Supporting Shellfish Aquaculture in the Chesapeake

Using Artificial Intelligence to Detect Poor Water Quality through Sampling and Remote Sensing

Stephanie Schollaert Uz[1]

Troy Ames[1], Nargess Memarsadeghi[1], Amita Mehta[1,2], Guangming Zheng[2,3], Shannon McDonnell[2], Neil Blough[2], John McKay[4]

[1]Goddard Space Flight Center
[2]University of Maryland
[3]National Oceanic and Atmospheric Administration
[4]Maryland Department of Environment

Advanced Information Systems Technology project began in 2020

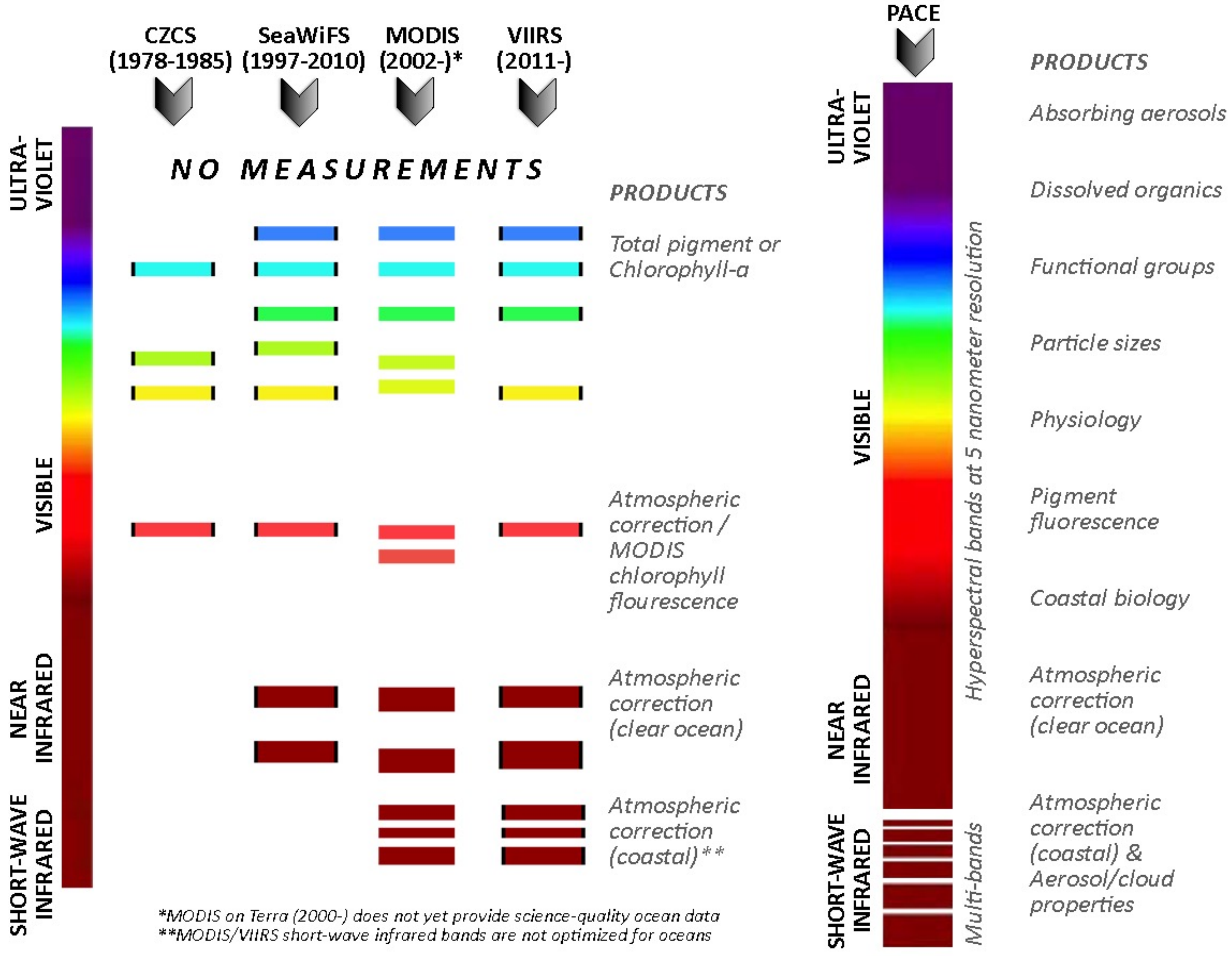
Aquaculture is a growing industry around the Bay and world-wide – resources managers monitor water quality from boats to spot-check for problems. Remote sensing may provide early warning of harmful algal blooms and polluted run-off

Photo credit: John 'Rusty' McKay/MDE

Optical techniques are explored to exploit new technology in complex environment

Developing AI combining observations, models, with satellite data

Future space sensors will provide information about phytoplankton communities, ecosystem health



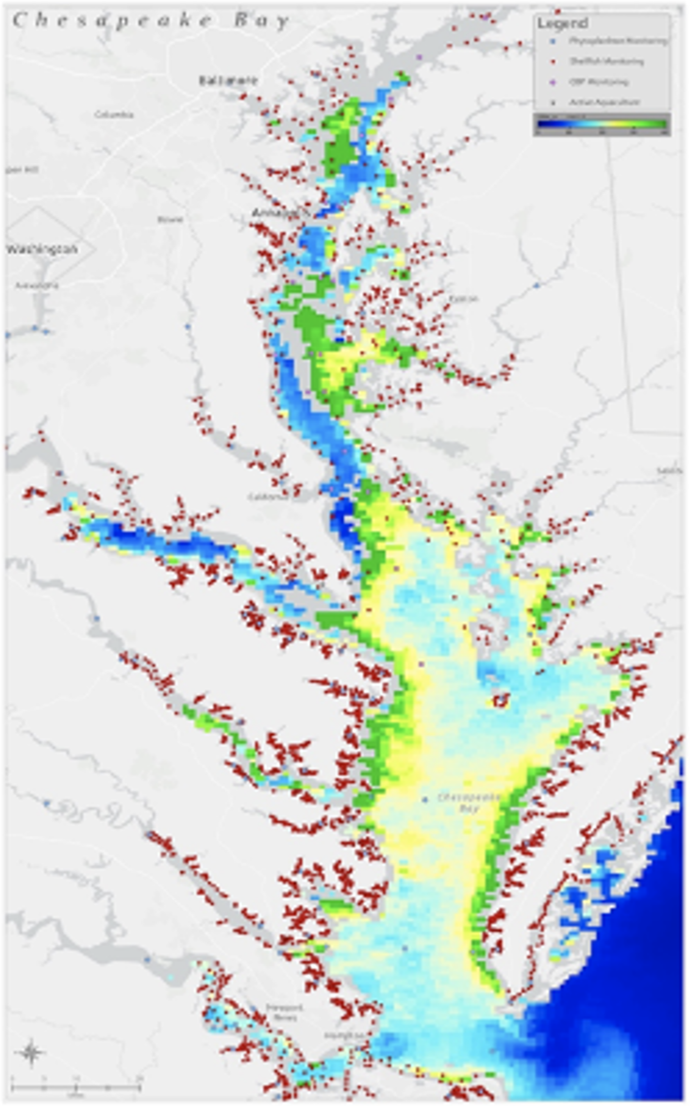
Source: https://pace.gsfc.nasa.gov

Maryland water quality criteria for shellfish harvesting

Parameter name	Water Quality Threshold
Fecal coliform	<14 MPN median per100ml
Bacteriological *Escherichia coli*	< 410 count per 100ml
Dissolved oxygen	> 5 mg/l
Temperature	< 90°F/32°C
pH	6.5 - 8.5
Turbidity	<150 nephelometer turbidity units
Color	< 75 platinum cobalt units
Water clarity	> 13% (tidal fresh)

http://www.dsd.state.md.us/comar/comarhtml/26/26.08.02.03-3.htm

- National Water Quality Monitoring Council (NWQMC), Chesapeake Bay Program (CBP) data acquired (1984-present) for Chl-a, DO, Kd, pH, temperature, salinity, bacteria



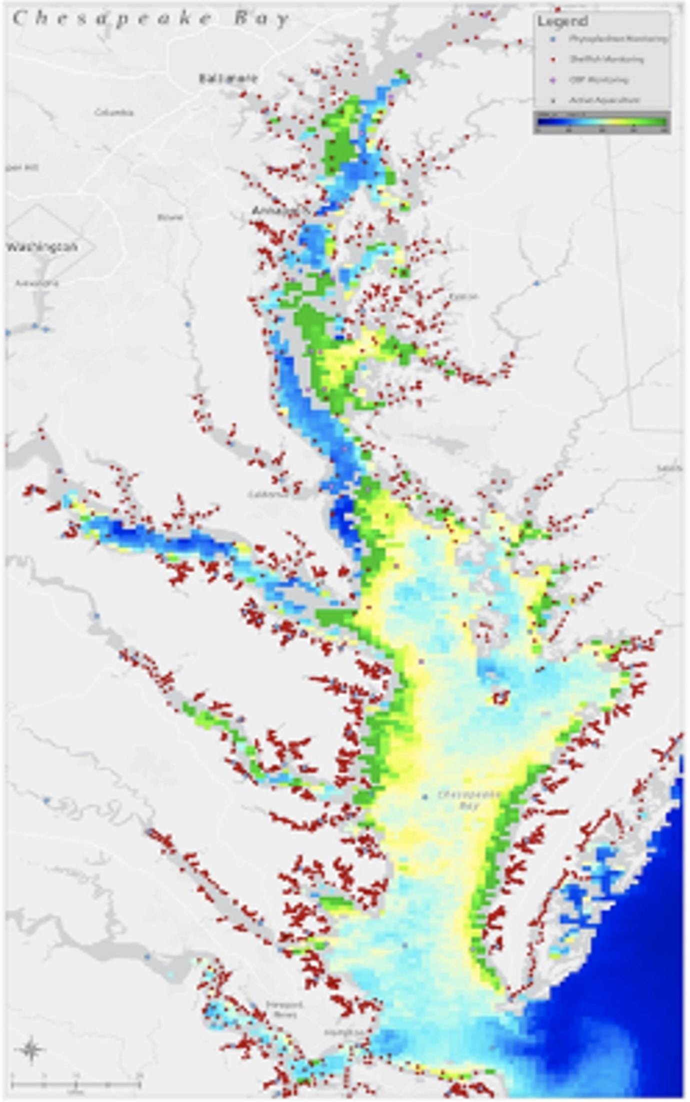
MODIS chlorophyll-a map from July 2, 2019 with routine sampling sites by Maryland and Virginia superimposed

Water quality criteria for shellfish harvesting

Parameter name	Water Quality Threshold
Fecal coliform	<14 MPN median per100ml
Bacteriological *Escherichia coli*	< 410 count per 100ml
Dissolved oxygen	> 5 mg/l
Temperature	< 90ºF/32ºC
pH	6.5 - 8.5
Turbidity	<150 nephelometer turbidity units
Color	< 75 platinum cobalt units
Water clarity	> 13% (tidal fresh)

http://www.dsd.state.md.us/comar/comarhtml/26/26.08.02.03-3.htm

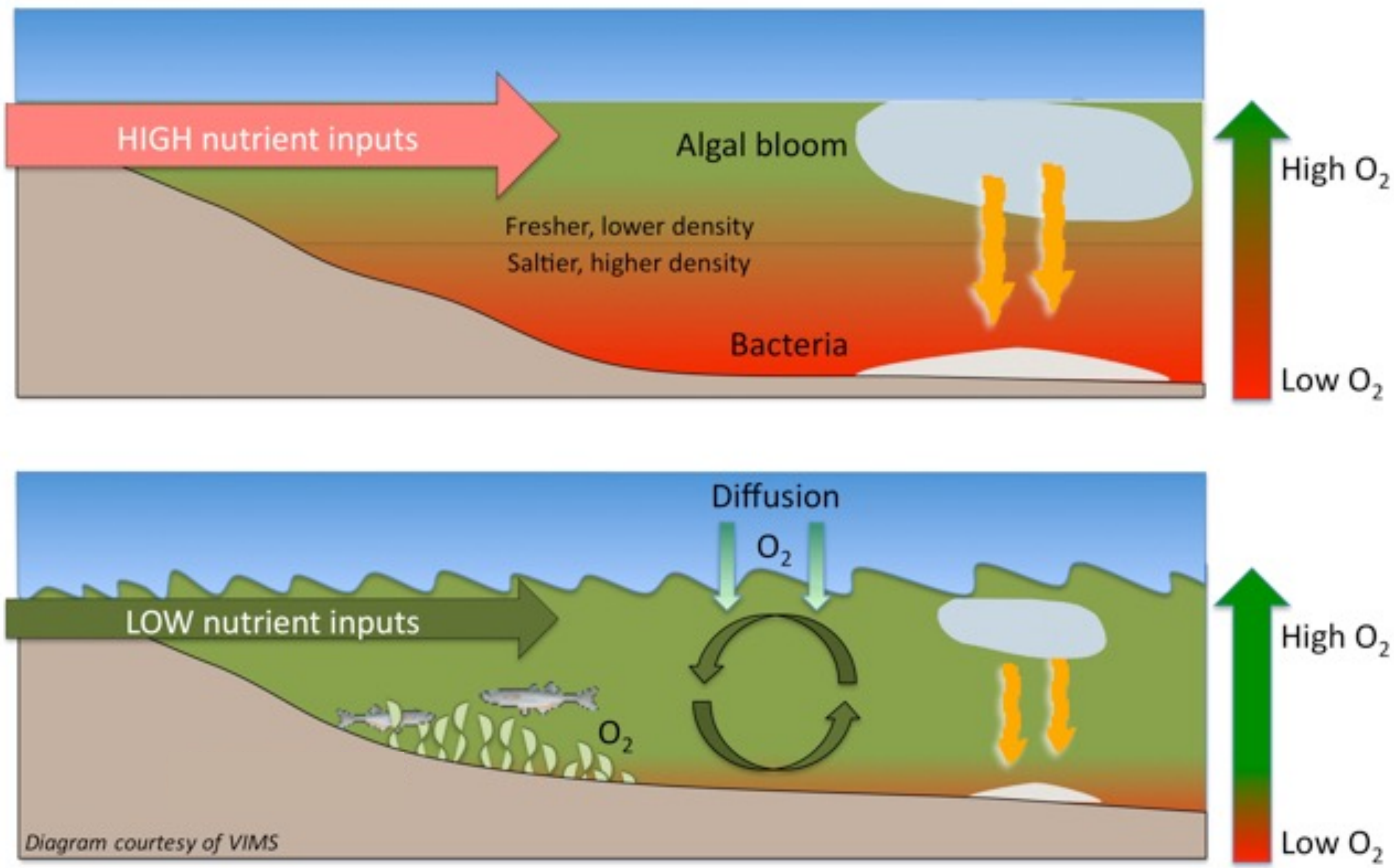
- National Water Quality Monitoring Council (NWQMC), Chesapeake Bay Program (CBP) data acquired (1984-present) for Chl-a, DO, Kd, pH, temperature, salinity, bacteria
- Satellite data processed (Sentinel 3a&3b OLCI, Aqua MODIS) to derive Rrs, Chl-a, Kd, Rhos
- Exploring relationship between satellite spectral bands, low oxygen, high chlorophyll, turbidity
- Field and lab work impacted by COVID-19: future work will incorporate absorption and fluorescence of samples collected around the Bay



MODIS chlorophyll-a map from July 2, 2019 with routine sampling sites by Maryland and Virginia superimposed

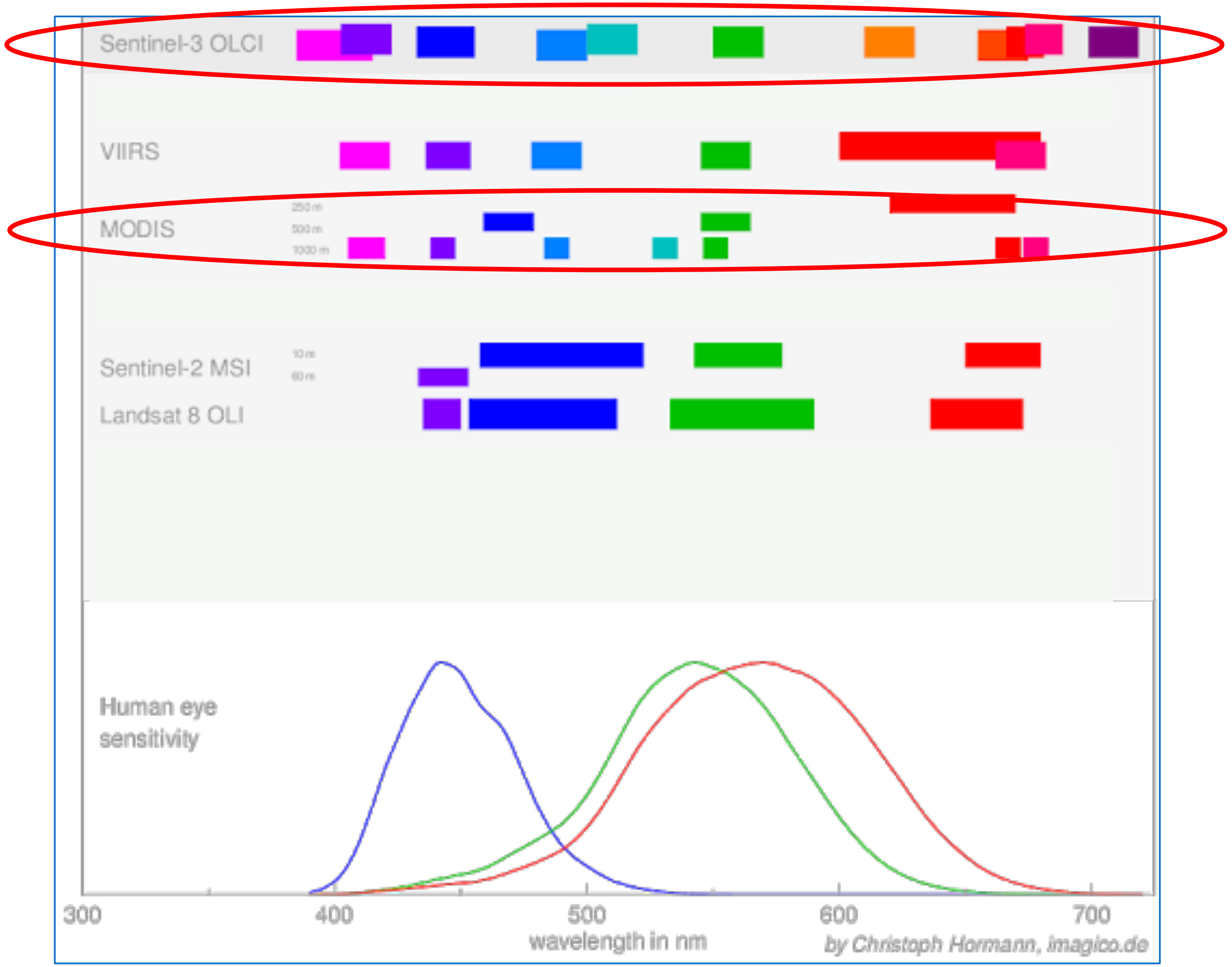
Low dissolved oxygen (hypoxia) as a proof-of-concept

- Establish method for using optical remote sensing for water quality
- Low Dissolved Oxygen (Hypoxia) is large scale, seasonal problem related to biological and physical mechanisms, i.e. algal biomass decay with stratification
- Although not visible in remote sensing, train AI to recognize its precursors, test feasibility for water quality

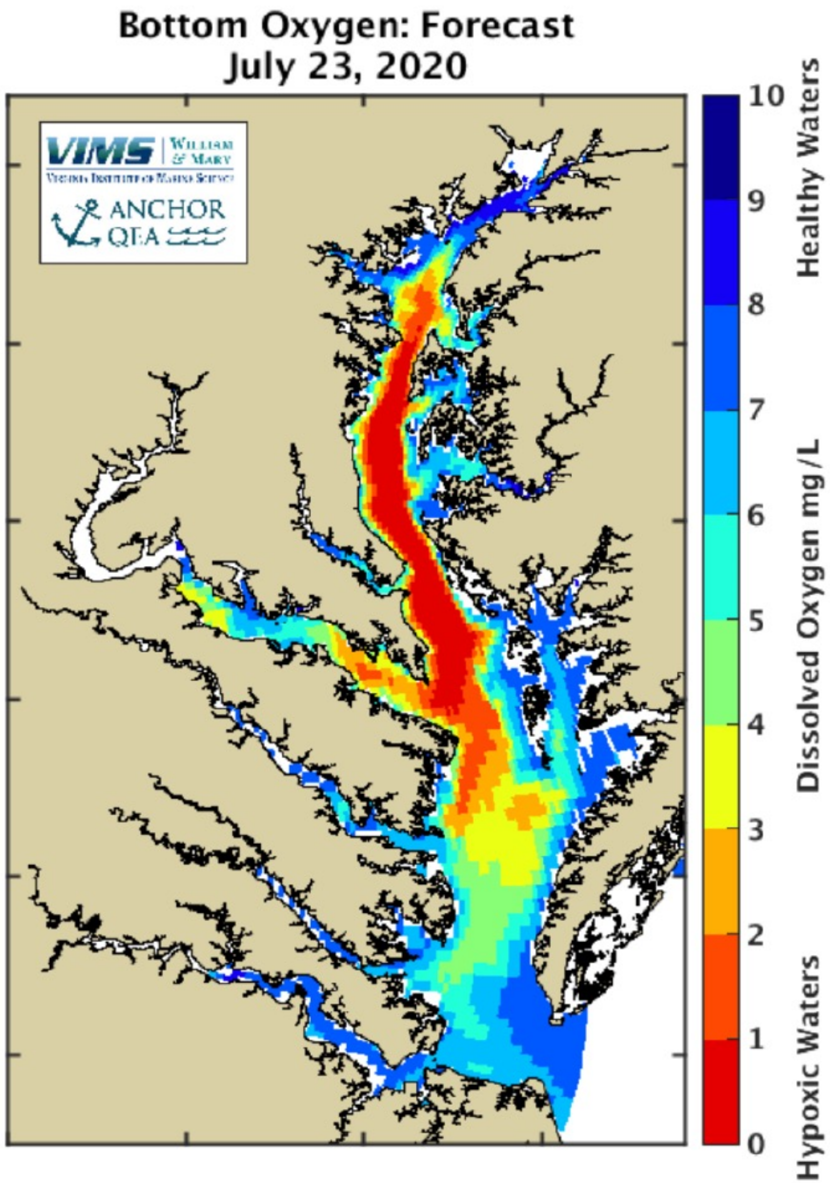


Large phytoplankton blooms, seen at surface, precede hypoxia

- Nutrient run-off, large phytoplankton blooms, warming and stratification as precursors to low oxygen conditions
- Sentinel 3 OLCI (2016-present) currently best spectral resolution, esp. red edge
- Aqua MODIS (2002-present) has good spectral resolution, longest time series
- Chesapeake Bay is challenging for atmospheric correction or Rrs, Chl-a:
  - AERONET OC will soon be installed north of Bay Bridge for calibration
  - Use top of atmosphere reflectances minus Rayleigh scattering, Rhos
  - New CSDAP project will use acolite software to derive better options



Train and validate machine learning with VIMS hypoxia model



Blue -> **high bottom oxygen**

Yellow/green -> **marginal oxygen**

Red -> **very low bottom oxygen (hypoxia)**

Feature variables:

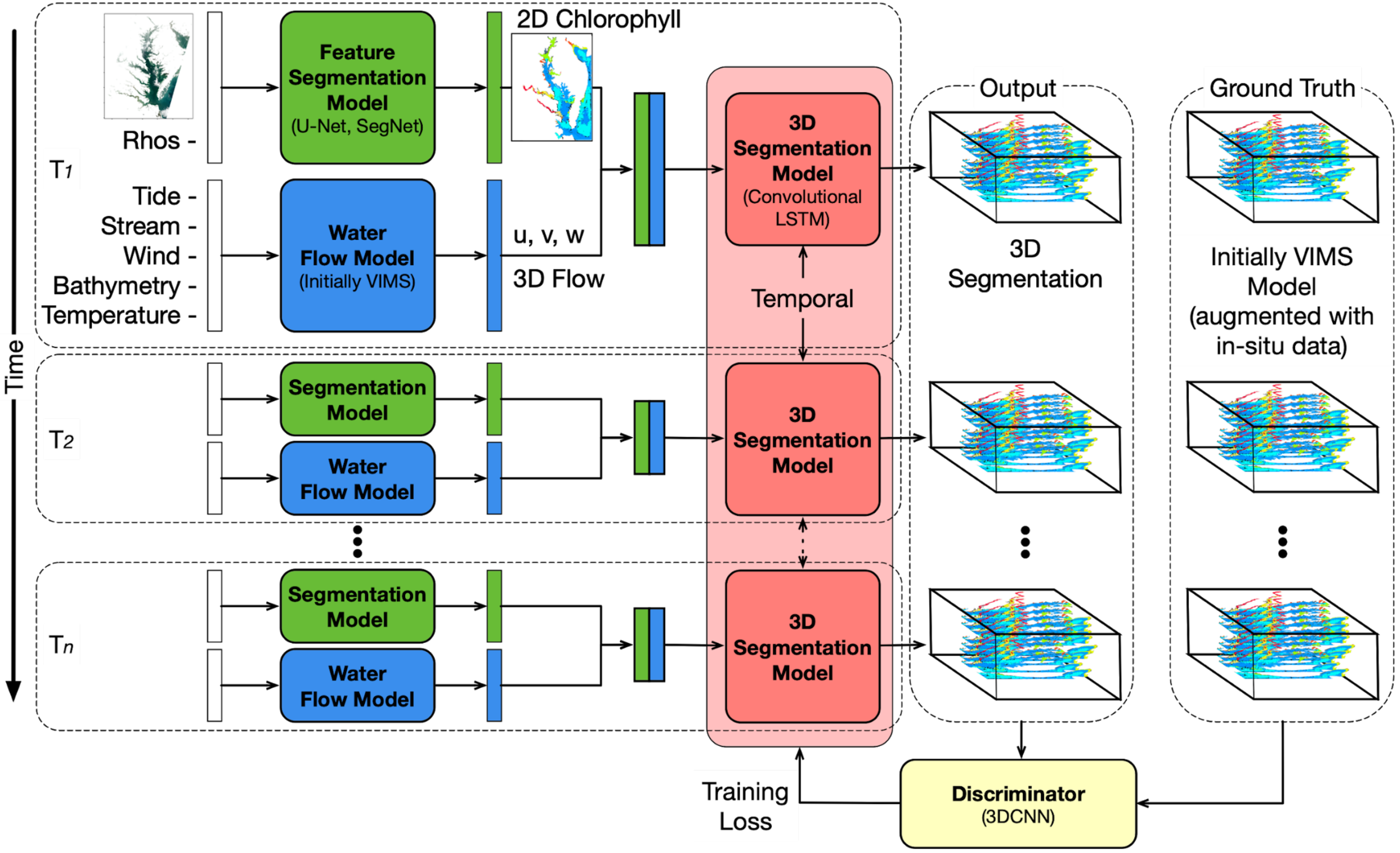
- **Satellite**: Rhos to indicate organic matter
- **ERA5**: wind (u & v)
- **VIMS**: currents (u, v, w), temperature (T), salinity(S)
- **Ancillary**: day of year for seasonal variations

Label variable: DO (from VIMS and CBP in situ data)

- 3-D array prediction for a region

Source: https://www.vims.edu/research/topics/dead_zones/forecasts/cbay

Integration of remote sensing into ML for hypoxia



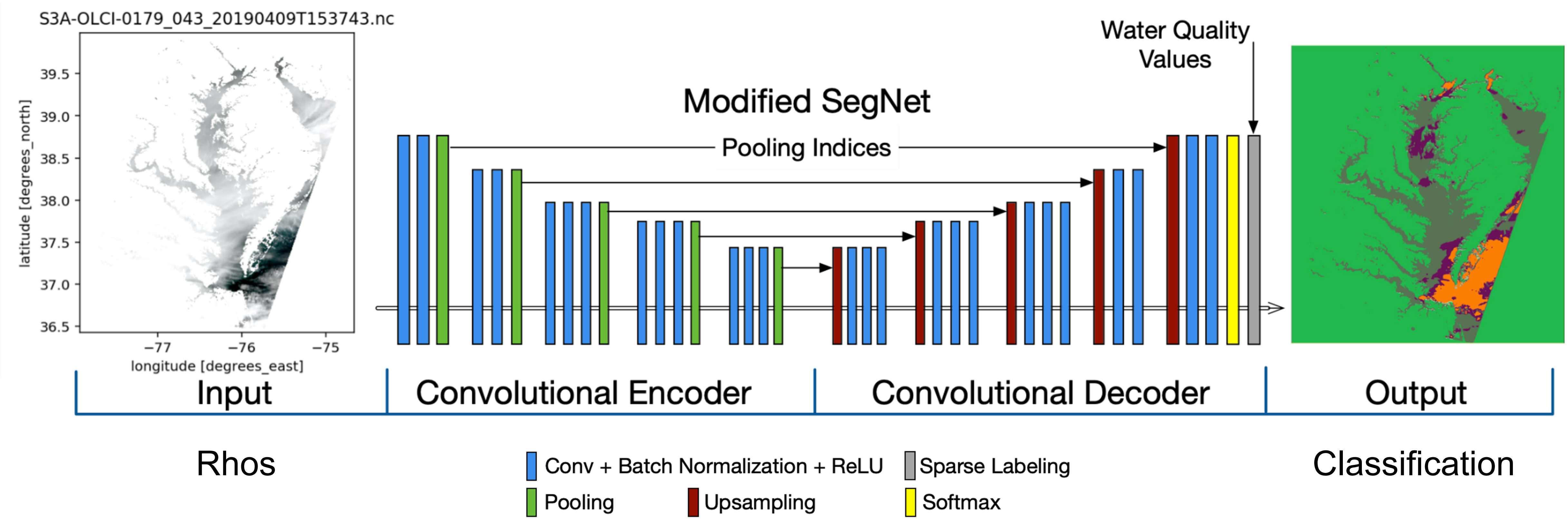
Credit: Troy Ames

Integration of remote sensing into ML for hypoxia

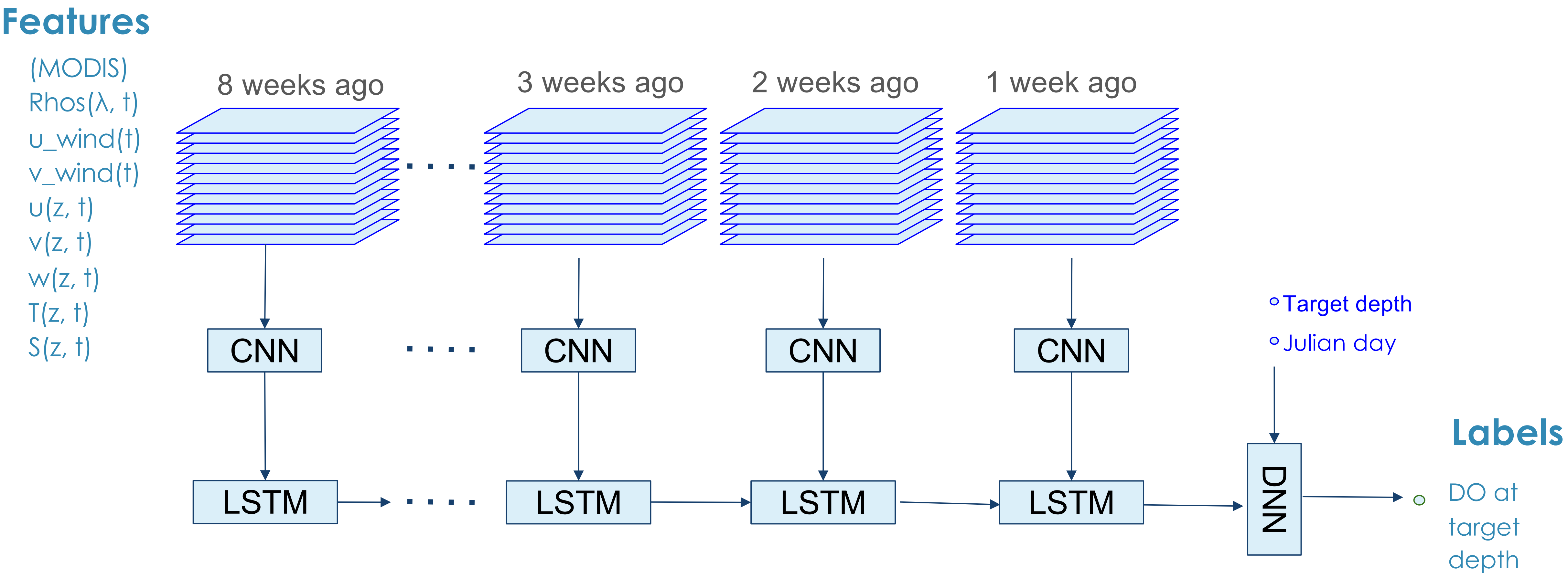
Segmentation Model

Water Quality Parameter Segmentation
Feature Segmentation Model

Feature matching utilizing weakly annotated data (i.e., in situ measurements)



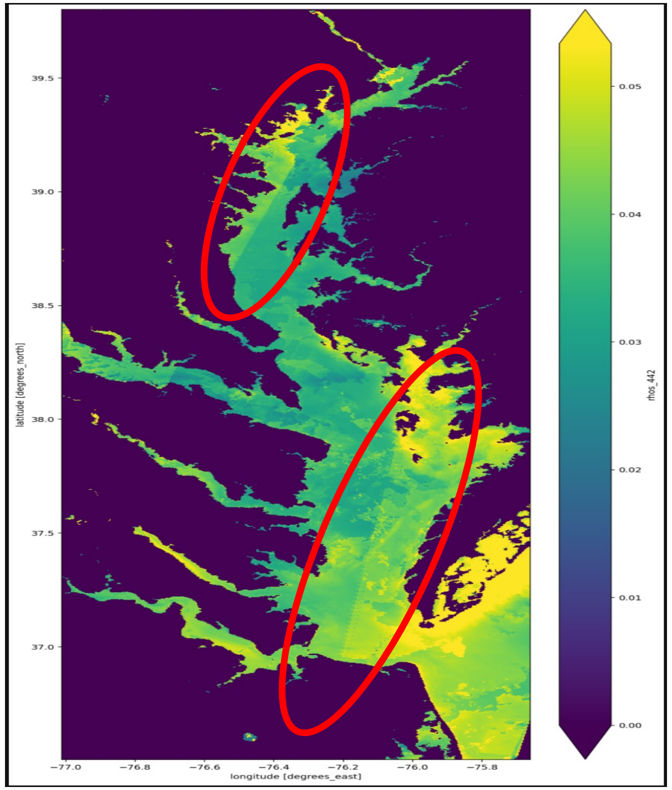
ML architecture around a point of interest to root cause



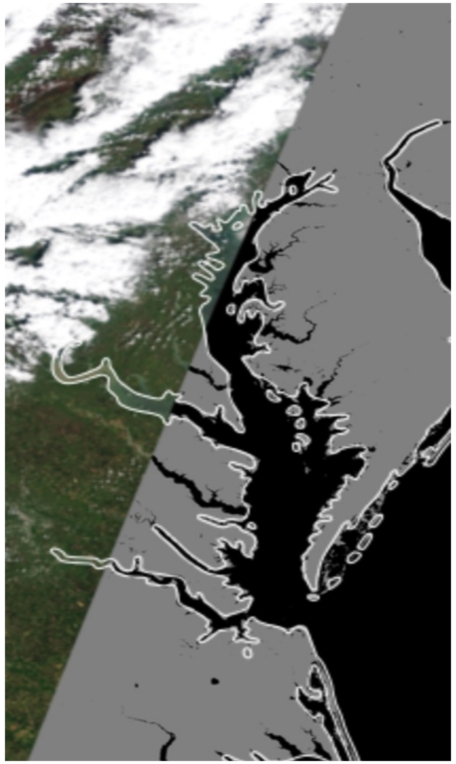
Credit: Guangming Zheng

Preparation of satellite data for bloom detection within ML

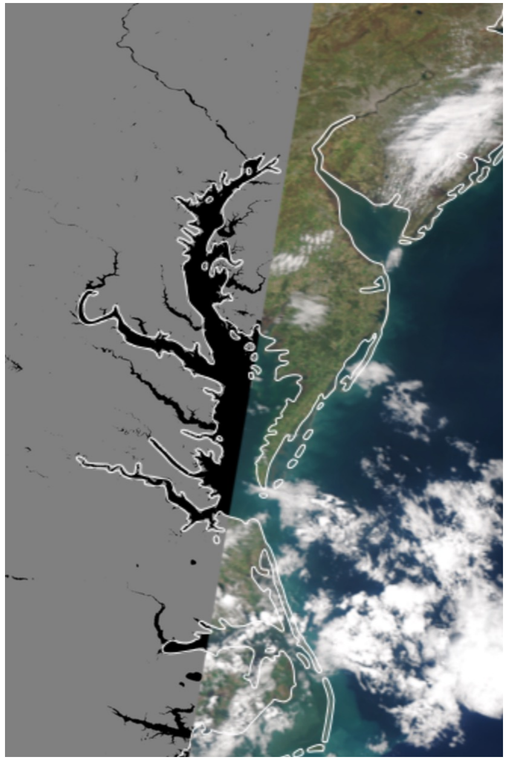
- First look: weekly composites of cloud-free Rhos
- Partial scenes create edges in composites
- Atmospheric gas and aerosol differences between scenes
- Daily scenes to resolve features



Sentinel-3A OLCI image
2020/04/06



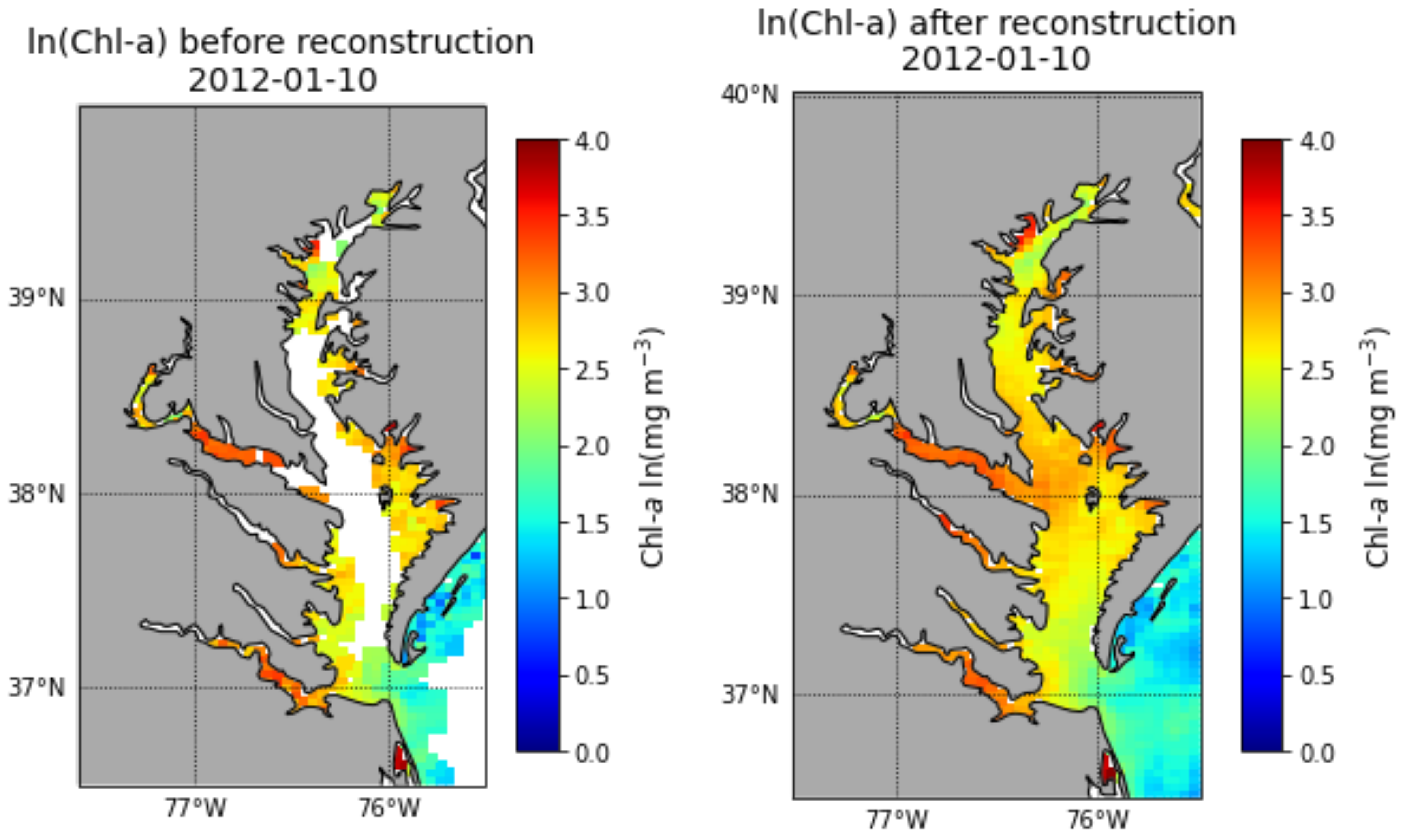
Sentinel-3B OLCI image
2020/04/09



Credit: Amita Mehta

Preparation of satellite data for integration into ML

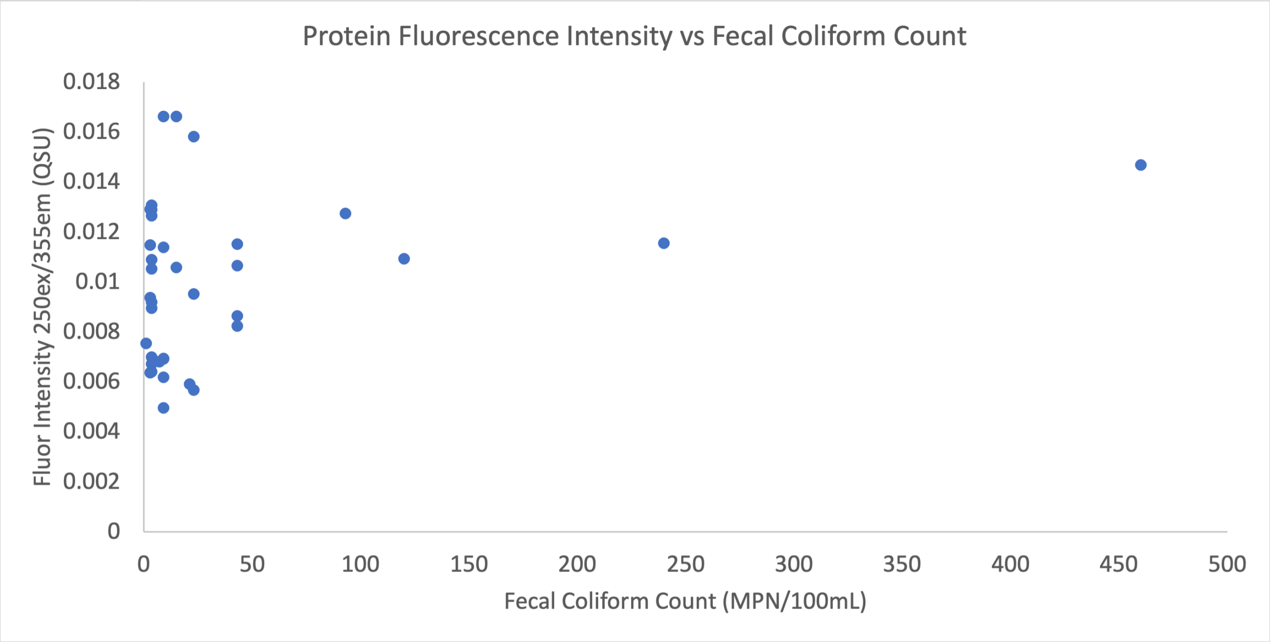
Fill missing satellite data using best interpolation method, i.e. DINEOF



Credit: Nikolay Nezlin

Resume field & lab work for optical indicators of bacteria

- Literature shows correlation between E-coli enumeration and enhanced protein-like fluorescent peaks
- Protein-like peak, tryptophan-like fluorescence, at 250nm excitation and 355 emission

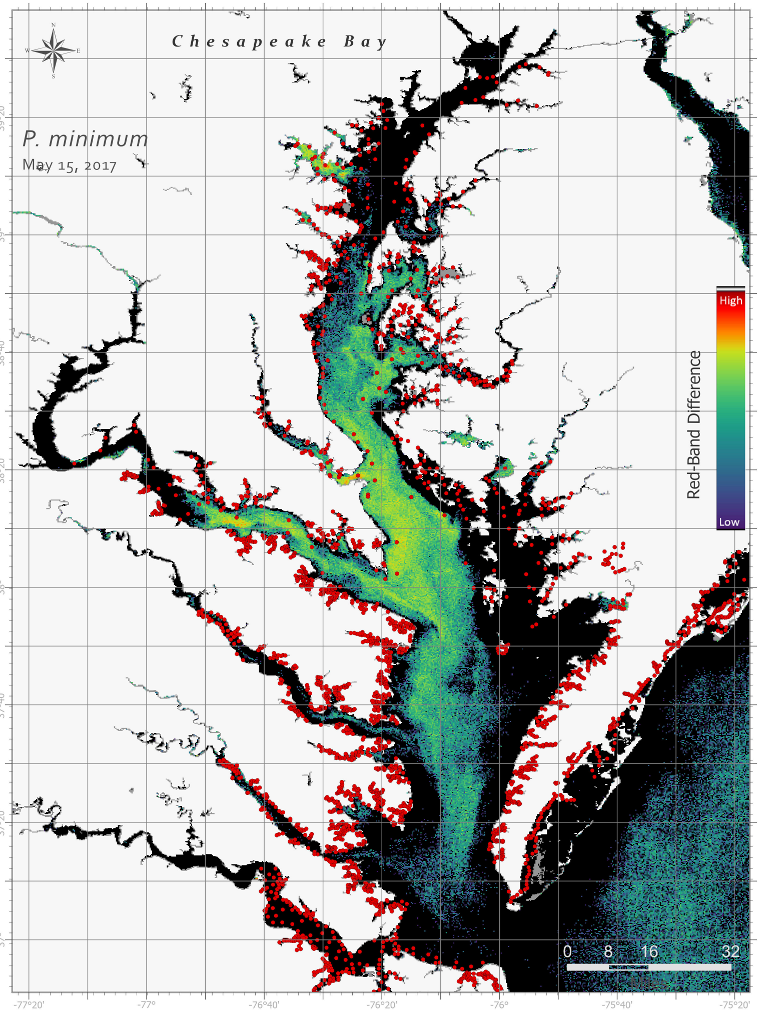


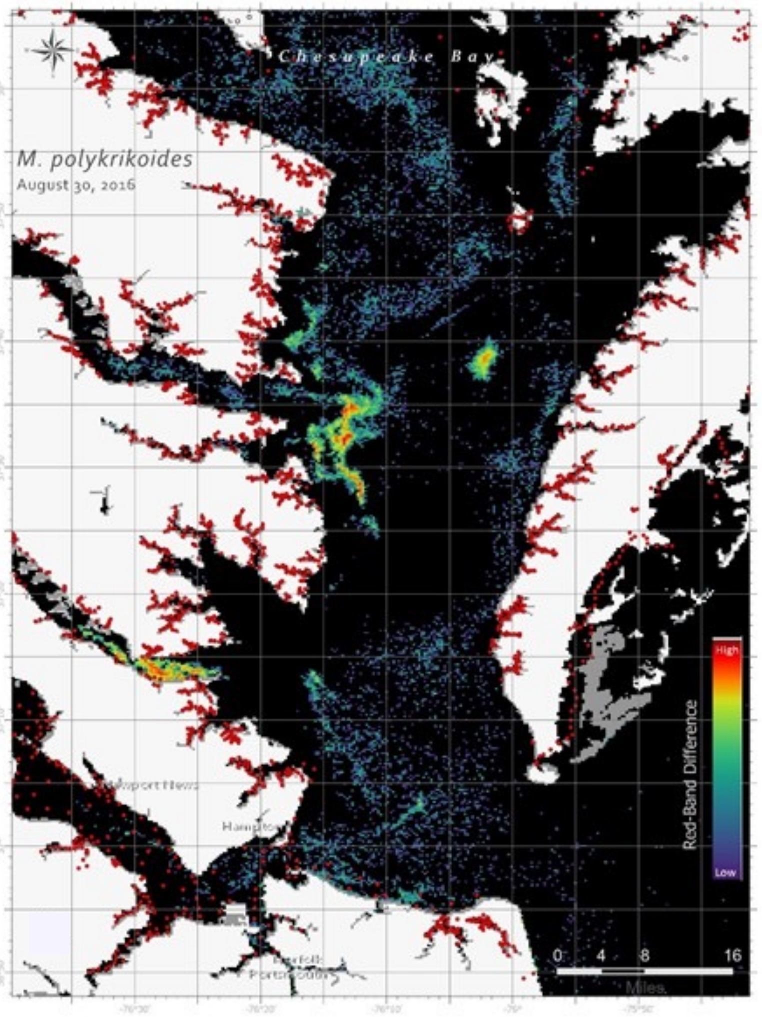
Credit: Shannon McDonnell

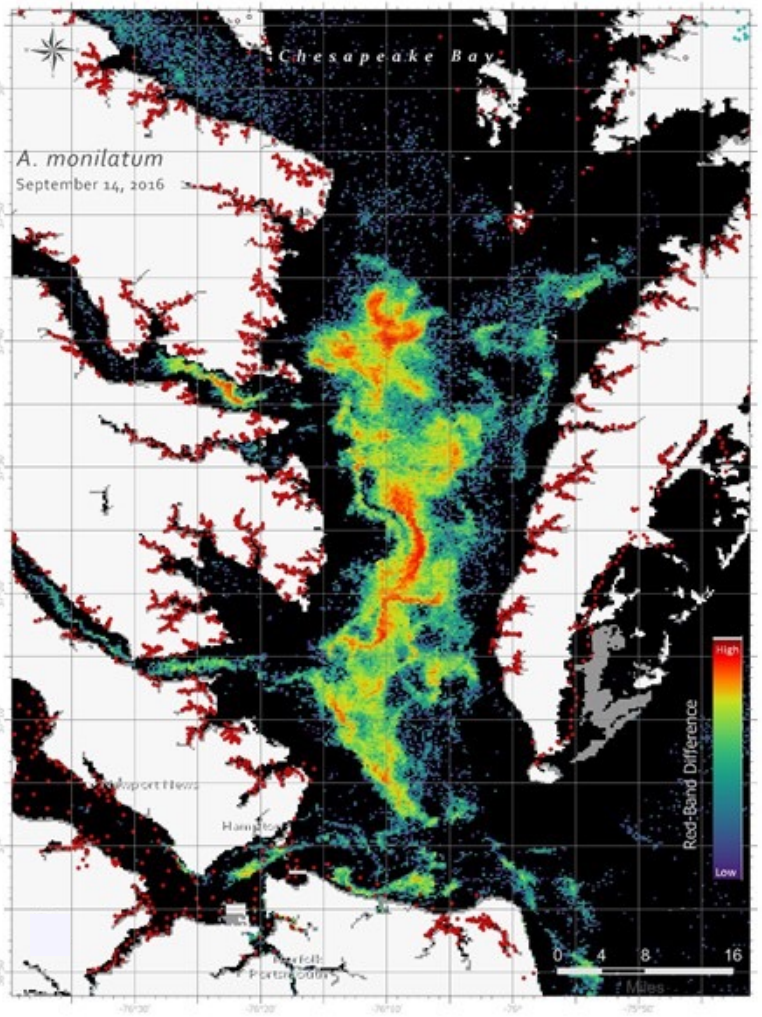
Early results and next steps

- Implementing ML semantic segmentation pipeline modules (U-Net CNN, SegNET, LSTM)
- Evaluating applicable data augmentation strategies for training, e.g. GAN for generating extra training examples utilizing weakly annotated data, i.e. in situ measurements, physical conditions, run-off
- Next: resume data collection and analysis: target poor water quality; distinguish pertinent amino acids, e.g. tyrosine/tryptophan, brighteners.
- Analyze spectral ratios of phytoplankton pigments and bacteria regions within complex CDOM absorption and fluorescence spectra
- Analyze high resolution commercial satellite data in combination with sampling, including DESIS (hyperspectral, 30m with 5m PAN band, launched 2019)

Harmful Algal Bloom detection in the Chesapeake Bay

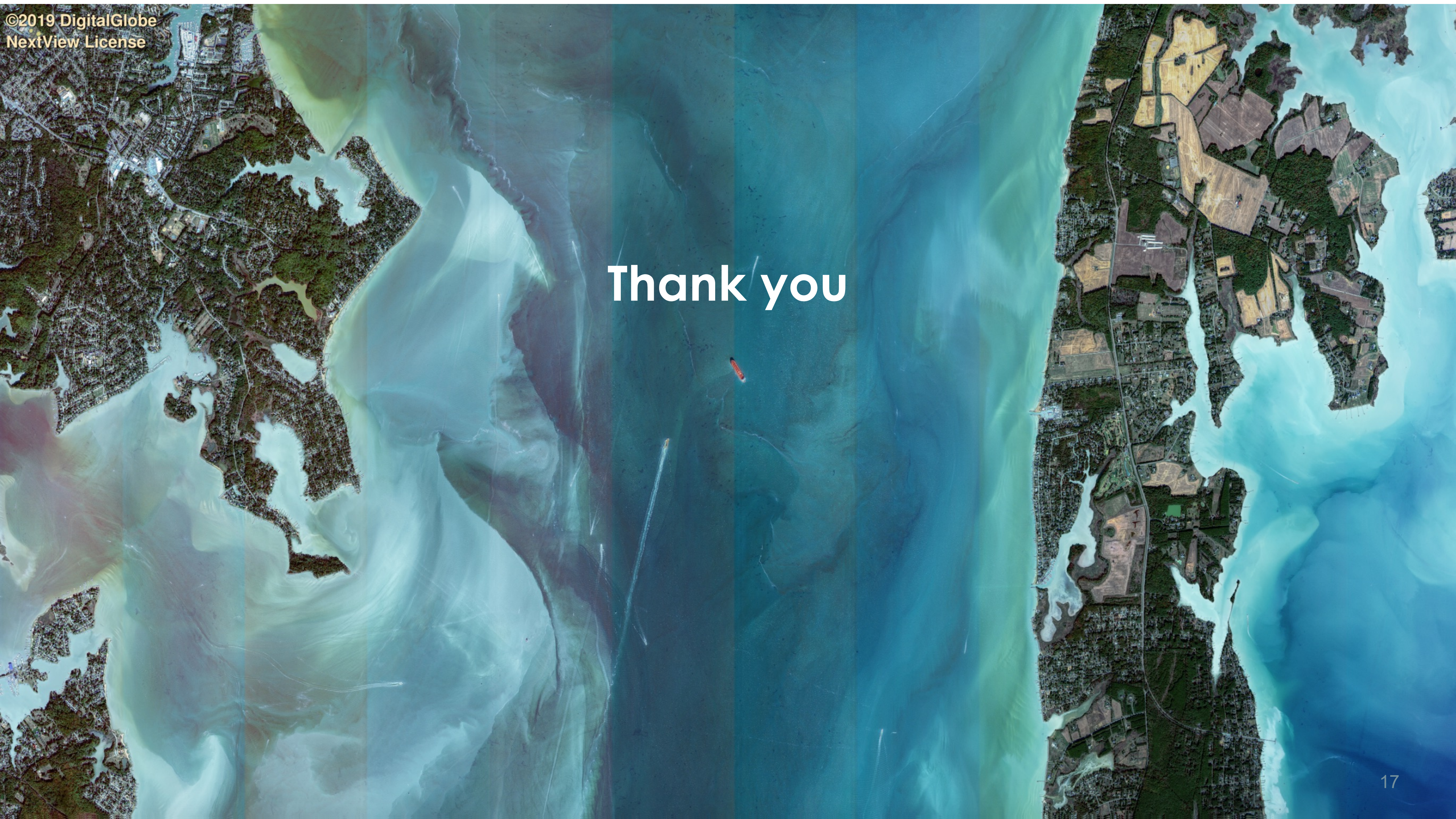






https://coastwatch.noaa.gov/cw_html/NCCOS.html

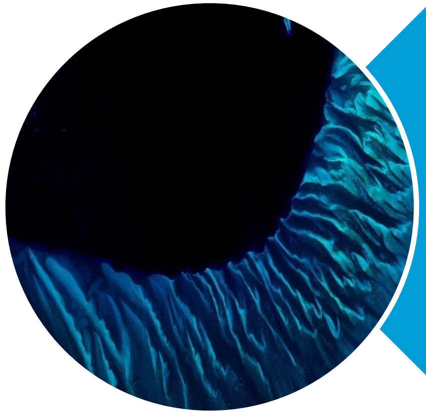
Wolny, J.L., M.C. Tomlinson, S. Schollaert Uz, T.A. Egerton, J.R. McKay, A. Meredith, K.S. Reece, G.P. Scott, and. R.P. Stumpf, 2020, Current and Future Remote Sensing of Harmful Algal Blooms in the Chesapeake Bay to Support the Shellfish Industry, *Front. Mar. Sci., doi:10.3389/fmars.2020.00337*



Thank you

Backup slides

Trade-offs in Satellite Technology



Trade-off in spectral vs spatial resolution

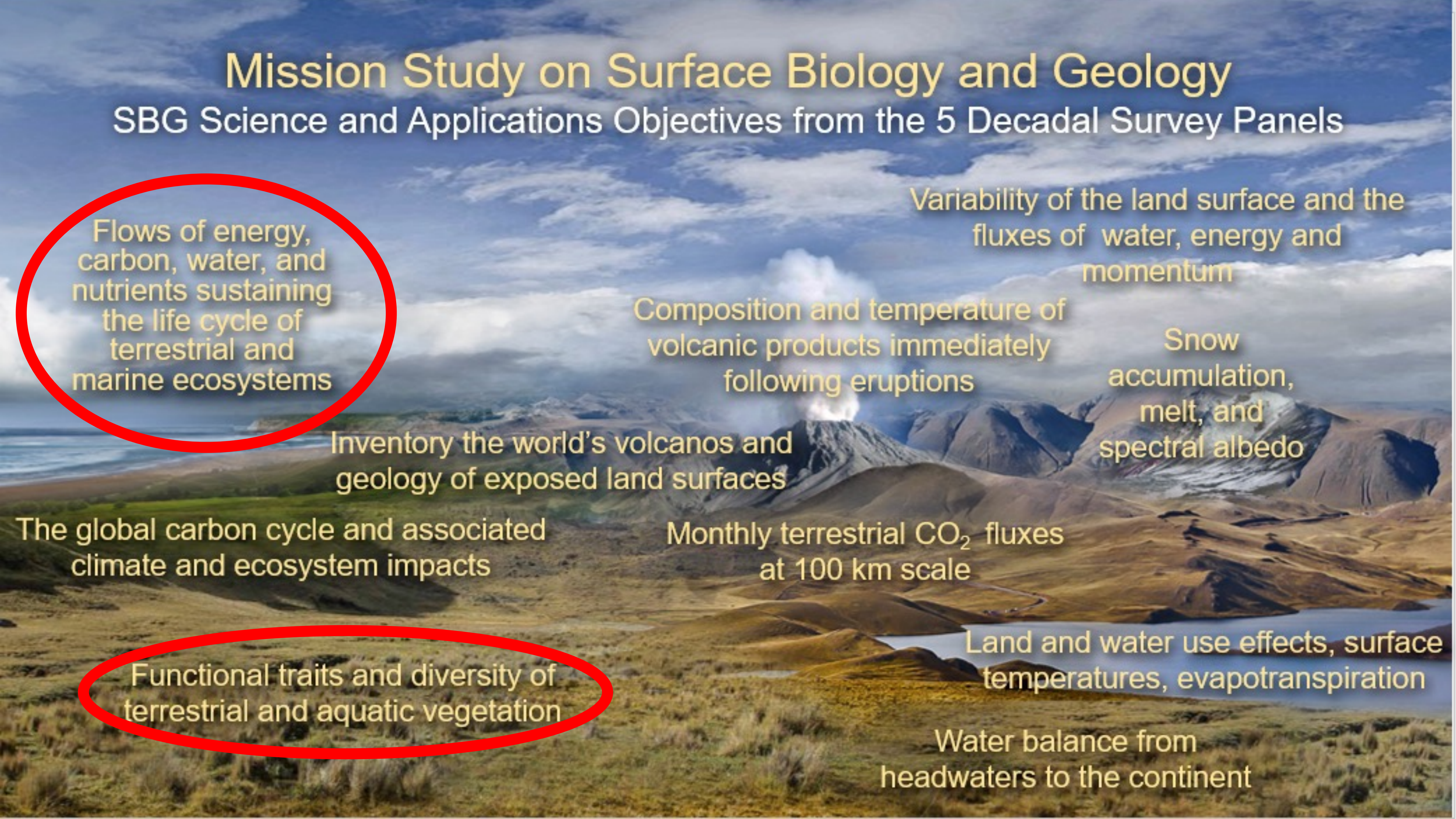
- More narrow spectral bands → Larger bins or pixels
- Few broad spectral bands → Smaller pixel

Trade-off in frequency vs spatial resolution

- Larger pixel → More frequent revisit
- Smaller pixel → Less frequent revisit

Digital Globe, Planet, etc.	• High spatial <1 m • Low spectral & uncalibrated • Low temporal, upon tasking (not guaranteed)
Landsat OLI, Sentinel2 MSI	• Medium spatial (10-30 m), Global coastal • Low spectral (3 channels & sun glint issues) • Low temporal (8-16 day revisit)
MODIS, S3 OLCI Ocean Color	• Low spatial (300-1000 m), Global • High spectral (5 nm bands) • Medium temporal (2 day revisit)
GLIMR Geostationary	• Medium spatial (30 m), Regional over GoMex, U.S. • High spectral (10 nm bands) • High temporal (Several looks per day)
Aircraft and Drones	• High spatial (1-10 m) Local • Potentially high spectral (5 nm bands) • Event and permission-specific (i.e. away from airports and > 50mi from U.S. Capitol)

NASA EARTH FLEET
OPERATING & FUTURE THROUGH 2023
SWOT (CNES)
LANDSAT-9 (USGS)
SENTINEL-6 Michael Freilich/B (ESA)
TROPICS (6)
GEOCARB
NISAR (ISRO)
MAIA
TSIS-2
TEMPO
PREFIRE (2)
PACE (NSO)
GLIMR
ICESAT-2
GRACE-FO (2) (DLR)
CYGNSS (8)
NISTAR, EPIC (DSCOVR/NOAA)
CLOUDSAT (CSA)
TERRA (JAXA, CSA)
AQUA (JAXA, AEB)
AURA (NSO, FMI, UKSA)
CALIPSO (CNES)
GPM (JAXA)
LANDSAT 7 (USGS)
LANDSAT 8 (USGS)
OCO-2
SMAP
SUOMI NPP (NOAA) (JAXA)
ISS INSTRUMENTS
EMIT
CLARREO-PF
GEDI
SAGE III
OCO-3
TSIS-1
ECOSTRESS
LIS
JPSS-2, 3 & 4 INSTRUMENTS
OMPS-Limb
LIBERA
03.24.20
INVEST/CUBESATS
RainCube
CSIM-FD
CubeRRT
TEMPEST-D
CIRiS
HARP
CTIM
HyTI
SNoOPI
NACHOS
(PRE) FORMULATION
IMPLEMENTATON
PRIMARY OPS
EXTENDED OPS


Mission Study on Surface Biology and Geology
SBG Science and Applications Objectives from the 5 Decadal Survey Panels
Flows of energy, carbon, water, and nutrients sustaining the life cycle of terrestrial and marine ecosystems
Variability of the land surface and the fluxes of water, energy and momentum
Composition and temperature of volcanic products immediately following eruptions
Snow accumulation, melt, and spectral albedo
Inventory the world's volcanos and geology of exposed land surfaces
The global carbon cycle and associated climate and ecosystem impacts
Monthly terrestrial CO_2 fluxes at 100 km scale
Land and water use effects, surface temperatures, evapotranspiration
Functional traits and diversity of terrestrial and aquatic vegetation
Water balance from headwaters to the continent