AI/ML Integration of Satellite Remote Sensing

Data Harmonization Challenges and Gaps

Dr. Stephanie Schollaert Uz

Applied Sciences Manager | Earth Sciences Division

NASA Goddard Space Flight Center

stephanie.uz@nasa.gov

Literature Review: Chesapeake Bay Living Resources use of AI/ML

Tidal wetlands:

Lamb et al., 2021, A Fused Radar–Optical Approach for Mapping Wetlands and Deepwaters of the Mid–Atlantic and Gulf Coast Regions of the United States, *Remote Sens.*, doi: 10.3390/rs13132495

Submerged aquatic vegetation:

Coffer et al., 2023, Providing a framework for seagrass mapping in United States coastal ecosystems using high spatial resolution satellite imagery, *J. Env. Manag.,* doi: 10.1016/j.jenvman.2023.117669

Phytoplankton pigment chlorophyll-a:

Yu et al., 2022, Chlorophyll-a in Chesapeake Bay based on VIIRS satellite data: Spatiotemporal variability and prediction with machine learning, *Journal of Ocean Modeling,* doi: 10.1016/j.ocemod.2022.102119

Pahlevan et al., 2022, Simultaneous retrieval of selected optical water quality indicators from Landsat-8, Sentinel-2, and Sentinel-3, *Remote Sensing of Environment*, doi: 10.1016/j.rse.2021.112860

Water clarity:

Schollaert Uz et al., 2024, DEEP-VIEW integration of coastal observations and models to inform water quality resource managers and decisions, *2024 IEEE Intl Geosci. and Remote Sens. Symposium,* doi: 10.1109/IGARSS53475.2024.10642274

Clark et al., 2024, Non-Euclidean Water Distance Based Interpolation for Increased Mapping of Coastal Water Clarity, *2024 IEEE Intl Geosci. and Remote Sens. Symposium,* doi: 10.1109/IGARSS53475.2024.10642967

Low oxygen zone:

Zheng, G. et al., 2024, Hypoxia Forecasting for Chesapeake Bay Using Artificial Intelligence, *Artificial Intelligence for the Earth Systems*, doi: 10.1175/AIES-D-23-0054.1

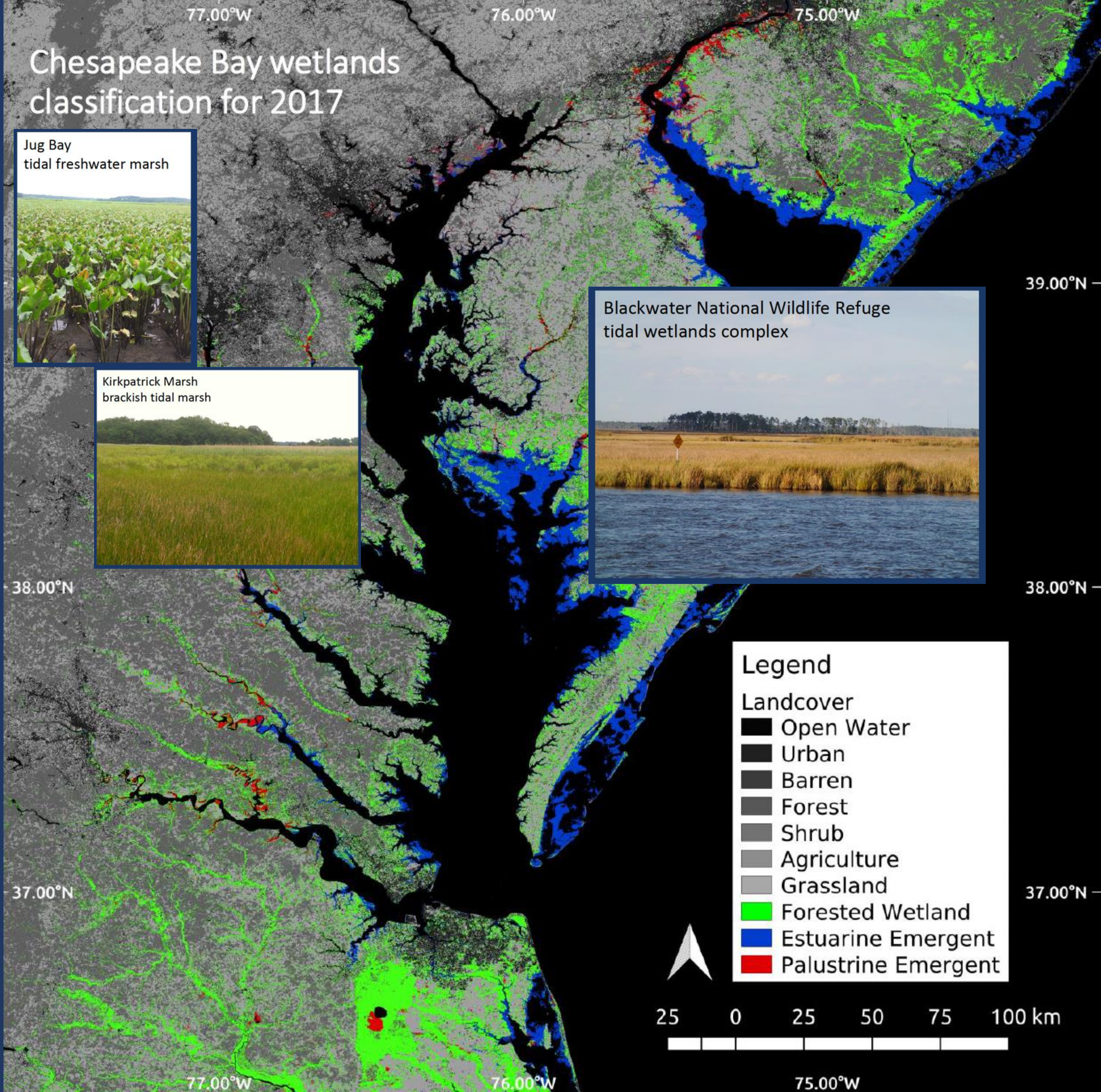
Satellite remote sensing challenges and gaps in the Chesapeake Bay

- Cloud cover, sun glint, challenging atmosphere, i.e. gases, particles, thin clouds
- Narrow waterways, i.e. land adjacency or missed aquatic signal wherever sensor saturates on brighter land, missing darker water
- Optically complex water (not clear)
- Spatial resolution vs signal, frequency trade-offs aquatic (1-km or 300m, high SNR, daily) vs terrestrial sensors (30m, low SNR, 8-16 day revisit)
- Despite robust sampling program, few above-water radiance measurements available for cal/val
- How can AI/ML help harmonize data and fill gaps?

Chesapeake Bay Wetlands Mapping

- Lamb, B.T., Tzortziou, M.A., McDonald, K.C. Evaluation of Approaches for Mapping Tidal Wetlands of the Chesapeake and Delaware Bays. Remote Sensing. 2019, 11(20), 2366. https://doi.org/10.3390/rs11202366
  - Landsat 8 seasonal vegetation indices (NDVI, TVI)
  - Sentinel-1 SAR seasonally averaged backscatter layers
  - Elevation layer from SRTM DEM
- 2016 NLCD and NWI overlap for training/validation data
- **Random Forest ML: estuarine wetlands (tidal marshes) classified with 83% producer's accuracy; elevation, summer & fall vegetation indices as most important predictors**

Source: Brian Lamb, USGS

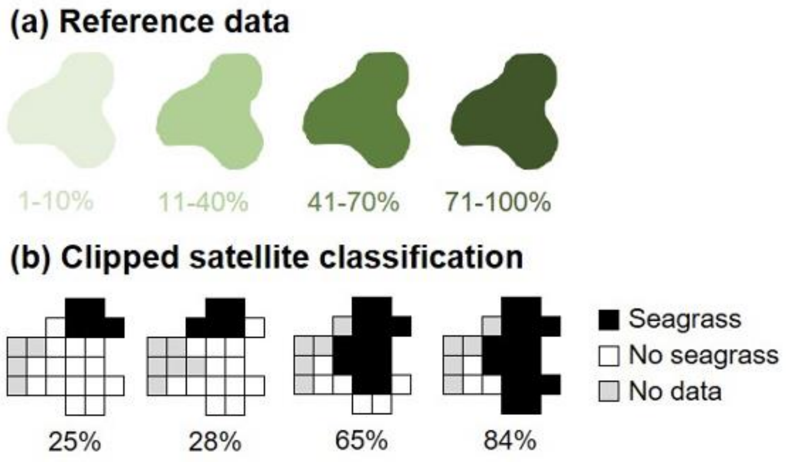


Submerged aquatic vegetation classification and extent

Coffer et al., 2023, Providing a framework for seagrass mapping in United States coastal ecosystems using high spatial resolution satellite imagery, *J. Env. Management,* doi: 10.1016/j.jenvman.2023.117669

WorldView-2 imagery at spatial resolution of 1.84m directly below the satellite, 6 visible wavelengths

Satellite imagery classification using deep convolutional neural network (DCNN) – performed best in dense, continuous seagrass compared to areas of sparse, discontinuous seagrass



Source: Coffer et al., 2023

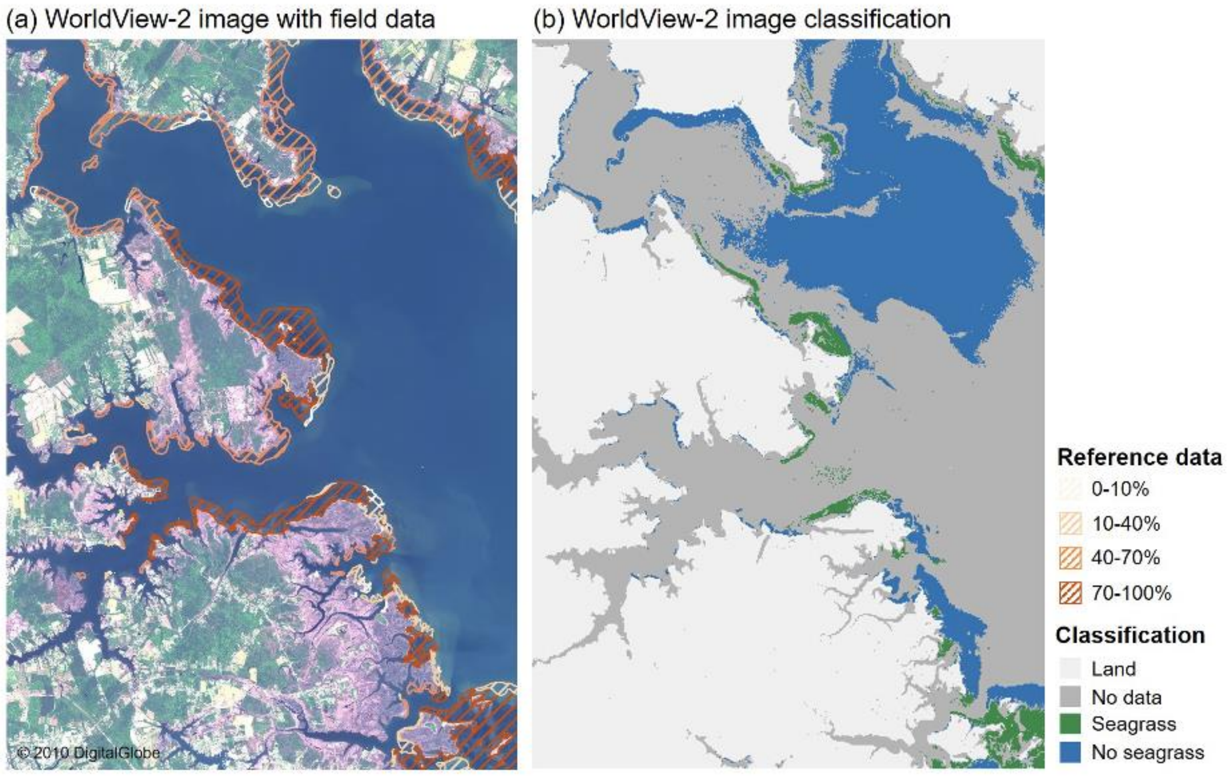


Figure S5. (a) A WorldView-2 satellite image acquired for Mobjack Bay, VA, on 4 May 2015 overlaid with reference data delineating seagrass percent cover obtained from the Virginia Institute of Marine Science in May through November 2015, and (b) results of an image classification with classes for land, no data, seagrass, and no seagrass.

Limitations for submerged aquatic vegetation classification

- biomass characterization requires image acquisition at consistent tidal stage, ideally low tide
- commercial image tasking does not guarantee a request will be fulfilled
- missing data due to clouds
- satellite signal attenuates in deep water: seagrass deep edge mischaracterized at increasing water depth
- multispectral satellite data are insufficient for identifying species of seagrass communities, but hyperspectral imagery can distinguish plant types by pigment discrimination

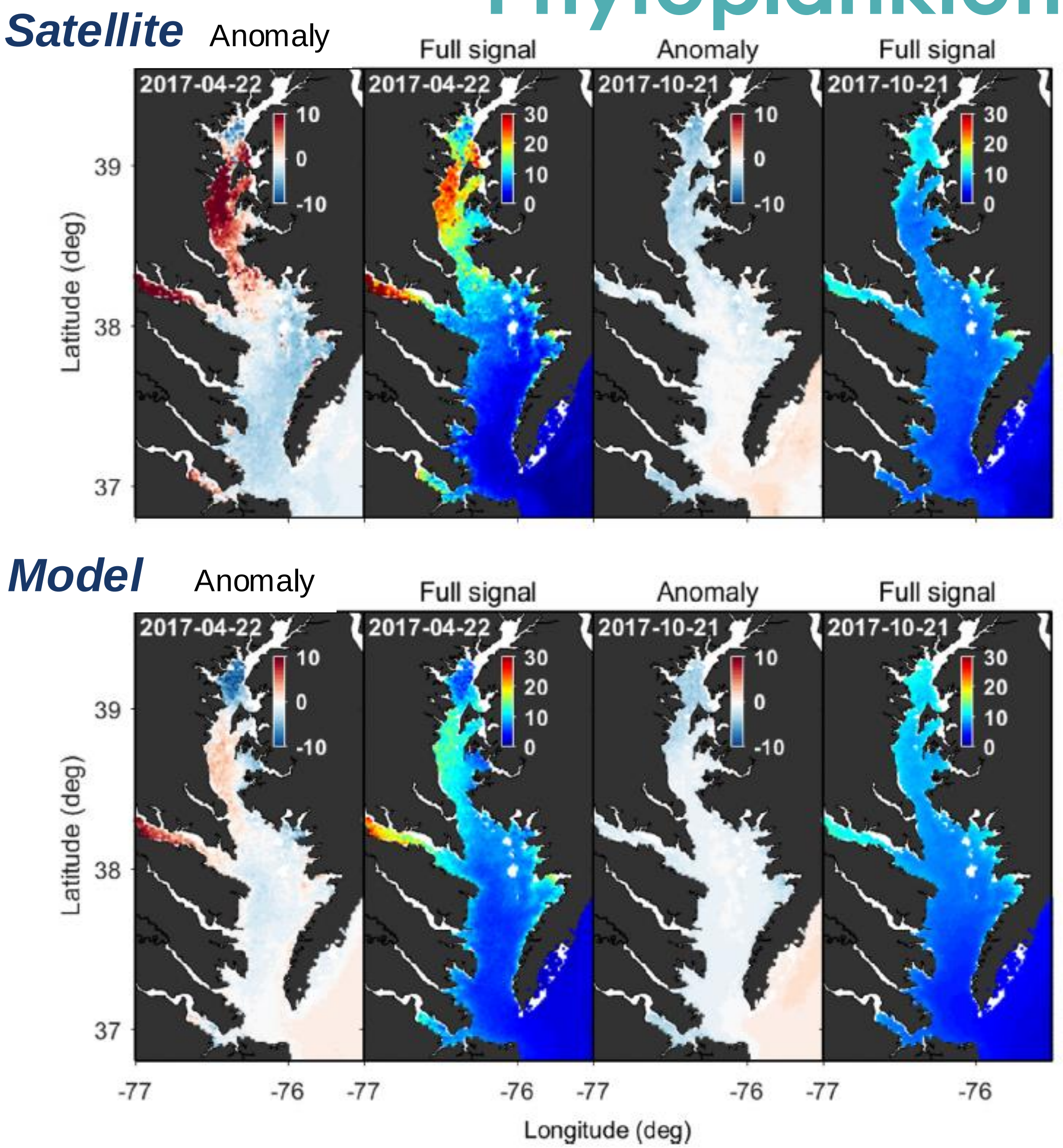
Source: Coffer et al., 2023

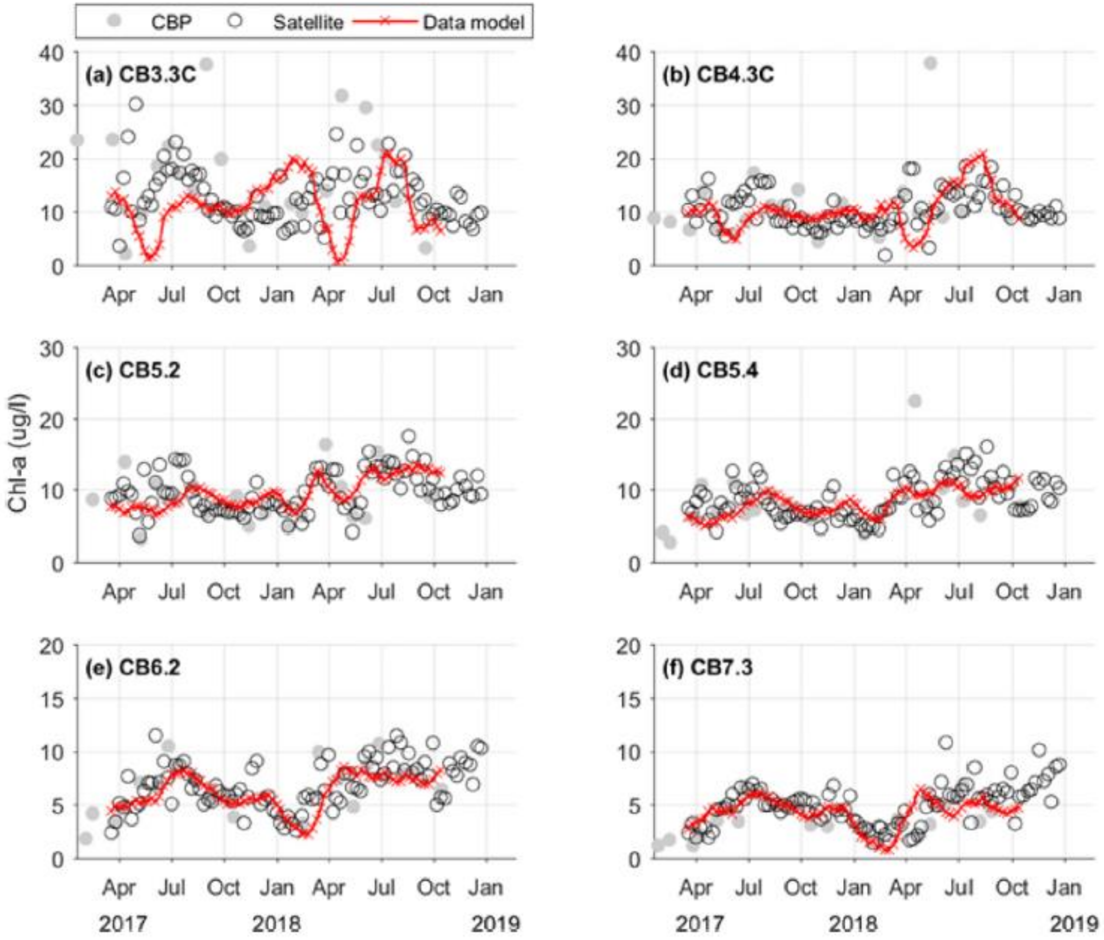
Phytoplankton pigment chlorophyll-a

Yu et al., 2022, Chlorophyll-a in Chesapeake Bay based on VIIRS satellite data: Spatiotemporal variability and prediction with machine learning, *Journal of Ocean Modeling,* doi: 10.1016/j.ocemod.2022.102119

- High resolution chl-a variation: VIIRS 750m, 7-day composites 2011-2018 chl-a calculated from absorption a_{ph}(670nm) per Zheng & DiGiacomo, 2017
- Data Interpolating Empirical Orthogonal Functions (DINEOF), applied to fill missing data
- **Machine learning data-driven model:**
  - **ten EOF modes from satellite data**
  - **artificial neural network**
  - **external forcing: river flow, nutrient loading, air temperature, solar radiation, 10m wind speed**
- Machine learning training period (2011-2016); testing/prediction (2017-2018)

Phytoplankton pigment chlorophyll-a





Model dampens chl-a mean and variability, esp. in shallows; different timing of seasonal peaks between the upper, mid bay

Source: Yu et al., 2022

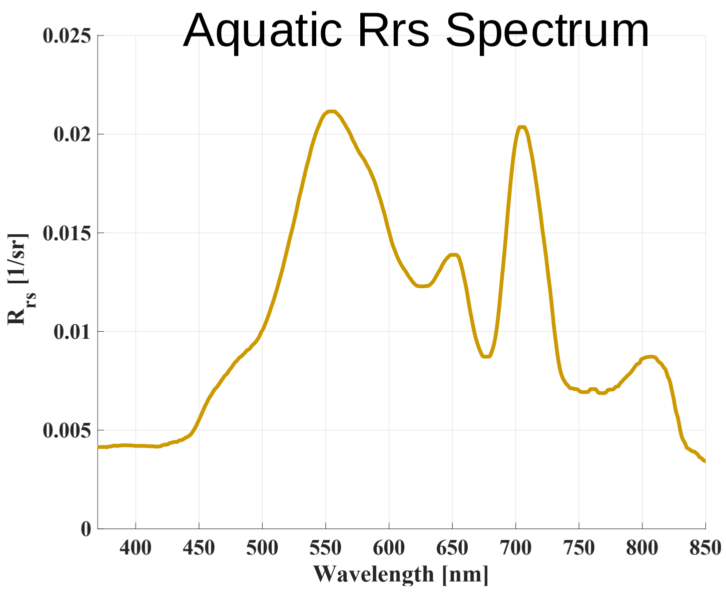
Water Quality indicators: chlorophyll-a, turbidity, CDOM

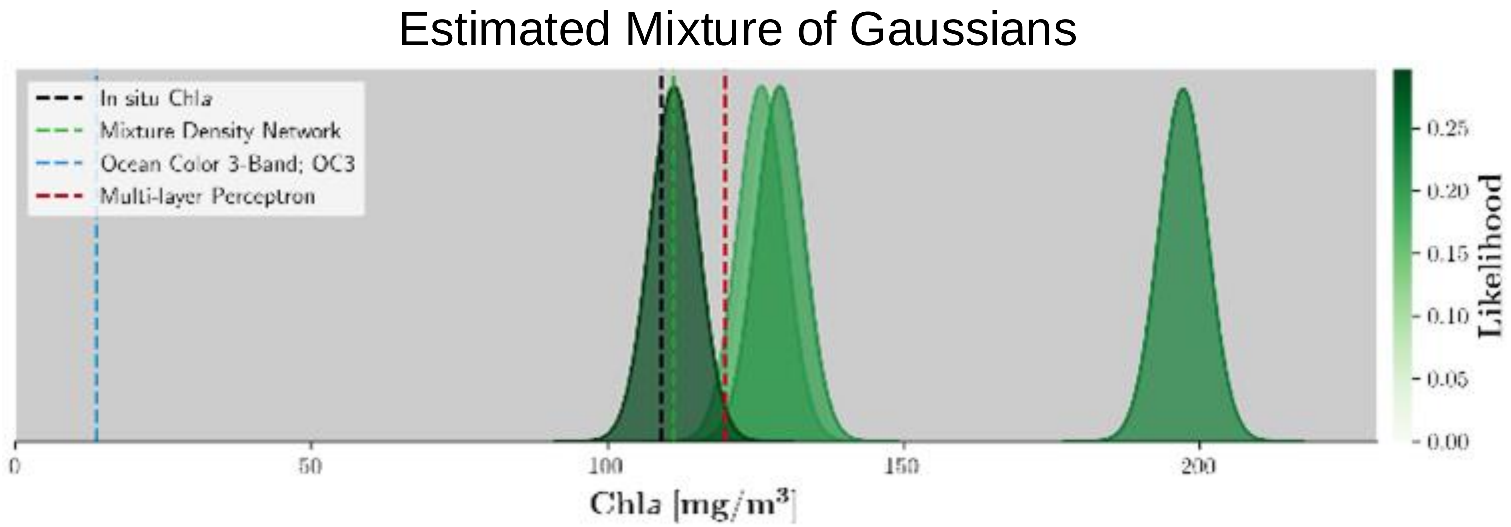
Pahlevan et al., 2022, Simultaneous retrieval of selected optical water quality indicators from Landsat-8, Sentinel-2, and Sentinel-3, *Remote Sensing of Environment*, doi: 10.1016/j.rse.2021.112860

Mixture Density Networks

Learn probability distribution over output space for multimodal target distributions with non-unique relationships between input and output features, i.e. aquatic remote sensing reflectance (Rrs) spectra

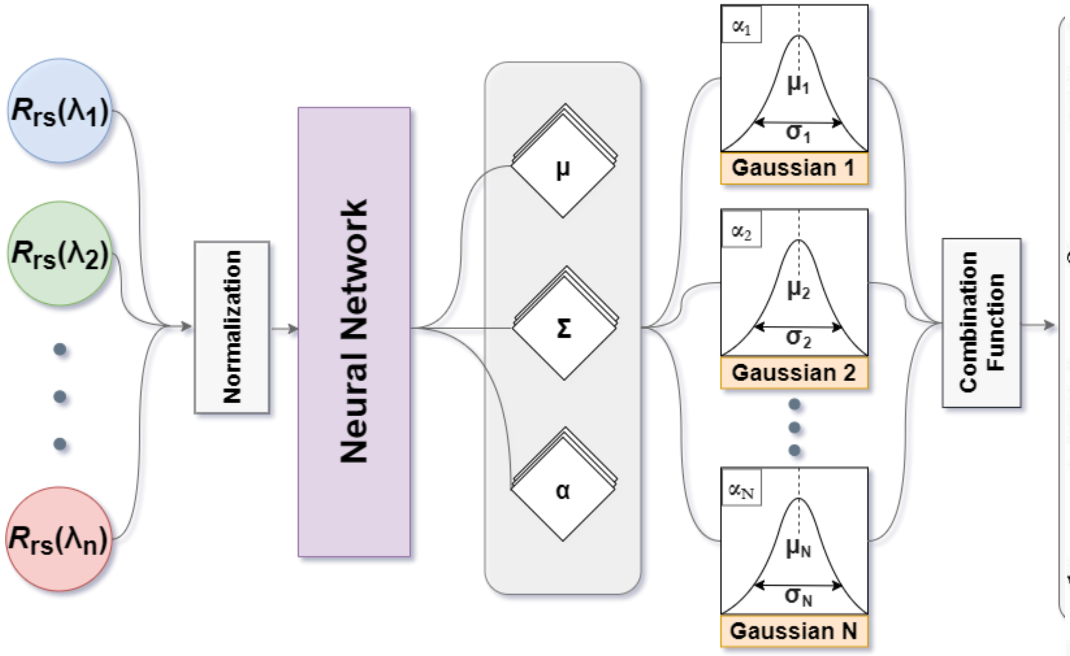
For each prediction, can use the maximum likelihood estimate or weighted average of all predictions

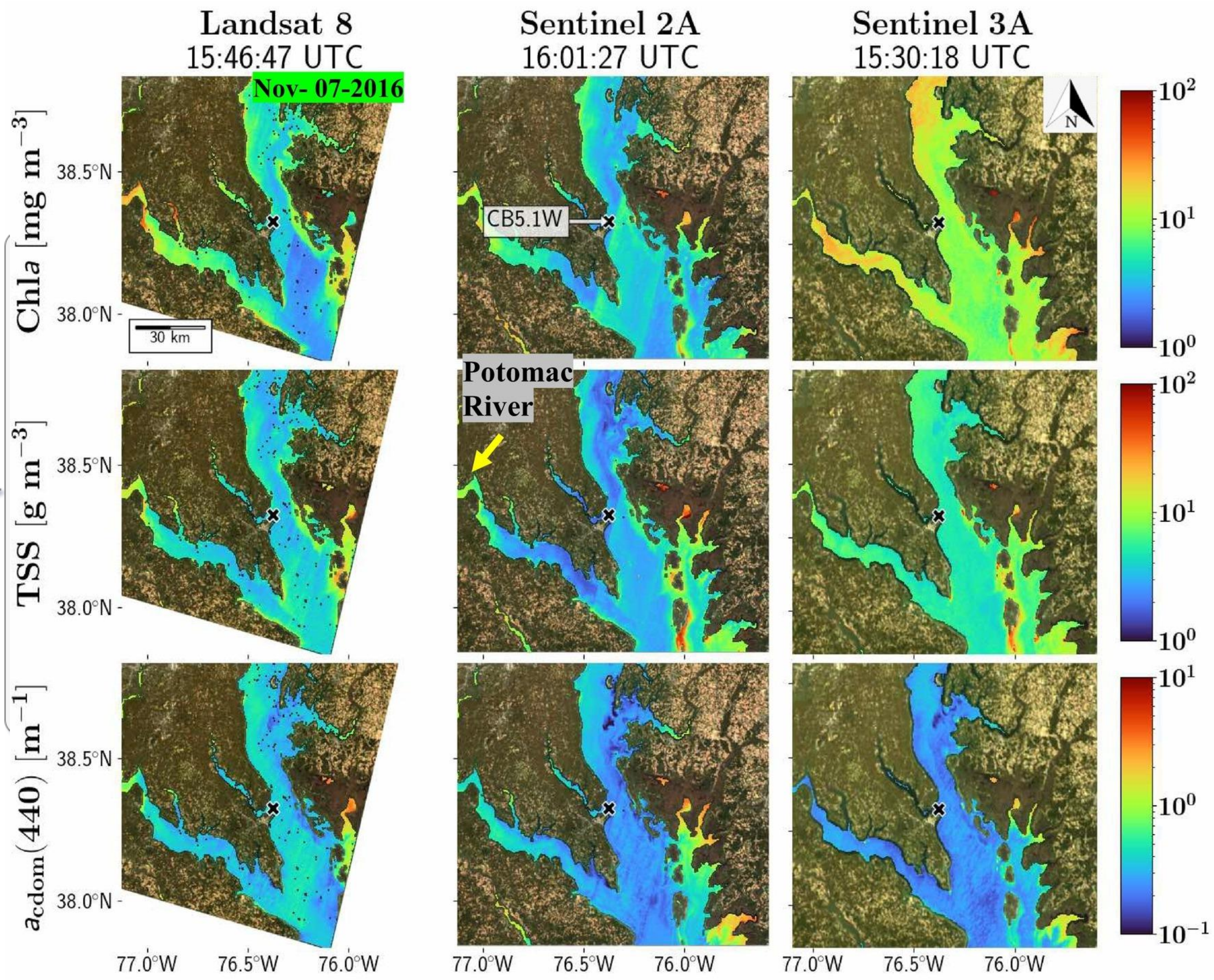




Source: Pahlevan et al., 2022

Water Quality indicators: chlorophyll-a, turbidity, CDOM





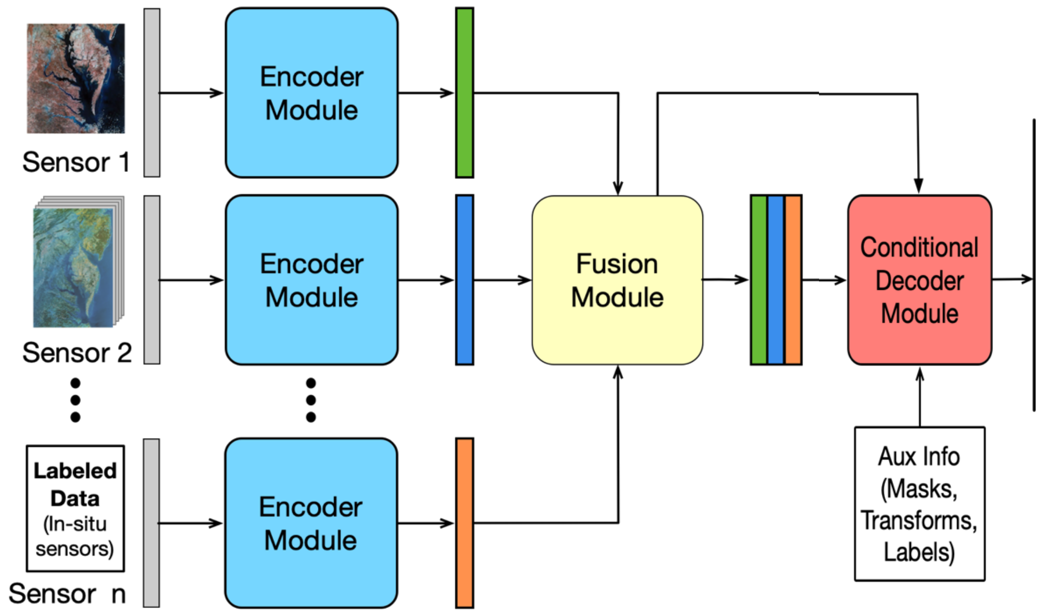
Code availability: https://github.com/STREAM-RS

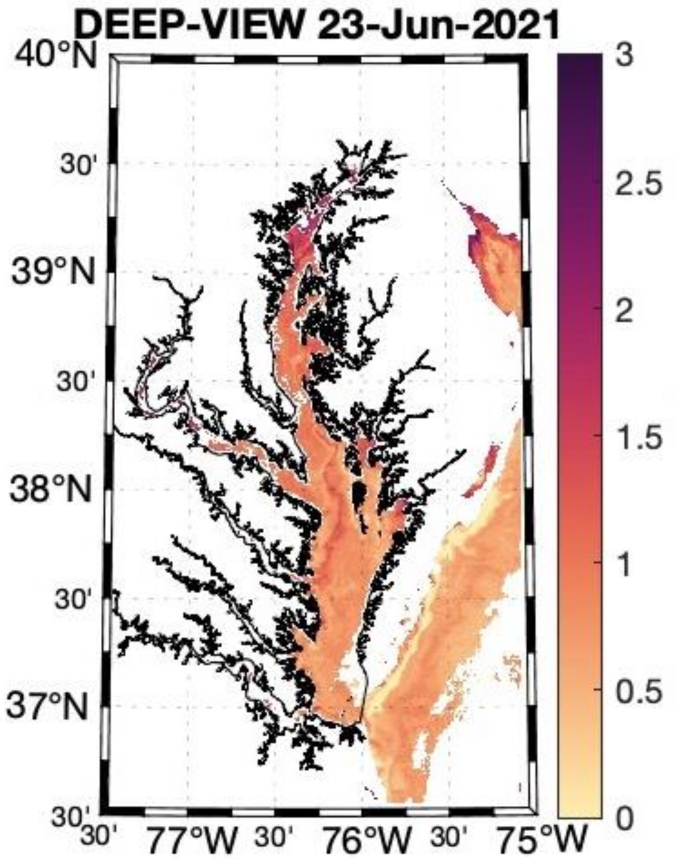
Data availability: https://doi.pangaea.de/10.1594/PANGAEA.948492

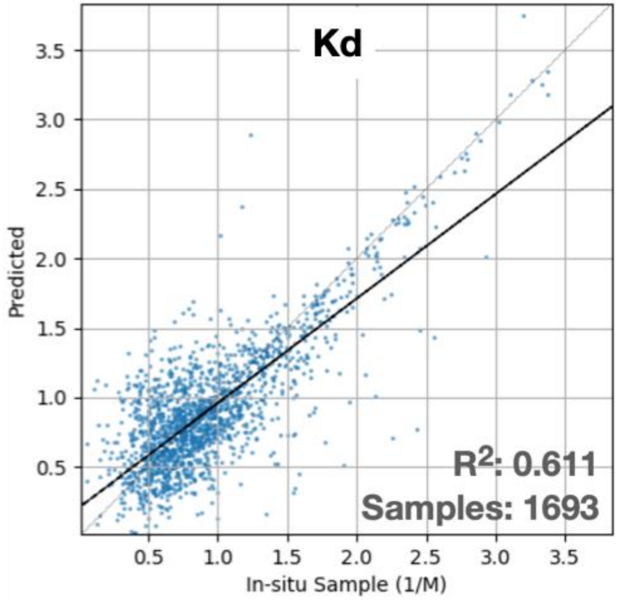
Source: Pahlevan et al., 2022

Water Quality indicator: clarity (diffuse attenuation coefficient)

Schollaert Uz et al., 2024, DEEP-VIEW integration of coastal observations and models to inform water quality resource managers and decisions, *2024 IEEE Intl Geosci and Remote Sensing Symposium,* doi: 10.1109/IGARSS53475.2024.10642274



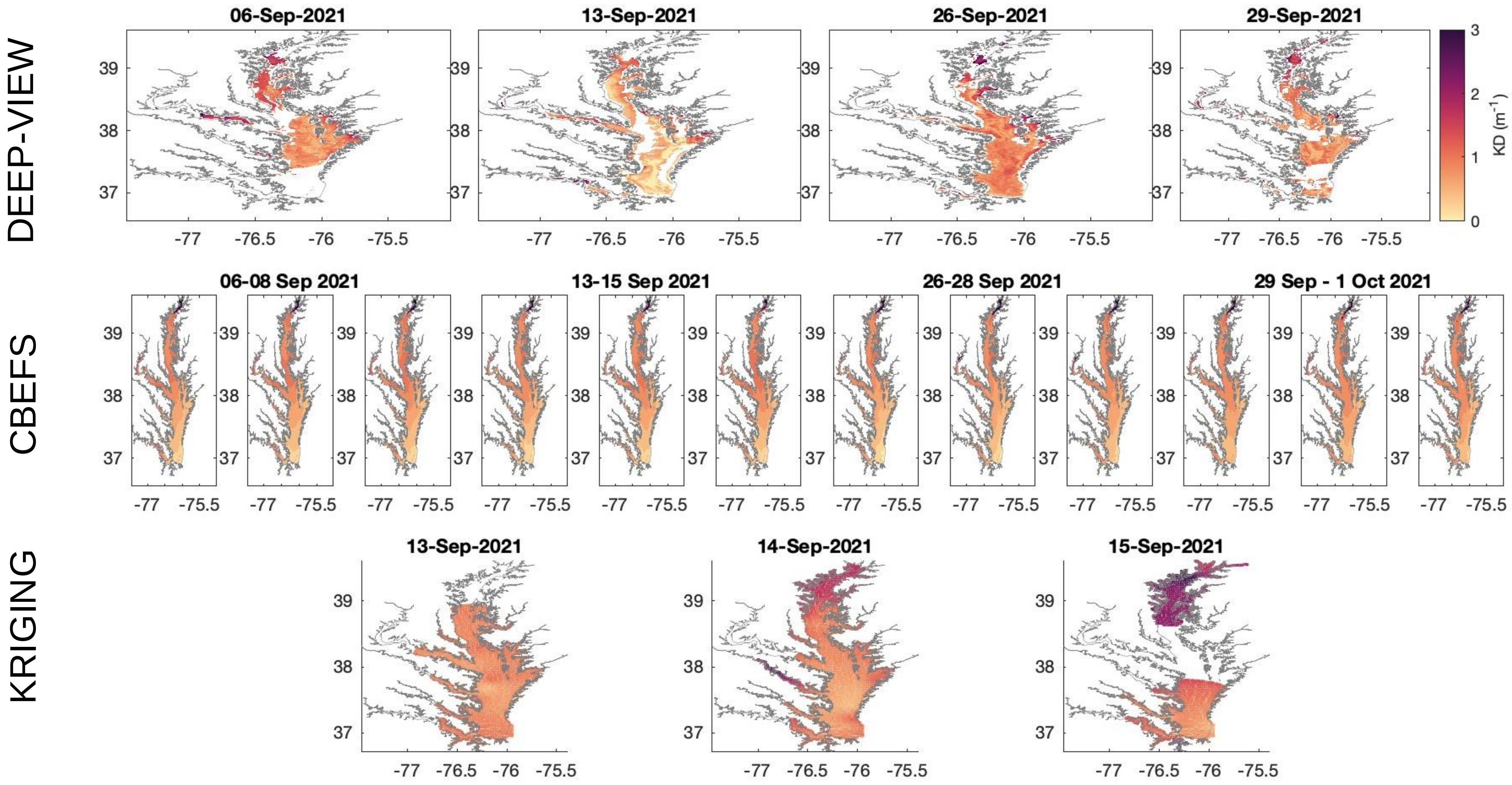




- Feature training utilizes long time series of MODIS (2003-2023)
- Spectral training data is done for MODIS, OLCI, and VIIRS (multi-layer fully connected neural net)
- Transfer learning validation utilizes the MODIS model
- Ongoing training of the Fusion and Conditional Decoder Modules

Water Quality indicator: clarity (diffuse attenuation coefficient)

Clark, J.B., S. Schollaert Uz, T.J. Ames, 2024, Non-Euclidean Water Distance Based Interpolation for Increased Mapping of Coastal Water Clarity, *2024 IEEE Intl Geosci and Remote Sensing Symposium*, doi: 10.1109/IGARSS53475.2024.10642967



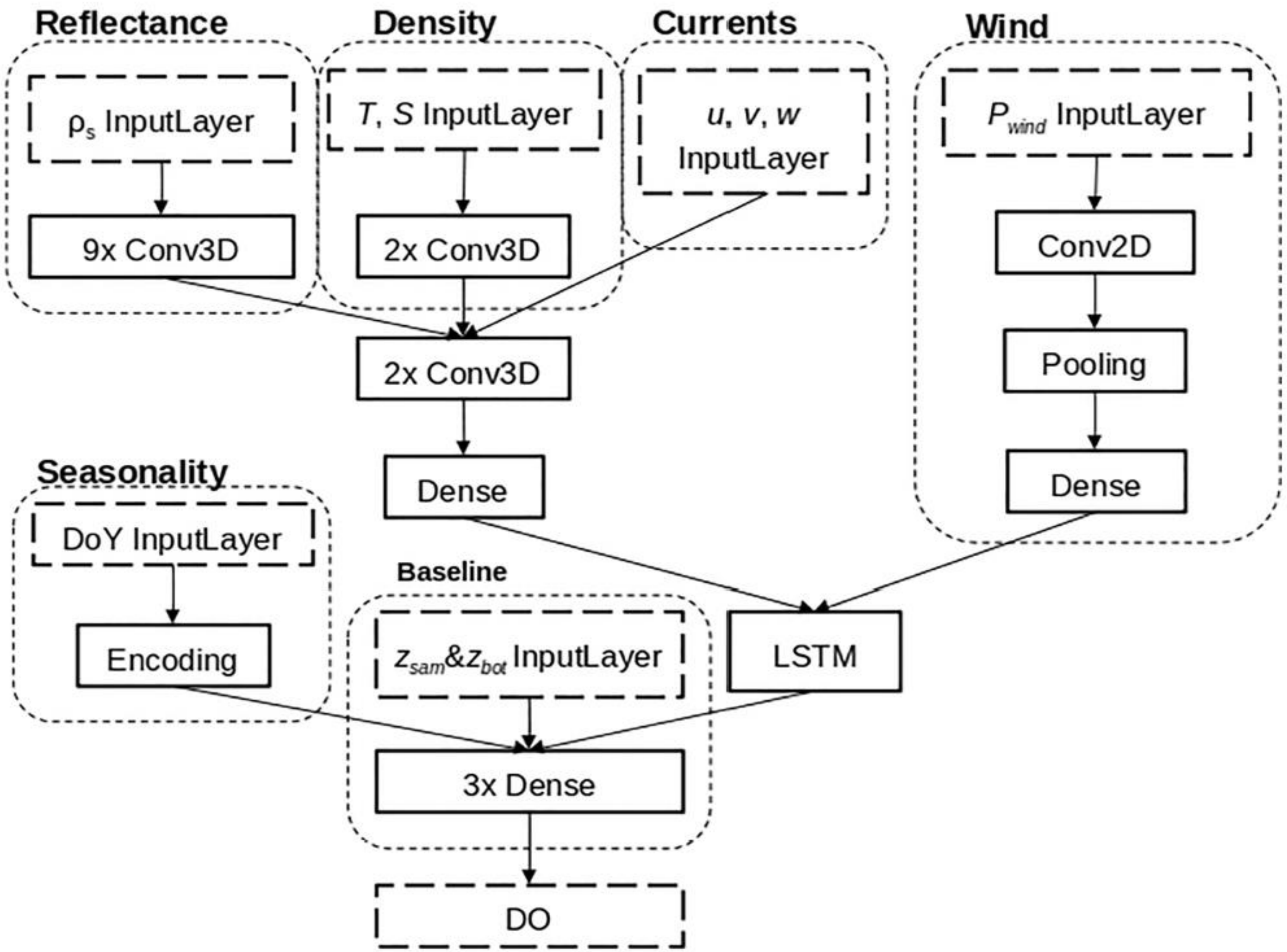
Source: Schollaert Uz et al., 2024 & Clark et al., 2024

Low oxygen zone

Zheng, G., et al., 2024, Hypoxia Forecasting for Chesapeake Bay Using Artificial Intelligence, *Artificial Intelligence for the Earth Systems*, doi: 10.1175/AIES-D-23-0054.1

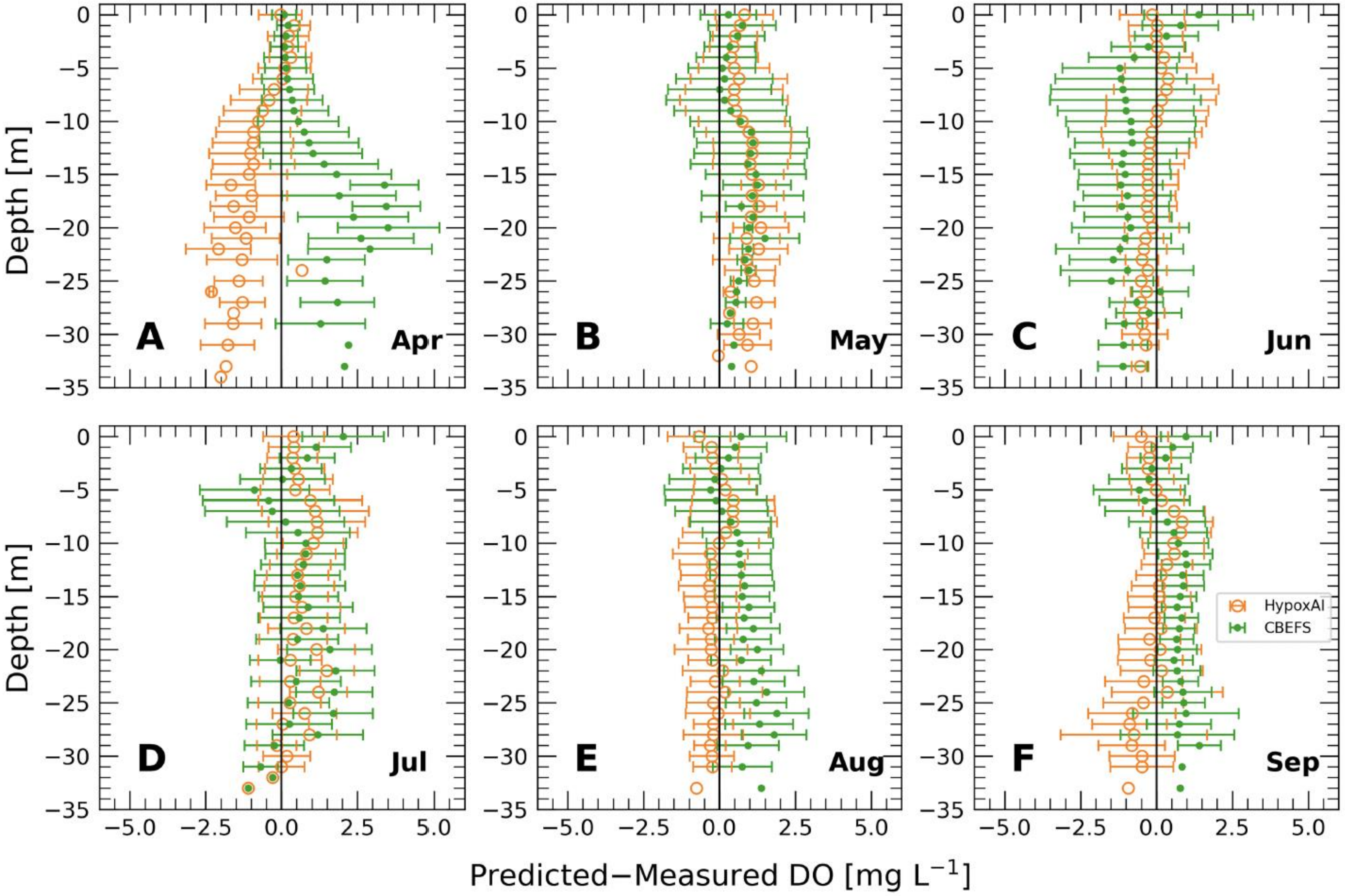
HypoxAI oxygen predictions from 8 weekly composites of daily data prior to oxygen observation

- Biogeochemical input:
  - Satellite-derived spectral reflectance data (MODIS-Aqua) as proxy for phytoplankton abundance from pigment chl-a
- Physical input:
  - CBEFS estuarine hydrodynamic model: water temperature, salinity, currents
  - wind velocity reanalysis information
- Model architecture:
  - Convolutional neural networks
  - LSTM networks with eight week time steps

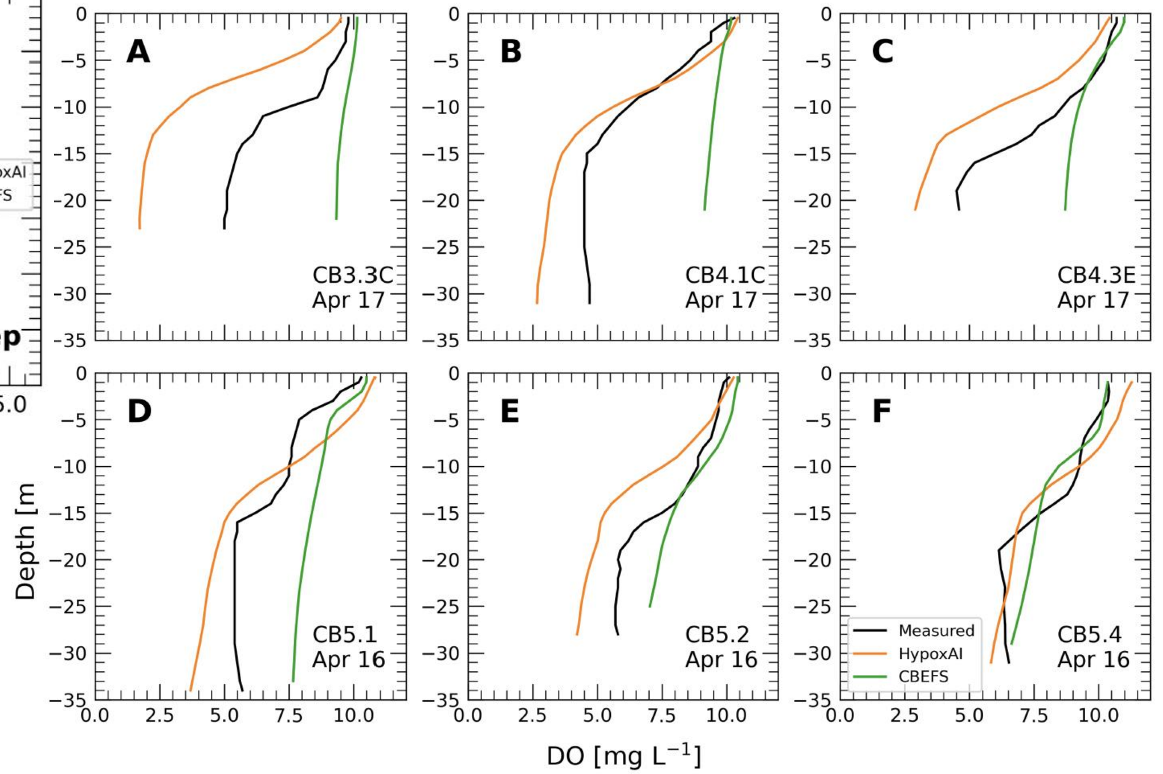


Ablation and cross-validation tests find strongest predictor is the 3D temperature field, which HypoxAI improves by 20% median absolute error.

Low oxygen zone



- Greatest difference between HypoxAI and CBEFS in April
- HypoxAI has difficulty predicting hypoxia onset
- HypoxAI performs best in June, August overall
- Stability of validation results, regardless of year used



Source: Zheng et al., 2024

Interagency Chesapeake Bay Working Group

- Satellite data providers and users meet on regional issues: run-off, water quality, algal blooms, carbon fluxes
- Linking current and future satellite data to observations for products to fill gaps, e.g. water temperature, clarity, algal blooms, CDOM, air quality, flooding
- Field work with resource managers, lab work at Goddard, University of Maryland
- AERONET-OC installed on USCG Tolchester navigation light tower in 2021 for cal/val; HYPERNET in 2023 – first in North America!
- Improving atmospheric correction, machine learning approaches, spectral libraries, phytoplankton classification
- In person workshops in 2018, 2019, 2023, perhaps summer 2025…