ICESat-2 data access and analysis with icepyx

GitHub: https://github.com/icesat2py/icepyx
Documentation: https://icepyx.readthedocs.io

Jessica Scheick
University of New Hampshire

Earthdata Webinar
19 November 2025
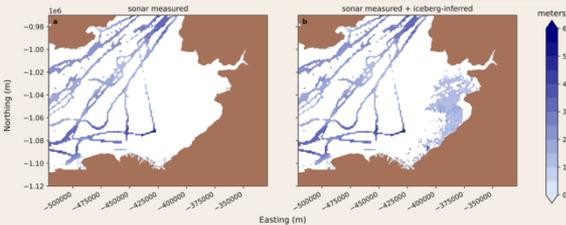
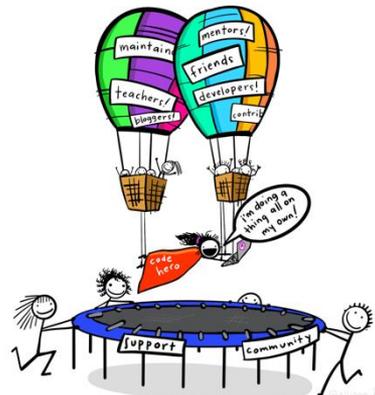# Who am I?

Glaciologist
(icebergs + ice-ocean interactions

Open-source Scientific Software Developer

Open Science Advocate + Educator

↔ ↔

Artwork by @allison_horst

icepyx: ICESat-2⁺ data tool
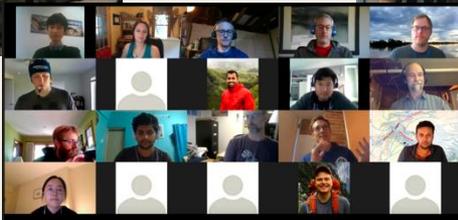
❖ *A community* of ICESat-2 data users, developers, data managers, and researchers

❖ An open-source Python *software library*

icepyx Origin Story: 2019 ICESat-2 Hackweek

A hackweek is a welcoming learning environment designed to build an *open and collaborative research community* while introducing participants to new software tools

A critical community tool

Client Comparison: Granules & Users

Note: OpenAltimetry granule count (2,688) is too small to display relative to other clients.

Harmony Metrics
1 Oct 2024 – 23 Sept 2025

Percentage of all On-Premises Downloads Using icepyx

On-premise Metrics
Feb 2021 – Oct 2024

Success Stories

I was able to bring on a master's student to work with ICESat-2 data because they wouldn't need an entire year just to learn how to access the data.
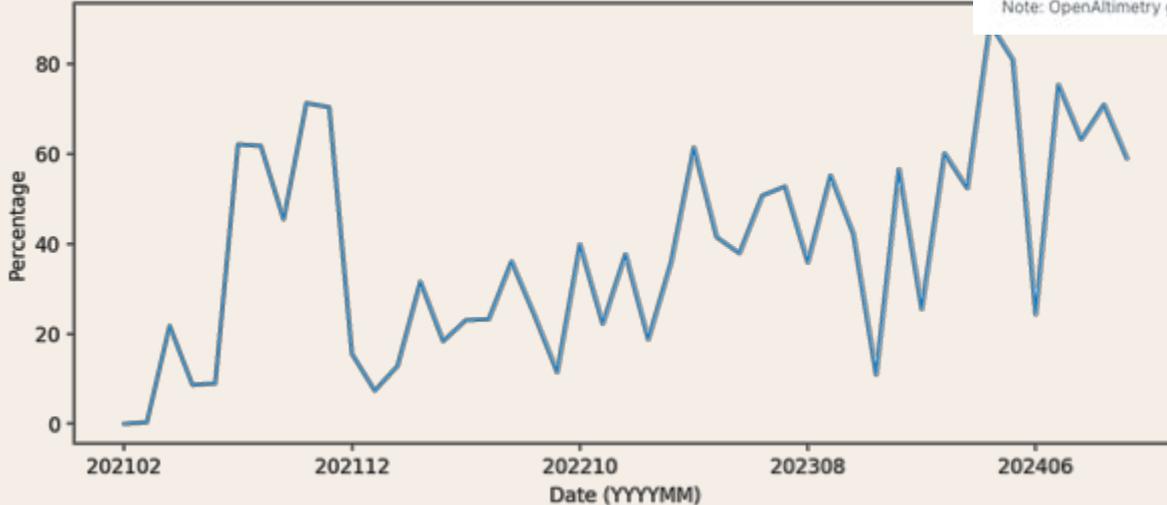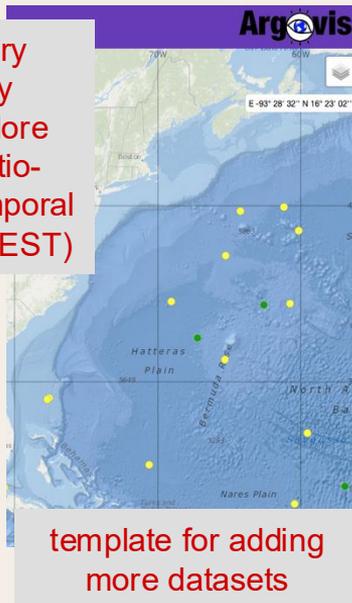
~ professor (cryosphere)

Query
Unify
Explore
Spatio-**T**emporal
(QUEST)

E -93° 28' 32" N 16° 23' 02"

Hatteras Plain

Bermuda

Nares Plain

template for adding more datasets

icepyx made it possible for me to understand how to work with ICESat-2 datasets. The other information out there was overwhelming.

~ software developer

Over 20 publications and presentations* using or featuring icepyx (*excludes tutorials)

Spatially Continuous Mapping of Forest Canopy Height in Canada by Combining GEDI and ICESat-2 with PALSAR and Sentinel

by Camile Sothe, Alemu Gonsamo, Ricardo B. Lourenço, Werner A. Kurz and James Snider

[1] School of Earth, Environment & Society, McMaster University, Hamilton, ON L8S 4L8, Canada
[2] Canadian Forest Service, Natural Resources Canada, Victoria, BC V8Z 1M5, Canada
[3] World Wildlife Fund Canada, Toronto, ON M5V 1S8, Canada
* Author to whom correspondence should be addressed.

Remote Sens. 2022, 14(20), 5158; https://doi.org/10.3390/rs14205158

Mapping the Diversity of Agricultural Systems in the Cuellaje Sector, Cotacachi, Ecuador Using ATL08 for the ICESat-2 Mission and Machine Learning Techniques

Garrido Fernando

Conference paper | First Online: 10 September 2021
828 Accesses | 1 Citations
Part of the Lecture Notes in Computer Science book series (LNTCS,volume 12957)

Geophysical Research Letters

Research Letter | Open Access | CC (i) (S)

How Are Under Ice Phytoplankton Related to Sea Ice in the Southern Ocean?

K. M. Bisson, B. B. Cael

First published: 21 October 2021 | https://doi.org/10.1029/2021GL095051 | Citations: 1

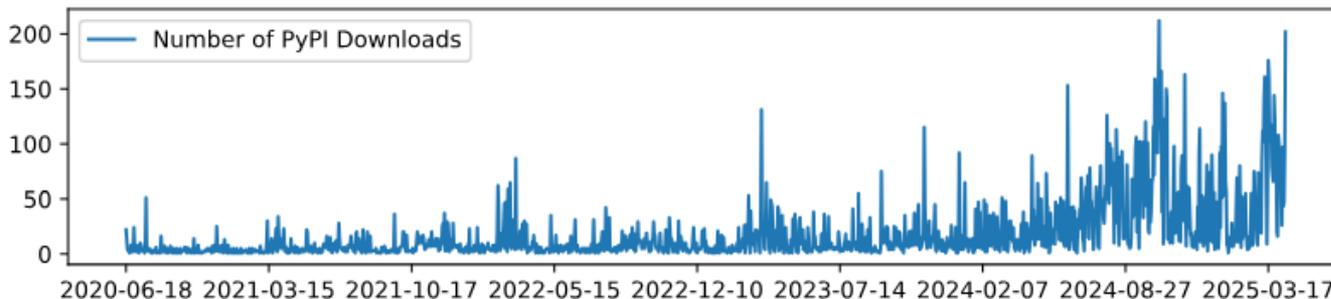Mapping the grounding zone of Larsen C Ice Shelf, Antarctica, from ICESat-2 laser altimetry

Tian Li, Geoffrey J. Dawson, Stephen J. Chuter, and Jonathan L. Bamber
Bristol Glaciology Centre, School of Geographical Sciences, University of Bristol, Bristol, BS8 1SS, UK

Correspondence: Tian Li (tian.li@bristol.ac.uk)

Received: 10 Apr 2020 – Discussion started: 25 May 2020 – Revised: 24 Aug 2020 – Accepted: 25 Sep 2020 – Published: 02 Nov 2020

Number of PyPI Downloads
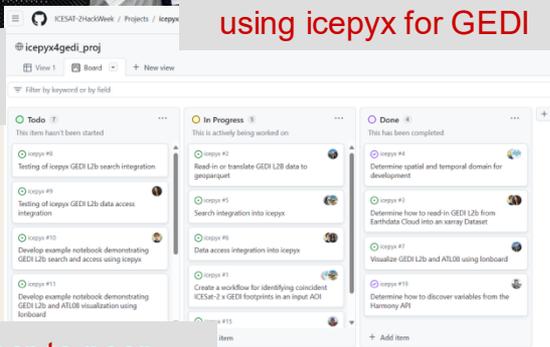
200
150
100
50
0

2020-06-18 2021-03-15 2021-10-17 2022-05-15 2022-12-10 2023-07-14 2024-02-07 2024-08-27 2025-03-17

Support all user types

Examples

Documentation and ICESat-2 Resources

Support

Easy Install

Software best practices and UML diagrams

Underlying tool development and community collaborations

icepyx: in the community

Hackweek Project

using icepyx for GEDI

Problem solving with DAAC data engineers and managers

Improved documentation

Tutorials and Cloud Computing

Practicing GitHub

Peer-to-peer teaching + learning

New projects are OPEN

INV51D-01 On a QUEST (Query, Unify, Explore SpatioTemporal) to Accelerate ICESat-2 Applications in Ocean Science via icepyx

Presentations and publications

Friday, 15 December 2023
08:30 - 08:40
301-302 - South (Level 3, South, MC)

Abstract

Collaboration and inclusivity will be critical for tackling modern environmental challenges resulting from increased anthropogenic stressors. Even as we celebrate 2023 as the Year of Open Source Science, many scientific disciplines - and thus their technologies, tools, and expertise - remain siloed. The ICESat-2 (Ice

icepyx in the ICESat-2 tool ecosystem: a confusing landscape
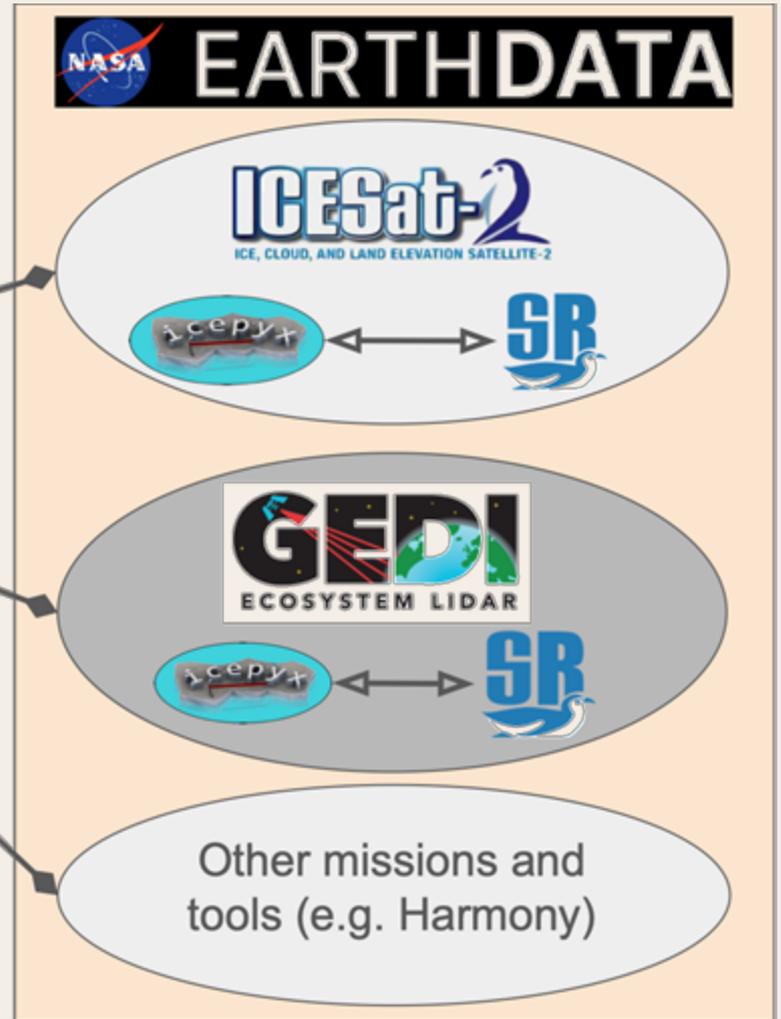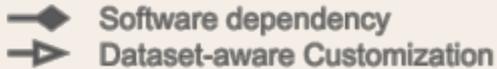
Extensible communities and tooling

icepyx: feature summary (part 1)

Query
Unify
Explore
Spatio-
Temporal
(QUEST)

Documentation and ICESat-2 Resources

Visualization

Adding a new dataset to QUEST

Step 1. Create new branch off development on icepyx

Step 2. Import libraries for analysis

```
from icepyx.quest.dataset_scripts.dataset import DataSet
from icepyx.core.spatial import geodataframe
import requests
import pandas as pd
import os
import numpy as np
```

icesat2py / icepyx Public

<> Code ⊙ Issues 49 ⏱ Pull requests

ᵖ development ▾ ᵖ 30 branches

Find or create a branch...

Branches Tags

✓ development default

Step 3. Create class name for the variable of interest

```
class Argo(DataSet):
    """
    Initialises an Argo Dataset object
    Used to query physical Argo profiles
```

Step 4. Initialize dataset with bounding box and timeframe

```
def __init__(self, boundingbox, timeframe):
    super().__init__(boundingbox, timeframe)
    assert self._spatial._ext_type == "bounding_box"
    self.profiles = None
```

Step 5. Add code specific to API query for your dataset

```
# builds URL to be submitted
baseURL = "https://argovis-api.colorado.edu/argo"
payload = {
    "startDate": self._temporal._start.strftime("%Y-%m-%dT%H:%M:%S.%fZ"),
    "endDate": self._temporal._end.strftime("%Y-%m-%dT%H:%M:%S.%fZ"),
    "polygon": [self._fmt_coordinates()],
}
```
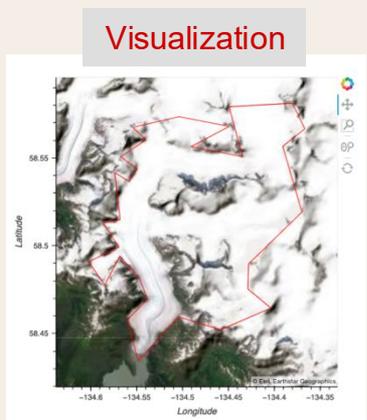
Step 6. Add quality control code specific to your dataset

Step 7. Add functionality to format coordinates from the icepyx call to the API for data
    (e.g., convert spatial extent into string format needed by API server)

Step 8. Add functionality to parse data into dataframe

icepyx

icepyx version 2.0.0 JOSS 10.21105/joss.04912 DOI 10.5281/zenodo.7729

Python tools for obtaining and working with ICESat-2 data

Quick Links: Installation | Citation | Examples | Source Code | Contact

icepyx is both a software library and a community composed of ICESat-2 data users, developers, and the scientific community. We are working together to develop a shared library of resources - including existing resources, new code, tutorials, and use-cases/examples - that simplify the process of querying, obtaining, analyzing, and manipulating ICESat-2 datasets to enable scientific discovery. To further enhance data discovery, we have developed the QUEST module to facilitate querying of ICESat-2 data and complimentary Argo oceanographic data, with additional dataset support expected in the future.

GETTING STARTED
Origin and Purpose
Installation
Citing icepyx

EXAMPLES
Accessing ICESat-2 Data
Subsetting ICESat-2 Data
ICESat-2's Nested Variables
Visualizing ICESat-2 Elevations
Reading ICESat-2 Data in for Analysis
ICESat-2 AWS cloud data access
QUEST Example: Finding Argo and ICESat-2 data

USER GUIDE

Getting Started   User Guide   QUEST

Meet open science and software requirements

template for adding more datasets

Easy Install

Anaconda.org 2.0.0   pypi package 2.0.0

Download and Cloud data access

Minimal Code

```
import icepyx as ipx
region = ipx.Query(short_name, spatial_extent, date_range)
region.download_granules(path)
```

earth access

earthaccess
A Python Library for NASA Earthdata

Easy authentication and token management

icepyx: feature summary (part 2)

Examples

Peer-reviewed, citable software

JOSS 10.21105/joss.04912

Easily make your workflow reproducible or transition to the cloud

Under-the-hood transition to Harmony services

Support

Community

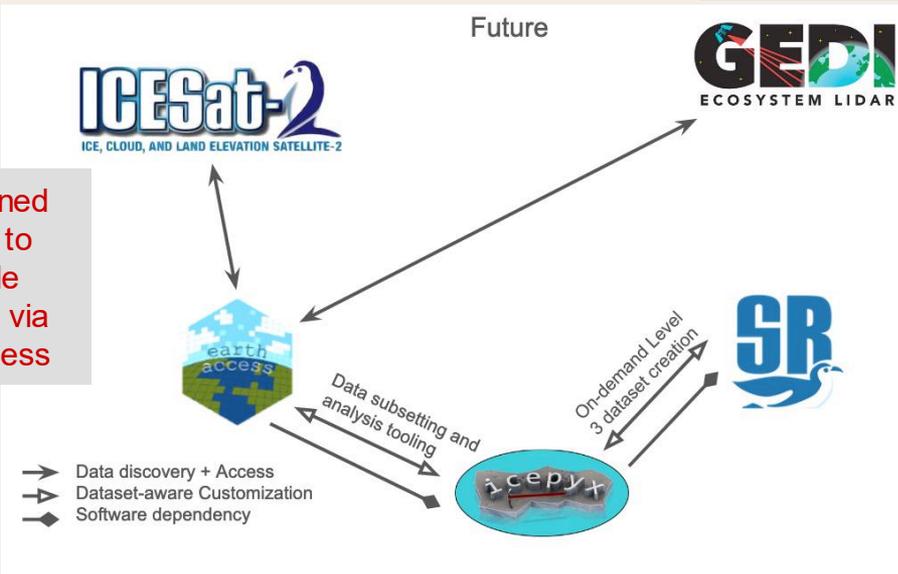Data read-in

Nested variable management (can be used for ANY hdf5 files)

icepyx: ongoing

Subsetting of cloud data

Optimized Cloud data read-in

Streamlined access to multiple datasets via earthaccess

Future

ICESat-2
ICE, CLOUD, AND LAND ELEVATION SATELLITE-2

GEDI
ECOSYSTEM LIDAR

SR

Data subsetting and analysis tooling

On-demand Level 3 dataset creation

icepyx

→ Data discovery + Access
→ Dataset-aware Customization
➜ Software dependency

Handling of QuickLook data products

xarray.Dataset

▸ Dimensions: (gran_idx: 26, spot: 6, photon_idx: 17478)

▾ Coordinates:

gran_idx	(gran_idx)	uint64	19705 19713 73801 ... 118012 ...
spot	(spot)	uint64	1 2 3 4 5 6
photon_idx	(photon_idx)	int64	1 2 3 4 ... 17475 17476 17477 17...
source_file	(gran_idx)	object	'./is2-download/processed_ATL...
delta_time	(photon_idx)	datetime64[ns]	2020-12-07T23:14:42.1079928...

▾ Data variables:

sc_orient	(gran_idx)	float64	1.0 0.0 1.0 0.0 ... 0.0 1.0 1.0 0.0
cycle_number	(gran_idx)	float64	5.0 13.0 1.0 3.0 ... 11.0 12.0 14.0
rgt	(gran_idx)	float64	197.0 197.0 ... 1.18e+03 1.18e+...
atlas_sdp_gps_...	(gran_idx)	datetime64[ns]	2018-01-01T00:00:18 ... 2018-...
data_start_utc	(gran_idx)	datetime64[ns]	2019-10-09T19:31:32.837255 ....
data_end_utc	(gran_idx)	datetime64[ns]	2019-10-09T19:32:23.365738 ...
h_li	(spot, gran_idx, photon_idx)	float32	nan nan nan nan ... nan nan na...
latitude	(spot, gran_idx, photon_idx)	float64	nan nan nan nan ... nan nan na...
longitude	(spot, gran_idx, photon_idx)	float64	nan nan nan nan ... nan nan na...
gt	(gran_idx, spot)	object	nan nan nan ... 'gt1l' 'gt1r'

▾ Attributes:

data_product :	ATL06
Description :	The land_ice_height group contains the primary set of derived ATL06 products. This includes geolocation, height, and standard error and quality measures for each segment. This group is sparse, meaning that parameters are provided only for pairs of segments for which at least one beam has a valid surface-height measurement.
data_rate :	Data within this group are sparse. Data values are provided only for those ICESat-2 20m segments where at least one beam has a valid land ice height measurement.

CryoCloud

birdaro
POWERED BY CSCCE

Sustainability and governance

YOUR contributions!!

Maintained JupyterHub cloud compute environment

Parallelized workflows

dask

More built-in data processing (e.g. strong/weak beams, analysis)

Why should you use icepyx?

- ❖ save time and effort

- ❖ easy to install and use

- ❖ reproducible

- ❖ always get the latest data

- ❖ easily transition to the cloud

- ❖ works on the cloud and locally

- ❖ "knows" about ICESat-2 data

- ❖ be part of an expert community

How do I get involved?

❖ Ask questions/post a discussion

❖ contribute – it doesn't have to be code!

   o Report bugs

   o Improve documentation

   o Request features

   o Share your example

❖ Attend an upcoming coworking session

❖ cite your data and software

GitHub:
https://github.com/icesat2py/icepyx
Documentation:
https://icepyx.readthedocs.io

Demo: use icepyx to access and plot ATL13

date: 2019-07-20

```
spatial_extent = [-114.6, 36.06, -114.36, 36.2]
dates = ['2019-07-01', '2025-07-30']
track = ['341']
```

```
mead = ipx.Query('ATL13', spatial_extent, dates, tracks=
```

date

2019-07-20

https://github.com/icesat2py/icepyx/blob/hydrology/doc/source/example_notebooks/IS2_atl13_hydrology.ipynb