



Jet Propulsion Laboratory
California Institute of Technology

podaac

Physical Oceanography Distributed Active Archive Center

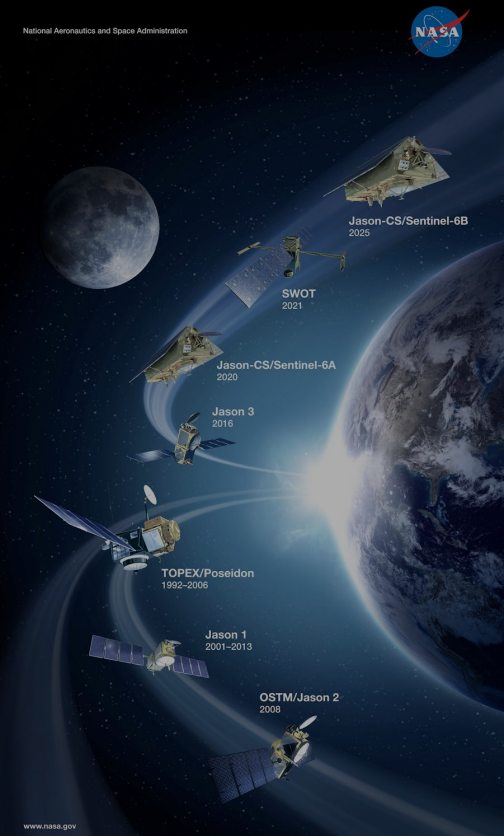


SWOT Data Access from NASA PO.DAAC

Suresh Vannan*, Catalina Taglialatela*,
Cassie Nickles, Jack McNelis, PO.DAAC Team
Jet Propulsion Laboratory, California Institute of Technology

<https://podaac.jpl.nasa.gov/>

19 September 2023
SWOT Science Team Meeting - Toulouse, France



National Aeronautics and Space Administration



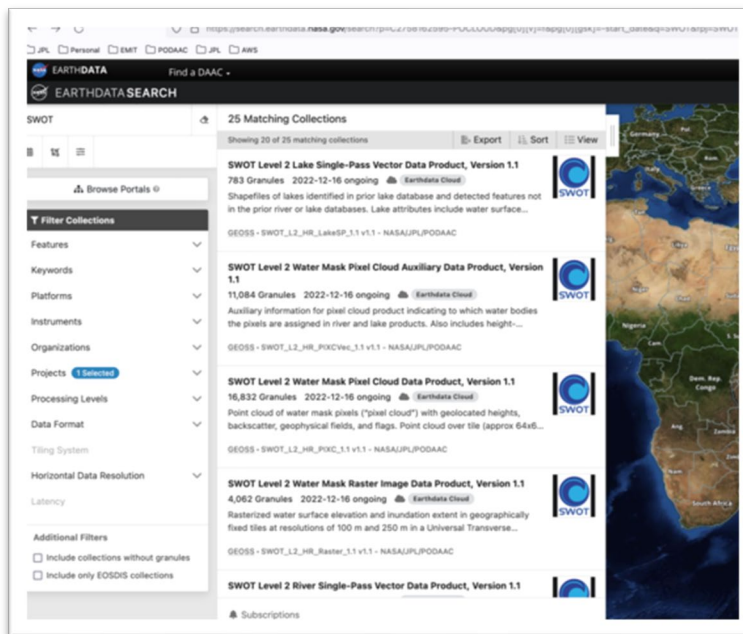
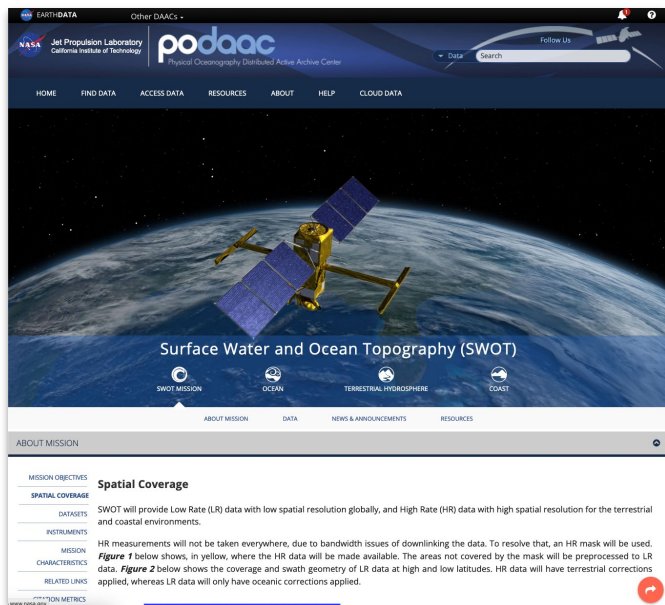
www.nasa.gov

© 2023 California Institute of Technology. Government sponsorship

*Presenting

PO.DAAC Access to SWOT data

→ Discover
→ Access



PO.DAAC Website

- Dataset landing page
- Announcements
- ATBDs, PDDs, User Guide
- Data Stories



<https://podaac.jpl.nasa.gov/SWOT?sections=data>

Earthdata Search

- Search and Discover SWOT and related data
- Faceted search to identify dataset per needs
- Metadata
- **https** and **s3** data access links



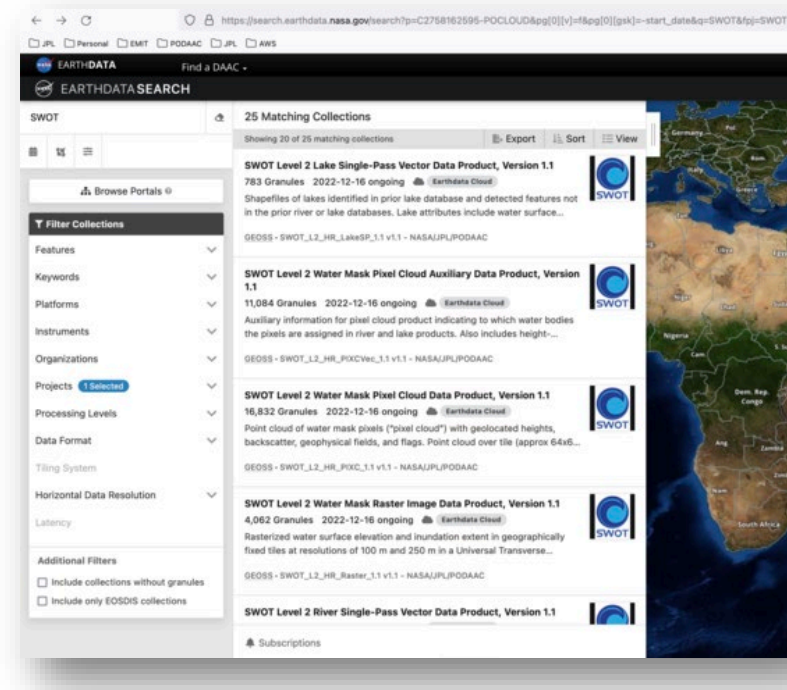
<https://search.earthdata.nasa.gov/search?fpj=SWOT>

PO.DAAC Access to Beta PreValidated SWOT data

As a reminder, to access the beta pre-validated KaRIn data products from the NASA PO.DAAC please do the following:

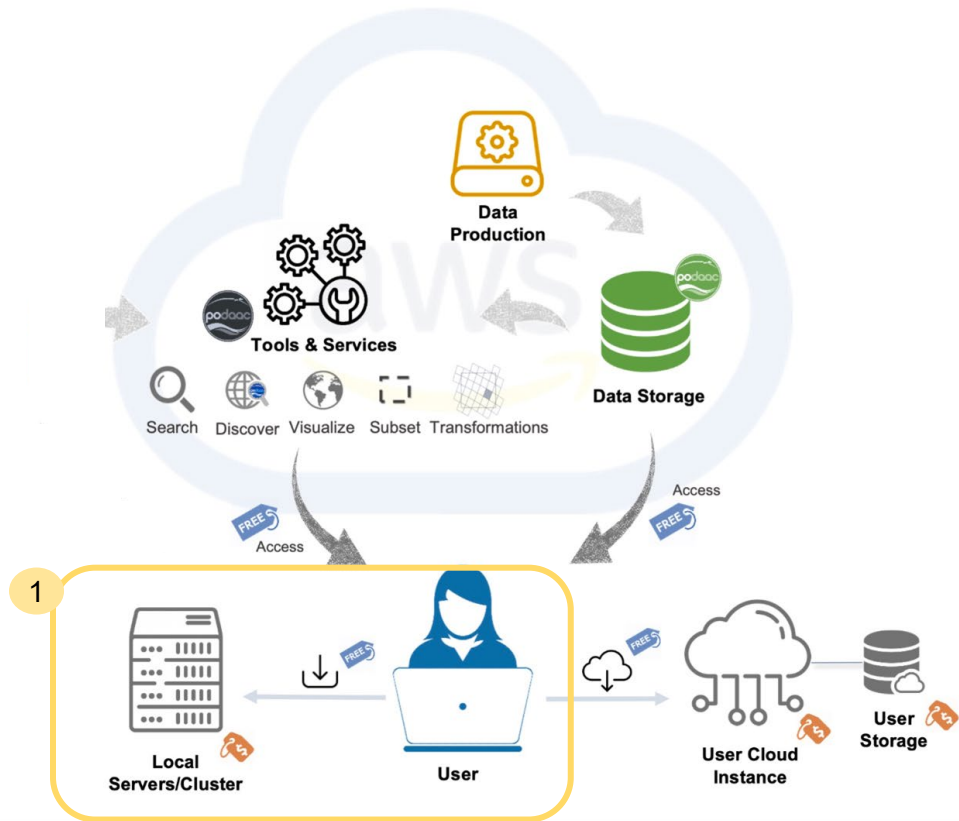
1. Register for a NASA Earthdata login at: <https://urs.earthdata.nasa.gov>
2. Visit the following Google form and provide your Earthdata credentials. These credentials will enable your access to the reprocessed data. Please do not share this link. <https://forms.gle/R96YLceswU7GqfyA6>

See Shailen's Release Note for details.



<https://search.earthdata.nasa.gov/search?fpj=SWOT>

PO.DAAC Access to SWOT dataDownload



Two main access pathways:

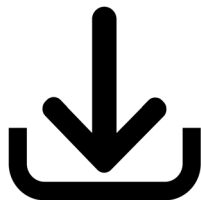
1. **Download** to local computer, laptop, server
2. **In-cloud** workflows (no download)

1 Download /Subscribe to SWOT data

→ Download
→ Subscribe

Downloader

The Downloader is useful if you need to download PO.DAAC data once in a while or prefer to do it “on-demand”. The Downloader makes no assumptions about the last time run or what is new in the archive, it simply uses the provided requests and downloads all matching data.



Example:

```
podaac-data-downloader -c SWOT_L2_HR_RiverSP_1.1 -d ./SWOT_L2_HR_RiverSP_1.1  
--start-date 2023-04-08T00:00:00Z --end-date 2023-04-08T23:59:59Z
```

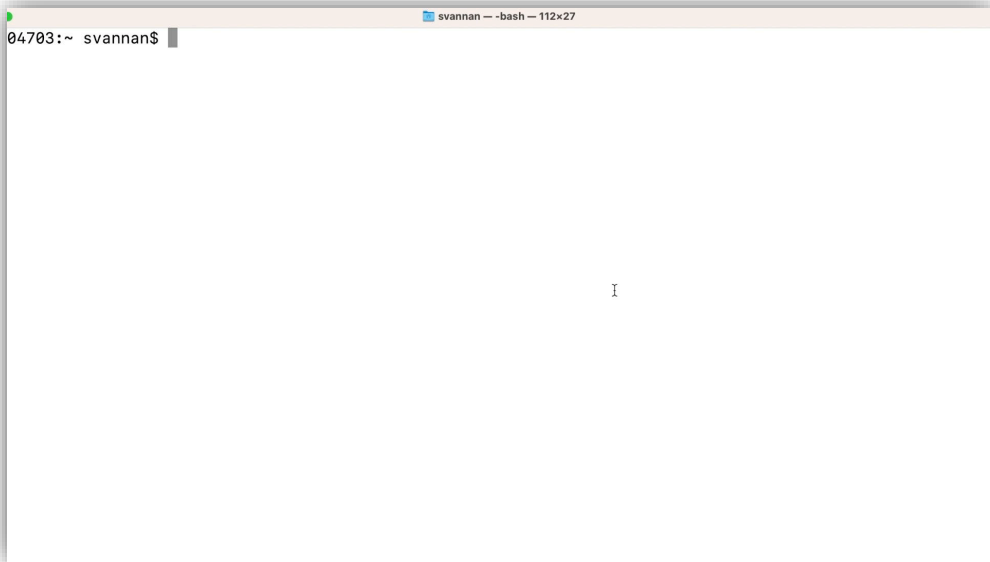
https://podaac.github.io/tutorials/quarto_text/DataSubscriberDownloader.html

Subscriber

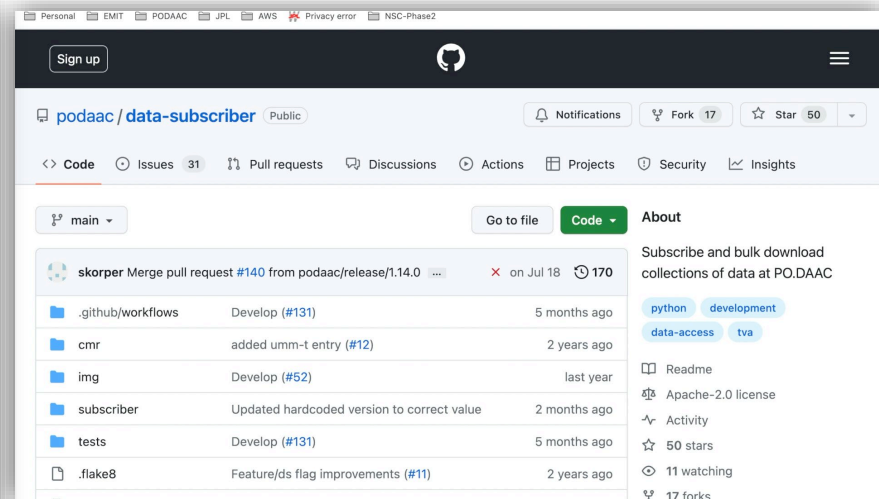
The subscriber is useful for users who need to continuously pull the latest data from the PO.DAAC archive. If you feed data into a model or real time process (e.g. data assimilation), the subscriber allows you to repeatedly run the script and only download the latest data.



DEMO



PO.DAAC Data Subscriber/Downloader Demo



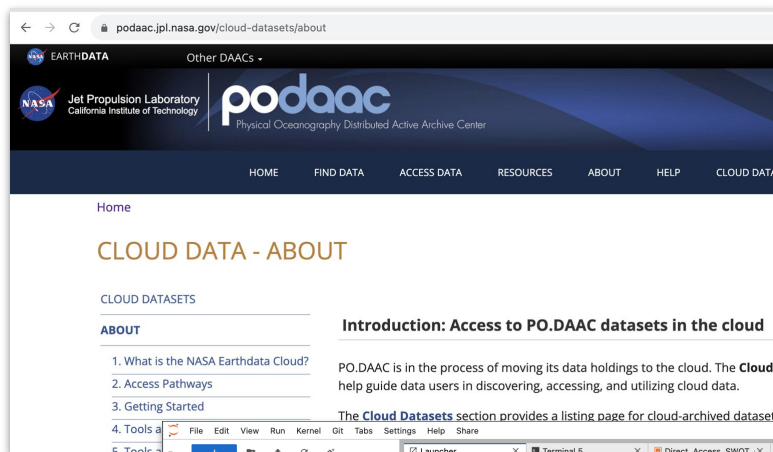
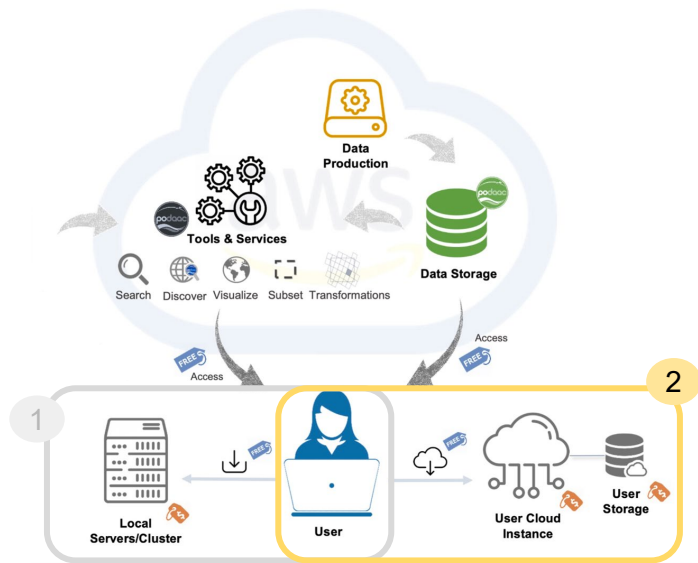
PO.DAAC Data Subscriber/Downloader Installation Demo

Data Download Cheat Sheet



Dry Run before downloading	<code>podaac-data-downloader -c SWOT_L2_HR_RiverSP_1.1-d ./SWOT_L2_HR_RiverSP_1.1--start-date 2023-04-08T00:00:00Z --end-date 2023-04-08T23:59:59Z --dry-run</code>
Download by Time extent	<code>podaac-data-downloader -c SWOT_L2_HR_RiverSP_1.1-d ./SWOT_L2_HR_RiverSP_1.1 --start-date 2023-04-08T00:00:00Z --end-date 2023-04-08T23:59:59Z</code>
Download by Spatial Extent	<code>podaac-data-downloader -c SWOT_L2_HR_RiverSP_1.1-d ./SWOT_L2_HR_RiverSP_1.1--start-date 2023-04-08T00:00:00Z --end-date 2023-04-08T23:59:59Z -b="-97,12,-96.5,13"</code>
Download specific file	<code>podaac-data-downloader -c SWOT_L2_HR_RiverSP_1.1-d ./SWOT_L2_HR_RiverSP_1.1 --gr SWOT_L2_HR_RiverSP_Node_484_004_AU_20230407T233306_20230407T233309_PIB0_01.zip</code>
Subscribe to data	<code>podaac-data-subscriber -c SWOT_L2_HR_RiverSP_1.1-d ./SWOT_L2_HR_RiverSP_1.1 --start-date 2023-04-08T1:00:00Z -m 60</code>
Download specific pass	<code>podaac-data-downloader -c SWOT_L2_HR_RiverSP_1.1-d ./SWOT_L2_HR_RiverSP_1.1 --gr SWOT_L2_HR_RiverSP_Node_??_013*.zip</code>

2 Accessing SWOT data in the cloud



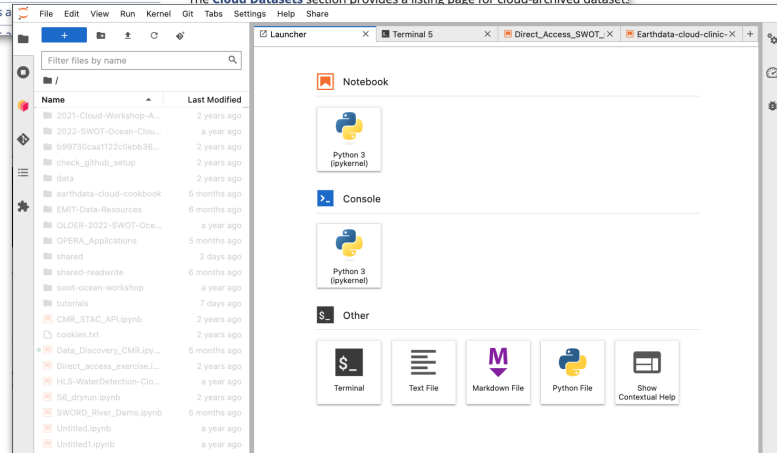
PO.DAAC provides help to understand access to data in the cloud:

<https://podaac.jpl.nasa.gov/cloud-datasets/about>

https://podaac.github.io/tutorials/quarto_text/cheatsheet.html →

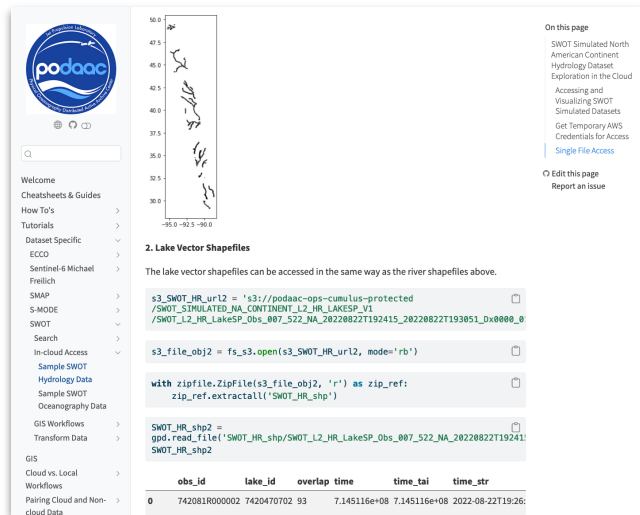


The **Earthdata Cloud Playground** - cloud compute environment through a managed Jupyter Hub interface will be available late 2023 to get started in the cloud.



Accessing SWOT data in the cloud

- In-cloud data access available
 - Advantage: **no data download**, no data movement, can access (free) directly from the archive
 - Data access free, but user needs cloud computing environment to work with data - *recall Cloud Playground coming soon*
 - For direct streaming data, us-west-2 region in AWS, where PO.DAAC SWOT data is archived
- Example jupyter notebooks for how to access SWOT simulated data directly in the cloud (no download)
 - HR data*: https://podaac.github.io/tutorials/notebooks/meetings_workshops/swot_ea_workshop_sept2022/SWOTHHR_s3Access.html
 - LR data: https://podaac.github.io/tutorials/external/Direct_Access_SWOT_sim_Oceanography.html



2. Lake Vector Shapefiles

The lake vector shapefiles can be accessed in the same way as the river shapefiles above.

```
s3_SWOT_HR_ur12 = 's3://podaac-ops-cumulus-protected/SWOT_SIMULATED_NA_CONTINENT_L2_HR_LAKESP_V1/SWOT_L2_HR_LAKESP_Obs_007_522_NA_20220822T192415_20220822T193951_Dv0000_0'

s3_file_obj2 = fs_s3.open(s3_SWOT_HR_ur12, mode='rb')

with zipfile.ZipFile(s3_file_obj2, 'r') as zip_ref:
    zip_ref.extractall('SWOT_HR_shp')
```

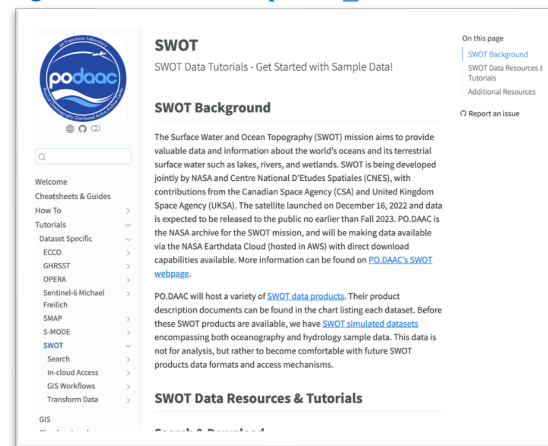
```
SWOT_HR_shp2 =
gpd.read_file('SWOT_HR_shp/SWOT_L2_HR_LakeSP_Obs_007_522_NA_20220822T192415_SWOT_HR_shp2')
```

obs_id	lake_id	overlap	time	time_tai	time_str
0	742081R000002	7420470702	93	7.145116e+08	7.145116e+08 2022-08-22T19:26:

More SWOT tutorials in the **PO.DAAC Cookbook SWOT chapter**:
https://podaac.github.io/tutorials/quarto_text/SWOT.html

*example today
with real
(restricted) SWOT
data ...

For a copy of the
notebook example,
contact PO.DAAC
speakers.



SWOT
SWOT Data Tutorials - Get Started with Sample Data!

SWOT Background

The Surface Water and Ocean Topography (SWOT) mission aims to provide valuable data and information about the world's oceans and its terrestrial surface water such as lakes, rivers, and wetlands. SWOT is being developed jointly by NASA and Centre National D'Etudes Spatiales (CNES), with contributions from the Canadian Space Agency (CSA) and United Kingdom Space Agency (UKSA). The satellite launched on December 16, 2022 and data is expected to be released to the public no earlier than Fall 2023. PO.DAAC is the NASA archive for the SWOT mission, and will be making data available via the NASA Earthdata Cloud (hosted in AWS) with direct download capabilities available. More information can be found on [PO.DAAC's SWOT webpage](#).

PO.DAAC will host a variety of [SWOT data products](#). Their product description documents can be found in the chart listing each dataset. Before these SWOT products are available, we have [SWOT simulated datasets](#) encompassing both oceanography and hydrology sample data. This data is not for analysis, but rather to become comfortable with future SWOT products data formats and access mechanisms.

SWOT Data Resources & Tutorials





Firefox File Edit View History Bookmarks Tools Window Help Mon Sep 18 2:37 PM

SWOTHR_s3Acc... (3) - Jupyter x PO.DAAC Cookbook - SWOT River x +

https://openscapes.212c.cloud/user/sciencecat18/lab/tree/SWOT_STM_2023/SWOTHR_s3Access_v0-Copy1.ipynb 170%

File Edit View Run Kernel Git Tabs Settings Help Share

Terminal 5 DirectCloud_Access_SWO SWOTHR_s3Access_v0-Copy1

Python 3 (ipykernel)

From the PO.DAAC Cookbook, to access the GitHub version of the notebook, follow [this link](#). - *currently does not link to correct GitHub. Will be updated once SWOT data is public*

SWOT Hydrology Dataset Exploration in the Cloud

Accessing and Visualizing SWOT Datasets

Requirement:

This tutorial can only be run in an **AWS cloud instance running in us-west-2**: NASA Earthdata Cloud data in S3 can be directly accessed via `earthaccess` python library; this access is limited to requests made within the US West (Oregon) (code: `us-west-2`) AWS region.

Learning Objectives:

- Access all 5 products of SWOT HR data (archived in NASA Earthdata Cloud) within the AWS cloud, without downloading to local machine

Simple 3 Python 3 (ipykernel) ... Mem: 1.16 / 3.7... Mode: Comm... Ln 2, C... SWOTHR_s3Access_v0-Copy1... 1

Accessing SWOT data in the cloud

Demo quick preview

→ In-cloud Access (AWS)



Direct in-cloud access without moving, storing or downloading files

Learning Objectives:

- Access all 5 products of SWOT HR data (archived in NASA Earthdata Cloud) within the AWS cloud, without downloading to local machine
- Visualize accessed data for a quick check

SWOT Level 2 KaRIn High Rate Version 1.0 or 1.1 (where available) Datasets:

1. **River Vector Shapefile** - SWOT_L2_HR_RIVERSP_1.1
2. **Lake Vector Shapefile** - SWOT_L2_HR_LAKESP_1.1
3. **Water Mask Pixel Cloud NetCDF** - SWOT_L2_HR_PIXC_1.1
4. **Water Mask Pixel Cloud Vector Attribute NetCDF** - SWOT_L2_HR_PIXCVec_1.0
5. **Raster NetCDF** - SWOT_L2_HR_Raster_1.0

0

Libraries Needed

Use this while SWOT data has restricted access: need to install the dev version for earthaccess

```
import glob
import os
import requests
import s3fs
import fiona
import netCDF4 as nc
import h5netcdf
import xarray as xr
import pandas as pd
import geopandas as gpd
import numpy as np
import matplotlib.pyplot as plt
import hvplot.xarray
import earthaccess
from earthaccess import Auth, DataCollections, DataGranules, Store
```

1

Earthdata Login

An Earthdata Login account is required to access data, as well as discover restricted data, from the NASA Earthdata system. Thus, to access NASA data, you need Earthdata Login. If you don't already have one, please visit <https://urs.earthdata.nasa.gov> to register and manage your Earthdata Login account. This account is free to create and only takes a moment to set up. We use `earthaccess` to authenticate your login credentials below.

```
#auth = earthaccess.login(strategy="interactive", persist=True) #if you do not have a ne
auth = earthaccess.login(strategy="netrc") #if you have created a netrc prior with you
```

2

Search for the data of interest

```
#Retrieves granule from the day we want, in this case by passing to `earthdata.search_data` function
river_results = earthaccess.search_data(short_name = 'SWOT_L2_HR_RIVERSP_1.1',
temporal = ('2023-04-08 00:00:00', '2023-04-25 23:59:59'),
granule_name = '*Reach*_013_NA*', # here we filter by Reach
count=2000) #for restricted datasets, need to specify count
```

3

Set up an s3fs session for Direct Cloud Access

`s3fs` sessions are used for authenticated access to s3 bucket and allows for typical file-system style operations. Below we create session by passing in the data access information.

```
fs_s3 = earthaccess.get_s3fs_session(results=river_results)
```

4

Create Fiona session to work with zip and embedded shapefiles in the AWS Cloud

The native format for this data is a `.zip` file, and we want the `.shp` file within the `.zip` file, so we will create a Fiona AWS session using the credentials from setting up the `s3fs` session above to access the shapefiles within the zip files. If we don't do this, the alternative would be to download the data to the cloud environment (e.g. EC2 instance, user S3 bucket) and extract the `.zip` file there.

```
fiona_session=fiona.session.AWSSession(
aws_access_key_id=fs_s3.storage_options["key"],
aws_secret_access_key=fs_s3.storage_options["secret"],
aws_session_token=fs_s3.storage_options["token"]
)
```

```
# We use the zip+ prefix so fiona knows that we are operating on a zip file
river_shp_url = f"zip+{river_data[0]}"
```

```
with fiona.Env(session=fiona_session):
    SWOT_HR_shp1 = gpd.read_file(river_shp_url)
```

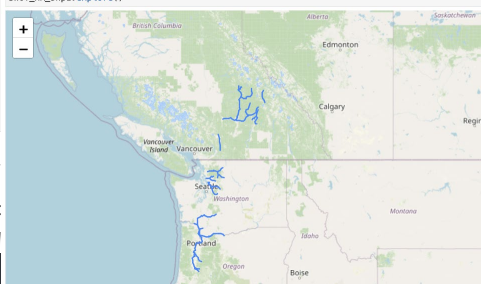
```
#view the attribute table
SWOT_HR_shp1
```

reach_id	time	time_tai	time_str	p_lat	p_lon	river_name
100900135	-1.0000000e+12	-1.0000000e+12	no_data	46.159638	-123.070171	Columbia River

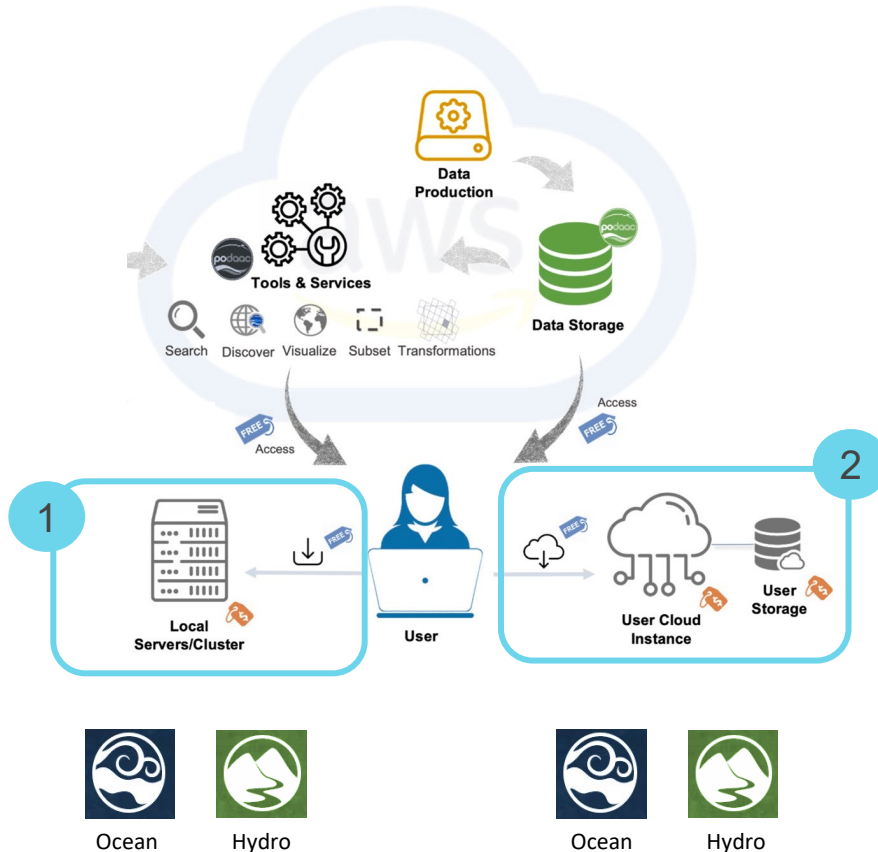
5

Quickly plot the SWOT river data

```
## Another way to plot geopandas dataframes is with `explore`, which also plots a basemap
SWOT_HR_shp1.explore()
```



Accessing SWOT data from PO.DAAC



Two main access pathways **applicable to either SWOT LR or HR data**:

1. **Download** to local computer, laptop, server
 - a. *E.g. subscriber/downloader*
2. **In-cloud direct access workflows** (no download)
 - a. *E.g. Python/Jupyter Lab*

Tools for accessing SWOT HR data

Timeseries

I would like SWOT timeseries data for my [river/lake/region] of interest.

SWOT hydrology timeseries API - Hydrocron (under development)



Current Data Workflow:

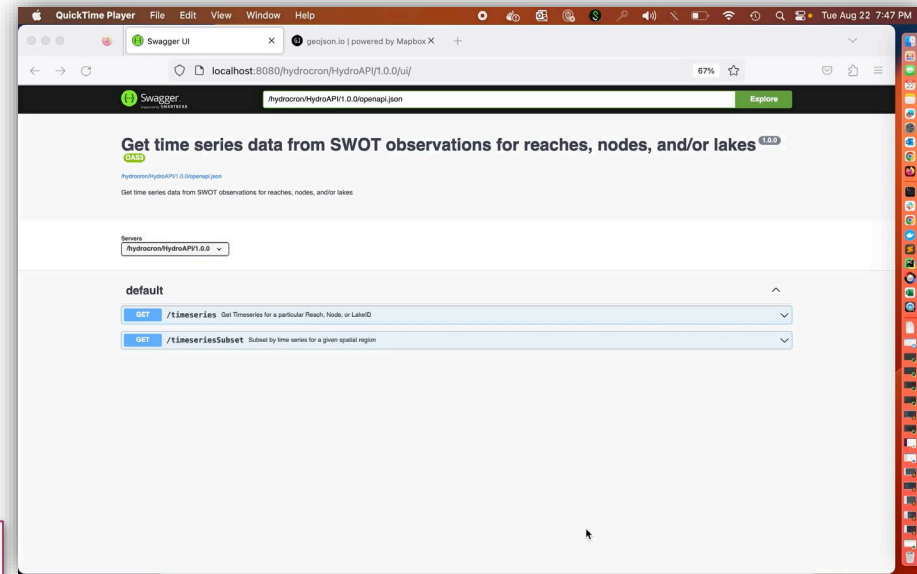
- Download or access individual data files (one per timestamp) and merge in some fashion (often code) to create a time series of a specific area

Hydrocron API Workflow

- Input river reach, node, lake ID, or geospatial region of interest (bounding box) → get one file of timeseries in csv or GeoJSON!

Potential uses

- Programmatically, in code or jupyter notebooks
- Connect data in user dashboard



We welcome feedback on whether this meets your needs, what else might be helpful

Tools for accessing SWOT HR data

On Demand Raster

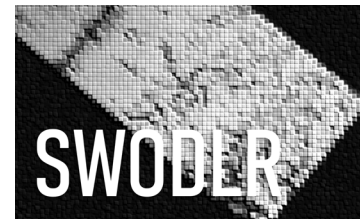
I would like SWOT raster data with a custom spatial resolution.

SWOT on-demand raster generation (SWODLR) (under development)

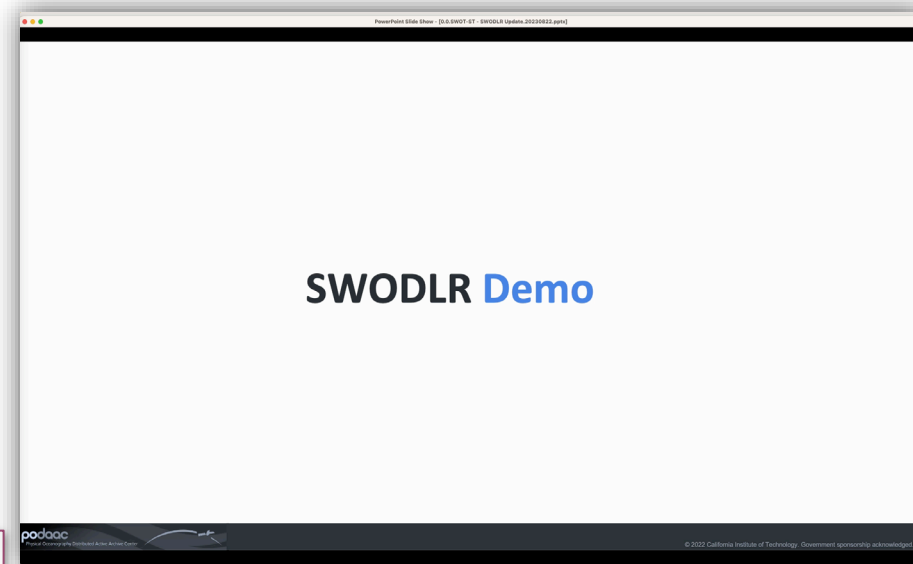
- SWOT L2 Raster Standard Data Products are limited to 100m and 250m resolutions
- PO.DAAC will offer On-demand Data Product raster generation allowing projection, resolution, and overlap customization

Use Cases examples:

- Match resolutions with different data products beyond standard 100m/250 m, aiding models or comparisons.
- Analyze SWOT data at finer spatial resolutions (ex: studying floodplain and delta details in specific areas)
- Seamlessly combine data tiles using lat/lon coordinates, simplifying stitching compared to UTM coordinates.
- Explore edge points of interest with a more accurate view using the 256 x 128 km extent, enhancing raster analysis.



ETA Oct 2023



We welcome feedback on whether this meets your needs, what else might be helpful

Tools for accessing SWOT data- Cheatsheet



Ocean



Hydro



Coast



Learn/Information

- PO.DAAC Dataset Mission Page and Landing Pages <https://podaac.jpl.nasa.gov/SWOT?sections=data>



Find Data - Map GUI interface

- Search & Access in *Earthdata Search* <https://search.earthdata.nasa.gov/search?q=SWOT%20HR&long=0.0703125>



Access - Command line/automated scripts

- Subscriber/Downloader https://podaac.github.io/tutorials/quarto_text/SWOT.html



Access & Subset - GUI

- HiTIDE <https://hitide.podaac.earthdatacloud.nasa.gov/>





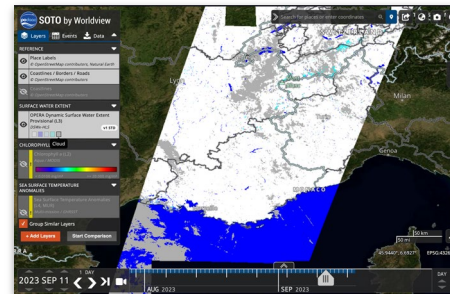
Access - Cloud native, Big data, ML

- in-cloud access also available: [example for LR ocean](#), [example for HR hydro](#)



Access & Explore - In development or planning phase:

- Hydrocron Timeseries API - *in development* 
- SWODLR On-demand Raster - *in development* 
- Exploratory Analysis in [SOTO by Worldview](#) - *planning*
- QGIS and ArcGIS - local: download and open - *works now*
- Web services (e.g. WFS “next gen” OGC) - *scoping*





Thank you! Questions

If you have specific needs, how can we help?

**Come chat with us! at the PO.DAAC “Help Desk”
in Salon Grand Toulouse during breaks and poster sessions**

Suresh Vannan - PO.DAAC Manager
suresh.vannan@jpl.nasa.gov

Catalina Tagliatela – PO.DAAC Applied Science Systems Engineer Lead
catalina.oaida@jpl.nasa.gov

Cassie Nickles – PO.DAAC Applied Science System Engineer
cassandra.l.nickles@jpl.nasa.gov

Jack McNelis – PO.DAAC Data Publication Engineer Lead
jack.mcnelis@jpl.nasa.gov

Upcoming Planned Events:

- PO.DAAC at SWOT Applications Early Adopters workshop - Dec 2023
- PO.DAAC at AGU in San Francisco, CA - Dec 2023
- NASA Earthdata Webinar – Dec 2023
- PO.DAAC Open Ocean Data in the Cloud – Workshop – Spring 2024
- AGU Chapman (hydro workshop) – Feb 2024; WaterSciCon – Jun 2024



<https://podaac.jpl.nasa.gov/swot>



[PO.DAAC Cookbook - SWOT Chapter](#)
Tutorials and Resources