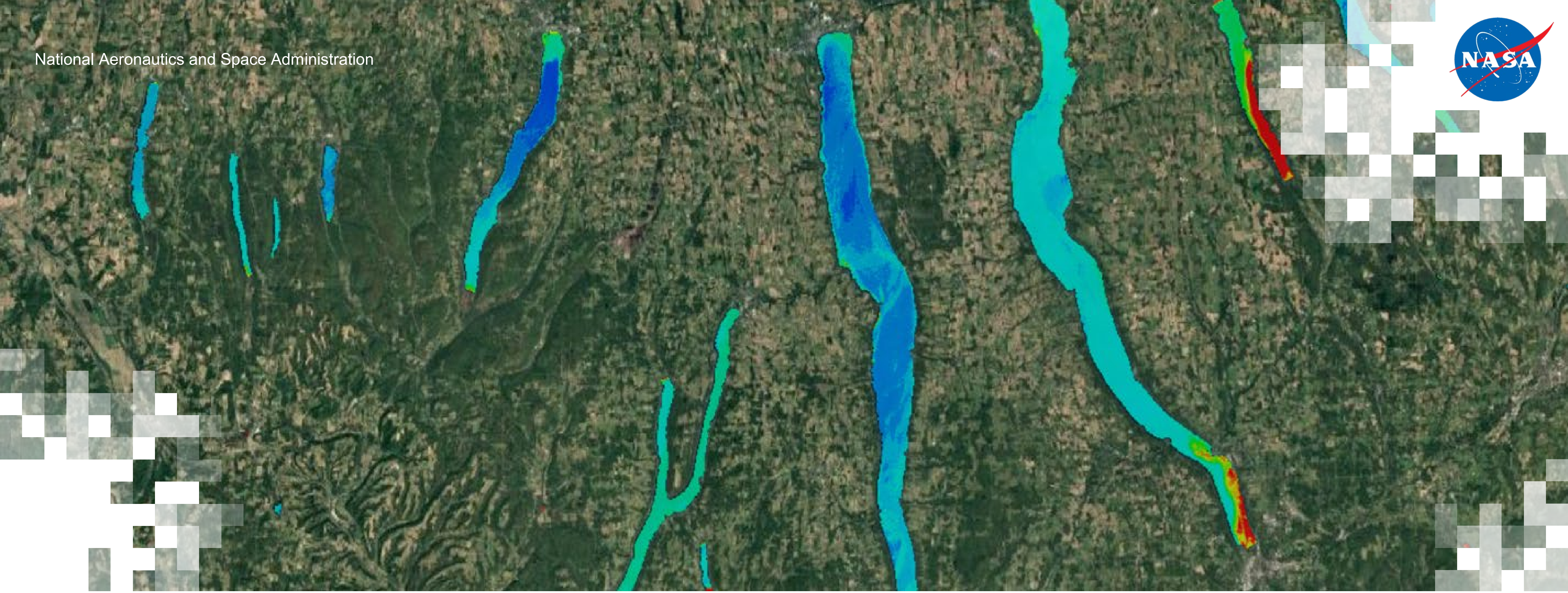
National Aeronautics and Space Administration



Monitoring Water Quality in Lakes and Coastal Regions Using STREAM

Part 2: Introduction to a Machine Learning Model to Estimate Water Quality Parameters Based on Satellite Observations

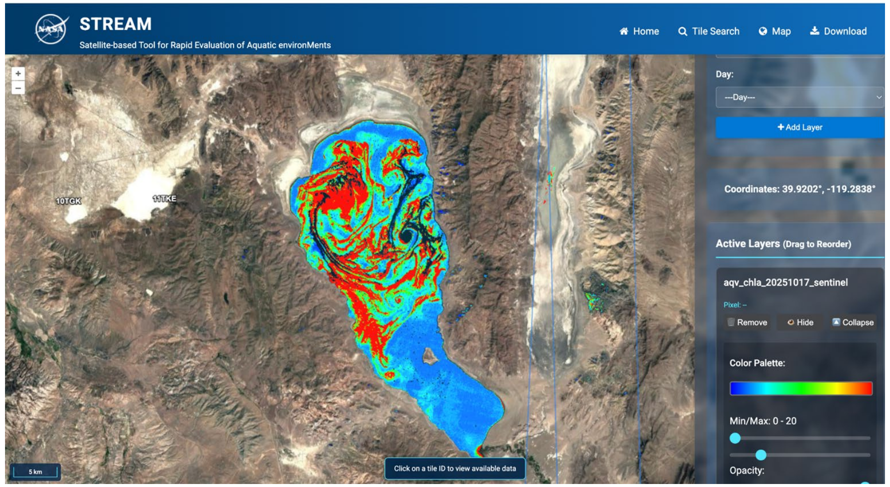
Amita Mehta (GESTAR II, 612), Ryan O'Shea (SSAI, 619)

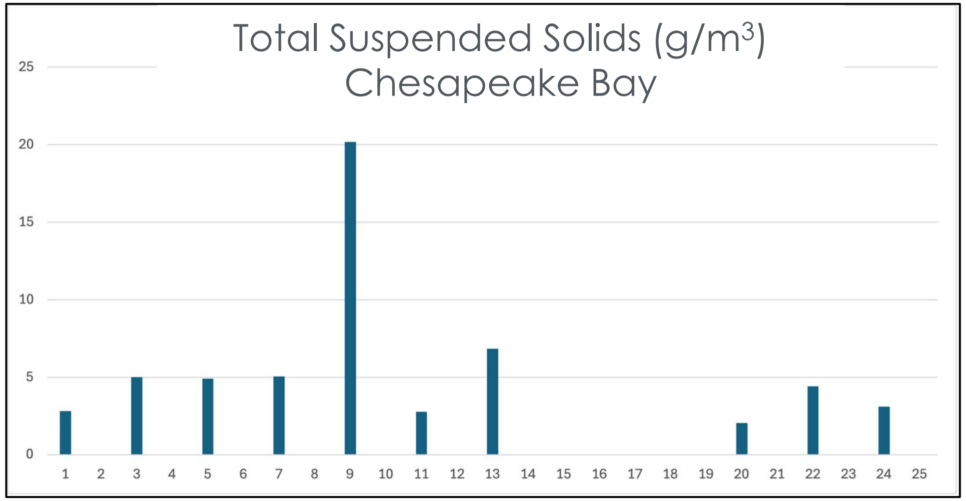
February 17, 2026

Part 1 Review

- Background, overview, and demonstration of STREAM web tool and API:
  - Uses Landsat 8 & 9, and Sentinel 2 a, b, and c to obtain water quality parameters, including chlorophyll a concentration, Total Suspended Solids, and Secchi Disk Depth in coastal estuaries and inland lakes in the US.
  - Provides the water quality parameters at 20 m to 30 m spatial resolution.
  - API allows search and download of multiple images of an area of interest.
- Example of selection of the water quality parameters using STREAM:
  - Map the water quality parameters.
  - Access and download multiple water quality data using the STREAM API.
  - Make time series of area-averaged water quality parameters using QGIS.

Chlorophyll-a concentration in Pyramid Lake
From STREAM (Sentinel 2/MSI)





Training Outline

Part 1

Introduction and Demonstration of STREAM

February 10, 2026

Part 2

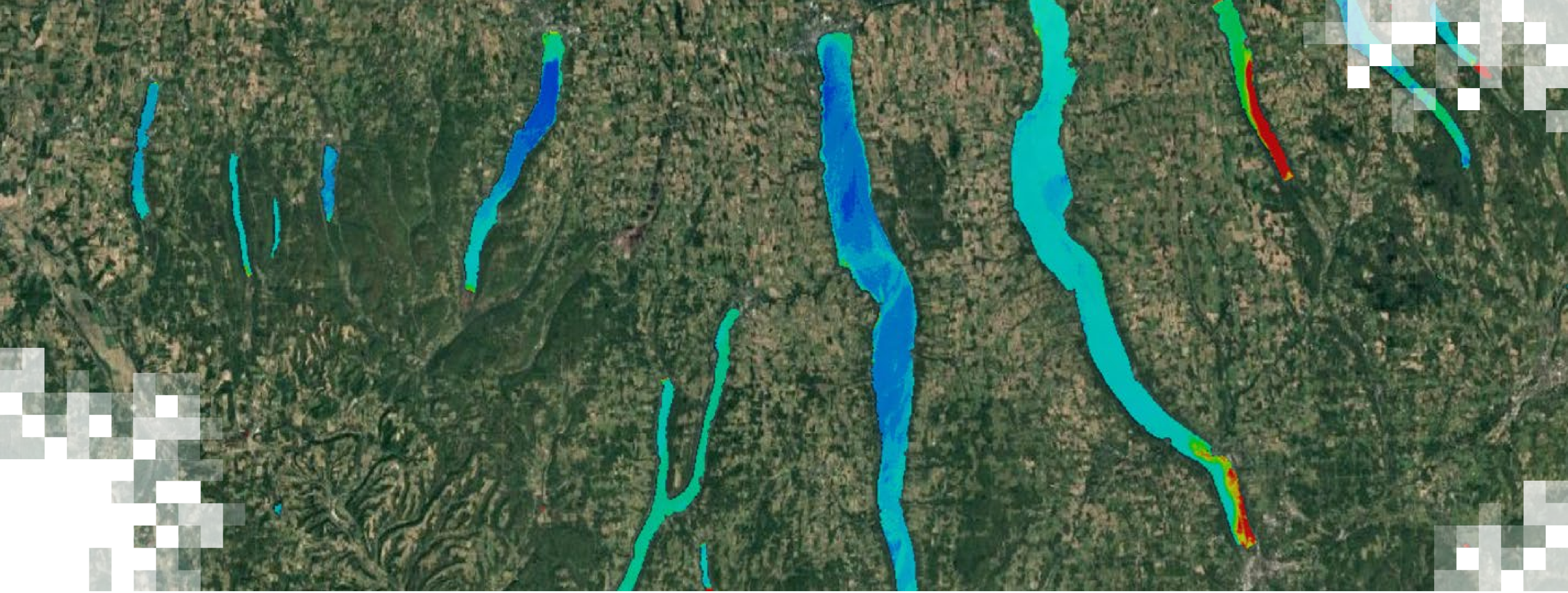
Introduction to a Machine Learning Model to Estimate Water Quality Parameters Based on Satellite Observations

February 17, 2026

Homework

Opens February 17 – Due March 10 – Posted on Training Webpage

A certificate of completion will be awarded to those who attend all live sessions and complete the homework assignment(s) before the given due date.



Monitoring Water Quality in Lakes and Coastal Regions Using STREAM

Part 2: Introduction to a Machine Learning Model to Estimate Water Quality Parameters Based on Satellite Observations

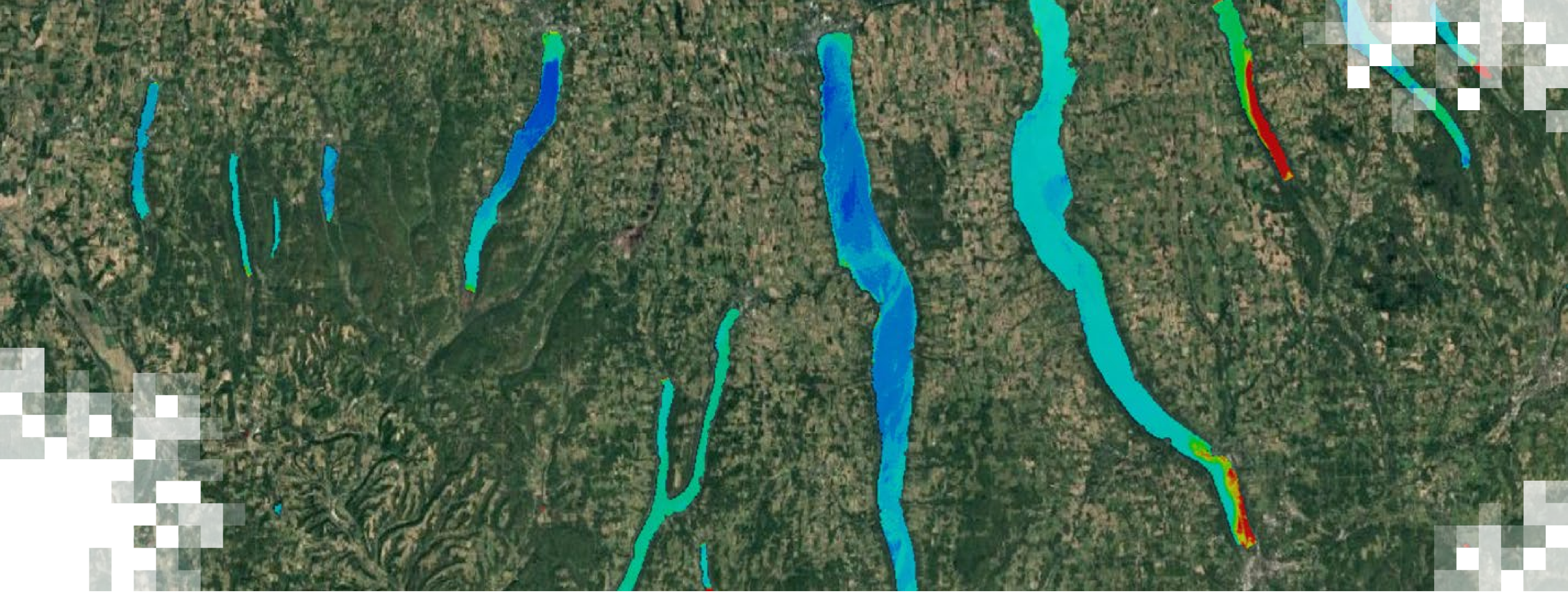
Part 2 Objectives

By the end of Part 2, participants will be able to:

- Become familiar with a Mixture Density Network (MDN) based model used for deriving water quality parameters from satellite observations in STREAM.
- Recognize how to select and process satellite images for deriving water quality parameters in a waterbody of interest.
- Recognize how to access the MDN model code for deriving water quality parameters in the waterbody of interest.

Part 2 Outline

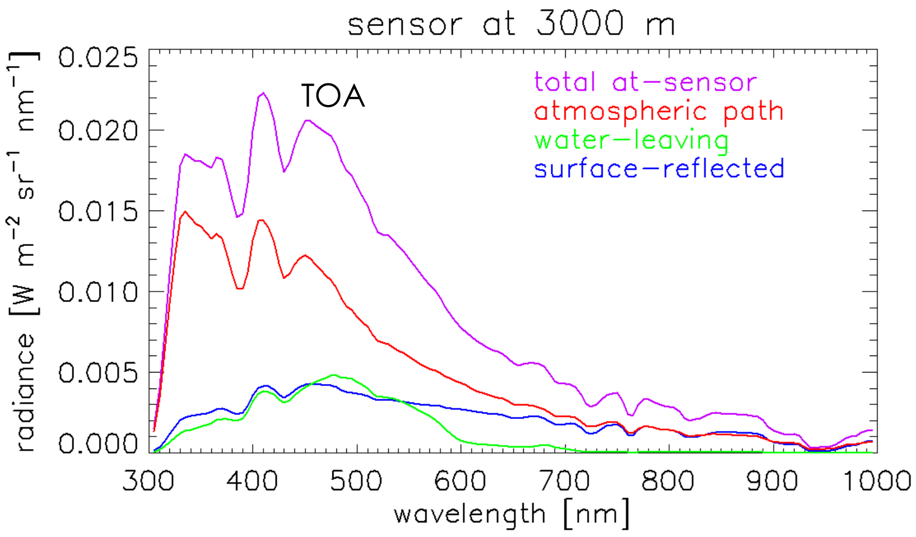
- Pre-processing satellite images for inputs to the Mixture Density Network (MDN)-based model
- Overview of the MDN-based model used in STREAM
- Demonstration: Access and apply the MDN model code for deriving water quality parameters from satellite data
  - Examples: water quality parameters in San Francisco Bay using Sentinel-3 OLCI images and in Lake Erie using PACE OCI

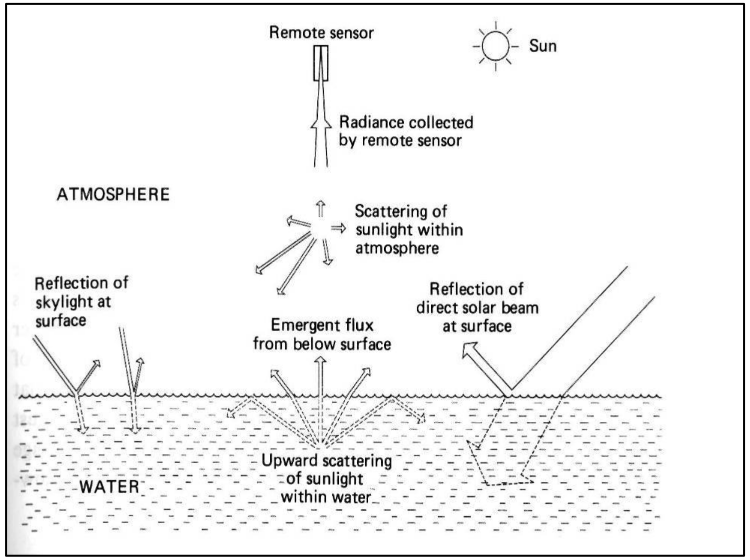


Pre-Processing Satellite Images for Inputs to the MDN-Based Model

Satellite Measurements for Water Quality

- Satellite sensors measure top-of-atmosphere (TOA) radiances.
- The TOA radiances result from a combination of surface and atmospheric conditions, including the effects of clouds and aerosol particles.
- Water-leaving reflectance depends on backscattering and absorption of radiation due to materials in water that is optically active.
- Monitoring water-leaving reflectance from surface is used to estimate the quality of the water.





Pre-processing Satellite Images for MDN: Atmospheric Correction

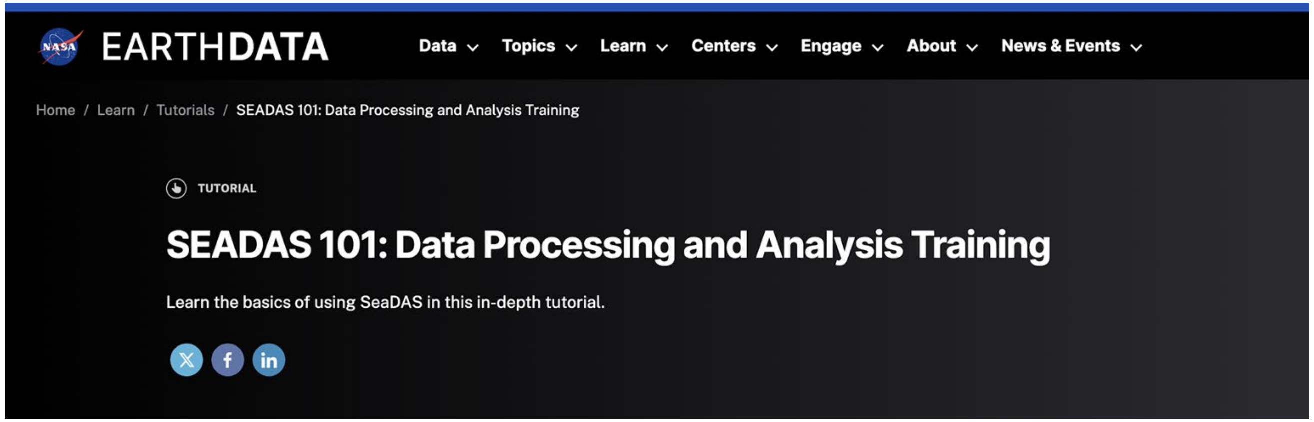
- Satellite observations of TOA radiances must be corrected for atmospheric effects for acquiring water-leaving or remote sensing reflectances (Rrs_{λ}).
- Various algorithms exist for the atmospheric corrections:
  - NASA Ocean Biology Processing Group Algorithm
  - Remote Sensing and Ecosystem Monitoring (REMSEM): ACOLITE
  - HYGEOS Polymer
  - NASA Aquaverse
- Level-1 TOA_{λ} reflectance from satellites are used.
- Atmospherically corrected water-leaving, remote sensing reflectances Rrs_{λ} are Level-2 data and are used to estimate water quality parameters.

Pre-processing Satellite Images for MDN: Water-leaving Reflectances from TOA Reflectances using SeaDAS/OCSSW

1. Obtain Level 1 satellite Data.
2. Use NASA Sea, earth, atmosphere Data Analysis System (SeaDAS), Ocean Color Science Software (OCSSW) to get Level-2 water-leaving remote sensing reflectances.
3. The remote sensing reflectances are used as input to the MDN model.

The following ARSET Trainings describe using SeaDAS/OCSSW to get remote sensing reflectances:

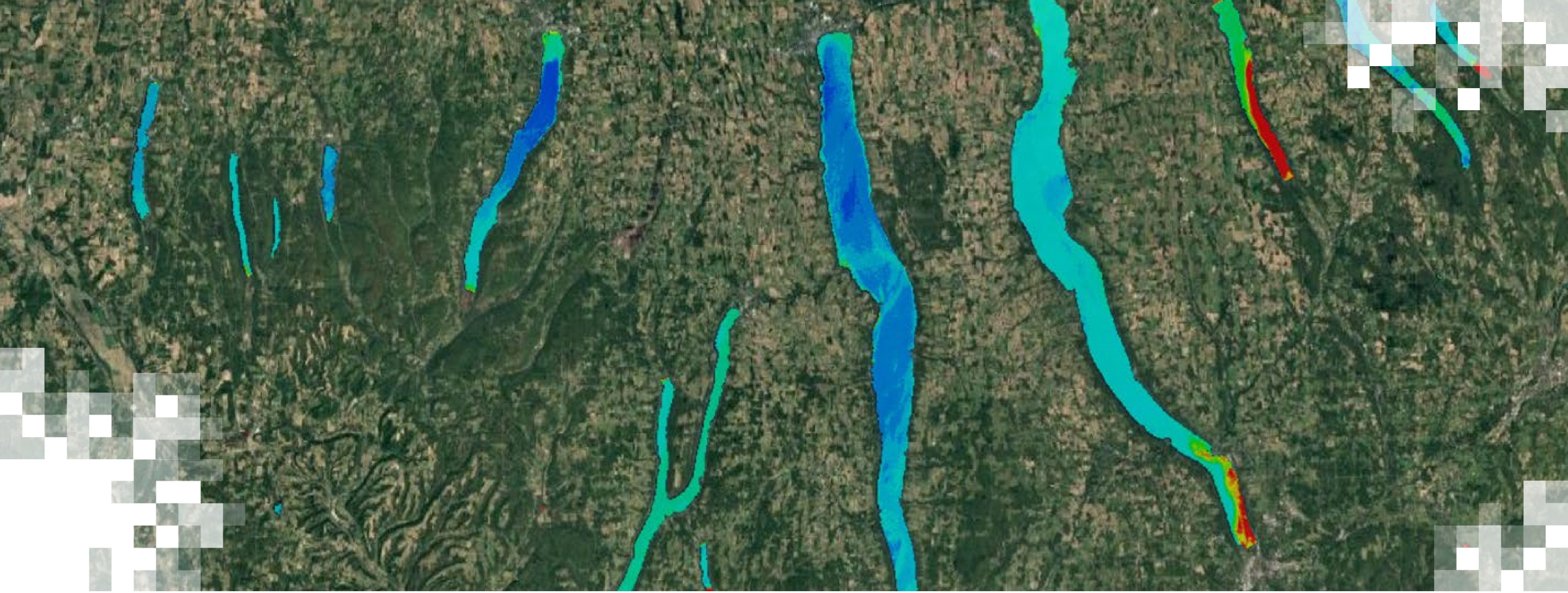
- Part 2: Image Processing using SeaDAS
- Overview of SeaDAS 8.4.1 for the Processing, Analysis, and Visualization of Optical Remote Sensing Data for Water Quality Monitoring



SeaDAS Tutorial

Where to Get Level-1 Data?

Satellites Orbits	Sensors	Level-1 Data Access
Landsat 8 Landsat 9 polar	TIRS, OLI TIRS-2, OLI-2	USGS Earth Explorer
Terra Aqua Polar	MODIS	NASA Ocean Data
SNPP JPSS polar	VIIRS	NASA Ocean Data
Sentinel 2A Sentinel 2B & 2C Polar	MSI	Copernicus Browser
Sentinel 3A & 3B Polar	OLCI	Copernicus Browser
PACE Polar	OCI	NASA Ocean Data



Overview and Demonstration
Mixture Density Network (MDN) Model for STREAM

How to Ask Questions

- Please put your questions in the Questions box and we will address them at the end of the webinar.
- Feel free to enter your questions as we go. We will try to get to all of the questions during the Q&A session after the webinar.
- The remainder of the questions will be answered in the Q&A document, which will be posted to the training website about a week after the training.

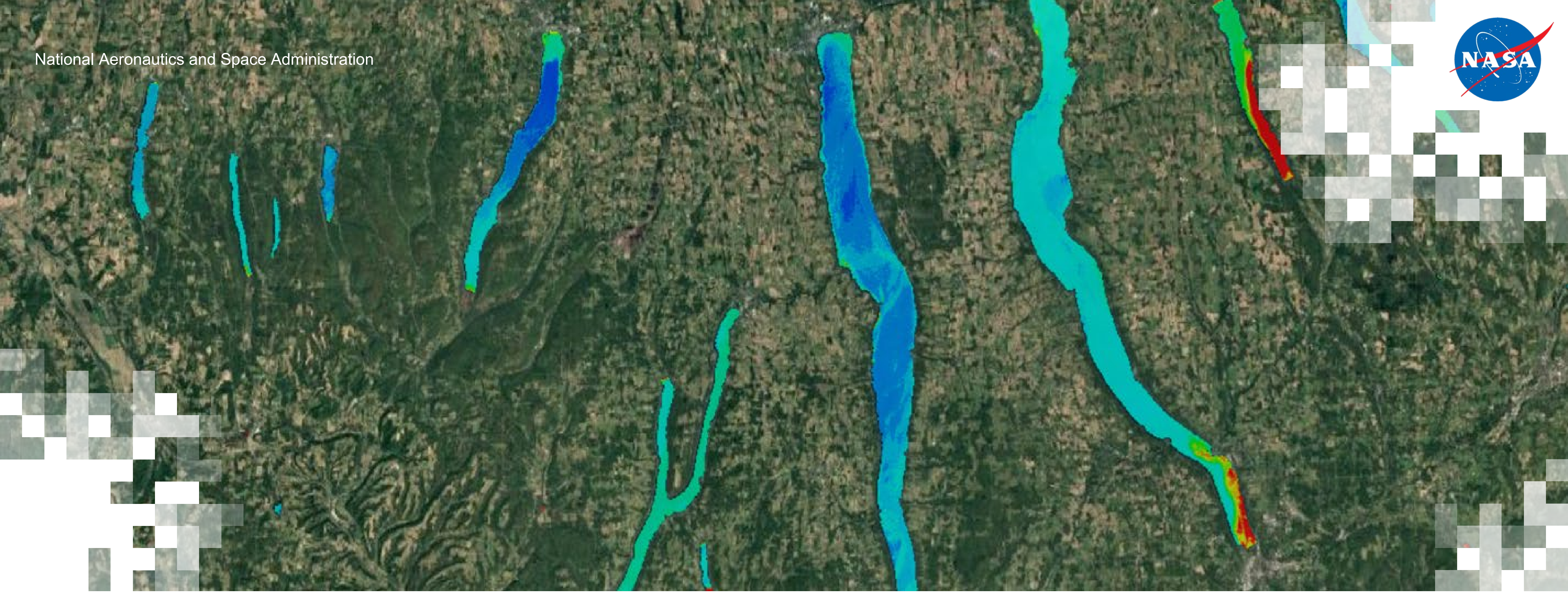
Part 2 – Trainer

Dr. Ryan O'Shea

Senior Research Scientist

SSAI, 619, NASA-GSFC

National Aeronautics and Space Administration



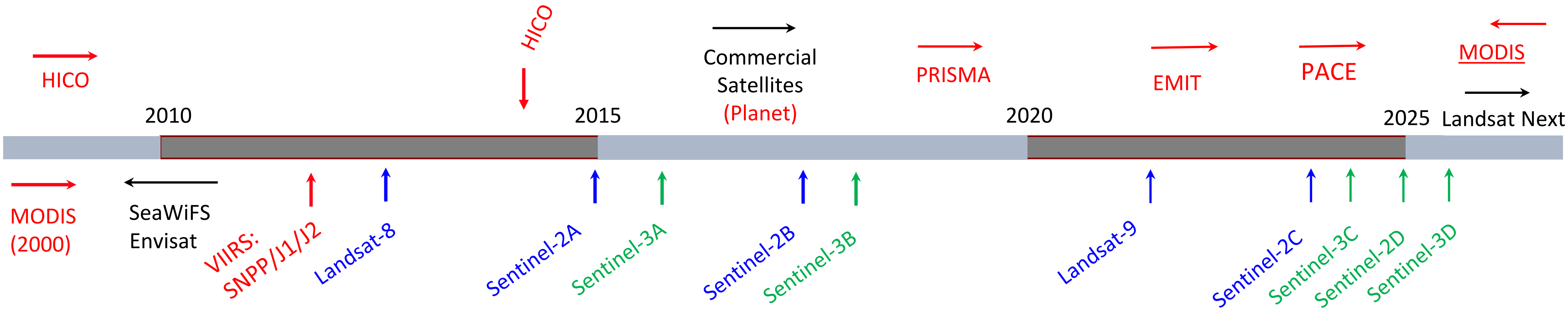
Monitoring Water Quality in Lakes and Coastal Regions Using STREAM

Part 2: Introduction to a Machine Learning Model to Estimate Water Quality Parameters Based on Satellite Observations

Ryan E. O'Shea[1,2], Arun M. Saranathan[1,2], Akash Ashapure[1,2], William Wainwright[1,2], & Brandon Smith[1,2]

February 17, 2026

Freshwater and Littoral Imaging from Space: Mixture Density Network (MDN) Supported Sensors



- Radiometrically capable sensors
- Daily observations
- Different spatial/spectral capabilities

Robust workflow for producing seamless products over fresh and coastal estuaries

Satellite data available:

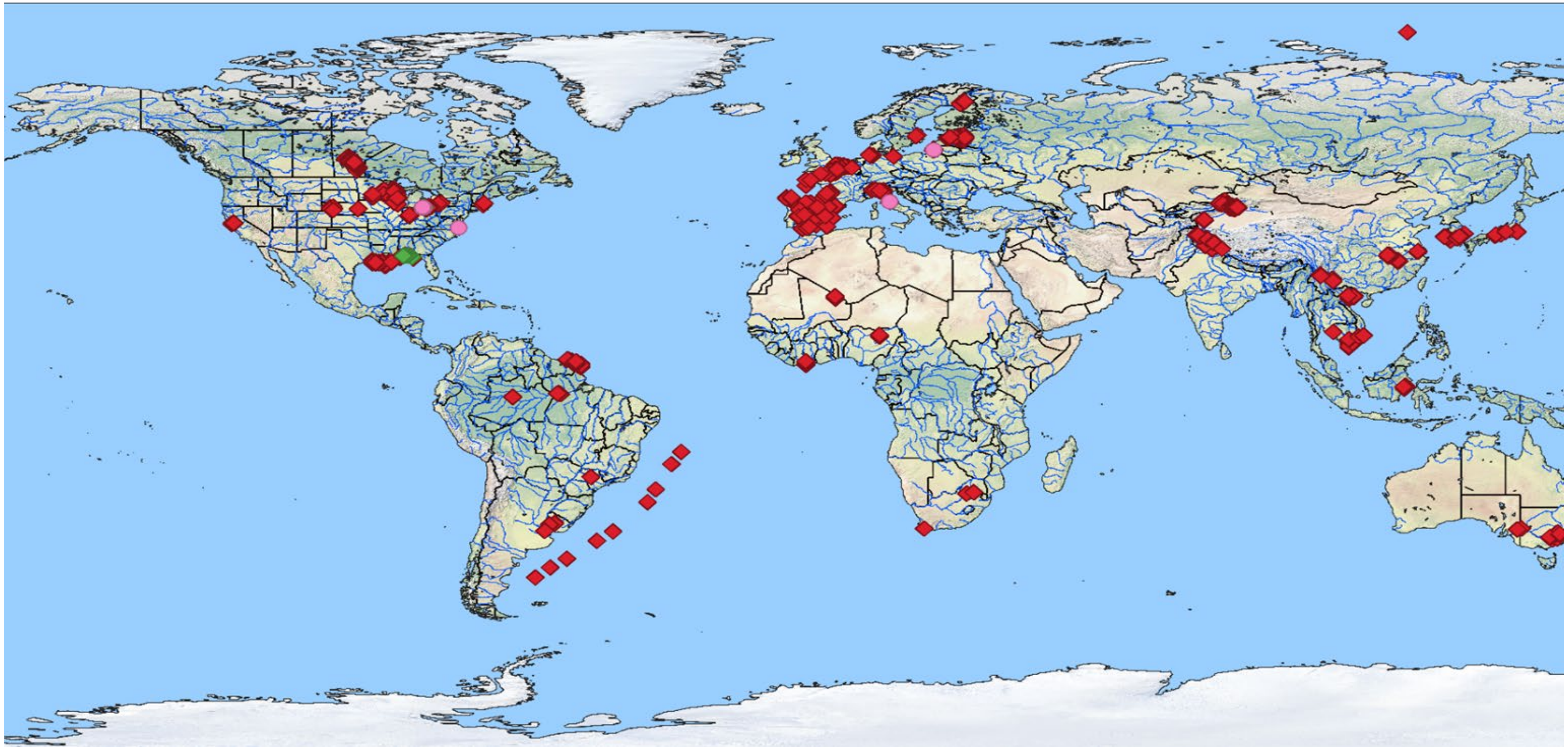
Coming Soon
- SeaWiFS …..

Just MDN
- MODIS …..

STREAM
- Landsat-8

Coming Soon to STREAM
- Sentinel-3A ….

To Train a Globally Applicable Machine Learning Database, We Need a Representative Dataset:

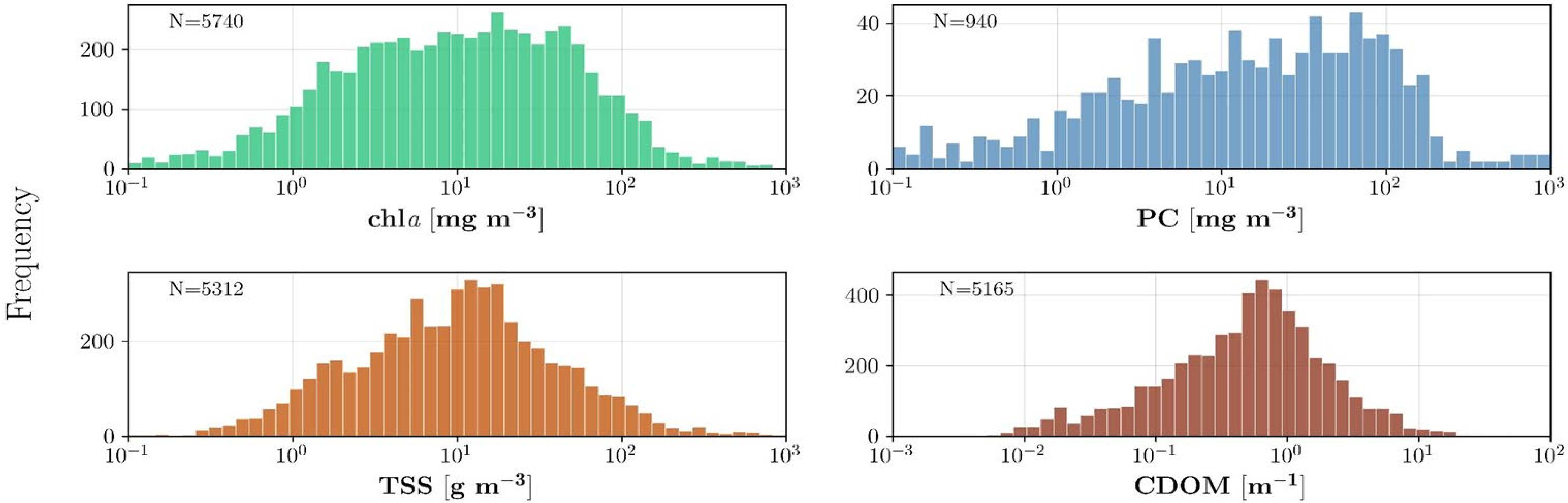


Training data: Large (N>~10,000) Globally Distributed In situ Dataset[1,2]:
Inherent Optical Properties (IOPs) & biogeochemical parameters from coastal and inland waters

[1]Lehmann, M.K., et al. **GLORIA - A globally representative hyperspectral in situ dataset for optical sensing of water quality**. *Scientific Data*. https://doi.org/10.1038/s41597-023-01973-y.

[2]*O'Shea et al. 2023 (Remote Sensing of Environment)*

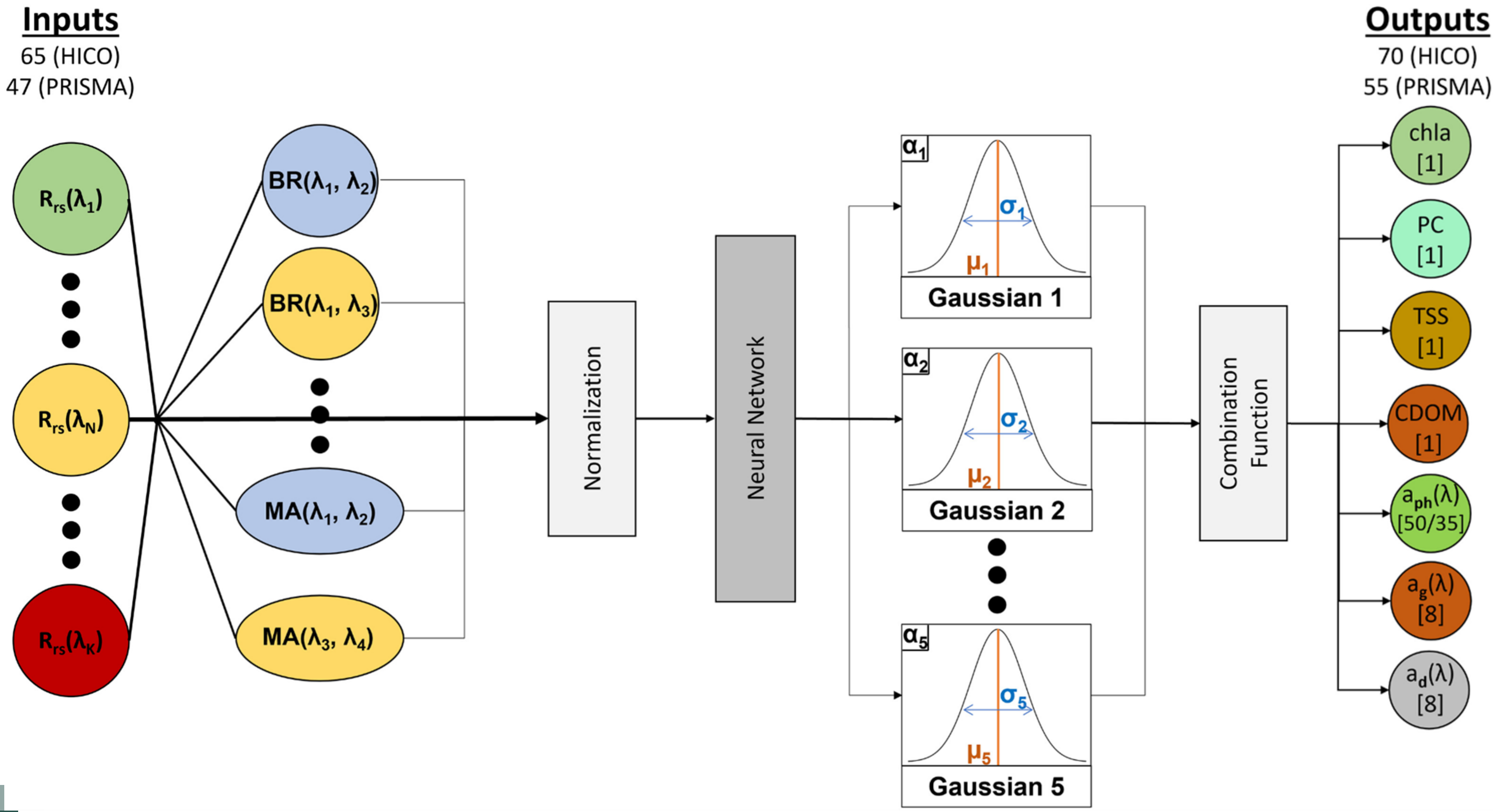
Aquaverse Training Data: In Situ Chla, PC, TSS, & CDOM Each Span ~4 Orders of Magnitude Concentration Range with Thousand(s) of Samples



Chla: Chlorophyll-a, PC: Phycocyanin, TSS: Total Suspended Solids, CDOM, Colored Dissolved Organic Matter

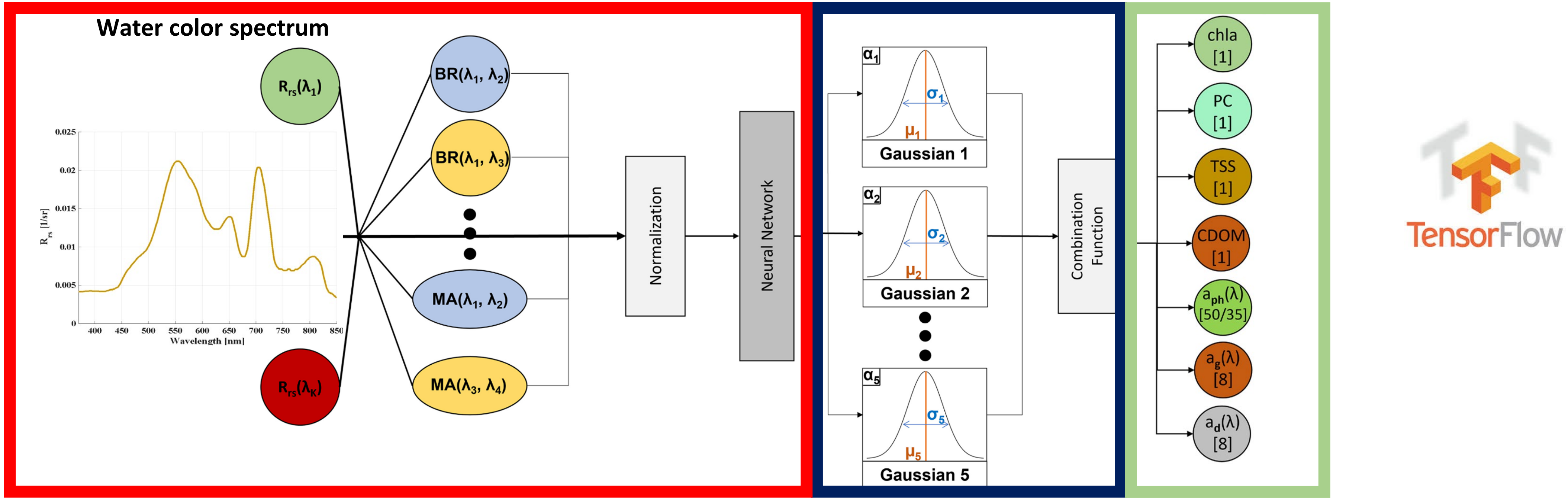
O'Shea et al. 2023 (Remote Sensing of Environment)

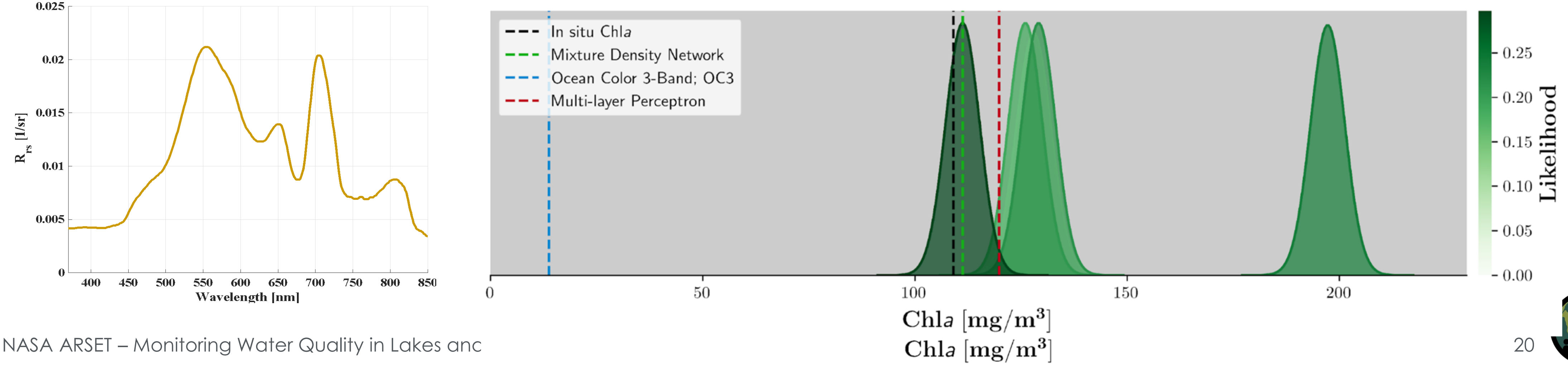
Training data: Large (N>~10,000) Globally Distributed In Situ Dataset[1]: IOPs & Biogeochemical Parameters from Coastal and Inland Waters



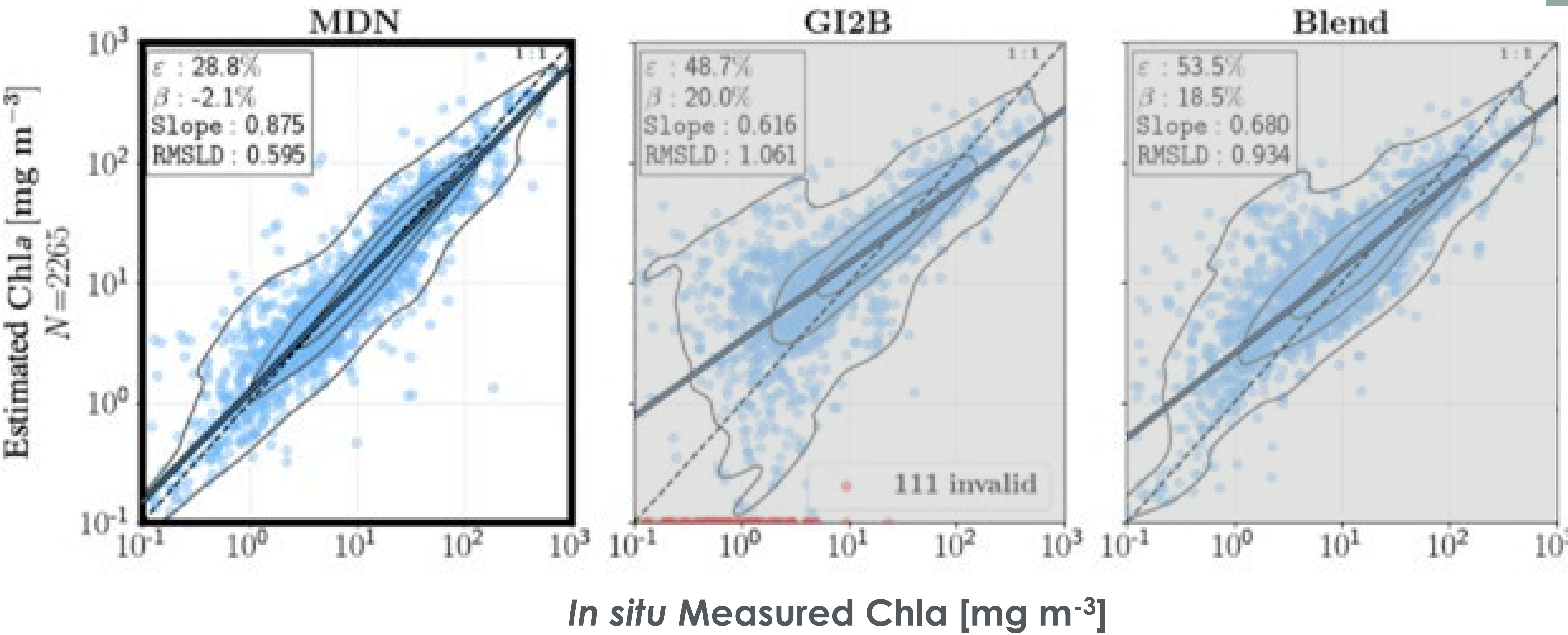
Aquaverse Core Algorithm: Mixture Density Networks (MDNs)

Non-unique inverse problem





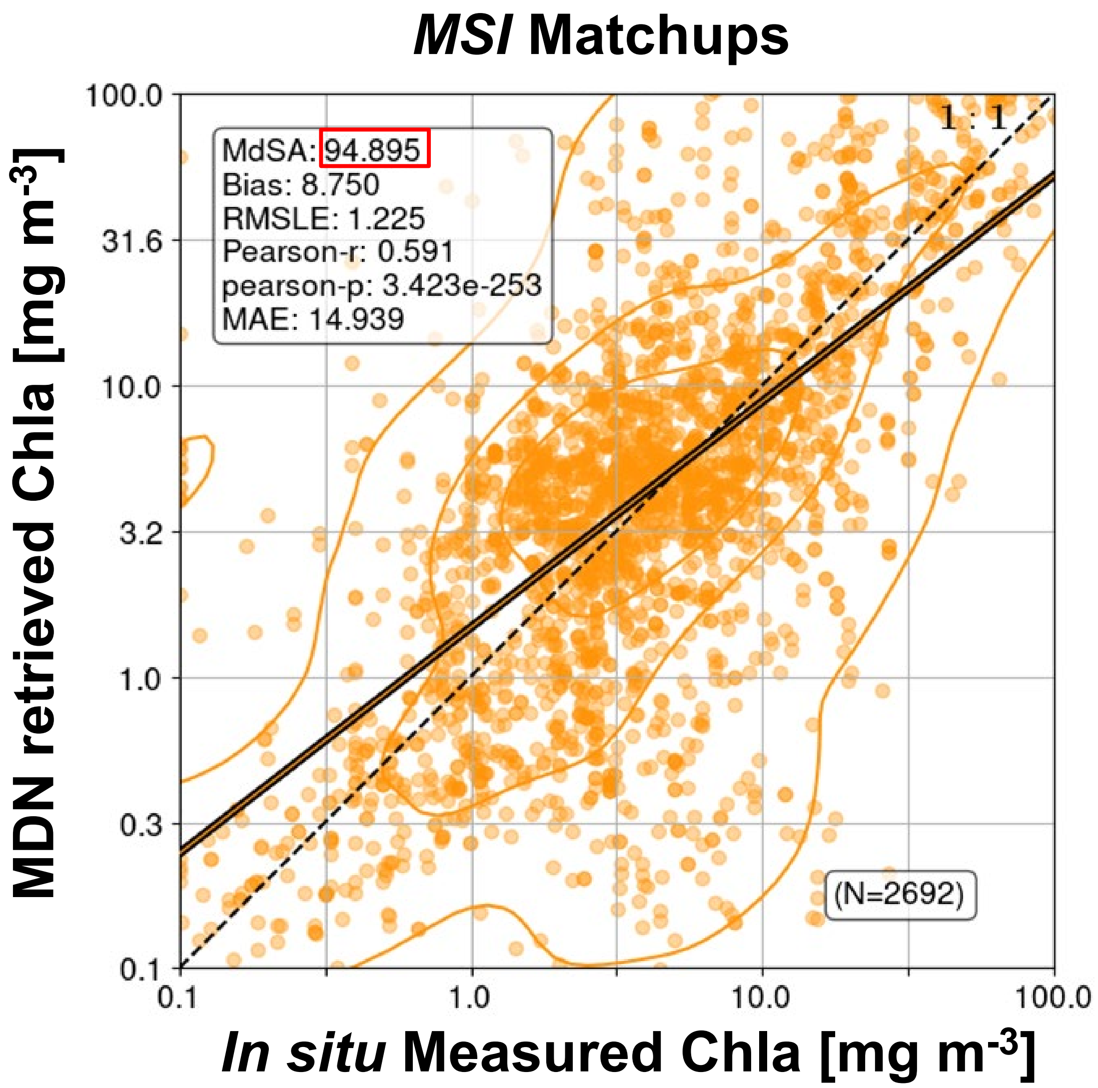
The 50/50 Training Split of In Situ Data Shows Ideal Model Performance



MDN outperforms operational algorithms in ideal case, across a wide dynamic range!

Pahlevan et al. RSE, 2022

STREAM MSI Global Chlorophyll-a Validation

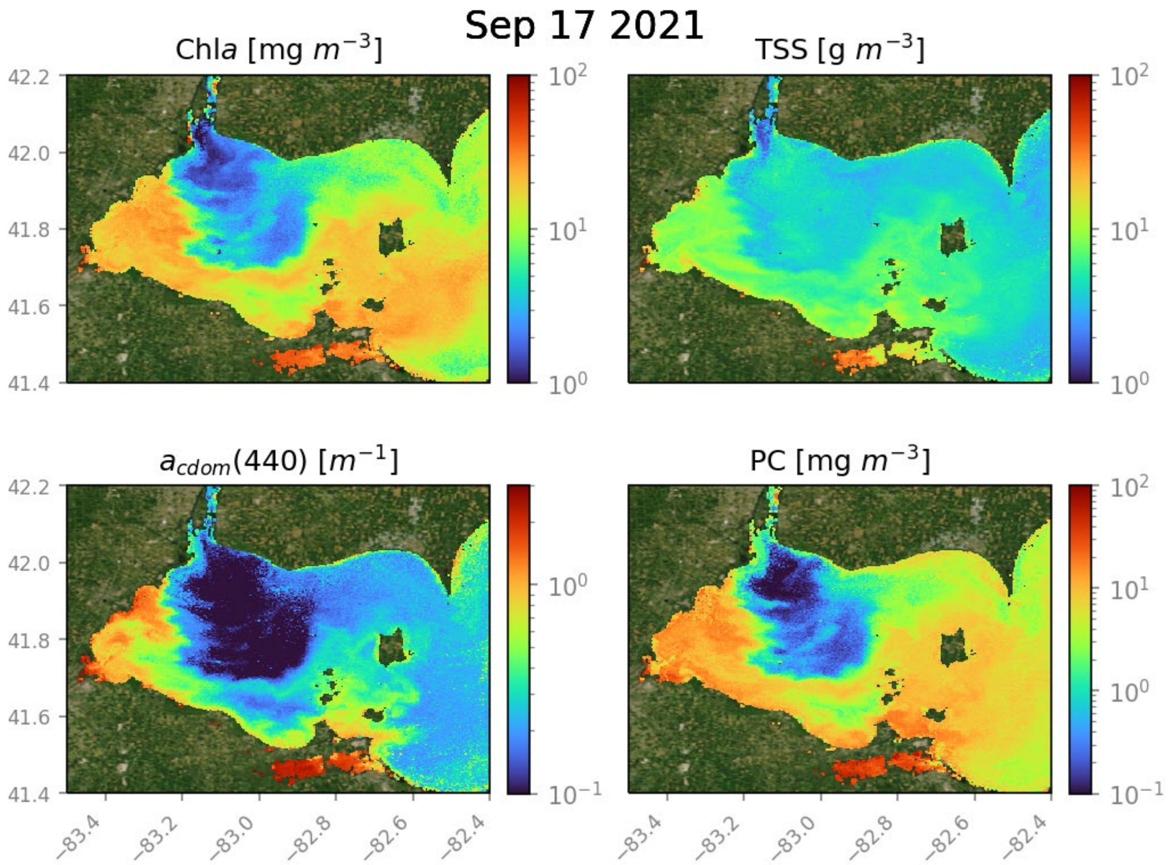


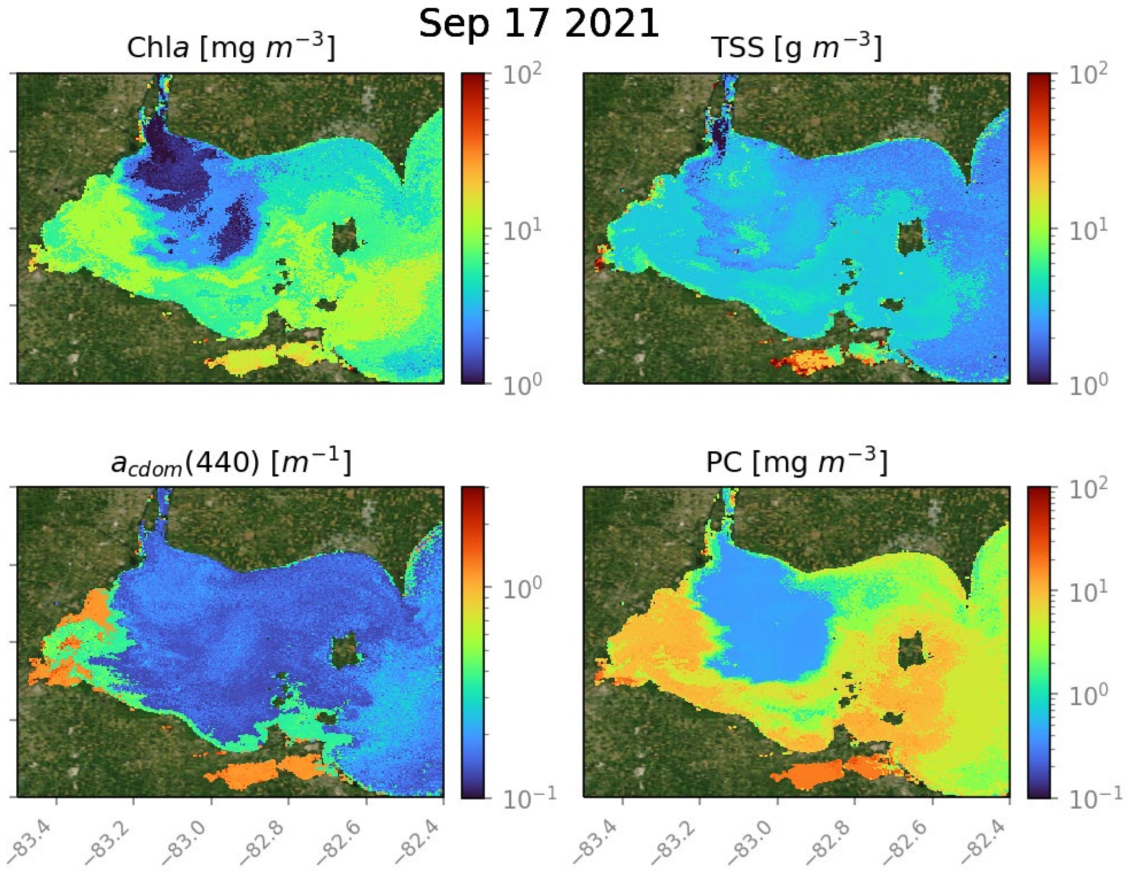
Saranathan et al. 2023:
https://ieeexplore.ieee.org/abstract/document/11243435

Uncertainties from atmospheric correction substantially reduce accuracy!

MDN Derived Products for Lake Erie HAB are Spatially Consistent

OLCI



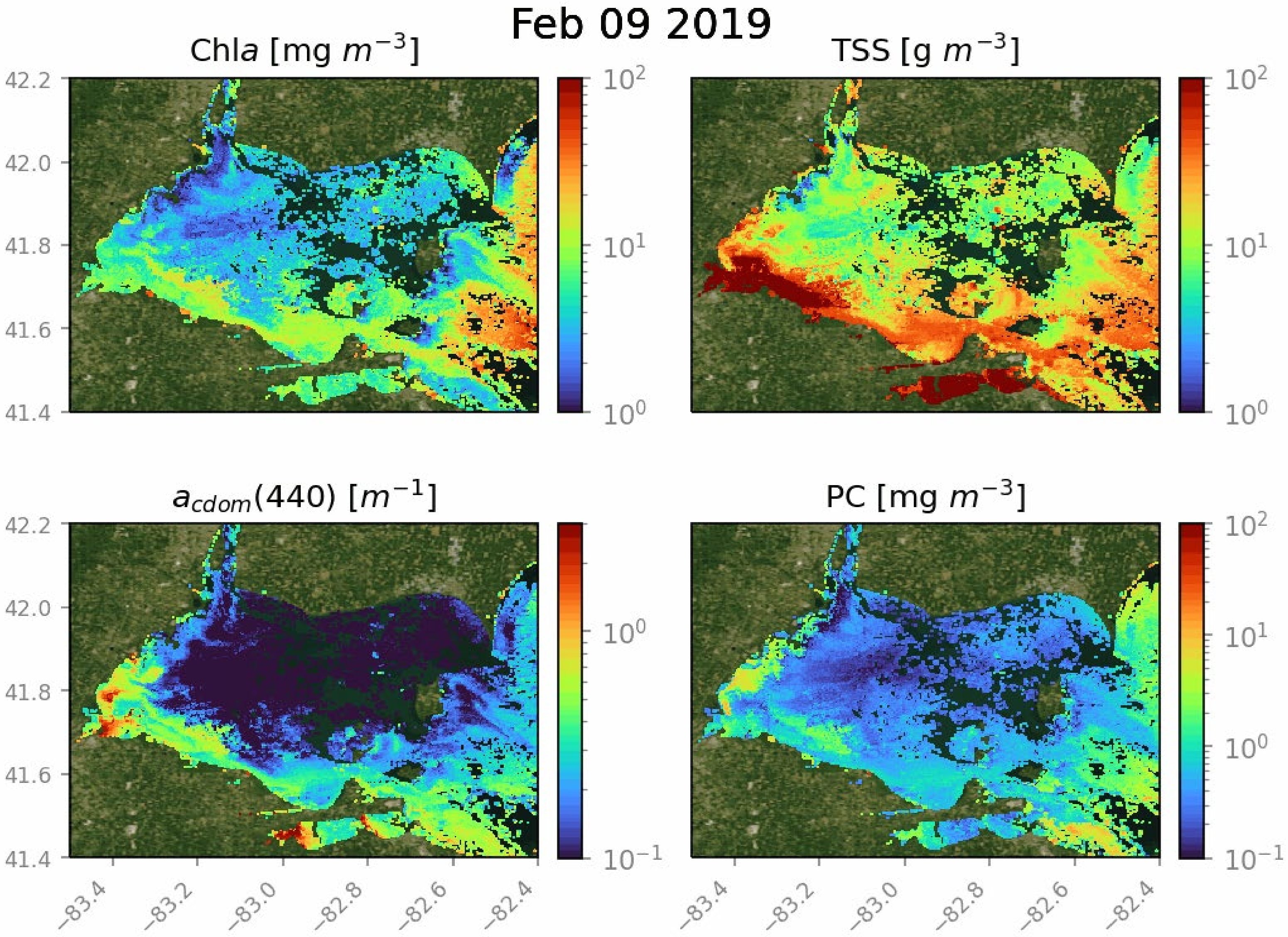


Uncertainties demonstrate models' confidence in their own estimates.

SeaDAS-processed imagery

MDN Derived Products for Lake Erie HAB Development are Temporally Consistent

OLCI



Saranathan et al. 2023: https://ieeexplore.ieee.org/abstract/document/11243435

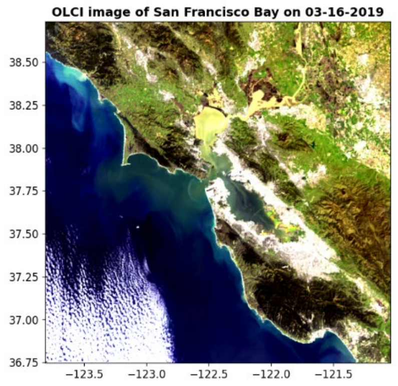
Tutorials for Applying MDNs to Satellite Imagery

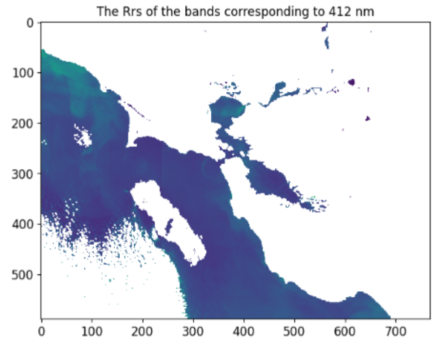
SC_3_MDNs_with_satellite_imagery.ipynb

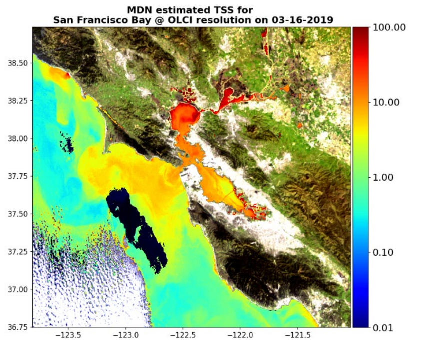
1. **Import package functions**
2. **Visualize pre-corrected (via acolite) OLCI image**
   - *display_sat_rgb*
   - *User's need to supply AC image (e.g., l2gen)*
3. **Retrieve Rrs from AC tile**
   - *Pull Rrs from l2gen, acolite, or polymer corrected tiles*
4. **Generate predictions from Rrs**
   - *Predictions should be comparable to published maps*
5. **Generate predictions & uncertainties from Rrs**
   - *Identify regions with higher uncertainty (near clouds and land, and outlier estimates)*
6. **Generate retrievals from l2gen AC PACE imagery**
   - *Image is from 05/31/25, prior to cyanobacteria bloom formation (~06/12/24)*

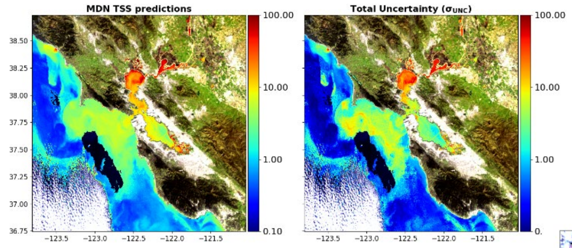
Section 1: MDN Package Import and Verification

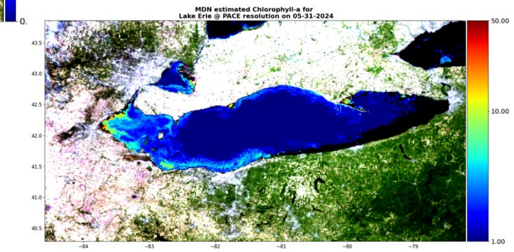
```
'Import the base python packages needed for this notebook'
import numpy                                                as np
import os
from   pathlib                  import Path
from   urllib.request           import urlretrieve
import zipfile
import matplotlib                                           as mpl
import matplotlib.pyplot                                    as plt
from   matplotlib.colors        import LogNorm

'Import some MDN specific functions and packages'
from   MDN                      import get_sensor_bands, get_tile_data
from   MDN                      import current_support
from   MDN                      import get_gloria_trainTestData, get_args
from   MDN                      import get_mdn_preds, create_scatterplots_trueVsPred, performance
from   MDN                      import OC, resample_Rrs
```



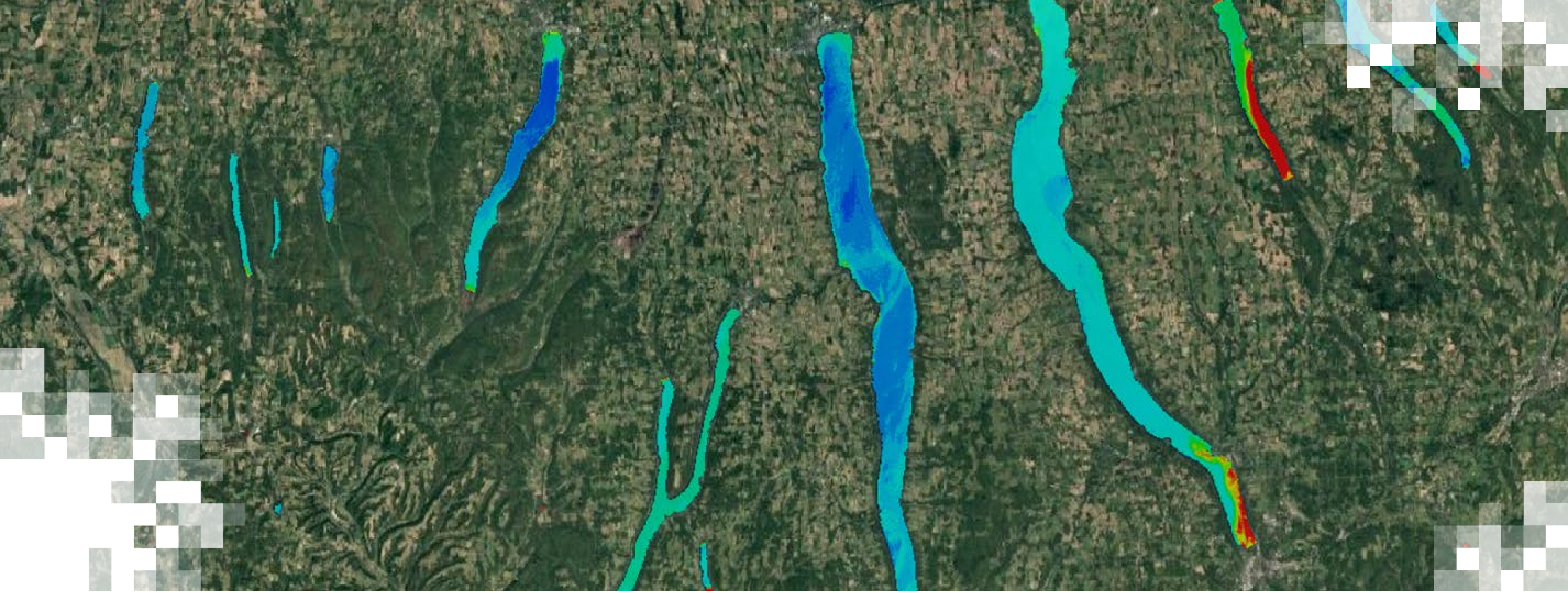




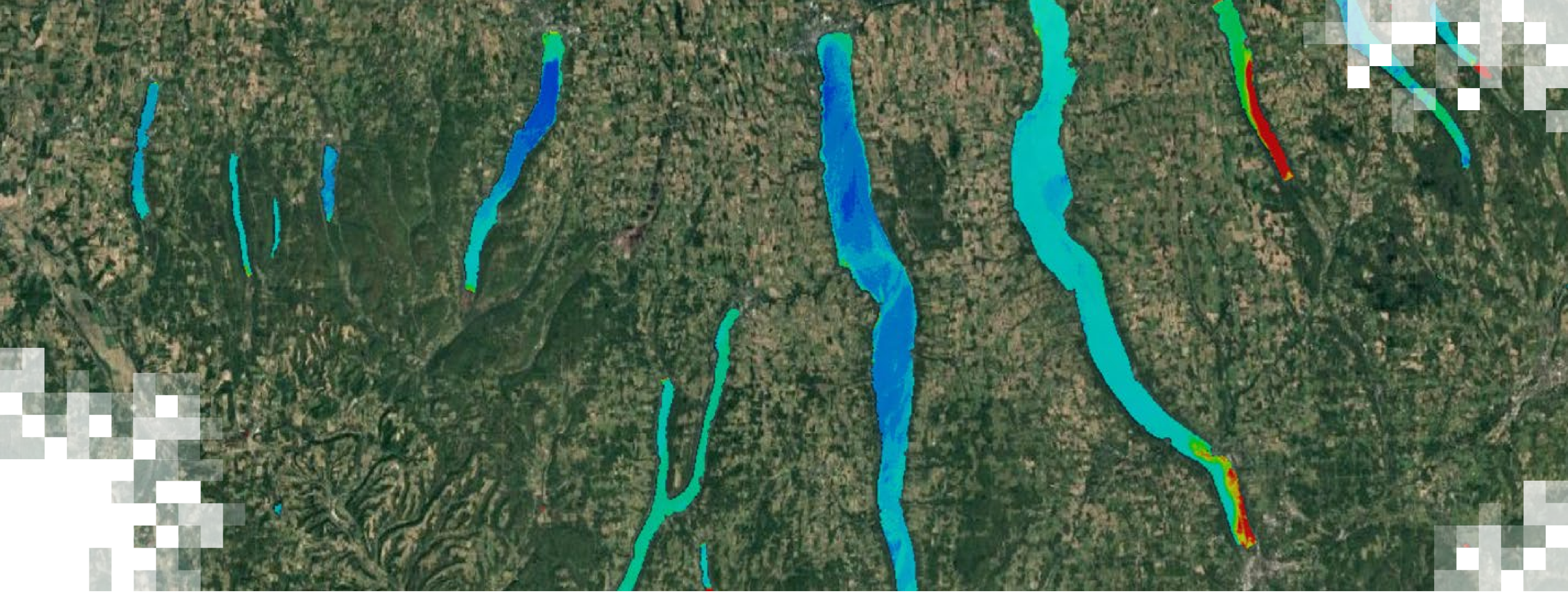




For issue tracking please comment on Github (or email ryan.oshea@ssaihq.com)



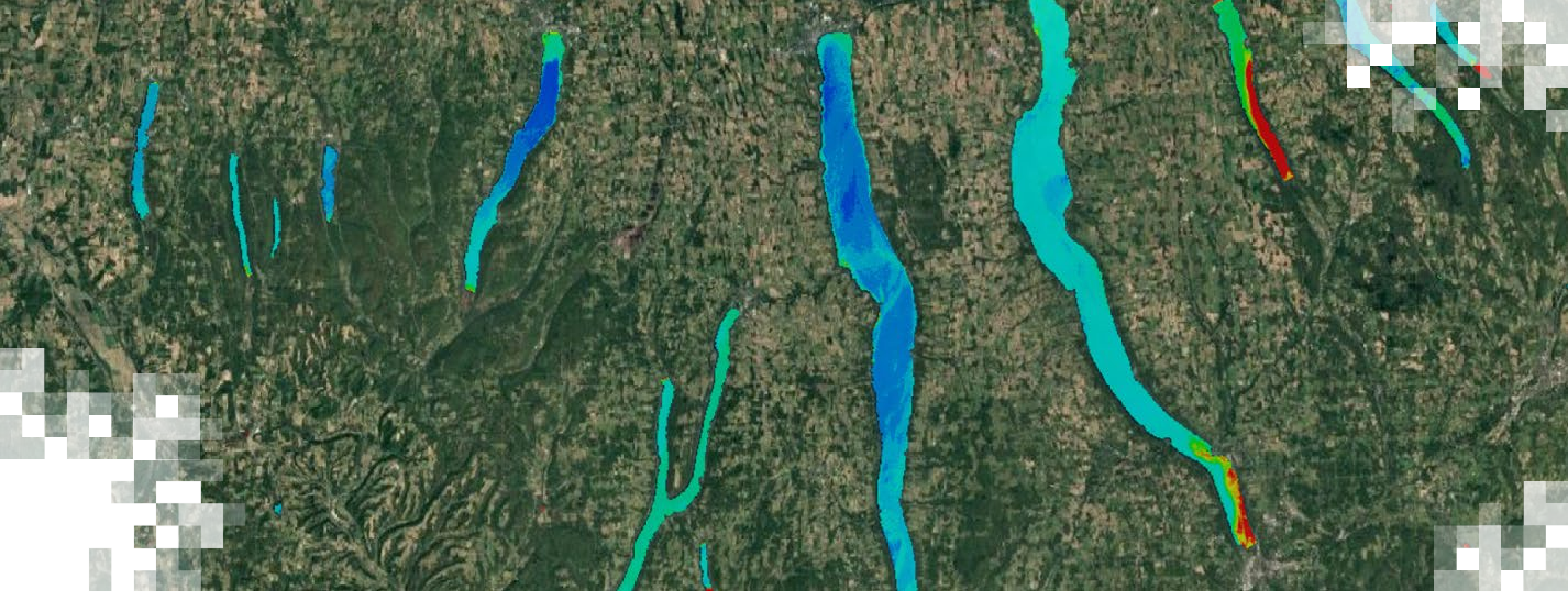
Demonstration
Application of MDNs to Satellite Imagery



Part 2:
Summary

Summary

- Overview of pre-processing top of atmosphere (TOA) radiances from satellite images.
  - Obtain Level-1 satellite data.
  - Get atmospherically corrected Level-2 remote sensing reflectances from the TOA data.
- Overview and demonstration of using Mixture Density Network (MDN) models to derive water quality parameters
  - The models are trained using a large globally distributed dataset of in situ water quality measurements and co-located hyperspectral measurements from a wide range of optically complex regions (sharing many of the samples with GLORIA).
  - The MDN approach is well suited for addressing non-unique inverse problems like ocean color remote sensing due to its output layer.
  - The models retrieve key water quality parameters, depending on spectral availability, including — chlorophyll-a, total suspended solids (TSS), colored dissolved organic matter (CDOM), and phycocyanin (PC)—along with their associated uncertainties.
  - Similar models were developed to support multiple multi- and hyperspectral satellite imagers, including MODIS, VIIRS, Landsat-8/9, Sentinel-2, Sentinel-3, and PACE.



Monitoring Water Quality in Lakes and Coastal Regions Using STREAM
Summary

Training Summary

- **Overview and Demonstration of the Satellite-based Tool for Rapid Evaluation of Aquatic Environments (STREAM) Web Tool and API**
  - Map water quality parameters—chlorophyll-a concentration, total suspended solids, and Secchi disk depth — from 2024 to near-real time in coastal estuaries and inland lakes across the United States at 20–30 m spatial resolution using Landsat 8 & 9 and Sentinel-2A, 2B, and 2C imagery.
  - Search and download water quality products using the STREAM web tool and API.
  - Examine time series of water quality parameters in a region of interest using QGIS.

- **Overview and Demonstration of the Aquaverse Algorithm**
  - Introduction to the Mixture Density Network (MDN) model to derive water quality parameters.
  - Access and application of MDN-based model code to retrieve water quality parameters.

- **STREAM web tool/API and the MDN code are open source**

Homework and Certificates

- **Homework:**
  - One homework assignment
  - Opens on 02/17/2026
  - Access from the training webpage
  - Answers must be submitted via Google Forms
  - **Due by 03/10/2026**

- **Certificate of Completion:**
  - Attend all live webinar sessions (attendance is recorded automatically)
  - Complete the homework assignment by the deadline
  - You will receive a certificate via email approximately two months after completion of the course.

Acknowledgements

William Wainwright

Senior Scientific Programmer
NASA-GSFC 619, SSAI

Ryan O'Shea

Senior Research Scientist

SSAI, 619, NASA-GSFC

Contact Information

Trainers:

- Ryan O'Shea
  - ryan.oshea@ssaihq.com
- William Wainwright
  - william.wainwright@ssaihq.com
- Arun M. Saranathan
  - arun.saranathan@ssaihq.com
- Akash Ashapure
  - akash.ashapure@ssaihq.com
- Amita Mehta
  - Amita.v.mehta@nasa.gov

- ARSET Website
- ARSET YouTube

For questions, comments, or to share how you have applied our trainings to your work or studies, email nasa.arset@gmail.com.

Join our quarterly newsletter to stay up-to-date on our latest trainings:

1. Send an email with no subject line to arset-join@lists.nasa.gov.
2. Follow the instructions sent in response.

Resources

- https://github.com/ryan-edward-oshea/MDN_tutorials
- https://ladsweb.modaps.eosdis.nasa.gov/stream/
- https://doi.org/10.1016/j.rse.2023.113706

Thank You!