

Confluence Dataset: SWORD of Science (SoS)

Nikki Tebaldi, Suresh Vannan*

Jet Propulsion Laboratory, California Institute of Technology

Colin Gleason, Travis Simmons

University of Massachusetts Amherst

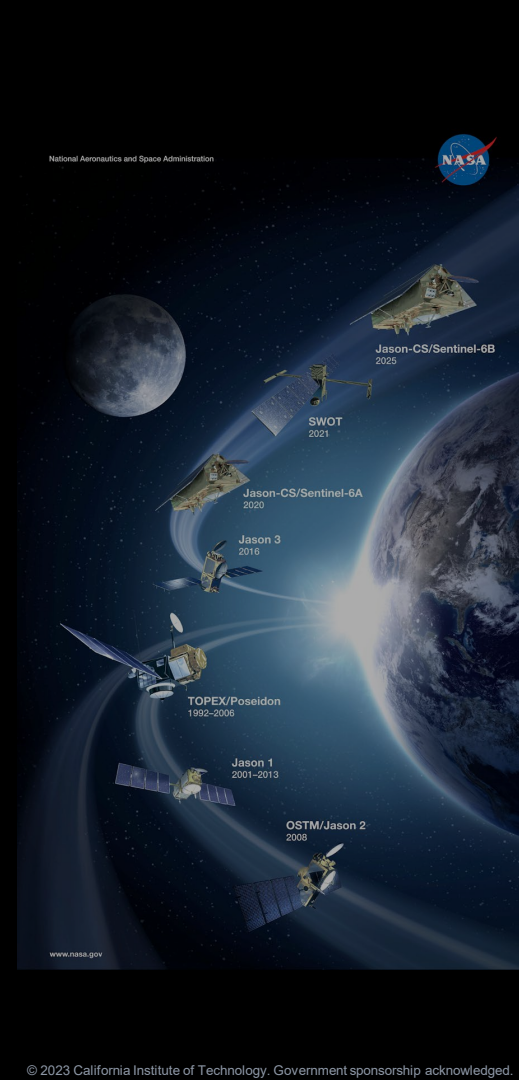
Mike Durand, Steve Coss

Ohio State University

September, 2023



*Funded by the Advanced Information Systems
Technology (AIST) Program*



Presentation Overview

- History of the SoS and where we are now
- Top-level overview of the dataset structure
- Tools to access and use the SoS

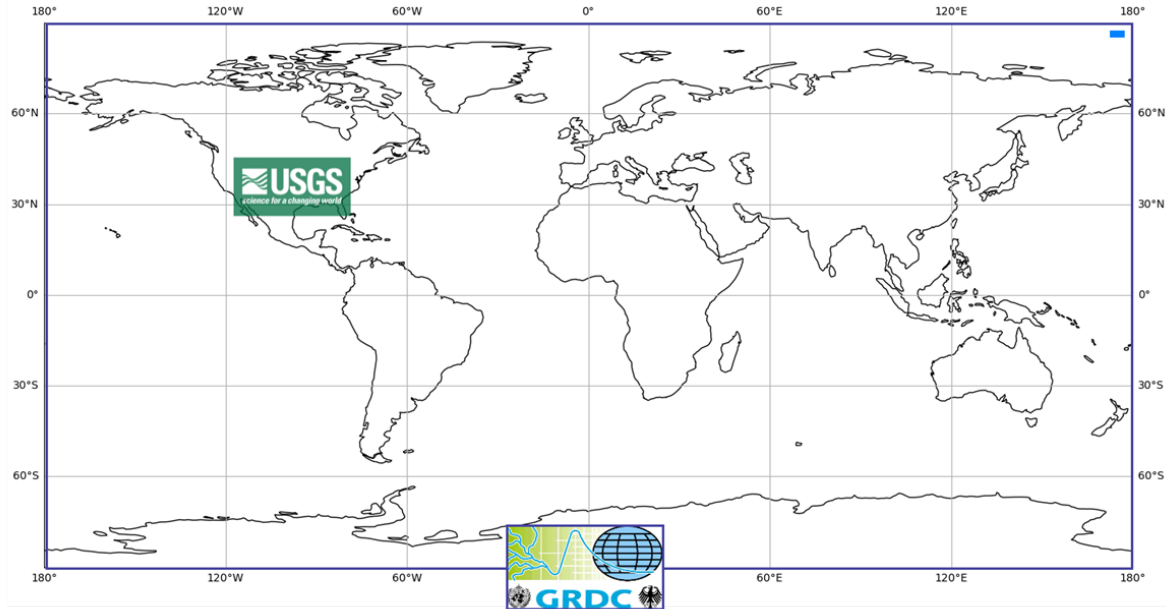


SWORD of Science (SoS)

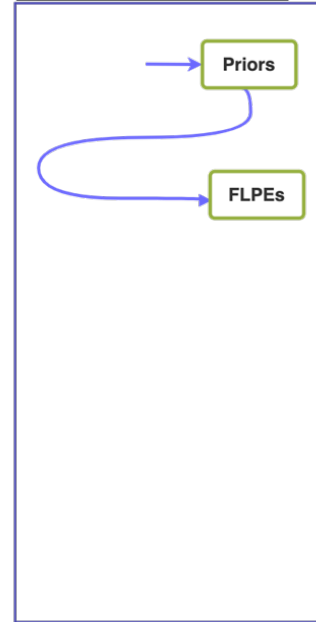
- The SWORD of Science (SoS) is a community-driven dataset produced for and from the execution of the Confluence workflow in the cloud which enables quick data access and compute on SWOT data.
- The SoS stores **prior data** used in the execution of the Confluence workflow and includes hydrology community data products such as: Agency gauge data from around the world, the Water Balance Model (WBM) and Global Reach-scale A priori Discharge Estimates (GRADES).
- The SoS also stores the **final results** of the execution of each module in the Confluence workflow and provides the required discharge parameters for the SWOT mission.

Evolution of the SoS: Version 0

Gauge Agencies

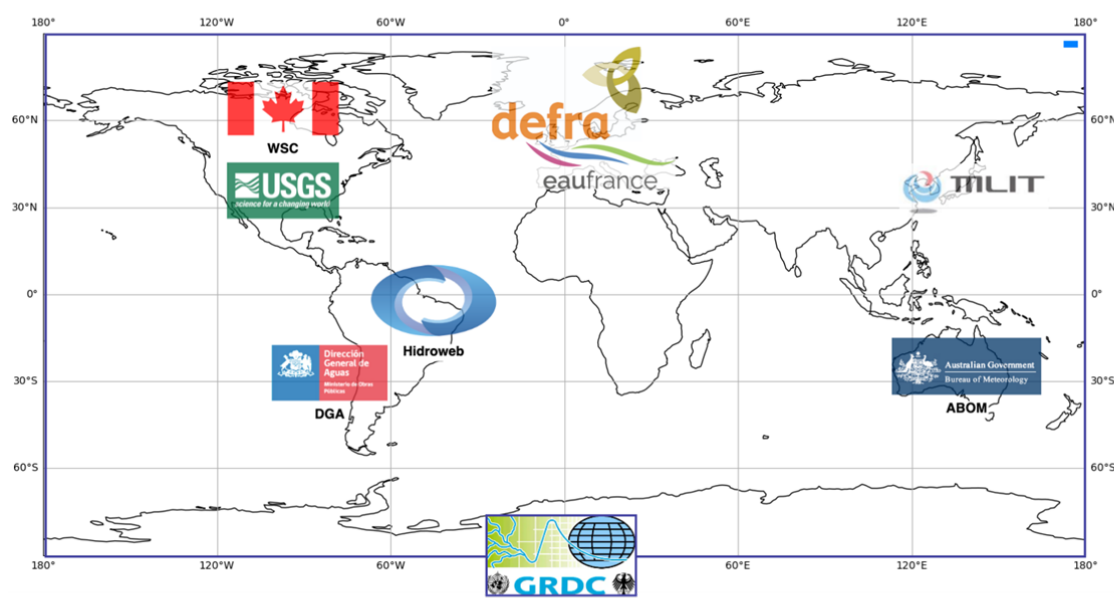


Confluence Workflow Modules

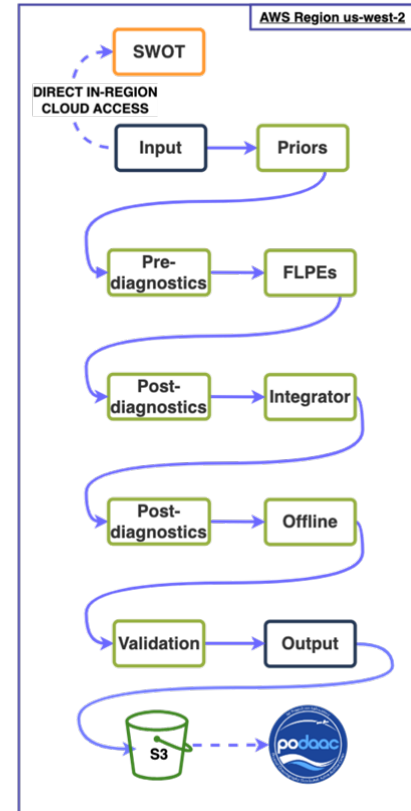


Evolution of the SoS: Current Version

Gauge Agencies



Confluence Workflow Modules



Organization & Structure

- The SoS is organized by continent following SWOT River Database (SWORD) structure and naming conventions.
 - It is indexed on the same reach and node identifier dimensions found in SWORD.
 - Time series data is stored by cycle and pass on an observation dimension.
- There is a constrained (with GRADES data) and unconstrained (with WBM data) version.
- NetCDF Groups separate data into logical areas and can contain their own attributes and dimensions.
 - Prior data is organized by gauge agency and model data source.
 - Results data is organized by Confluence workflow module.

Tools

An [Earthdata Search](#) or [CMR query](#) will allow users to search, locate and download or access the dataset based on temporal coverage and/or continental coverage.



| EARTH**DATA**



EARTHDATA

CMR Search



Tools

[PO.DAAC Data Subscriber](#) to download data via:

[data-subscriber](#)

- Continuously pull the latest data from the PO.DAAC archive and repeatedly run the tool to only download the latest data

[data-downloader](#)

- Packaged with the subscriber this tool allows you to download PO.DAAC data "on-demand". Run the tool once to download all matching data.

Tools



Hydocron (*under active development and not yet released*) is a new tool that will allow hydrologists to have direct access to filtered data from our newest satellites.

- Provides a way to filter data by feature ID, date range, polygonal area, and more.
- This data will be returned in formats such as CSV and geoJSON.

