne Wright, Tonya D Clayton, and Amar Nayegandhi (2004). "LIDAR optical rugosity of coral reefs in Biscayne National Park, Florida". *Coral Reefs* 23, pp. 48–59.

Corcoran, Forrest, Christopher E. Parrish, Lori A. Magruder, and J. P. Swinski (2024). "A Scalable, Cloud-Based Workflow for Spectrally-Attributed ICESat-2 Bathymetry With Application to Benthic Habitat Mapping Using Deep Learning". *Earth and Space Science* 11.11. e2024EA003735 2024EA003735, e2024EA003735. DOI: https://doi.org/10.1029/2024EA003735. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024EA003735. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2024EA003735.

Costa, BM, TA Battista, and SJ Pittman (2009). "Comparative evaluation of airborne LiDAR and ship-based multibeam SoNAR bathymetry and intensity for mapping coral reef ecosystems". *Remote Sensing of Environment* 113.5, pp. 1082–1100.

Dietrich, James, Ann R Rackley Reese, Aimée Gibbons, Lori A. Magruder, and Christopher Parrish (Sept. 11, 2023). *Analysis of ICESat-2 Data Acquisition Algorithm Parameter Enhancements to Improve Worldwide Bathymetric Coverage*. preprint. Preprints. DOI: 10.22541/essoar.169447414.45310708/v1. URL: https://essopenarchive.org/users/657809/articles/662942-analysis-of-icesat-2-data-acquisition-algorithm-parameter-enhancements-to-improve-worldwide-bathymetric-coverage?commit=4518ff33a68d2d866f5fd7756d6b2892ead8cff0 (visited on 10/14/2023).

Dietrich, James T and Christopher E Parrish (2025). "Development and Analysis of a Global Refractive Index of Water Data Layer for Spaceborne and Airborne Bathymetric Lidar". *Earth and Space Science* 12.3, e2024EA004106.

Eren, Firat, Jaehoon Jung, Christopher E. Parrish, Nicholas Sarkozi-Forfinski, and Brian R. Calder (Aug. 1, 2019). "Total Vertical Uncertainty (TVU) Modeling for Topo-Bathymetric LIDAR Systems". *Photogrammetric Engineering & Remote Sensing* 85.8, pp. 585–596. DOI: 10.14358/PERS.85.8.585. URL: https://www.ingentaconnect.com/content/10.14358/PERS.85.8.585 (visited on 08/24/2022).

Forfinski-Sarkozi, Nicholas and Christopher Parrish (Sept. 2016). "Analysis of MABEL Bathymetry in Keweenaw Bay and Implications for ICESat-2 ATLAS". *Remote Sensing* 8.9, p. 772. DOI: 10.3390/rs8090772. URL: http://www.mdpi.com/2072-4292/8/9/772 (visited on 12/17/2021).

Forfinski-Sarkozi, Nicholas A. and Christopher E. Parrish (Apr. 1, 2019). "Active-Passive Spaceborne Data Fusion for Mapping Nearshore Bathymetry". *Photogrammetric Engineering & Remote Sensing* 85.4, pp. 281–295. DOI: 10.14358/PERS.85.4.281. URL: https://www.ingentaconnect.com/content/10.14358/PERS.85.4.281 (visited on 06/15/2023).

Gao, Jay (2009). "Bathymetric mapping by means of remote sensing: methods, accuracy and limitations". *Progress in Physical Geography* 33.1, pp. 103–116.

Goodman, Lowell R (1979). "LASER HYDGROGRAPHY". *Technical Papers on Airborne Laser Hydrography:(March-December 1978)* 25, p. 17.

Guenther, Gary C (1985). *Airborne laser hydrography: System design and performance factors*. Tech. rep. National Oceanic and Atmospheric Administration.

— (2007). "Airborne lidar bathymetry". *Digital elevation model technologies and applications: the DEM users manual* 2, pp. 253–320.

Guinehut, S, AL Dhomps, G Larnicol, and PY Le Traon (2021). *Multi Observation Global Ocean 3D Temperature Salinity Height Geostrophic Current and MLD [data set]*.

Guinehut, Stephanie, A-L Dhomps, Gilles Larnicol, and P-Y Le Traon (2012). "High resolution 3-D temperature and salinity fields derived from in situ and satellite observations". *Ocean Science* 8.5, pp. 845–857.

Hart, Peter, Nils Nilsson, and Bertram Raphael (1968). "A Formal Basis for the Heuristic Determination of Minimum Cost Paths". *IEEE Transactions on Systems Science and Cybernetics* 4.2, pp. 100–107. DOI: 10.1109/TSSC.1968.300136. URL: http://ieeexplore.ieee.org/document/4082128/ (visited on 04/09/2024).

Hickman, G Daniel and John E Hogg (1969). "Application of an airborne pulsed laser for near shore bathymetric measurements". *Remote sensing of Environment* 1.1, pp. 47–58.

Hoge, Frank E, Wallops Flight Center, and Robert N Swift (1980). "AIRBORNE OCEANOGRAPHIC LIDAR (AOL) DATA ACQUISITION MODES AND RECENT RESULTS". *Director, PMEL*, p. 75.

Jasinski, Michael F, Jeremy D Stoll, William B Cook, Michael Ondrusek, Eric Stengel, and Kelly Brunt (2016). "Inland and near-shore water profiles derived from the high-altitude Multiple Altimeter Beam Experimental Lidar (MABEL)". *Journal of Coastal Research* 76, pp. 44–55.

Jerlov, Nils Gunnar (1976). *Marine optics*. Elsevier.

Kastdalen, Leif, Morten Stickler, Christian Malmquist, and Jan Heggenes (2024). "Evaluating methods for measuring in-river bathymetry: Remote sensing green LIDAR provides high-resolution channel bed topography limited by water penetration capability". *River Research and Applications* 40.4, pp. 467–482.

Lambert, Selina E and Christopher E Parrish (2023). "Refraction Correction for Spectrally Derived Bathymetry Using UAS Imagery". *Remote Sensing* 15.14, p. 3635.

Leigh, Holly W. and Lori A. Magruder (Oct. 6, 2016). "Using dual-wavelength, full-waveform airborne lidar for surface classification and vegetation characterization". *Journal of Applied Remote Sensing* 10.4, p. 045001. DOI: 10.1117/1.JRS.10.045001. URL: http://remotesensing.spiedigitallibrary.org/article.aspx?doi=10.1117/1.JRS.10.045001 (visited on 07/07/2024).

Ma, Yue, Nan Xu, Zhen Liu, Bisheng Yang, Fanlin Yang, Xiao Hua Wang, and Song Li (2020). "Satellite-derived bathymetry using the ICESat-2 lidar and Sentinel-2 imagery datasets". *Remote Sensing of Environment* 250, p. 112047.

Magruder, Lori, Thomas Neumann, and Nathan Kurtz (May 2021). "ICESat-2 Early Mission Synopsis and Observatory Performance". *Earth and Space Science* 8.5. DOI: 10.1029/2020EA001555. URL: https://onlinelibrary.wiley.com/doi/10.1029/2020EA001555 (visited on 09/03/2022).

Magruder, Lori A. (Jan. 1, 2010). "Lidar waveform stacking techniques for faint ground return extraction". *Journal of Applied Remote Sensing* 4.1, p. 043501. DOI: 10.1117/1.3299657. URL: http://remotesensing.spiedigitallibrary.org/article.aspx?doi=10.1117/1.3299657 (visited on 07/07/2024).

Magruder, Lori A., Sinead L. Farrell, Amy Neuenschwander, Laura Duncanson, Beata Csatho, Sahra Kacimi, and Helen A. Fricker (Jan. 30, 2024). "Monitoring Earth's climate variables with satellite laser altimetry". *Nature Reviews Earth & Environment* 5.2, pp. 120–136. DOI: 10.1038/s43017-023-00508-8. URL: https://www.nature.com/articles/s43017-023-00508-8 (visited on 02/22/2024).

Magruder, Lori A., Tom Neumann, Nathan Kurtz, and Tyler Sutterley (2024). *Assessment of the Ice, Cloud and Land Elevation Satellite-2 Performance Against Prime Mission Science Requirements*. DOI: 10.2139/ssrn.4819375. URL: https://www.ssrn.com/abstract=4819375 (visited on 05/21/2024).

Magruder, L., Parrish, C., Perry, J., Swinski, J.P, Holwill, M., Kief, K., and Corcoran, F. (2025). *Ice, Cloud and Land Elevation Satellite (ICESat-2) Project Algorithm Theoretical Basis Document (ATBD) for Coastal and Along-track Bathymetry Product (ATL24)*. 1.0. NASA. DOI: 10.5067/PXJMCZDOMYLN.

Markus, Thorsten, Tom Neumann, Anthony Martino, Waleed Abdalati, Kelly Brunt, Beata Csatho, Sinead Farrell, Helen Fricker, Alex Gardner, David Harding, Michael Jasinski, Ron Kwok, Lori Magruder, Dan Lubin, Scott Luthcke, James Morison, Ross Nelson, Amy Neuenschwander, Stephen Palm, Sorin Popescu, Ck Shum, Bob E. Schutz, Benjamin Smith, Yuekui Yang, and Jay Zwally (Mar. 2017). "The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation". *Remote Sensing of Environment* 190, pp. 260–273. DOI: 10.1016/j.rse.2016.12.029. URL: https://linkinghub.elsevier.com/retrieve/pii/S0034425716305089 (visited on 01/30/2022).

Mobley, Curtis et al. (2022). "The oceanic optics book".

Mobley, Curtis D (2001). "Radiative transfer in the ocean". *Encyclopedia of ocean sciences* 4, pp. 2321–2330.

— (Dec. 1999). "Estimation of the remote-sensing reflectance from above-surface measurements". *Applied Optics* 38.36, p. 7442. DOI: 10.1364/AO.38.007442. URL: https://www.osapublishing.org/abstract.cfm?URI=ao-38-36-7442 (visited on 01/17/2022).

Moore, J Grange (1947). "The determination of the depths and extinction coefficients of shallow water by air photography using colour filters". *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 240.816, pp. 163–217.

National Academies of Sciences, Engineering and Medicine (2018). *Thriving on Our Changing Planet: A Decadal Strategy for Earth Observation from Space*. Washington, DC: The National Academies Press. DOI: 10.17226/24938. URL: https://nap.nationalacademies.org/catalog/24938/thriving-on-our-changing-planet-a-decadal-strategy-for-earth.

Neumann, Thomas A., Anthony J. Martino, Thorsten Markus, Sungkoo Bae, Megan R. Bock, Anita C. Brenner, Kelly M. Brunt, John Cavanaugh, Stanley T. Fernandes, David W. Hancock, Kaitlin Harbeck, Jeffrey Lee, Nathan T. Kurtz, Philip J. Luers, Scott B. Luthcke, Lori Magruder, Teresa A. Pennington, Luis Ramos-Izquierdo, Timothy Rebold, Jonah Skoog, and Taylor C. Thomas (Nov. 1, 2019a). "The Ice, Cloud, and Land Elevation Satellite – 2 mission: A global geolocated photon product derived from the Advanced Topographic Laser Altimeter System". *Remote Sensing of Environment* 233, p. 111325. DOI: 10.1016/j.rse.2019.111325. URL:

https://www.sciencedirect.com/science/article/pii/S003442571930344X (visited on 08/07/2023).

Neumann, Thomas A., Anthony J. Martino, Thorsten Markus, Sungkoo Bae, Megan R. Bock, Anita C. Brenner, Kelly M. Brunt, John Cavanaugh, Stanley T. Fernandes, David W. Hancock, Kaitlin Harbeck, Jeffrey Lee, Nathan T. Kurtz, Philip J. Luers, Scott B. Luthcke, Lori Magruder, Teresa A. Pennington, Luis Ramos-Izquierdo, Timothy Rebold, Jonah Skoog, and Taylor C. Thomas (Nov. 2019b). "The Ice, Cloud, and Land Elevation Satellite – 2 mission: A global geolocated photon product derived from the Advanced Topographic Laser Altimeter System". *Remote Sensing of Environment* 233, p. 111325. DOI: 10.1016/j.rse.2019.111325. URL: https://linkinghub.elsevier.com/retrieve/pii/S003442571930344X (visited on 11/06/2021).

NOAA (2021). *Hydrographic Surveys Specifications and Deliverables*. https://www.nauticalcharts.noaa.gov/publications/docs/standards-and-requirements/specs/HSSD_2022.pdf. Accessed: 2024-07-05.

Ohlwiler, Ellery E., Charles E. Ghartey, Ruth M. McCullough, Christopher E. Parrish, Lori A. Magruder, James T. Dietrich, Matthew Holwill, and Jonathan Markel (2023). *ICESat-2 Bathymetry Training and Testing Database*. In collab. with Larry Ward and Christopher E. Parrish. Version 1. DOI: 10.7267/J3860G66D. URL: https://ir.library.oregonstate.edu/concern/datasets/j3860g66d (visited on 07/06/2024).

Parrish, Christopher E., Lori Magruder, Ute Herzfeld, and et al. (Oct. 2022). "ICESat-2 Bathymetry: Advances in Methods and Science". OCEANS Conference 2022 Hampton Roads. Virginia Beach, VA.

Parrish, Christopher E., Lori A. Magruder, Amy L. Neuenschwander, Nicholas Forfinski-Sarkozi, Michael Alonzo, and Michael Jasinski (Jan. 2019). "Validation of ICESat-2 ATLAS Bathymetry and Analysis of ATLAS's Bathymetric Mapping Performance". *Remote Sensing* 11.14, p. 1634. DOI: 10.3390/rs11141634. URL: https://www.mdpi.com/2072-4292/11/14/1634 (visited on 11/06/2021).

Plant, Nathaniel G, K Todd Holland, and Jack A Puleo (2002). "Analysis of the scale of errors in nearshore bathymetric data". *Marine Geology* 191.1-2, pp. 71–86.

Polcyn, FABIAN C and DAVID R Lyzenga (1975). *Remote bathymetry and shoal detection with ERTS: ERTS water depth*. Tech. rep.

Quan, Xiaohong and Edward S Fry (1995). "Empirical equation for the index of refraction of seawater". *Applied optics* 34.18, pp. 3477–3480.

Ranndal, Heidi, Philip Sigaard Christiansen, Pernille Kliving, Ole Baltazar Andersen, and Karina Nielsen (2021a). "Evaluation of a Statistical Approach for Extracting Shallow Water Bathymetry Signals from ICESat-2 ATL03 Photon Data". *Remote Sensing* 13.17. DOI: 10.3390/rs13173548. URL: https://www.mdpi.com/2072-4292/13/17/3548.

— (2021b). "Evaluation of a statistical approach for extracting shallow water bathymetry signals from ICESat-2 ATL03 photon data". *Remote Sensing* 13.17, p. 3548.

Rice, Glen, Katrina Wyllie, Barry Gallagher, and Phuntsok Geleg (2023). "The National Bathymetric Source". *OCEANS 2023-MTS/IEEE US Gulf Coast*. IEEE, pp. 1–7.

Schutz, B. E., H. J. Zwally, C. A. Shuman, D. Hancock, and J. P. DiMarzio (2005). "Overview of the ICESat Mission". *Geophysical Research Letters* 32.21, L21S01. DOI: 10.1029/2005GL024009. URL: http://doi.wiley.com/10.1029/2005GL024009 (visited on 09/03/2022).

Shean, David, J. P. Swinski, Ben Smith, Tyler Sutterley, Scott Henderson, Carlos Ugarte, Eric Lidwa, and Thomas Neumann (Jan. 18, 2023). "SlideRule: Enabling rapid, scalable, open science forthe NASA ICESat-2 mission and beyond". *Journal of Open Source Software* 8.81, p. 4982. DOI: 10.21105/joss.04982. URL: https://joss.theoj.org/papers/10.21105/joss.04982 (visited on 03/07/2023).

Thomas, Nathan, Brian Lee, Oliver Coutts, Pete Bunting, David Lagomasino, and Lola Fatoyinbo (2022). "A Purely Spaceborne Open Source Approach for Regional Bathymetry Mapping". *IEEE Transactions on Geoscience and Remote Sensing* 60, pp. 1–9. DOI: 10.1109/TGRS.2022.3192825. URL: https://ieeexplore.ieee.org/document/9834969/ (visited on 08/26/2022).

Thomas, T. C., S. B. Luthcke, T. A. Pennington, J. B. Nicholas, and D. D. Rowlands (Apr. 2021). "ICESat-2 Precision Orbit Determination". *Earth and Space Science* 8.4. Number: 4, e2020EA001496. DOI: 10.1029/2020EA001496. URL: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020EA001496 (visited on 04/04/2024).

Titov, Vasily V, Frank I Gonzalez, EN Bernard, Marie C Eble, Harold O Mofjeld, Jean C Newman, and Angie J Venturato (2005). "Real-time tsunami forecasting: Challenges and solutions". *Natural Hazards* 35.1, pp. 35–41.

Warne, Denis Keith (1978). "Bathymetric Mapping with Landsat: A Practical Approach". *Offshore Technology Conference*. OTC, OTC–3225.

Wedding, Lisa M, Alan M Friedlander, Matthew McGranaghan, Russell S Yost, and Mark E Monaco (2008). "Using bathymetric lidar to define nearshore benthic habitat complexity: Implications for management of reef fish assemblages in Hawaii". *Remote Sensing of Environment* 112.11, pp. 4159–4165.

Wilson, Nicholas, Christopher E. Parrish, Tim Battista, C. Wayne Wright, Bryan Costa, Richard K. Slocum, Jennifer A. Dijkstra, and Matthew T. Tyler (Nov. 2019). "Mapping Seafloor Relative Reflectance and Assessing Coral Reef Morphology with EAARL-B Topobathymetric Lidar Waveforms". *Estuaries and Coasts*. DOI: 10.1007/s12237-019-00652-9. URL: http://link.springer.com/10.1007/s12237-019-00652-9 (visited on 10/27/2021).

Wright, C Wayne, Christine Kranenburg, Timothy A Battista, and Christopher Parrish (2016). "Depth calibration and validation of the Experimental Advanced Airborne Research LIDAR, EAARL-B". *Journal of Coastal Research* 76, pp. 4–17.

Xu, Nan, Yue Ma, Hui Zhou, Wenhao Zhang, Zhiyu Zhang, and Xiao Hua Wang (Sept. 2020). "A Method to Derive Bathymetry for Dynamic Water Bodies Using ICESat-2 and GSWD Data Sets". *IEEE Geoscience and Remote Sensing Letters*, pp. 1–5. DOI: 10.1109/LGRS.2020.3019396. URL: https://ieeexplore.ieee.org/document/9186300/ (visited on 12/14/2021).

Zhang, Dongfang, Yifu Chen, Yuan Le, Yusen Dong, Guangming Dai, and Lizhe Wang (2022). "Refraction and coordinate correction with the JONSWAP model for ICESat-2 bathymetry". *ISPRS Journal of Photogrammetry and Remote Sensing* 186, pp. 285–300.