Supporting Shellfish Aquaculture in the Chesapeake
Using Artificial Intelligence to Detect Poor Water Quality through Sampling and Remote Sensing

Stephanie Schollaert Uz¹
Troy Ames¹, Nargess Memarsadeghi¹,
Amita Mehta¹,², Guangming Zheng²,³, Shannon McDonnell², Neil Blough², John McKay⁴

¹Goddard Space Flight Center
²University of Maryland
³National Oceanic and Atmospheric Administration
⁴Maryland 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.

Optical techniques are explored to exploit new technology in complex environment

Developing AI combining observations, models, with satellite data

Photo credit: John ‘Rusty’ McKay/MDE
Future space sensors will provide information about phytoplankton communities, ecosystem health.

Source: https://pace.gsfc.nasa.gov
### Maryland water quality criteria for shellfish harvesting

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

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

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

MODIS chlorophyll-a map from July 2, 2019 with routine sampling sites by Maryland and Virginia superimposed.
Water quality criteria for shellfish harvesting

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

- 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

http://www.dsd.state.md.us/comar/comarhtml/26/26.08.03-3.htm

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

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

**Features**

- MODIS
- $\text{Rhos}(\lambda, t)$
- $u_{\text{wind}}(t)$
- $v_{\text{wind}}(t)$
- $u(z, t)$
- $v(z, t)$
- $w(z, t)$
- $T(z, t)$
- $S(z, t)$

**Labels**

- DO at target depth

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

Credit: Amita Mehta
Preparation of satellite data for integration into ML

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

Thank you
Backup slides
Trade-offs in Satellite Technology

Trade-off in spectral vs spatial resolution
- More narrow spectral bands $\rightarrow$ Larger bins or pixels
- Few broad spectral bands $\rightarrow$ Smaller pixel

Trade-off in frequency vs spatial resolution
- Larger pixel $\rightarrow$ More frequent revisit
- Smaller pixel $\rightarrow$ Less frequent revisit
<table>
<thead>
<tr>
<th>Dataset Type</th>
<th>Spatial Resolution</th>
<th>Spectral Resolution</th>
<th>Temporal Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital Globe, Planet, etc.</td>
<td>High spatial &lt;1 m</td>
<td>Low spectral &amp; uncalibrated</td>
<td>Low temporal, upon tasking (not guaranteed)</td>
</tr>
<tr>
<td>Landsat OLI, Sentinel2 MSI</td>
<td>Medium spatial (10-30 m), Global coastal</td>
<td>Low spectral (3 channels &amp; sun glint issues)</td>
<td>Low temporal (8-16 day revisit)</td>
</tr>
<tr>
<td>MODIS, S3 OLCI Ocean Color</td>
<td>Low spatial (300-1000 m), Global</td>
<td>High spectral (5 nm bands)</td>
<td>Medium temporal (2 day revisit)</td>
</tr>
<tr>
<td>GLIMR Geostationary</td>
<td>Medium spatial (30 m), Regional over GoMex, U.S.</td>
<td>High spectral (10 nm bands)</td>
<td>High temporal (Several looks per day)</td>
</tr>
<tr>
<td>Aircraft and Drones</td>
<td>High spatial (1-10 m) Local</td>
<td>Potentially high spectral (5 nm bands)</td>
<td>Event and permission-specific (i.e. away from airports and &gt; 50mi from U.S. Capitol)</td>
</tr>
</tbody>
</table>
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
- Monthly terrestrial CO₂ fluxes at 100 km scale
- The global carbon cycle and associated climate and ecosystem impacts
- Land and water use effects, surface temperatures, evapotranspiration
- Functional traits and diversity of terrestrial and aquatic vegetation
- Water balance from headwaters to the continent