

UC San Diego

UC San Diego Previously Published Works

Title

OpenAltimetry - rapid analysis and visualization of Spaceborne altimeter data

Permalink

<https://escholarship.org/uc/item/5fb5m15p>

Journal

Earth Science Informatics, 15(3)

ISSN

1865-0473

Authors

Khalsa, Siri Jodha S
Borsa, Adrian
Nandigam, Viswanath
[et al.](#)

Publication Date

2022-09-01

DOI

10.1007/s12145-020-00520-2

Peer reviewed

[Click here to view linked References](#)

OpenAltimetry - Rapid Analysis and Visualization of Spaceborne Altimeter Data

Siri Jodha S. Khalsa¹, Adrian Borsa², Viswanath Nandigam³, Minh Phan³, Kai Lin³, Christopher Crosby⁴, Helen Fricker², Chaitan Baru³, Luis Lopez¹

¹University of Colorado Boulder, CO, USA

²Scripps Institute of Oceanography, San Diego, CA, USA

³University of California San Diego, San Diego, CA, USA

⁴UNAVCO, Boulder, CO, USA

Corresponding Author:

Siri Jodha S. Khalsa

email: sjsk@nsidc.org

<http://orcid.org/0000-0001-9217-5550>

Revision Submitted 17 August 2020

Submitted 24 February 2020

to

Earth Science Informatics Special Issue

“Technical Innovations to Increase Scientific Access to Earth Observation Data”

Abstract

NASA's Ice, Cloud, and land Elevation Satellite-2 (ICESat-2) carries a laser altimeter that fires 10,000 pulses per second towards Earth and records the travel time of individual photons to measure the elevation of the surface below. The volume of data produced by ICESat-2, nearly a TB per day, presents significant challenges for users wishing to efficiently explore the dataset. NASA's National Snow and Ice Data Center (NSIDC) Distributed Active Archive Center (DAAC), which is responsible for archiving and distributing ICESat-2 data, provides search and subsetting services on mission data products, but providing interactive data discovery and visualization tools needed to assess data coverage and quality in a given area of interest is outside of NSIDC's mandate. The OpenAltimetry project, a NASA-funded collaboration between NSIDC, UNAVCO and the University of California San Diego, has developed a web-based cyberinfrastructure platform that allows users to locate, visualize, and download ICESat-2 surface elevation data and photon clouds for any location on Earth, on demand. OpenAltimetry also provides access to elevations and waveforms for ICESat (the predecessor mission to ICESat-2). In addition, OpenAltimetry enables data access via APIs, opening opportunities for rapid access, experimentation, and computation via third party applications like Jupyter notebooks. OpenAltimetry emphasizes ease-of-use for new users and rapid access to entire altimetry datasets for experts and has been successful in meeting the needs of different user groups. In this paper we describe the principles that guided the design and development of the OpenAltimetry platform and provide a high-level overview of the cyberinfrastructure components of the system.

Keywords

Cyberinfrastructure, data visualization, data discovery, altimetry, ICESat-2

Declarations

- Authors declare they have no Conflicts of interest or Competing interests
- Authors' contributions

Khalsa, Borsa, Nandigam, Crosby, Fricker and Baru participated in the conceptualization of the project and securing of funding. Development of software was the primary responsibility of Nandigam, Phan and Lin. Lopez developed the original Jupyter notebook that was linked to OpenTopography. Borsa, Khalsa and Fricker provided scientific guidance. Borsa provided overall project management.

- No new scientific data were generated by the project
- The OpenAltimetry project is supported by a grant from the National Aeronautics and Space Administration under award number NNX16AL89A
- Code is open sourced and available at <https://github.com/OpenAltimetry>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1. Introduction

The increasing complexity and size of Earth sciences datasets calls for better ways to access, visualize, analyze and understand the data that is being generated. In addition, Earth observation data is increasingly being used to inform decision making regarding pressing environmental and societal issues, driving the need for intuitive data interfaces that support the needs of non-expert users (Kavvada, et al., 2020). One of the challenges facing NASA's Earth Science Data and Information Systems (ESDIS) is that the innovative measurement techniques employed in new Earth observing missions often generate voluminous data products whose utilization is accompanied by steep learning curves and heavy demands on computational resources. NASA's Advancing Collaborative Connections for Earth System Science (ACCESS) Program is one way that the agency is attempting to address this challenge. ACCESS aims to improve and expand the use of NASA's Earth science data for scientific research and applications by leveraging modern techniques for discovering, managing and analyzing large and complex Earth science datasets.

This paper describes OpenAltimetry (OA), a cyberinfrastructure platform developed under the ACCESS 2015 program. OA targets a particular category of Earth-observing missions: satellite laser altimeters, and specifically the Ice, Cloud and land Elevation Satellite (ICESat) mission and its successor, ICESat-2. These missions launched in 2003 and in 2018, respectively, and produce data that are fundamentally different from the gridded data products that generated by most NASA Earth observing satellite missions. Polar orbiting satellite altimeters collect elevation data along densely sampled surface tracks that are widely separated in mid and lower latitudes. These data are only available as a set of file-based hierarchical data products, sequential in time but not spatially organized, with key parameters distributed among multiple data products. Non-expert users must therefore rely on data services provided by the mission and/or archive center, both of which are resource-limited and may not be attuned to the needs of all user constituencies.

The goal of OA is to provide an altimetry-specific data discovery, access and visualization platform that focuses on ease-of-use and transparent high-level functionality, with appeal to new users and experts alike.

2. The Design Drivers for OpenAltimetry

Here we describe the principles and requirements that guided the design of OpenAltimetry. OA was imagined as a cyberinfrastructure tool for easy access to satellite altimeter data from any source. Altimetry data measure surface topography (latitude, longitude, elevation, time), collected along densely-sampled ground tracks which, away from the poles, are separated by distances of km to 10s of km. For missions such as NASA's ICESat and ICESat-2, which were designed to provide repeated coverage of the same surface locations, altimetry data are not

1
2
3 posted at fixed points but are instead stochastically distributed around a predicted ground track
4 (Markus et al.,2017). This presents data visualization and discovery challenges that are not well
5 addressed by current tools for accessing rasterized satellite imagery data.
6

7
8 The initial development of the OA cyberinfrastructure targeted the ICESat dataset (Schutz, et al.,
9 2005), which had had its final reprocessing and represented an intermediate step to the more
10 complex and far larger dataset expected from ICESat-2. Existing tools for browsing ICESat data
11 were limited in several ways. NASA’s Reverb client (EOSDIS, 2009) allowed users to view the
12 coarse spatial coverage of individual ICESat data files but did not allow users to view
13 representations of data values. The ICESat mission’s internal Science Computing Facility (SCF)
14 visualizer was built for specialist users and prioritized granular control of data selection over
15 visualization functionality or ease of use. NASA’s WorldView (Murphy et al., 2015), which is an
16 excellent example of a user-focused data exploration tool, is nevertheless oriented toward raster
17 data and only shows ICESat/ICESat-2 ground tracks.
18
19
20
21

22 We considered the lack of an easy means of traversing the ICESat/ICESat-2 datasets and
23 visualizing their primary observables a missed opportunity for generating interest among new
24 users. More critically, access methods available at the time for ICESat (and planned for ICESat-
25 2) required a high level of familiarity with the data. At the time, few systems existed that were
26 able to serve expert users, who are often the first and most important target of any data
27 distribution system, while still being intuitive enough to accommodate the wider user
28 community.
29
30
31

32 The OpenTopography¹ project funded by the US National Science Foundation has demonstrated
33 that online access to data and analysis tools via easy-to-use interfaces can significantly increase
34 data usage across a wide range of users, from students to academic researchers, to those in the
35 commercial sector (Crosby et al., 2020). OA was therefore designed to democratize access to the
36 ICESat and ICESat-2 datasets by leveraging the cyberinfrastructure knowledge and design
37 philosophy behind the OpenTopography data distribution and processing system using state-of-
38 the-art tools and technologies while incorporating feedback from expert users and stakeholders in
39 the development process.
40
41
42
43

44 3. OpenAltimetry Cyberinfrastructure

45
46 The OA cyberinfrastructure is implemented using a service-oriented architecture composed of a
47 presentation tier, a services tier and a data tier. The data tier utilizes a hybrid solution for data
48 management, featuring a highly optimized PostgreSQL database with tiered storage and a
49 decoupled object-based system that stores data in HDF5 file format. The services or application
50 logic tier consist of programs and scripts that process data and are accessible as Web services
51
52
53
54

55
56 ¹ <https://opentopography.org/>
57
58
59
60
61
62
63
64
65

1
2
3 that can be invoked by applications in the presentation tier. Most users interact with the data via
4 the OA portal in the presentation/applications tier. Several components of this 3-tier architecture
5 are further described below.
6

7 The OA cyberinfrastructure was developed in multiple stages, with an initial objective of making
8 ICESat data access seamless and efficient. Since the ICESat mission ended prior to OA
9 development, the final version of the entire dataset (GLAS/ICESat LIB Global Elevation Data,
10 Version 34) was made available in OA. The ICESat-2 dataset (geolocated photon data and
11 surface-specific elevations) presented several additional challenges that required re-architecting
12 and modifying several components across all three tiers, with the biggest impact on the data tier.
13 First, the ICESat-2 mission is more complex than ICESat, simultaneously collecting data along
14 six ground tracks versus ICESat's single ground track. Second, the ICESat-2 laser fires at 10KHz
15 versus ICESat's 40Hz. Finally, ICESat-2 data is being collected continuously through the
16 mission, while ICESat collected data during several month-long campaigns each year. As a
17 result, ICESat-2 is generating several orders of magnitude more data than ICESat, requiring a
18 presentation and data management system that can handle discovery, access and processing of
19 massive volumes of data, as well as a highly streamlined data ingestion workflow.
20

21 One of the technical goals from onset was to design a service-oriented architecture that was
22 highly modular, to easily accommodate changing requirements. This enabled rapid prototyping
23 and modular development, lowering the time needed for everything from designing the
24 architecture to deploying changes.
25

26 3.1. Cyberinfrastructure Design Philosophy

27 OpenAltimetry was built as a web-based application so that end users would not be required to
28 install or download any applications to their local compute environment. In our experience, most
29 users want to quickly browse satellite datasets in an interactive and easy-to-use interface that
30 provides some ability to download data in their area of interest. A majority of users do not
31 require access to all the information contained in the standard data products generated by the
32 mission. However, they need the ability to conveniently acquire the authoritative source product
33 for their spatial and temporal area of interest if needed. Users also require the ability to quickly
34 visualize elevation profiles for the data products including the waveforms (from ICESat) and
35 photon clouds (from ICESat-2).
36

37 ICESat data are derived from individual laser shots acquired along specified reference
38 ground tracks, where each track represents the target path of the satellite over a single orbit of
39 earth. There are multiple unique tracks, each of which was traversed multiple times during the
40 mission. ICESat data were temporally organized into campaigns, which are demarcated
41 by start and stop times. Each laser shot is identified by a unique identifier and is described by a
42 set of attributes such as the laser footprint position (latitude/longitude/elevation), the shot time,
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3 and the histogram of reflected laser energy (the waveform) that is collected by the instrument.
4 Multiple shots along a track are aggregated into an elevation profile.
5

6 ICESat-2 operates similarly to ICESat, however instead of a single laser beam directed toward
7 the ground track, ICESat-2 has six beams organized into three widely spaced (10s km) sets of
8 closely spaced (100s m) beam pairs. Additionally, ICESat-2 records the time and position of
9 each surface-reflected photon from a laser shot, versus the time vs. energy waveform of
10 ICESat. For ICESat-2, photon data are aggregated into overlapping 40 m track segments, which
11 are somewhat analogous to ICESat footprints. Each segment is identified by a unique identifier
12 and described by a set of attributes (time, position, etc.), similar to that for ICESat. The 10 kHz
13 laser firing rate for ICESat-2 results in photons being returned continuously along the satellite
14 ground track. The aggregate of all collected photons is referred to as the photon cloud, whose
15 shape reveals fine topographic details of the surface.
16
17
18
19

20 OA leverages existing open source software projects for efficiency and reduction in development
21 time compared to building software solutions from scratch. Providing alternate pathways (e.g.
22 APIs) to the data was also deemed important, as we had seen with OpenTopography, since it
23 tends to bring an increase in usage of the data and the development of novel applications
24 independent of the portal environment.
25
26
27

28 While the OA cyberinfrastructure was initially installed and operated on commodity hardware, it
29 was designed to be readily deployable in a cloud environment. That is, it had to have the
30 capability of being migrated to the cloud without major changes or effort in order to potentially
31 leverage the mature commercial cloud offerings including high scalability, load balancing and
32 other optimizations.
33
34

35 Finally, central to OA's architectural design was the ability to enable continuous ingestion and
36 rapid availability of ICESat-2 data from NASA's NSIDC DAAC, as well as the ability to add
37 new data products from other missions in the future.
38
39

40 41 3.2. The OpenAltimetry Web Portal 42

43 Most users of OA interact with the data and visualization features via the application's portal.
44 The portal's data discovery interfaces were designed with NASA's popular EOSDIS Worldview
45 (Murphy, et al., 2015) application as a reference in order to provide a recognizable and
46 minimally disruptive user experience. Unlike the NASA EarthData search tool, which must work
47 with hundreds of different datasets within a single user interface, OA takes a more focused
48 approach to data access for a particular class of Earth-orbiting missions: satellite laser altimeters.
49
50
51

52 In the current implementation of OpenAltimetry, data from the ICESat and ICESat-2 missions
53 have their own map-based data discovery and elevation profile visualization interfaces. This
54 enables us to customize each interface specific to the mission while maintaining a core set of
55
56
57
58
59
60
61
62
63
64
65

1
2
3 functions that are common and applicable across all missions (e.g. base maps, geographic/polar
4 projection views, etc.). The interactive data discovery user interfaces were custom built using the
5 open source OpenLayers mapping library². OA uses ESRI ArcGIS World Imagery service for a
6 base layer in the geographic projection view and the MODIS Blue Marble Next Generation layer
7 for the North and South polar projection views. The ICESat and ICESat-2 data are delivered in
8 geographic coordinates. For mid-latitude views on the map interface, we use the geographic
9 projection (EPSG:4326). For polar views, we do projection conversions on the fly using ESRI
10 Web services for the polar projection views: WGS 84 / NSIDC Sea Ice Polar Stereographic
11 North (EPSG:3413) for the north pole and SWEREF99 15 45(EPSG:3031) for the south pole.
12
13

14
15
16 OA has implemented the NASA EarthData OAuth module which enables users to login using
17 their EarthData credentials. Once logged in, users are provided additional functions including the
18 capability to download the authoritative and original complete data (subsetting) in HDF5 format
19 from the NSIDC DAAC via their REST API. Users who are logged in can also create and share
20 annotations (referenceable areas of interest with a particular date/campaign selected).
21
22

23
24 The primary challenge with ICESat and ICESat-2 data is to provide a visual representation of the
25 laser footprints on the map interface without overwhelming the users with huge data volumes.
26 This must be achieved using commodity compute resources, optimized for web-based access.
27 Showing too much data is not useful to the users, overwhelms the client browser, and quickly
28 consumes limited server-side resources.
29

30
31 OA solves this problem by showing a small percentage of data initially and gradually increasing
32 the amount of data with increasing zoom levels on the map interface. The zoom levels, amount
33 of data shown, and functional capabilities vary by mission as well as products within the
34 mission. The amount of data shown at various zoom levels are also pre-determined based on
35 most common client screen resolution data gathered from web analytics. The common
36 capabilities across all interfaces are the rapid response times, availability of visualization tools
37 for plotting elevation data on demand, and ability to quickly download data from a user-selected
38 temporal and spatial area of interest.
39
40
41

42 43 3.3. ICESat Data Discovery Interfaces: 44 45 46 47 48 49 50 51 52 53 54

55
56 <https://opentopography.org>
57
58
59
60
61
62
63
64
65

Since the ICESat mission is complete, data for the entire mission is displayed in the map interface. In the ICESat data discovery interface, reference ground tracks are displayed after zooming into a predetermined zoom level, at which point a small percentage of the actual ICESat laser footprints start appearing as blue dots (Fig. 1). This percentage increases as the users zoom in to their area of interest, eventually displaying all of the actual footprints. The zoom slider

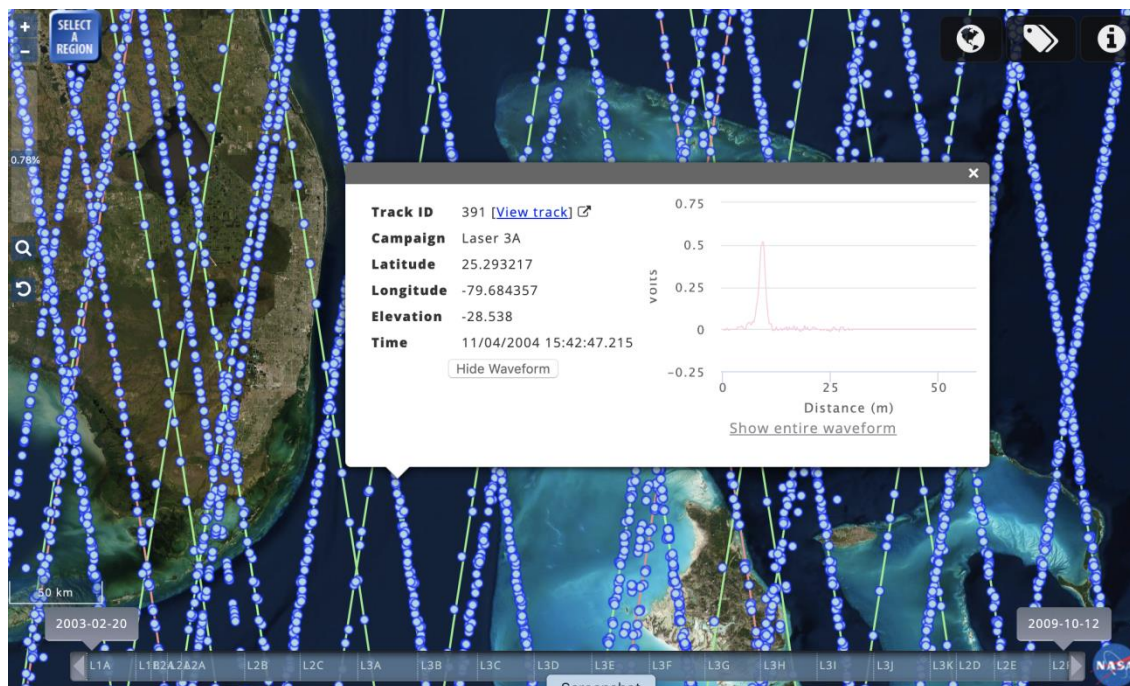


Fig. 1 OpenAltimetry user interface showing ICESat reference ground tracks (green) and location of laser footprints (blue dots). A window displaying information about an individual laser footprint and its associated waveform (see figure) opens when the user clicks on one of the blue dots.

(upper-left of Figure 1) displays the actual percentage of data being shown.

Footprints can be filtered on the fly by reference tracks or by ICESat observation periods (aka laser campaigns; Borsa et al., 2019). At sufficiently high zoom levels, users can draw a bounding box and download basic data attributes (latitude, longitude, elevation, etc.) from the boxed area in comma-separated-value (CSV) format. For levels of zoom that display 100% of the data, the dots corresponding to the footprints are color-coded by campaign (Fig. 2).

Users have the ability to view individual tracks and their laser footprints, which are color coded to represent their campaigns (Fig. 2). This is useful because the actual tracks of laser footprints do not all fall precisely the reference ground track due to the imprecisions in the orientation of the spacecraft. Users also have the ability to click on individual footprints to bring up a window with additional footprint metadata, including track ID, laser footprint position

(latitude/longitude), surface elevation, campaign and shot time (Fig 1). This window also shows the energy waveform (i.e. the profile of returned energy) for the footprint.

Finally, when users elect to view data from an individual track, they have the ability to select a region and view the elevation profile for a single or multiple campaigns for the track (Figure 2). Additional options are also available for viewing combined and individual waveforms of the laser footprints, along with the ability to download the presentation-quality plots as well as the underlying data. For all the on-the-fly visualizations and plots in OA, we use the open source (non-commercial use) Highcharts Javascript³, an SVG-based multi-platform charting library.

3.4. ICESat-2 Data Discovery Interfaces:

The ICESat-2 data discovery interface has a map-based interface similar to that for ICESat (Fig. 3). A small percentage of ICESat-2 segments are shown initially, increasing to 100% as the user zooms into an area of interest. ICESat-2 segments shown in OA vary with the data product from which they are derived. Segment elevations are provided for land ice height (the ATL06 product), sea ice height (ATL07), land and vegetation heights (ATL08), sea ice freeboard (ATL10), ocean surface height (ATL12) and inland water body heights (ATL13). The segments in these datasets have different spatial coverage and can yield different elevations. Documentation on all ICESat and ICESat-2 standard products are available from NSIDC⁴. Additionally, elevations are provided for the geolocated photons (ATL03) that serve as the raw data for the other data products.

Since the ICESat-2 mission is ongoing, with high volumes of data being continuously collected, we display ICESat-2 data in one day increments in the map interface. When users initially access the ICESat-2 user interface, data from the latest available date is shown for a default data product, which is currently ATL08. Users can select other data products at the top-left of the OA window, or other dates using the calendar in the sidebar (Fig 3). There are also controls for selecting particular tracks and individual beams. When a user requests a plot of the elevations from a surface-specific product, they are also given the option of viewing the photon data from which the product was derived (Fig. 4).

³ <https://www.highcharts.com/>

⁴ <https://nsidc.org/data/icesat-2>

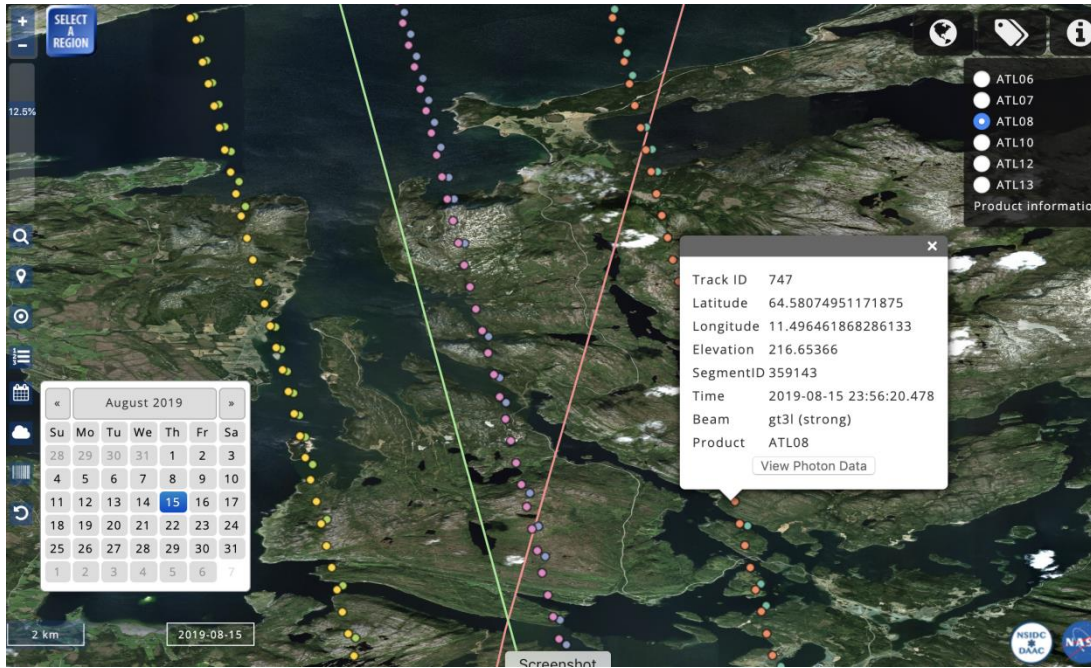


Fig. 3 OpenAltimetry display of ICESat-2 laser segments along a reference ground track (green line). Each of the six laser beams appear in different colors. When a user clicks on a segment, a window appears displaying segment information. The pink line in the figure shows a different reference ground track.

The percentage of total segments shown at different zoom levels is pre-determined based on performance testing that considers the most popular screen resolutions from web analytics data. It also varies by data product, since segment spacing can vary by product. Users have the ability to click on individual segments to reveal additional metadata, including its position (latitude/longitude/elevation), ID, beam number and track number. Users also have the ability to visualize the corresponding photon data on the fly for that segment (photon height vs latitude) (Fig. 4).

At a particular zoom level, users have the ability to draw a bounding box and view elevation profiles within the selected area. This opens up a new browser window that displays the elevation profiles of the selected data products within that area. In case of ATL08 canopy heights are also displayed and in the case of ATL10 sea ice freeboard is displayed in addition to ice surface elevation. Plots of the return signal photons (Fig. 4) are also available in the new browser

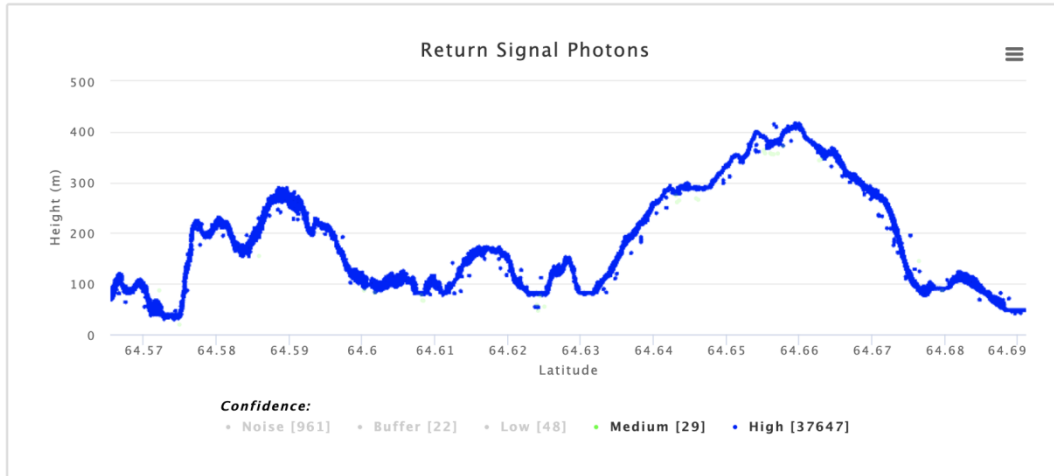
1
2
3 window. Since the number of photons can be extremely large, only a sampling of the photons is
4 typically displayed. Users can optionally request to visualize all photon data.
5
6

7 Select ATLAS beam [gt3r \(weak\)](#) | [gt3l \(strong\)](#) | [gt2r \(weak\)](#) | [gt2l \(strong\)](#) | [gt1r \(weak\)](#) | [gt1l \(strong\)](#)

8
9 Track ID: 747 - Beam: gt3l - Sample Rate: 100%

10 Total number of photons: 38,707 - Total segments: 704 - Segment range: [359,058 - 359,761]

11
12 Drag zoom on the plot below to view more detail.



32 **Fig. 4** Display of photon elevations from ICESat-2 for one beam along track 747. Points are color coded based
33 on confidence level. High and medium photons are displayed by default but a user can select lower confidence
34 levels by clicking on the labels.
35

36 Similar to the case with the ICESat interface, users have the ability to download data as comma-
37 separated values (CSV) for any of the available mission data products, subsetted by the spatial
38 and temporal area of interest. If logged in using EarthData login, users are allowed to download
39 the original HDF5 file from NSIDC DAAC via its REST API, subsetted to the user's area of
40 interest.
41
42

43 Another powerful capability of the OA portal is the ability for users to add annotations to the
44 datasets. Users can tag specific data views (spatial and temporal) as a persistent URL for easy
45 sharing and collaboration. Annotations also provide the capability to leverage community input
46 in identifying areas of interest and anomalies in the data. Finally, we have also integrated
47 MODIS daily true-color surface product from NASA GIBS (Murphy, et.al. 2015) to allow users
48 to see cloud and surface conditions for the day of any ICESat-2 acquisition. Users simply toggle
49 a satellite imagery basemap from a button in the sidebar, eliminating the need to access NASA
50 Worldview to view these images.
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Capture of user metrics and usage analytics are a key component of OA and have been built into the system from the very onset. These include details on spatial queries performed, data downloads, the number of elevation and photon visualization plots generated, including breakouts by individual products, the number of callbacks to the original HDF5 datasets at NSIDC and spatiotemporal “hot zones” of high data usage, among others. User metrics and usage analytics are vital not just for both architecture design and optimization but also as a measurement of success.

3.5. OA Data Management System

OA utilizes a hybrid solution for data management with a highly optimized PostgreSQL database featuring PostGIS for geographic objects support on tiered storage and a decoupled object-based storage system for storing HDF5 files. The ICESat waveform energy data files and ICESat-2 photon data files are stored in their original HDF5 format. We ingest only the few data elements from the original ICESat-2 HDF5 files needed for OA visualization. This dramatically shortens the data ingestion process, enabling rapid turnaround of newly available data. An added advantage of decoupling the object storage is the ease of porting this data management system to cloud distributed resources for scalability.

The core software component of OA that makes this storage strategy possible is the open source JHDF5 (HDF5 for Java), a Java binding for HDF5⁵, that extracts waveform and photon data from the HDF5 files on the fly. JHDF5 gives us the capability to fetch specific blocks of the data based on their index, unlike most other libraries (e.g. HDFQL, official HDF5 java), which require loading entire data into the memory and then extracting the needed data. This significantly improves memory usage and performance, vital for rapid response times in the application.

In order to further speed up response times for low zoom levels, we extract a decimated subset of the data products and place them into the PostgreSQL database. Sampling rates vary by product density, ranging from 1 in 2048 for ATL06 and ALT08 to 1 in 128 for ATL12. This eliminates the need to read multiple HDF5 files for showing data footprint locations at global zoom levels. OA performance is thereby improved while adding minimal overhead to the data loading process and minimal additional storage space.

3.6. ICESat-2 data pipeline from NSIDC

OA has a data pipeline established with the NSIDC DAAC to ingest the geolocated photon data product (ATL03) and other surface products (ATL06, ATL07, ATL08, ATL10, ATL12, ATL13)

⁵ <https://wiki-bsse.ethz.ch/pages/viewpage.action?pageId=26609113>

as they become available at NSIDC. The data pipeline is implemented by querying the NSIDC API service with the data product, temporal range and subset of data elements required. The NSIDC service endpoint utilizes URS authentication for their EOSDIS Service Interface results, which satisfies NASA's requirement to collect metrics on all data access and downloads. Once the request is received, the NSIDC services generate subsetted files for download. These data are then transferred and ingested into the OA data management system using several extract-transform-load (ETL) scripts. The data pipeline with NSIDC is further optimized by parallelizing the data pulls using asynchronous mode. This reduces the overall ETL time at least by half (this varies according to the computational demand at NSIDC) by simultaneously downloading different sections of the dataset in parallel.

3.7. The OpenAltimetry API

We have implemented application programming interfaces (APIs) to allow for programmatic access to all data products available in OA. API-based access to OA data products provide a flexible platform for researchers to design complex data access and processing workflows independent of the processing tools available within the OA portal.

APIs in OA are implemented using the open source Jersey RESTful Web Services framework⁶. The open source Swagger framework is also used for documenting and consuming these RESTful services⁷. We have also enabled pre-formatted API endpoints for every elevation plot, where users can conveniently copy and paste the API endpoint URL into their application and pull corresponding data with the subsetted spatial and temporal parameters.

Data outputs are available in multiple formats (including CSV and json) to boost the use of OA data directly in third party applications such as Jupyter notebooks, as well as to promote development of independent algorithms and functionality. To make it more convenient for users,

```

import pandas as pd
import json
import requests
import plotly.graph_objects as go
from plotly.offline import iplot

# Paste the OpenAltimetry API URL for Photon here:
OA_API_URL = 'https://openaltimetry.org/data/api/icesat2/atl103?minx=47.4

# Select the list of confidence to display: 'Noise', 'Buffer', 'Low', 'Me
OA_PHOTON_CONFIDENCE = ['Noise', 'Buffer', 'Low', 'Medium', 'High']

df_beams = []
oa_plots = []

# This function will request the 6 beams data using OpenAltimetry's API
def getPhotonData():
    series = []
    b_url = OA_API_URL + '&client=jupyter'

```

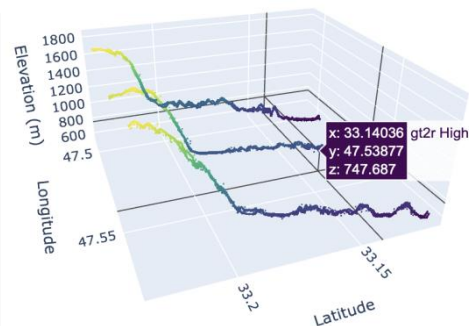


Fig. 5 Segment of Jupyter notebook that is linked from the plot pages of OpenAltimetry (left), and a view of the 3-D interactive display of photon elevation data that results when the code is executed.

⁶ <https://jersey.github.io/>

⁷ <https://openaltimetry.org/data/swagger-ui/>

1
2
3 we provide relevant API endpoint URLs under the elevation and photon plots in the user
4 interfaces, as well as links to sample Jupyter Notebooks (Fig. 5) that make use of these endpoints
5 to pull in data.
6

7 8 9 **4. Community Response and Adoption**

10
11 From its beginning, the OpenAltimetry project had a direct links to the ICESat-2 mission via two
12 of the OA Principal Investigators (PIs), one of whom was also a member of the ICESat-2 science
13 team. Our strategy was to engage with expert data users on the science team from the outset, to
14 understand their needs and to ensure that we were 1) aware of the nuances of the ICESat-2
15 mission and data products, and 2) informed of major developments impacting the ICESat-2
16 dataset. The dialogue took place in the form of in-person presentations by the OA PIs to the
17 science team as the interface was being developed for the original ICESat dataset, and occasional
18 updates at science team teleconferences and via email with interested science team members.
19 Since most of the ICESat-2 science team had extensive experience with ICESat, they were able
20 to inform our development efforts even before ICESat-2 was launched. Our development plan
21 was, in fact, predicated on implementing ideas and improvements from an expert cohort before
22 we expanded OA access to the broader community, and this proved to be an effective strategy
23 for building the early version of the OA interface.
24
25
26
27
28
29

30 One of the initial signs that we were on the right track with this project was the high level of
31 enthusiasm and support among the science team for OA, which has continued up to the present.
32 In parallel with our science team engagement, the OA team and NSIDC staff also conducted
33 outreach with the broader community of non-expert users of ICESat and ICESat-2 data. Our
34 initial efforts took place in partnership with the ICESat-2 Applications Program, wherein OA
35 quickly became the de facto interface for introducing new users to the ICESat-2 dataset due to its
36 rapid data visualization capabilities and its ease of use. Enthusiasm from this different cohort
37 was an independent sign that we were achieving our original design goal to serve users across all
38 levels of expertise. Since the launch of ICESat-2, OA is increasingly featured as a central point
39 of data access for our community. Examples include the use of OA in NSIDC's presentations at
40 town halls and workshops, in the annual ICESat-2 HackWeek⁸, and in presentations of PI-led
41 research at scientific conferences and in social media. The AGU 2019 Fall meeting was likely
42 the largest gathering of ICESat-2 data users since the data were publicly released. OA project
43 members and NSIDC DAAC staff all noted unsolicited and repeated references to OA in poster
44 sessions, town hall meetings and other events.
45
46
47
48
49
50
51
52
53

54
55 ⁸ <https://github.com/ICESAT-2HackWeek>
56
57
58
59
60
61
62
63
64
65

An important metric of success for any project is consistent growth in usage. Since its initial deployment, OA has been growing organically, with dramatically elevated usage after ICESat-2 data distribution commenced in March of 2019 (Fig. 6).

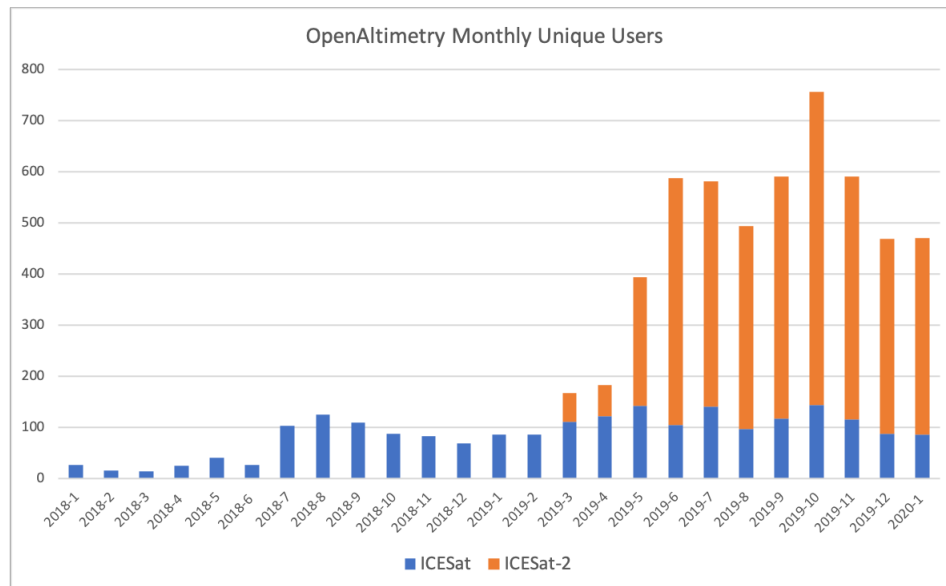


Fig. 6 Count of users each month since OpenAltimetry was released, with uses of the ICESat interface in blue and uses of the ICESat-2 interface in orange.

5. Discussion and Conclusions

The most compelling lessons that have been learned through the course of this project are that: 1) novel and complex datasets need tools that provide easy access to and visualization of those data and, more importantly, 2) development must be responsive to user needs and desires from the start. Having the opportunity to demonstrate early versions of OA to the scientists who were developing the ICESat-2 mission and its derived data products was also of great benefit.

OpenAltimetry's success can be attributed to the fact that it enables rapid and responsive exploration at a level of detail in the data that no other tool could provide. This created new use cases and paved the way for new users of the data.

Getting promising technologies into sustained operations is a common challenge. A disturbingly high percentage of cyberinfrastructure projects fail to attract a large enough user community to justify long-term sustainment. A report from the US National Science Foundation (Atkins, et al. 2003) suggests that sociological and culture barriers to technology adoption may be a root cause. What is often referred to as the "Valley of Death" (Mcintyre, 2014) in business, i.e. the gap

1
2
3 between basic research and commercialization of a new product, is, in the context of projects
4 funded with public money, such as NASA's investments via ACCESS, the gap between having a
5 working prototype and garnering a large enough of a demand for the service it provides that a
6 NASA data center is convinced of the value to its users and is willing to support it in the long
7 term.
8
9

10 OpenAltimetry appears to have succeeded in crossing the Valley of Death. NASA is working to
11 make OpenAltimetry an operational capability within the EOSDIS environment. As a first step
12 we have enabled a seamless handoff from NASA's EarthData Search Client to OA, whereby a
13 user's product selection, spatial and temporal bounds and preferred map projection are all passed
14 to OA and the users is able to immediately begin exploring ICESat-2 data in their area of
15 interest.
16
17
18

19 Based on the response to OpenAltimetry, we expect demand to grow for the inclusion of
20 additional data products. Potential candidates include data products from the Global Ecosystem
21 Dynamics Investigation (GEDI), the Airborne Topographic Mapper (ATM) from Operation
22 IceBridge, and the Land Vegetation and Ice Sensor (LVIS) Facility instrument from NASA
23 Goddard Space Flight Center.
24
25
26

27 6. References

28

- 29
30 Atkins, D., Droegemeier, K., Feldman, S., Garcia-Molina, H., Klein, M., Messerschmitt, D.,
31 Messina, P., Ostriker, J., Wright, M. (2003). Revolutionizing Science and Engineering
32 Through Cyberinfrastructure. Report of the National Science Foundation Blue-Ribbon
33 Advisory Panel on Cyberinfrastructure, January.
34
35
36 Borsa, A.A., Fricker, H.A. and Brunt, K.M. (2019). A Terrestrial Validation of ICESat Elevation
37 Measurements and Implications for Global Reanalyses. IEEE Transactions on
38 Geoscience and Remote Sensing, 57(9), pp.6946-6959.
39
40
41 Earth Observing System Data and Information System (EOSDIS).(2009). Earth Observing
42 System ClearingHouse (ECHO) / Reverb, [retired online application]. Greenbelt, MD:
43 EOSDIS, Goddard Space Flight Center (GSFC) National Aeronautics and Space
44 Administration (NASA). URL: <http://reverb.earthdata.nasa.gov>
45
46
47 Crosby, C.J., Arrowsmith, J.R. and Nandigam, V. (2020). Zero to a trillion: Advancing Earth
48 surface process studies with open access to high-resolution topography. In *Developments*
49 *in Earth Surface Processes* (Vol. 23, pp. 317-338). Elsevier.
50
51
52 Kavvada, A., Metternicht, G., Kerblat, F., Mudau, N., Haldorson, M., Laldaparsad, S., Friedl, L.,
53 Held, A., Chuvieco, E. (2020). Towards delivering on the sustainable development goals
54 using earth observations. Remote Sensing of Environment 247, 111930.
55 <https://doi.org/10.1016/j.rse.2020.111930>
56
57
58
59
60
61
62
63
64
65

1
2
3
4 Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker,
5 H., Gardner, A., Harding, D. and Jasinski, M. (2017). The Ice, Cloud, and land Elevation
6 Satellite-2 (ICESat-2): science requirements, concept, and implementation. *Remote*
7 *Sensing of Environment*, 190, pp.260-273.
8

9
10 Mcintyre, R. A. (2014). Overcoming “The Valley of Death.” *Science Progress*, 97(3), 234–248.
11 <https://doi.org/10.3184/003685014X14079421402720>
12

13
14 Murphy, K.J., Davies, D.K., Michael, K., Justice, C.O., Schmaltz, J.E., Boller, R., McLemore,
15 B.D., Ding, F., Vollmer, B. and Wong, M.M. (2015). LANCE, NASA’s land, atmosphere
16 near real-time capability for EOS. In *Time-Sensitive Remote Sensing* (pp. 113-127).
17 Springer, New York, NY.
18

19
20 Schutz, B. E., H. J. Zwally, C. A. Shuman, D. Hancock, and J. P. DiMarzio (2005). Overview of
21 the ICESat Mission. *Geophysical Research Letters*, Vol. 32, L21S01,
22 doi:10.1029/2005GL024009.
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

FIGURES

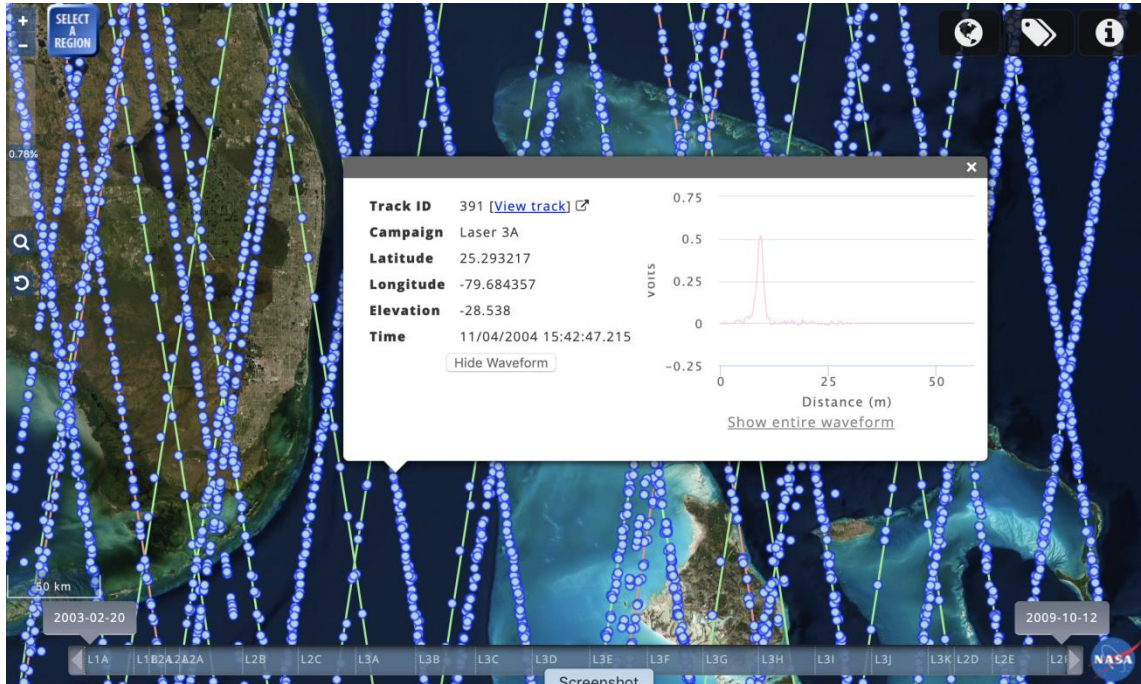


Fig. 1 OpenAltimetry user interface showing ICESat reference ground tracks (green) and location of laser footprints (blue dots). The embedded window displaying information about an individual shot pops up when the user clicks on one of the blue dots.

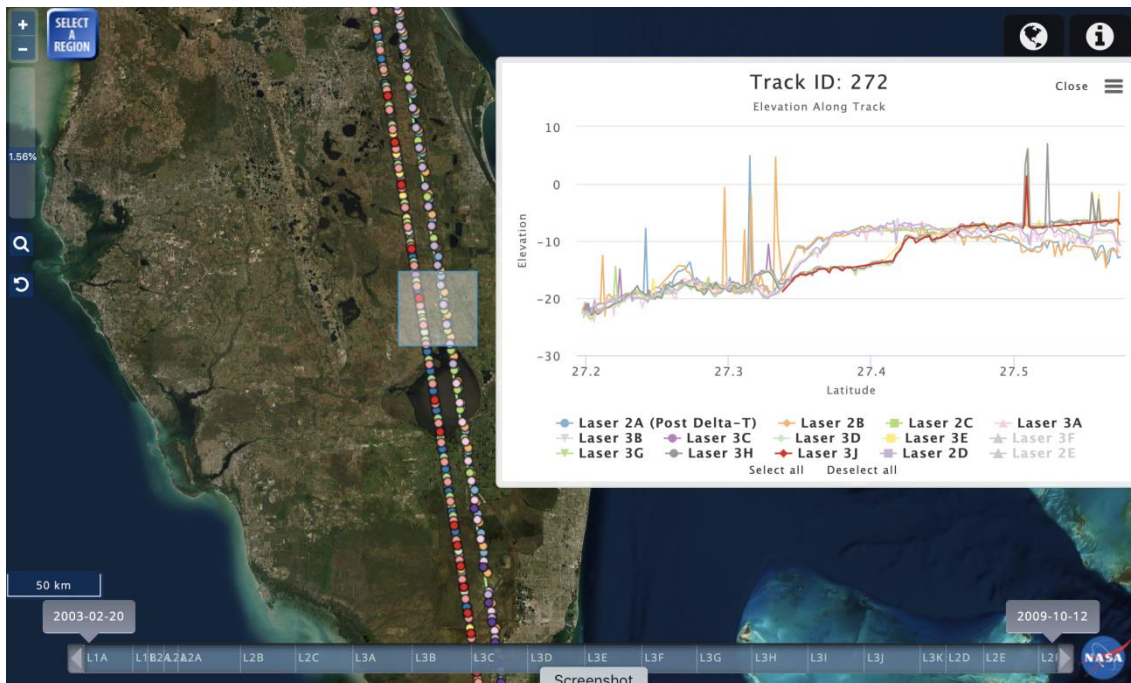


Fig. 2 OpenAltimetry display of ICESat laser shots along a single reference ground track, along with an elevation profile derived from all campaigns within the highlighted region. Campaigns are labeled at the bottom of the plot, and can be individually turned on and off in the display. Time slider at the bottom of the window allows restriction of data displayed by campaign.

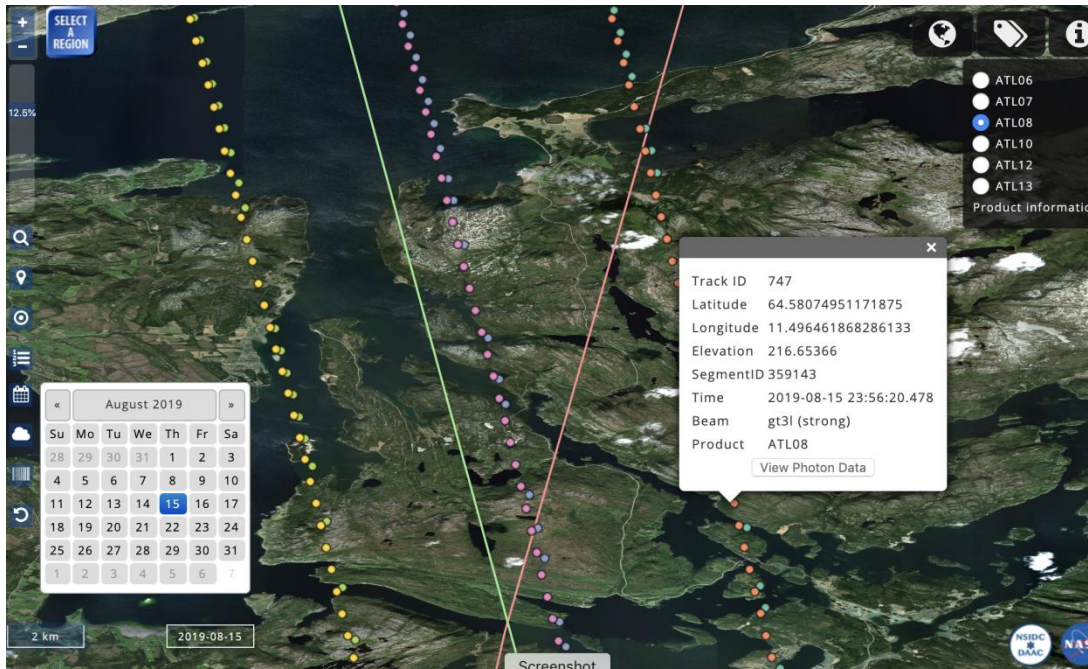


Fig. 3 OpenAltimetry display of ICESat-2 laser footprints along a reference ground track (green line). Each of the six laser beams appear in different colors. When a user clicks on a footprint a window appears displaying information about that shot. The pink line shows a different reference ground track.

Select ATLAS beam [gt3r \(weak\)](#) | [gt3l \(strong\)](#) | [gt2r \(weak\)](#) | [gt2l \(strong\)](#) | [gt1r \(weak\)](#) | [gt1l \(strong\)](#)

Track ID: 747 - Beam: gt3l - Sample Rate: 100%

Total number of photons: 38,707 - Total segments: 704 - Segment range: [359,058 - 359,761]

Drag zoom on the plot below to view more detail.

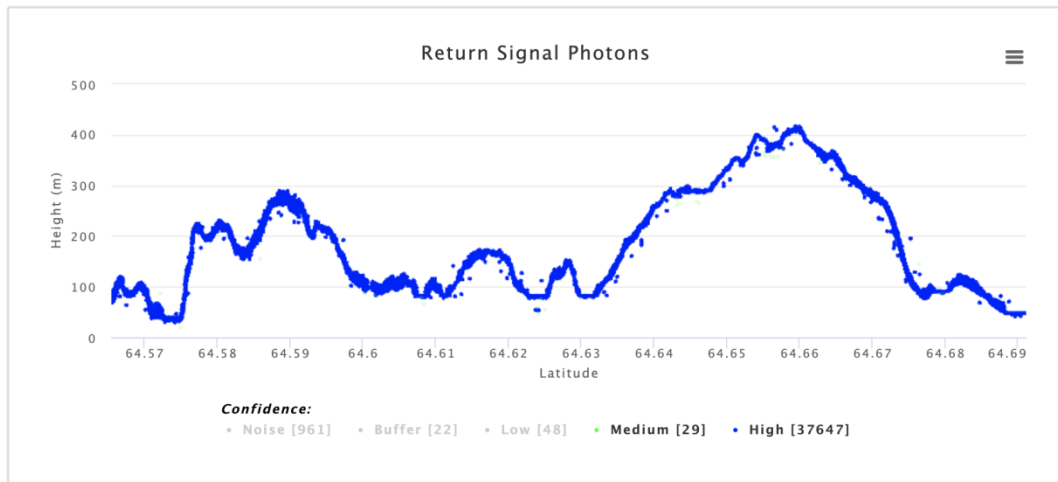


Fig. 4 Display of photon elevations from ICESat-2 for one beam along track 747. Points are color coded based on confidence level. High and medium photons are displayed by default but a user can select lower confidence levels by clicking on the labels.

```
import pandas as pd
import json
import requests
import plotly.graph_objects as go
from plotly.offline import iplot

# Paste the OpenAltimetry API URL for Photon here:
OA_API_URL = 'https://openaltimetry.org/data/api/icesat2/at103?&minx=47.4

# Select the list of confidence to display: 'Noise', 'Buffer', 'Low', 'Me
OA_PHOTON_CONFIDENCE = ['Noise', 'Buffer', 'Low', 'Medium', 'High']

df_beams = []
oa_plots = []

# This function will request the 6 beams data using OpenAltimetry's API
def getPhotonData():
    series = []
    b_url = OA_API_URL + '&client=jupyter'
```

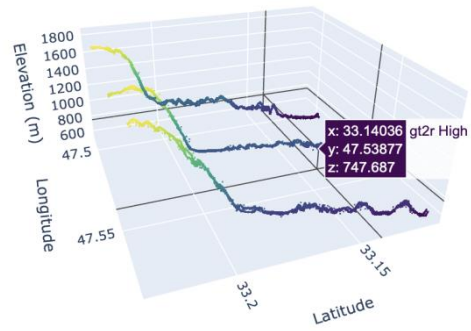


Fig. 5 Segment of Jupyter notebook that is linked from the plot pages of OpenAltimetry (left), and a view of the 3-D interactive display of photon elevation data that results when the code is executed.

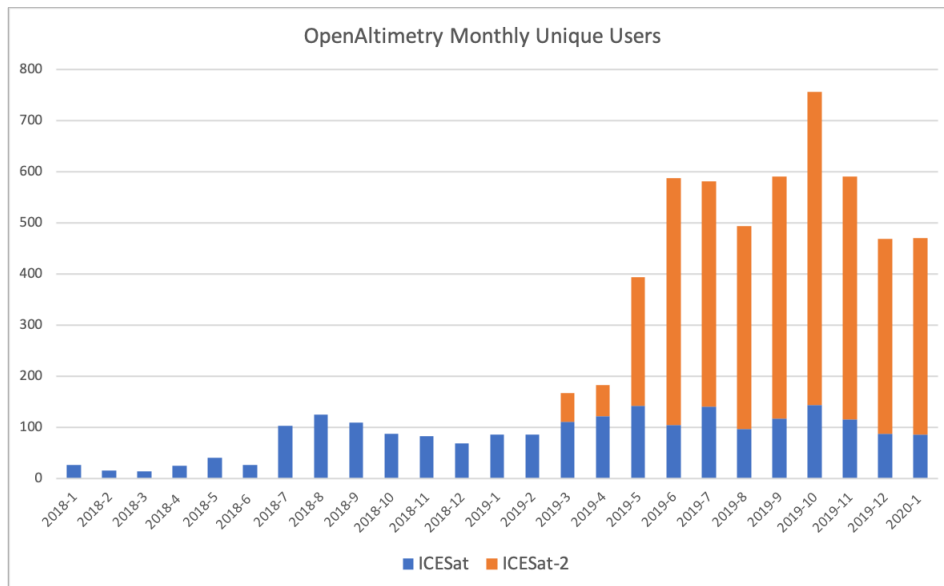


Fig. 6 Count of users each month since OpenAltimetry was released, with uses of the ICESat interface in blue and uses of the ICESat-2 interface in orange.