Tracking the Hunga Tonga-Hunga Ha'apai Eruption Stratospheric Aerosol and Trace Gas Plumes Using Machine Learning

Rhys Leahy, David Giles, Maria del Carmen Cazorla

June 5, 2024

[1]Science Systems and Applications, Inc.
[2]NASA Goddard Space Flight Center
[3]Universidad de San Francisco - Quito

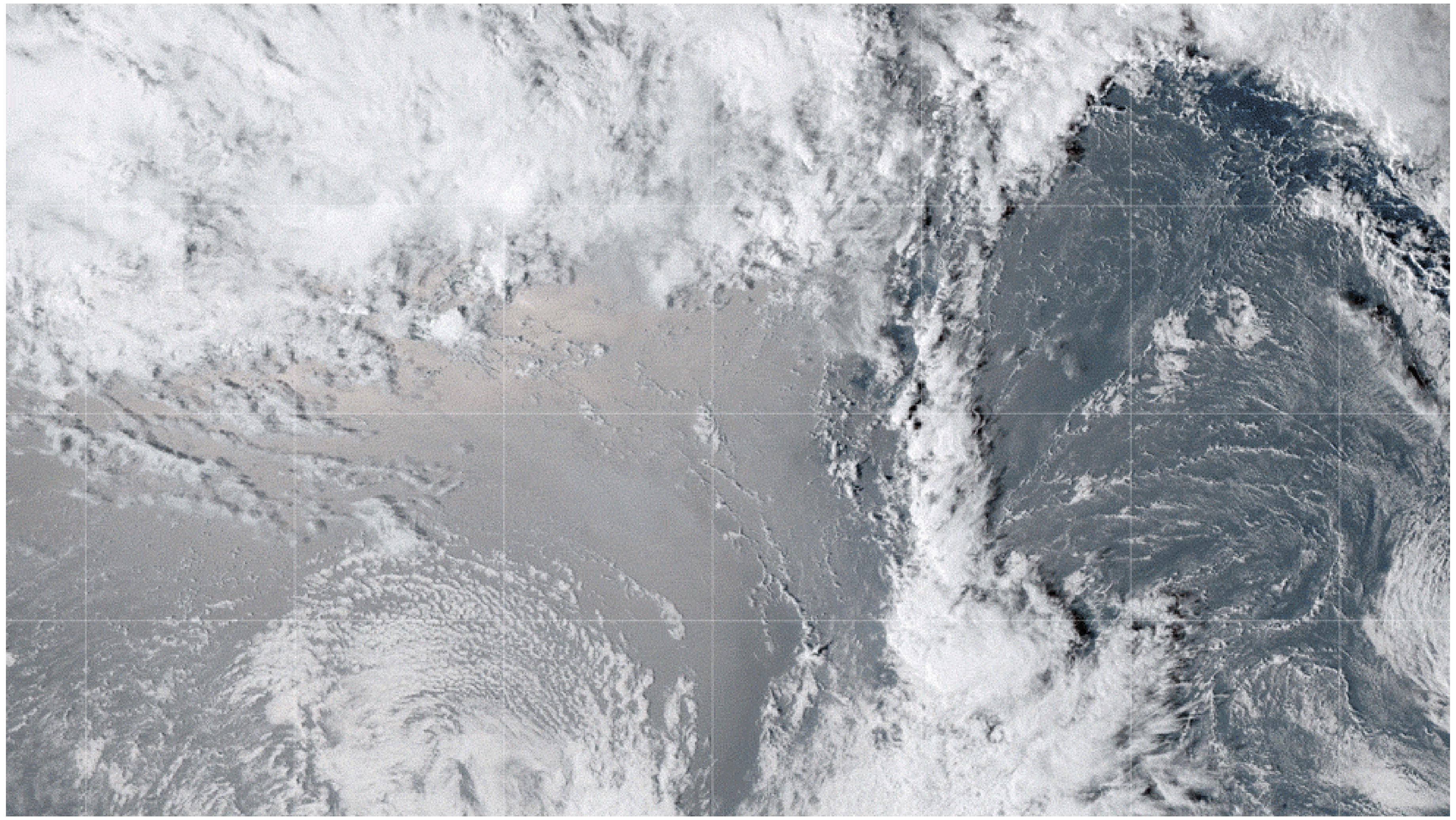
Agenda

1. Motivations
2. Methodology
   - Model and Dataset Selection
   - Extensible, Modular Data Processing Framework
3. Results and Model Evaluation
   - Predicting Plume Location
   - Understanding Inputs
4. Future Work and Applications

GODDARD EARTH SCIENCES

Motivations

- The most powerful volcanic eruption of the 21st century, and modern satellite era.
- NASA's Earth System Observatory collects and produces TBs a day.
- How can ML models analyze large volumes of remote sensing data around the HTHH eruption to map the plume and help measure its impact around the planet?



Gif from the GOES Advanced Baseline Imager, produced by NASA.

Methodology

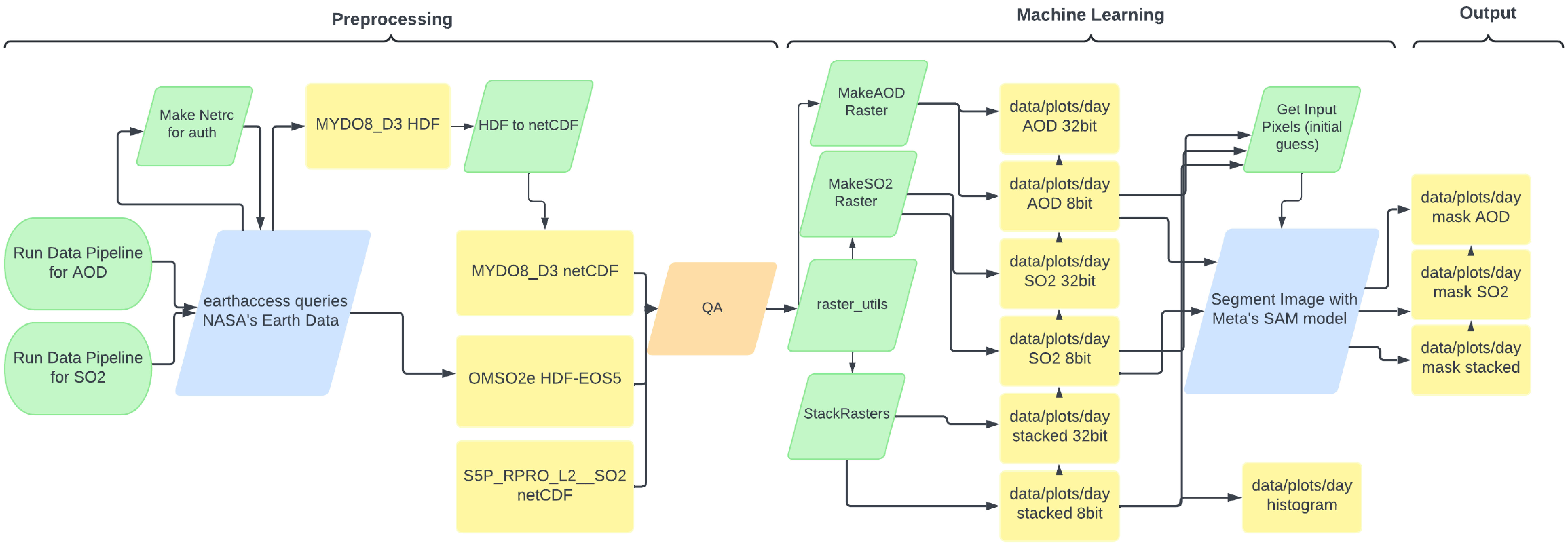
Model and Dataset Selection

Segment Anything (Meta)	Sulfur Dioxide (SO_2) OMI L3, TROPOMI L2	Lidar Features CALIPSO/CALIOP L2	Aerosol Optical Depth (MODIS L3)

• Model:
  ◦ Meta's open-source Segment Anything Model (SAM),
  ◦ the open-source segment-geospatial or **samgeo** Python package Semantic segmentation
• Categorization of every pixel
• Zero-shot predictions
  ◦ Predict classes not observed during training
• Built on Facebook AI Research Team's previous mask-RCNN
	• Trace gas common in volcanic eruptions.
	• SO2 concentration in atmospheric column.
	• Quality Control
	• HDF EOS5 and netCDF files read with Python rasterio.
	• Convert format to single band raster images.
		• Classification
		• Summarized stratospheric features
		• Quality Control
		• HDF files converted to netCDF and read with Python rasterio.
		• Convert format to single band raster images.
			• Total Column AOD (volcanic ash, smoke).
			• HDF with more than 600 statistical datasets.
			• Convert to netCDF and then single band raster images.

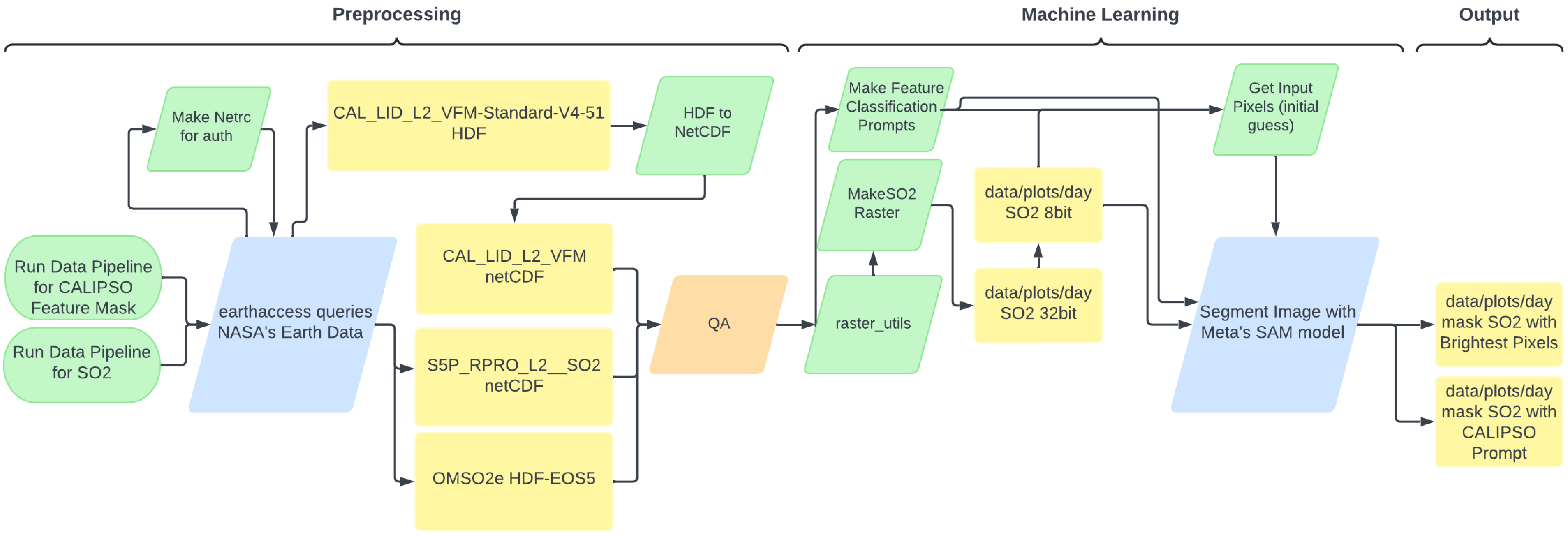
Modular Processing Framework

Brightest Pixels as Prompts

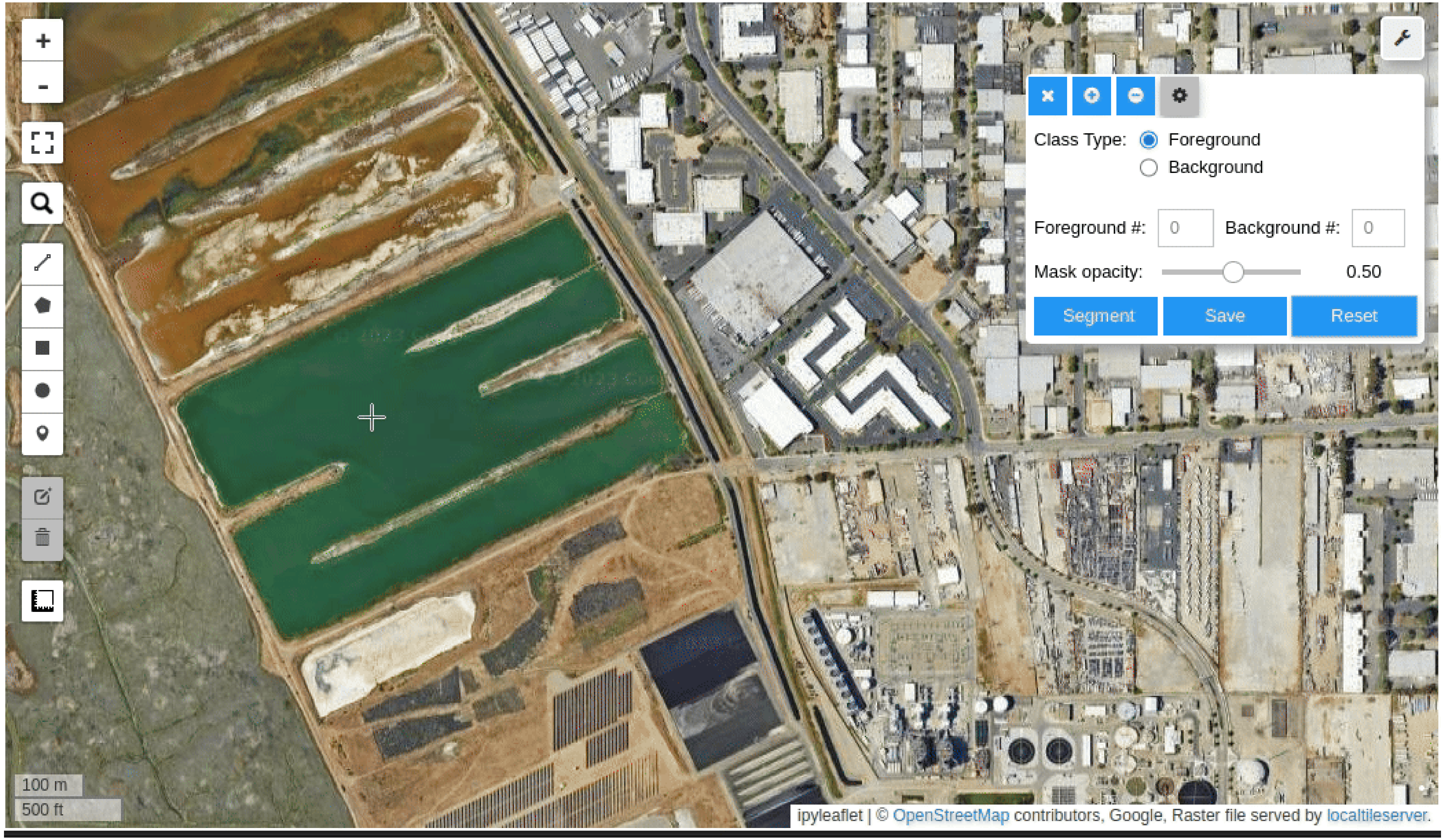


Modular Processing Framework

CALIPSO Retrievals as Prompts



Segment-Geospatial Input Prompts

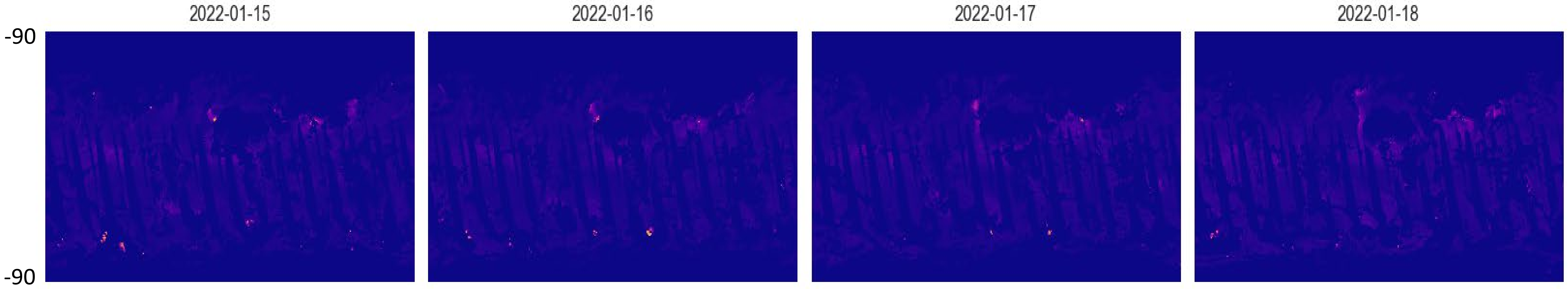


Input Prompt example from https://samgeo.gishub.org/examples/input_prompts/

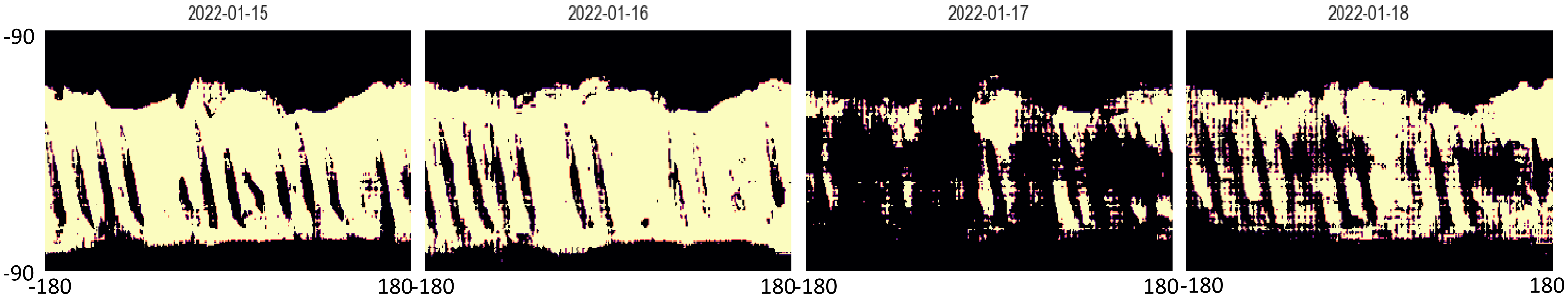
Results

Predicting Plume Location From MODIS L3 AOD Brightest Pixels

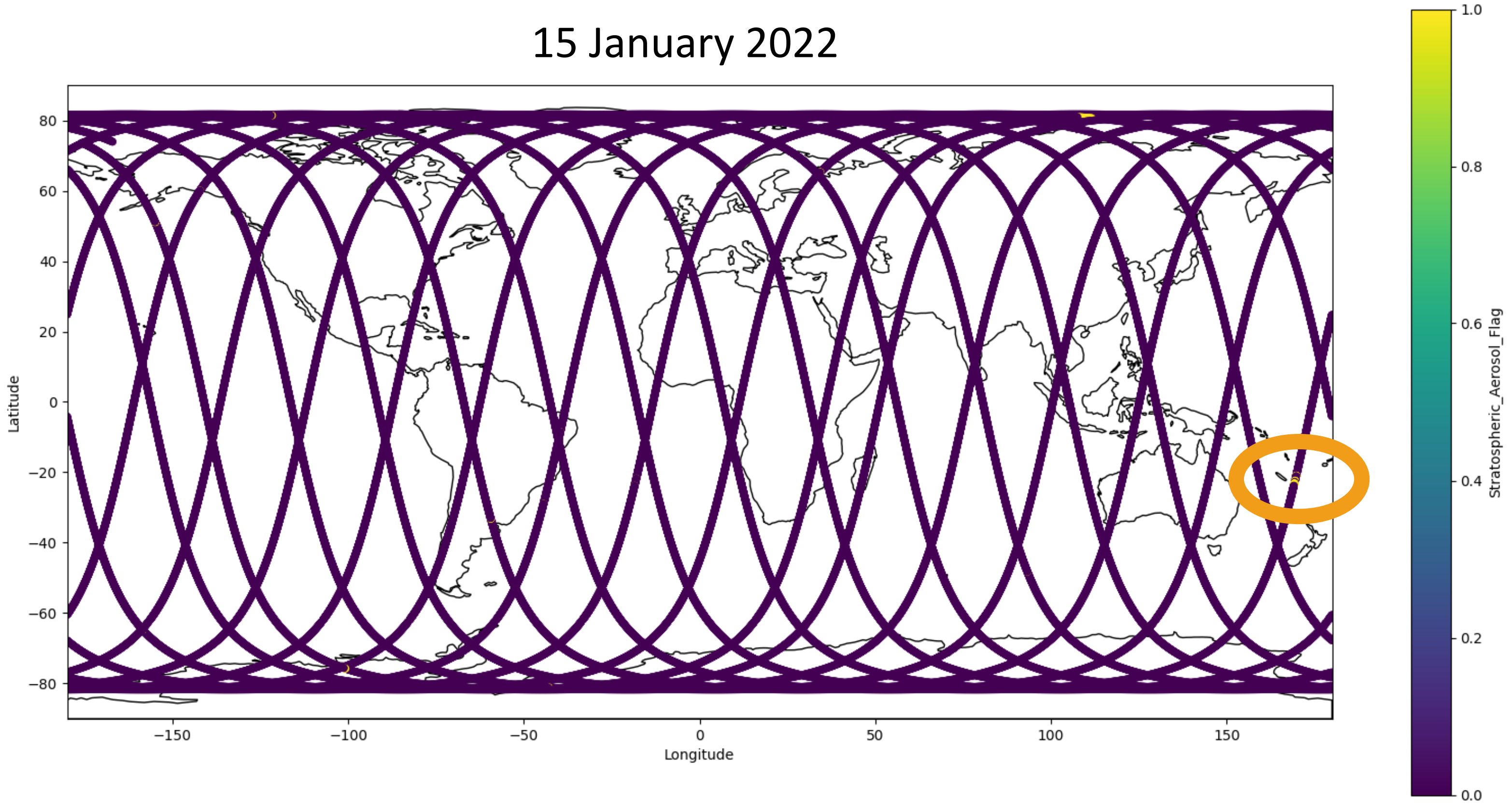
AOD Input Images

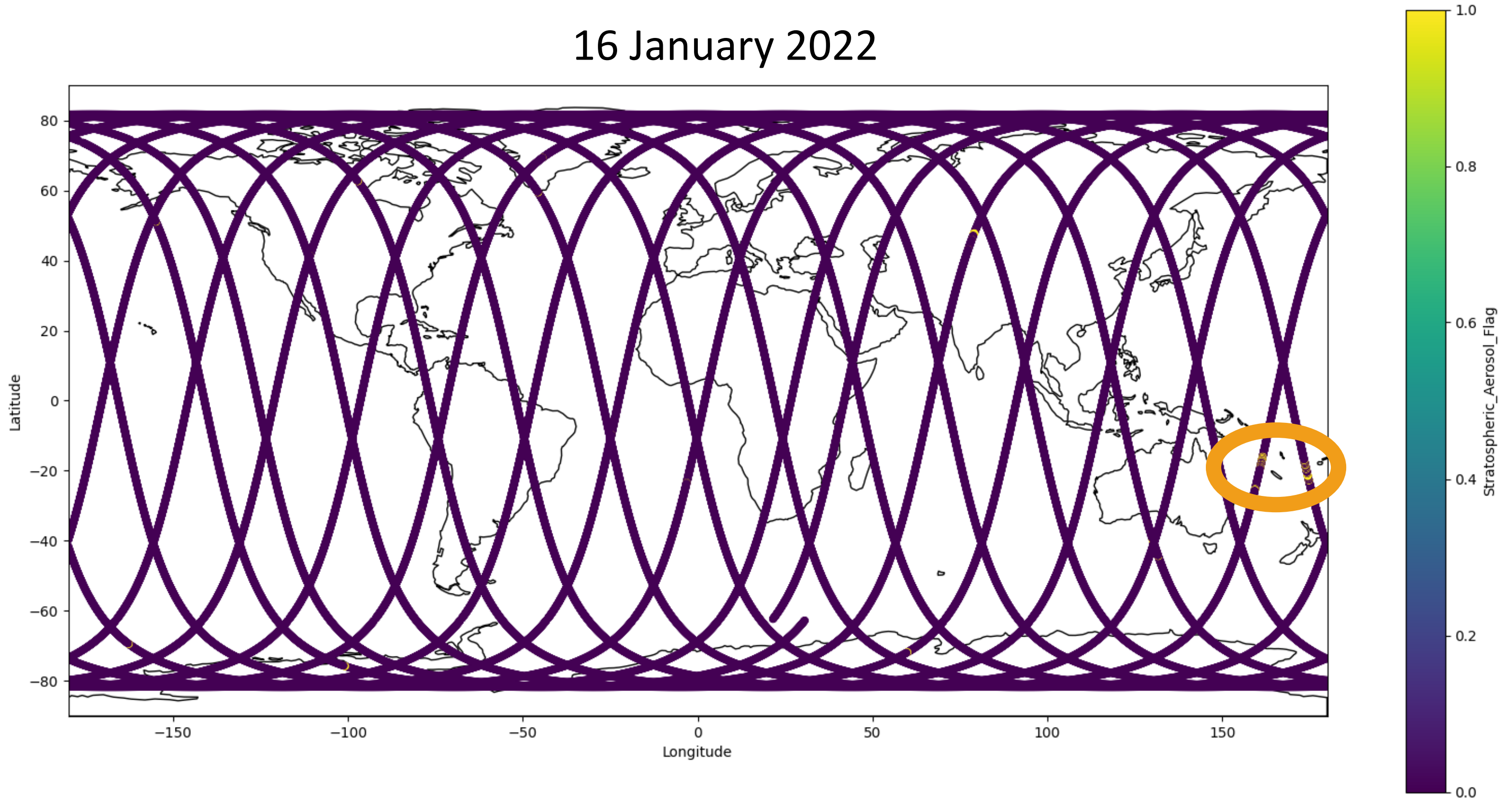


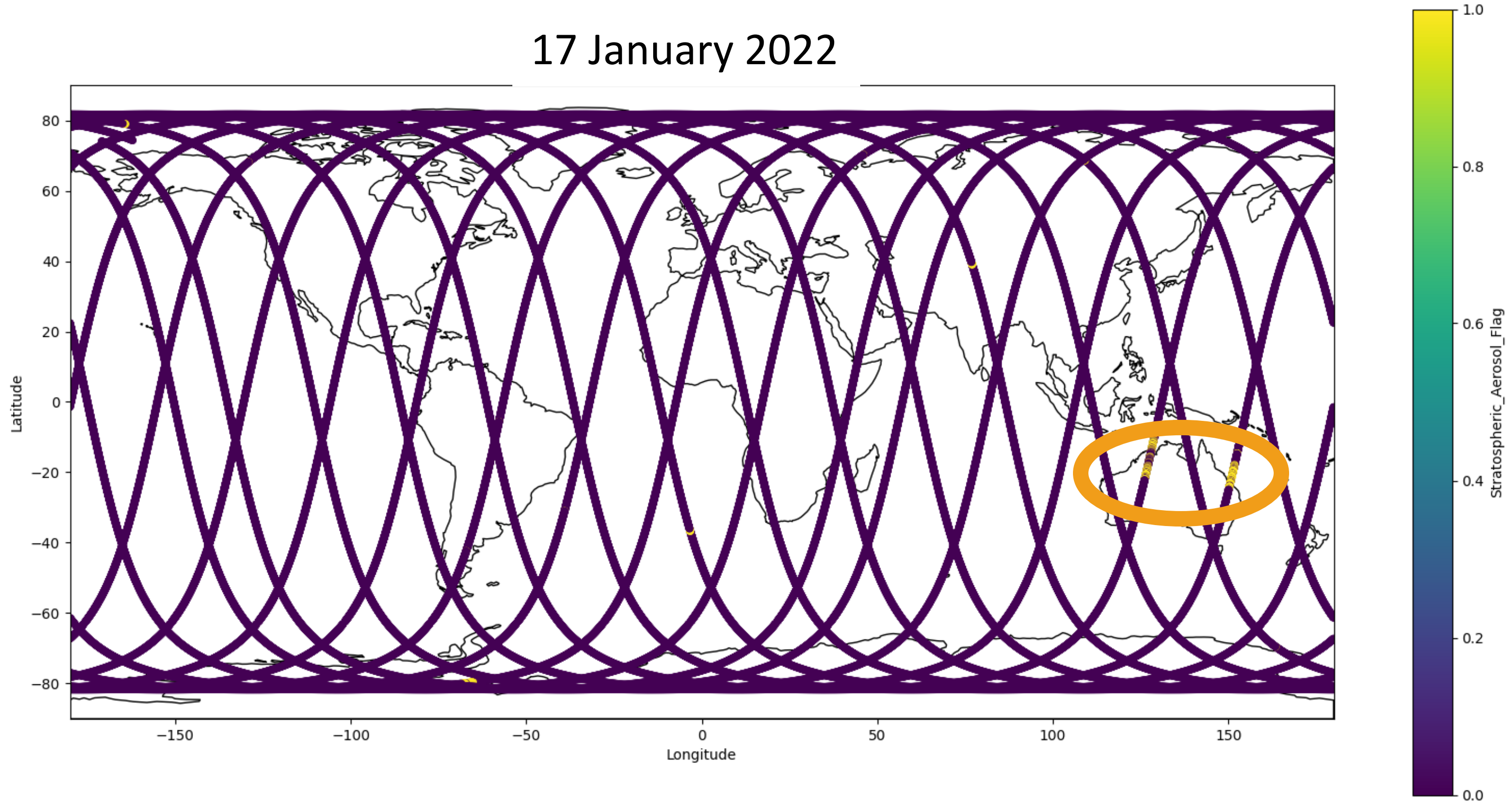
AOD Feature Masks

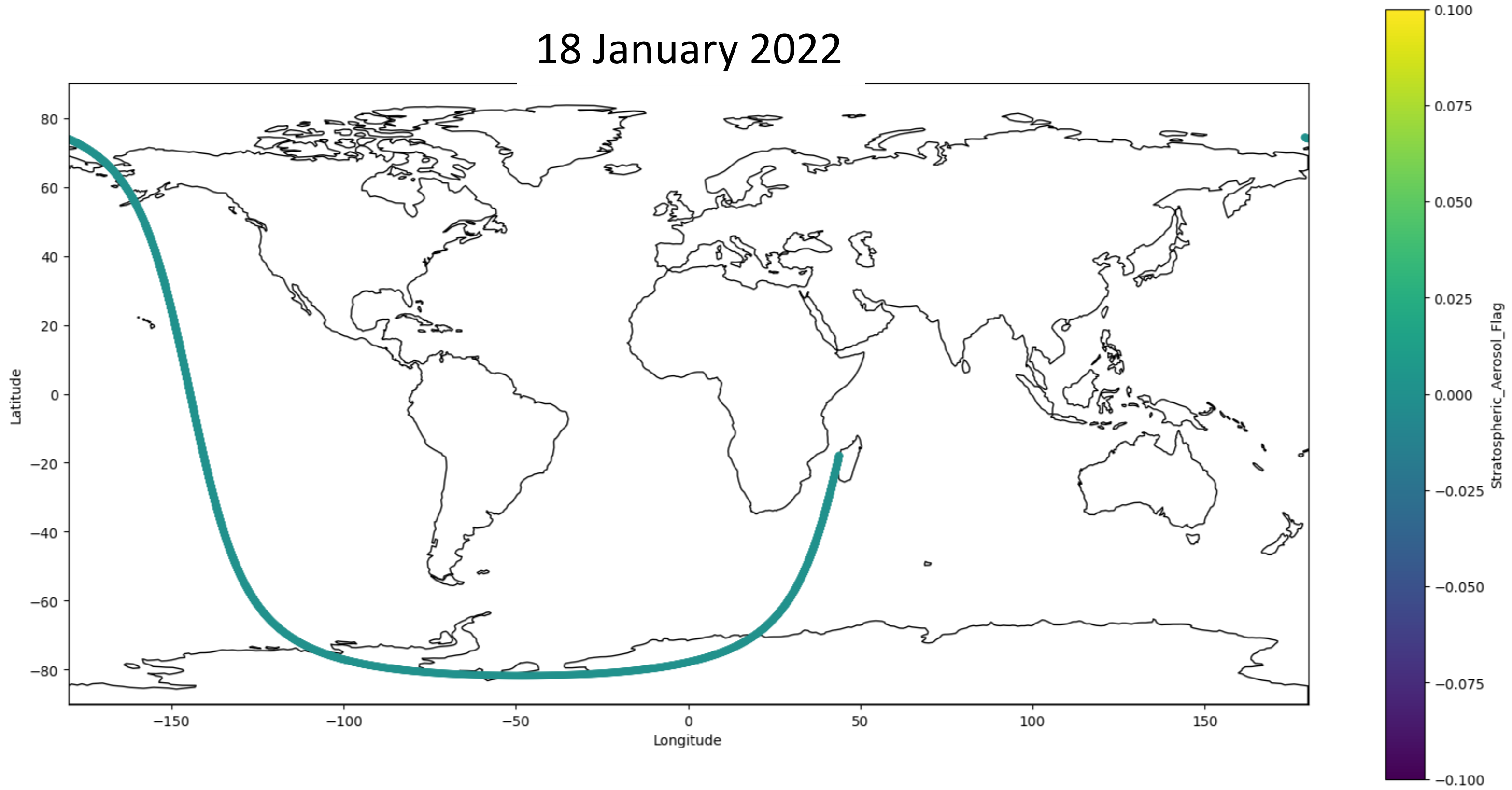


CALIPSO Prompts

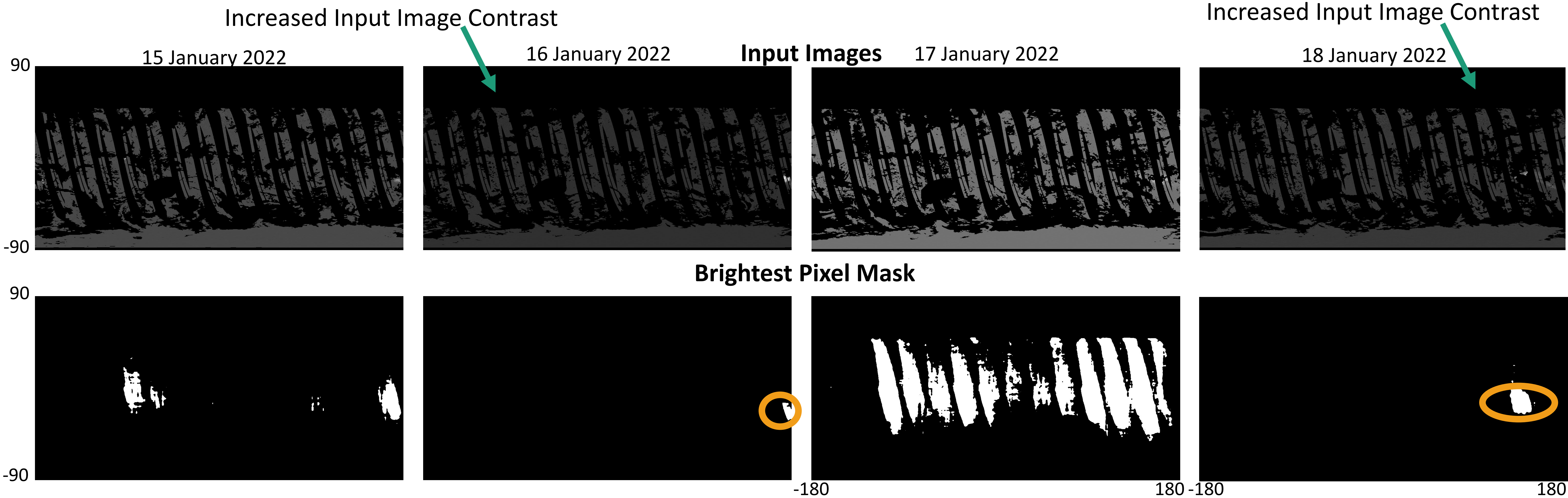




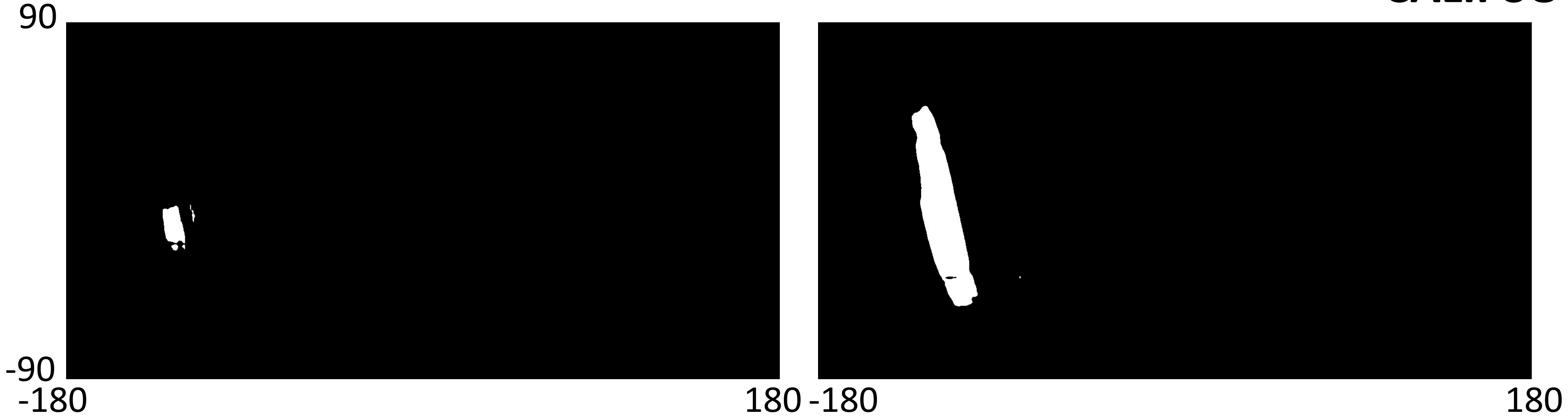




Predicting Plume Location From OMI L3 SO2 and CALIPSO



CALIPSO Prompt



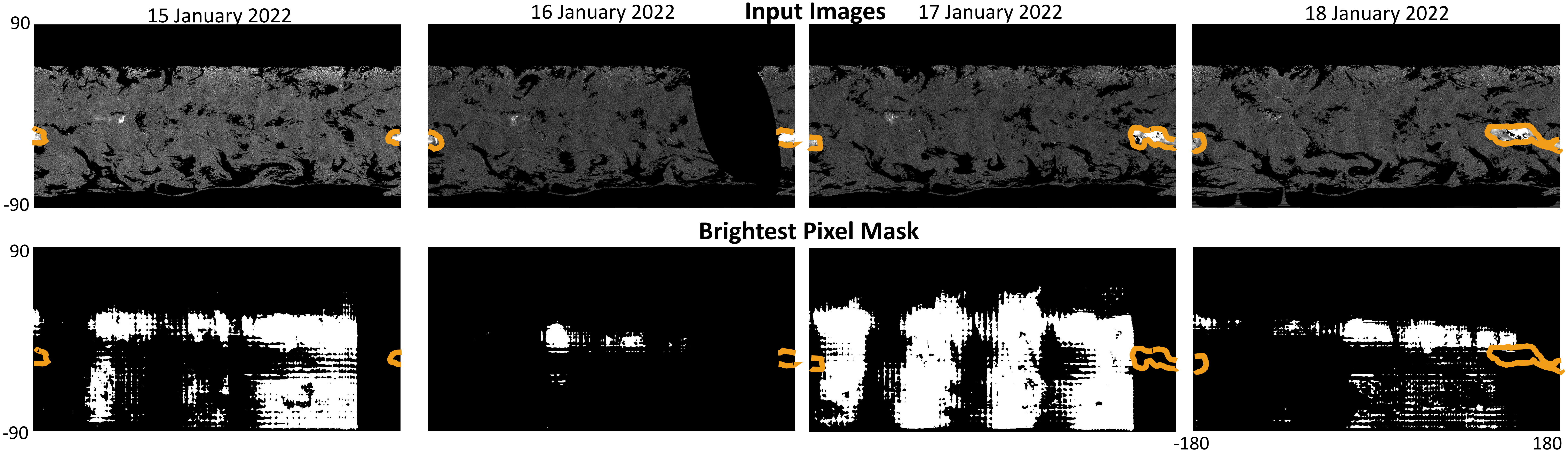
No CALIPSO/OMI Overlap

No CALIPSO Data 18 January 2022

- Improved performance for brightest pixel approach for stronger contrast on January 16 and 18

Predicting Plume Location from TROPOMI L2 SO2 and CALIPSO

Orange Outline is desired result from samgeo model



No CALIPSO Data 18 January 2022

- Poor performance for both approaches (Bright Pixel and Prompt)
- Desired result is orange outline region

Summary

- Novel use-case
  - Working on improving prompt engineering to setup a geophysical framework to predict the plume location in a systematic way.
- CALIPSO missing data due to the data dropout effects of the low laser power resulting from low canister pressure.
- Identification of features in an image and across multiple sensors
  - Use correlated geophysical variables to predict plume location
  - Use brightest pixels (simplest approach)
  - Use one variable to predict another (e.g., CALIPSO prompt for OMI SO2 concentration distribution)
  - Consider use of brightest pixels with layered raster images and prompts (more complex)
- Refining plume tracking technique lends to greater synergistic possibilities of future missions such as EarthCare and AOS/CALIGOLA

Future Work and Applications

Future Work

- Model performance depends on input data quality and availability.
- Improve understanding on implementing CALIPSO prompts
- Aim to integrate additional high-resolution datasets like MLS or TROPOMI stratospheric water vapor, OMPS stratospheric aerosol and GEO-LEO merged products and prompts such as ozone and water vapor soundings.
- Explore how spatial clustering can help select better input points.

Applications

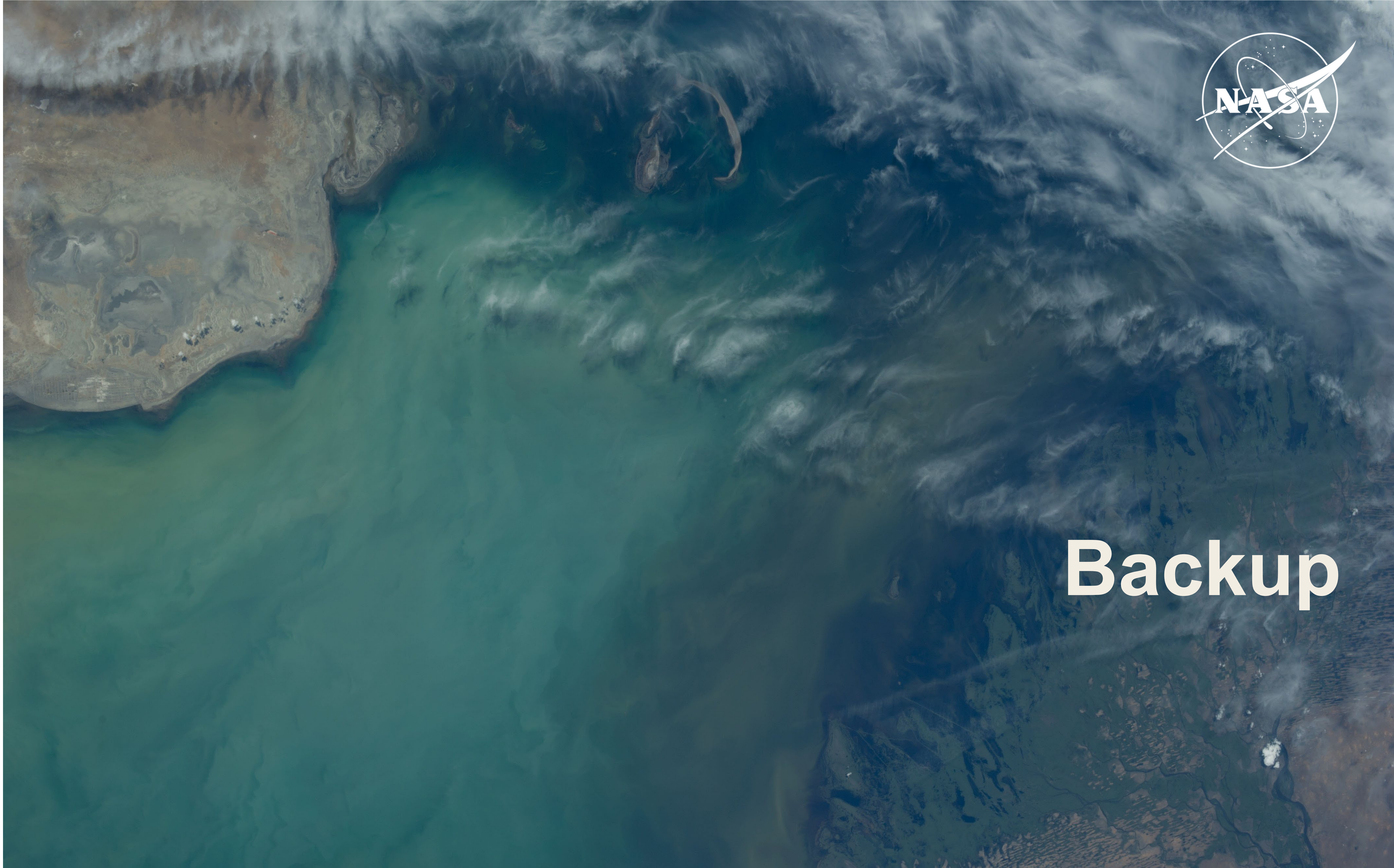
- Adapt this approach, codebase, and model can be adapted to study other disasters and extreme events.
- Additional volcanic eruptions like Mt. Ruang.
- Large-scale wildfires such as the 2020 California and Oregon wildfires.

References

- TROPOMI:
  - Copernicus Sentinel data processed by ESA, German Aerospace Center (DLR) (2020), Sentinel-5P TROPOMI Sulphur Dioxide SO2 1-Orbit L2 5.5km x 3.5km, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [***May 15, 2024***], 10.5270/S5P-74eidii
- OMI:
  - Can Li, Nickolay A. Krotkov, Peter Leonard, and Joanna Joiner (2020), OMI/Aura Sulphur Dioxide (SO2) Total Column 1-orbit L2 Swath 13x24 km V003,
    Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [***April 1, 2024***], 10.5067/Aura/OMI/DATA2022
- CALIPSO:
  - 10.5067/CALIOP/CALIPSO/CAL_LID_L2_VFM-Standard-V4-51
  - https://www-calipso.larc.nasa.gov/resources/calipso_users_guide/data_desc/cal_lid_l2_vfm_v4-51_desc.php
- MODIS:
  - Platnick, S., P. Hubanks, K. Meyer, and M. D. King, 2015: MODIS Atmosphere L3 Monthly Product (08_L3). NASA MODIS Adaptive Processing System, Goddard Space Flight Center
- samgeo:
  - Wu et al., (2023). samgeo: A Python package for segmenting geospatial data with the Segment Anything Model (SAM). Journal of Open Source Software, 8(89), 5663, https://doi.org/10.21105/joss.05663
  - https://samgeo.gishub.org/

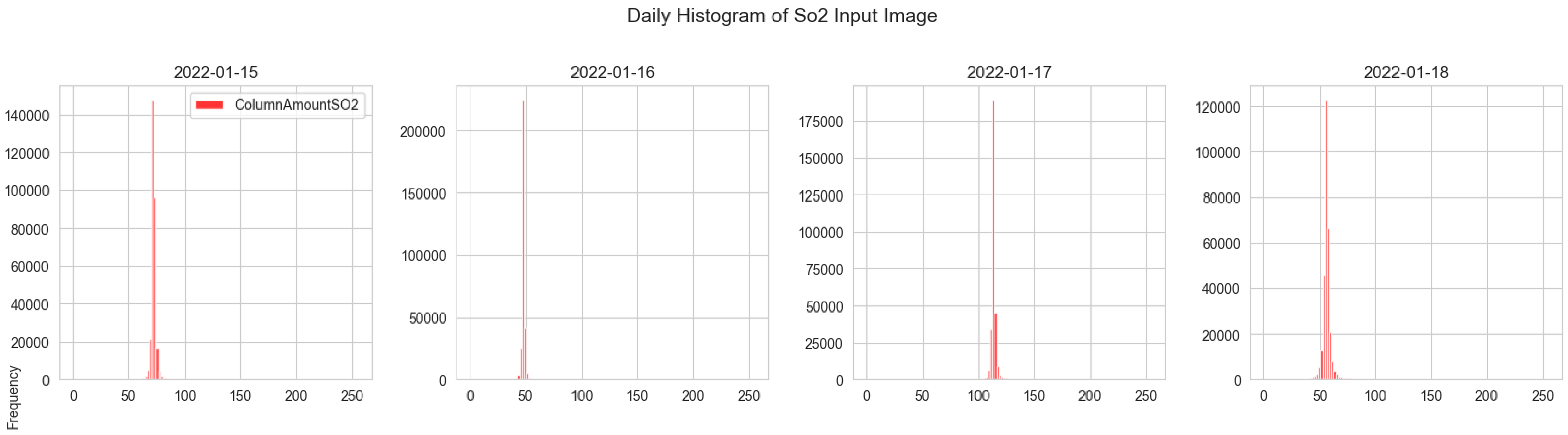
GODDARD EARTH SCIENCES

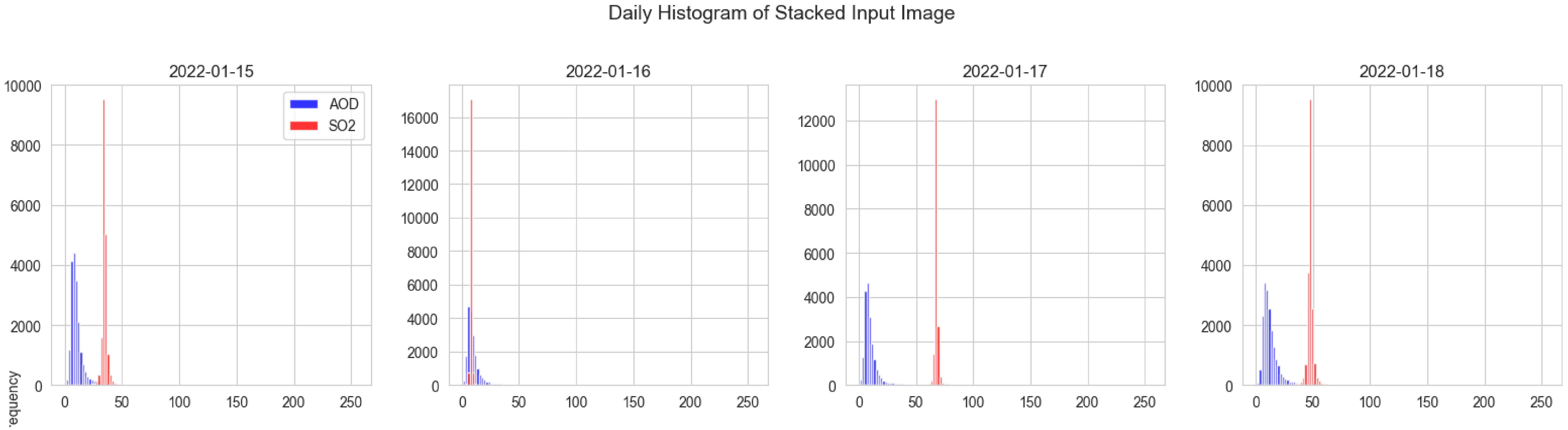

Thank you!



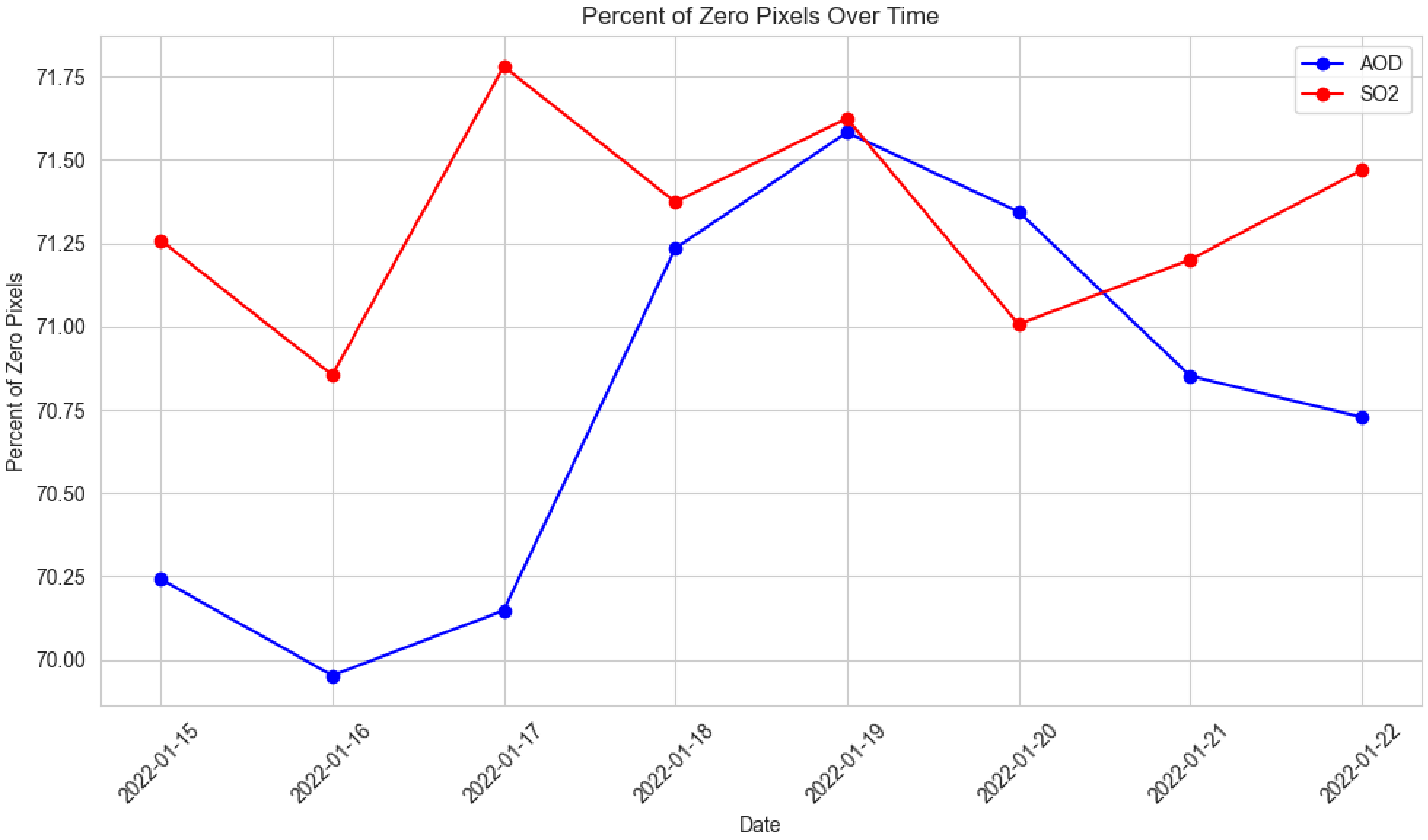
Backup

Evaluating Results





Evaluating Results



- Predictions were most accurate on January 16, 18, and 19.
- January 16 has the most complete data.
- January 18 and 19 have relatively complete data.