Hyperspectral Remote Sensing of Aquatic Microbes to Support Water Resource Management

Grace E. Kim
Strategic Innovation Group
Booz Allen Hamilton
Washington, DC, USA
kim_grace2@bah.com

Evan Poworoznek
Strategic Innovation Group
Booz Allen Hamilton
Rome, NY, USA
Poworoznek_Evan@bah.com

Susanne E. Craig
Ocean Ecology Laboratory
NASA Goddard Space Flight Center
Greenbelt, MD, USA
susanne.e.craig@nasa.gov

Abstract

Harmful algal blooms in drinking water supply and at recreational sites endanger human health. Excessive algal growth can result in low oxygen environments, making them uninhabitable for fish and other aquatic life. Harmful algae and algal blooms are predicted to increase in frequency and extent due to the warming climate, but microbial dynamics remain difficult to predict. Existing satellite remote sensing monitoring technologies are ill-equipped to discriminate harmful algae, while models do not adequately capture the complex controls on algal populations. This proposal explores the potential for Bayesian neural networks to detect phytoplankton pigments from hyperspectral remote sensing reflectance retrievals. Once developed, such a model could enable hyperspectral remote sensing retrievals to support decision making in water resource management as more advanced ocean color satellites are launched in the coming decade. While uncertainty quantification motivates the proposed use of Bayesian models, the interpretation of these uncertainties in an operational context must be carefully considered.

1 Introduction

Coastal waters support fisheries, which provide livelihoods and recreation for nearby communities. Reservoirs and lakes are major sources of drinking water. These aquatic systems are increasingly experiencing anthropogenic stresses from population growth and land development, along with longer term shifts in temperature from a changing climate. These water bodies are full of microscopic life, which serve as the primary producers for aquatic ecosystems and mediate biogeochemical processes. Over past decades, rising temperatures and increasing pollution have been linked to ecosystem shifts such as higher occurrences of harmful algal blooms (1) and low-oxygen waters caused by excessive algal growth (2). Harmful algae release toxins that can cause illness and even death when ingested directly or indirectly through shellfish poisoning. Low oxygen waters are uninhabitable to fish and other benthos, resulting in massive fish kills.

Due to the ecological importance of aquatic microbes, oceanographers have deployed advanced technologies to monitor and study their population dynamics. Polar-orbiting ocean color satellites provide observations of the world's oceans every few days. Geostationary ocean color satellites

Tackling Climate Change with Machine Learning workshop at NeurIPS 2020.

Figure 1: Harmful algal bloom at a drinking water intake in Lake Eerie. (Photo Credit epa.gov)

examine one region of the earth, providing multiple observations per hour. These satellites have historically carried multi-spectral spectrometers, which senses the Earth's radiance at a several discrete wavelengths from the ultraviolet to near infrared. Both types of sensors have been employed for monitoring harmful algae (3) and phytoplankton dynamics (4), but are limited by their coarse spectral resolution. The future generation of United States' ocean color sensors being developed will provide imagery at much greater spectral resolution, as the PACE (Plankton, Aerosol, Cloud, ocean Ecosystem) and GLIMR (Geosynchronous Littoral Imaging and Monitoring Radiometer) missions become operational in the coming decade.

Existing algorithms for remote sensing of the environment typically employ a two-step approach that seeks to first (1) remove the atmospheric interference between the sensor and the target and then (2) derive the geophysical parameter of interest from the atmosphere-corrected retrieval. In aquatic remote sensing, algorithms for atmospheric correction and parameter retrieval are largely derived from open ocean studies. However, these algorithms perform poorly in coastal and inland waters due to the complex mixtures of materials in the water and air (e.g. sediment, dissolved organic material, aerosols) specific to near-shore environments. Poor performance and the lack of a widely accepted uncertainty quantification have limited their use for decision making in water resource management. The improved spectral resolution of future ocean color sensors are promising hardware advances, but must be accompanied by advances in retrieval algorithms suitable for the increase in information available.

2 Methods

2.1 Model Description & Training Data

We propose the application Bayesian Neural Networks for monitoring microbial communities with hyperspectral sensors to support water resource management. Bayesian methodologies have been shown to improve accuracy in atmospheric correction (5) and in deriving biogeochemical information (6) for existing ocean color sensors over traditional algorithms. The appeal of the Bayesian approach for this problem is the ability to model the expectation and covariance of the posterior distribution, providing a measure of uncertainty for retrievals. In both (5; 6), models were applied to a discrete set of wavelengths to match existing ocean color sensors, although the training and validation data are available at higher spectral resolution. Within this Bayesian framework, we propose the implementation of neural networks due to the high dimensionality of the input space associated with hyperspectral imagery.

Specifically, the model proposed is a Bayesian neural network for predicting phytoplankton pigment concentrations $[mg\ m^{-3}]$ from remote sensing reflectance ($R_{rs}(\lambda)$ $[sr^{-1}]$). Field deployments of in situ spectroradiometers have traditionally been used to validate satellite sensors and atmospheric correction algorithms. Many of these measurements have been accompanied by coincident measurements of phytoplankton pigments, which are widely used as indicators of phytoplankton species in the aquatic environment. Harmful algae often shift the color of a water body when they are in full bloom, reflecting the unique pigment signatures associated with these organisms.

Data to train this model will draw from decades of ocean color measurements curated in NASA's SeaBASS database. Potential test datasets for this model include hyperspectral imaging studies in regions where harmful algal blooms cause recurring challenges to drinking water (Lake Eerie) or recreation (Gulf of Mexico). Additionally, regions within view of the South Korean Geostationary Ocean Color Imager (GOCI) are applicable for studying sub-daily scale coastal phytoplankton dynamics (e.g. KORUS-OC Campaign).

2.2 Data Pre-Processing

Machine learning based dimensionality reduction and data compression methods aid in the interpretation of images from hyperspectral sensors, which exhibit much higher resolution than their multispectral predecessors. Our proposed work builds on previous research that successfully derived phytoplankton proxies from hyperspectral radiometry using empirical orthogonal function (EOF) analysis to discriminate optical signals that are not apparent in higher dimensional space (7; 8). Pre-processing remote sensing reflectances by normalization and EOF analysis captures key modes of variability in the spectra, which were found to be leading predictors of an optical proxy of the phytoplankton community in (6).

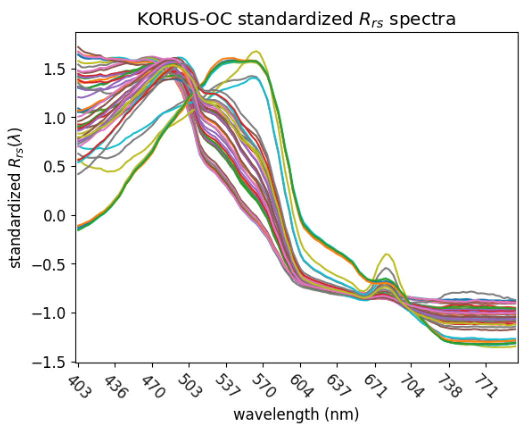

Figure 2: In situ remote sensing reflectance spectra (N=49) from the KORUS-OC Campaign. Data can be accessed via doi:10.5067/SeaBASS/KORUS/DATA001. EOF analysis in Appendix.

3 Broader Impact: Model Deployment and Decision Support

Recurring blooms of the cyanobacteria *Microcystis* in Lake Eerie can contaminate drinking water, while *Karenia brevis* blooms in Florida can cause shellfish poisoning and widespread fish mortality. These events are difficult to predict due to the myriad of environmental and biological controls on algal dynamics. As the frequency of harmful algal blooms (HABs) is predicted to increase with climate change, water resource managers are faced with the challenge of anticipating and monitoring for harmful levels of toxins. However, monitoring water quality is resource intensive, requiring frequent in-person water sample collection and analysis. The work in this proposal seeks to employ a combination of Bayesian and neural network approaches on hyperspectral imagery to aid decision support for closures and enable predictive modeling for HABs.

3.1 MLOps for water resource management

The model proposed for accurately detecting phytoplankton pigments in imagery alone will not address climate change related challenges for water resource management. The operational pipeline best suited for an application should be deployed based on the time scale and spatial extent of interest. For example, policy decisions related to land use change considered concurrently with climate stressors (e.g. (9; 10; 1)) or the contribution of coastal aquatic ecosystems for carbon sequestration (11) will require data in batches for intermittent studies and evaluation. Alternatively, a streaming or real-time deployment framework would be more appropriate for advisories. For example, a decision support tool for toxin and algal concentrations might be designed to ingest live monitoring streams from sensors to and satellites to send an alert when modeled toxin concentrations exceed management thresholds (e.g. microcystin concentration exceeding 0.3 ug/L triggers a drinking water advisory for young children, 1.6 ug/L for adults). Such an implementation framework might employ multiple machine learning models in the pipeline, ingesting multiple data streams and delivering actionable information to decision makers. While the uncertainty quantification motivates our proposed use of Bayesian models, our future work includes contemplation of the propagation of prior uncertainties through a neural network, and its appropriate interpretation in an operational context.

References

[1] H. W. Paerl, W. S. Gardner, K. E. Havens, A. R. Joyner, M. J. McCarthy, S. E. Newell, B. Qin, and J. T. Scott, “Mitigating cyanobacterial harmful algal blooms in aquatic ecosystems impacted by climate change and anthropogenic nutrients,” *Harmful Algae*, vol. 54, pp. 213–222, 2016.

[2] D. Breitburg, L. A. Levin, A. Oschlies, M. Grégoire, F. P. Chavez, D. J. Conley, V. Garçon, D. Gilbert, D. Gutiérrez, K. Isensee, G. S. Jacinto, K. E. Limburg, I. Montes, S. W. A. Naqvi, G. C. Pitcher, N. N. Rabalais, M. R. Roman, K. A. Rose, B. A. Seibel, M. Telszewski, M. Yasuhara, and J. Zhang, “Declining oxygen in the global ocean and coastal waters,” *Science*, vol. 359, no. 6371, 2018.

[3] J. M. Clark, B. A. Schaeffer, J. A. Darling, E. A. Urquhart, J. M. Johnston, A. R. Ignatius, M. H. Myer, K. A. Loftin, P. J. Werdell, and R. P. Stumpf, “Satellite monitoring of cyanobacterial harmful algal bloom frequency in recreational waters and drinking water sources,” *Ecological indicators*, vol. 80, pp. 84–95, 2017.

[4] J. Concha, A. Mannino, B. Franz, and W. Kim, “Uncertainties in the geostationary ocean color imager (goci) remote sensing reflectance for assessing diurnal variability of biogeochemical processes,” *Remote Sensing*, vol. 11, no. 3, p. 295, 2019.

[5] R. Frouin and B. Pelletier, “Bayesian methodology for inverting satellite ocean-color data,” *Remote Sensing of Environment*, vol. 159, pp. 332–360, 2015.

[6] S. E. Craig and E. M. Karaköylü, “Bayesian models for deriving biogeochemical information from satellite ocean color,” *EarthArXiv*, 2019.

[7] M. Soja-Woźniak, S. E. Craig, S. Kratzer, B. Wojtasiewicz, M. Darecki, and C. T. Jones, “A novel statistical approach for ocean colour estimation of inherent optical properties and cyanobacteria abundance in optically complex waters,” *Remote Sensing*, vol. 9, no. 4, p. 343, 2017.

[8] S. E. Craig, C. T. Jones, W. K. Li, G. Lazin, E. Horne, C. Caverhill, and J. J. Cullen, “Deriving optical metrics of coastal phytoplankton biomass from ocean colour,” *Remote Sensing of Environment*, vol. 119, pp. 72–83, 2012.

[9] I. D. Irby, M. A. M. Friedrichs, F. Da, and K. E. Hinson, “The competing impacts of climate change and nutrient reductions on dissolved oxygen in chesapeake bay,” *Biogeosciences*, vol. 15, no. 9, pp. 2649–2668, 2018.

[10] E. Sinha, A. Michalak, K. V. Calvin, and P. J. Lawrence, “Societal decisions about climate mitigation will have dramatic impacts on eutrophication in the 21 st century,” *Nature communications*, vol. 10, no. 1, pp. 1–11, 2019.

[11] C. S. Hopkinson, W.-J. Cai, and X. Hu, “Carbon sequestration in wetland dominated coastal systems—a global sink of rapidly diminishing magnitude,” *Current Opinion in Environmental Sustainability*, vol. 4, no. 2, pp. 186–194, 2012.

4 Appendix

4.1 Supplementary Figures

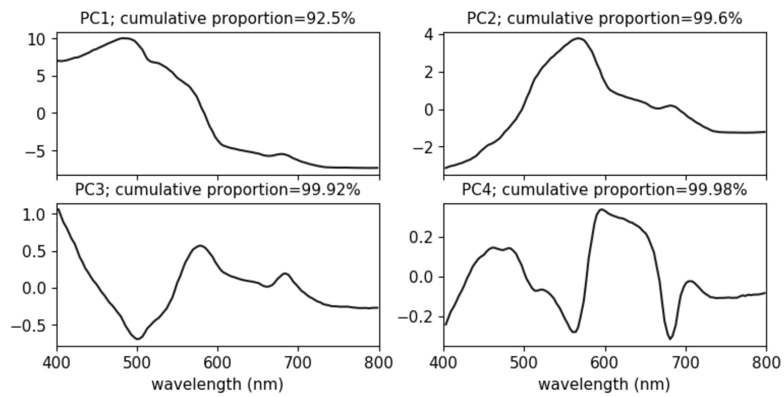

Figure 3: First 4 principal components from EOF analysis of remote sensing reflectances in Fig. 2.

4.2 Data and Code

Data from Fig. 2 can be accessed at `https://seabass.gsfc.nasa.gov/cruise/kr_2016`. Data cleanup and analysis available on Databricks community cloud, here.