Relationships between Volatile organic compounds and harmful algal blooms: Can VOCs be used as a time-sensitive and economical means for predicting bloom dynamics?

Lindsay Collart[1] , Kimberly Halsey[1]
[1]Department of Microbiology, Oregon State University

Introduction

Freshwater harmful algal blooms (HABs) are an expected occurrence during the summer months in the eutrophic Upper Klamath River Watershed. Agency Lake (AL) and Upper Klamath Lake (UKL) experience transitioning blooms that are dominated by the cyanobacterial genuses *Anabaena*, *Aphanizomenon*, *Gloeotrichia*, and *Microcystis*. HABs can negatively influence the ecology, potability, and recreation opportunities of these freshwater systems, and are themselves impacted by nutrient loading, temperature, pH, and a myriad of other environmental factors.

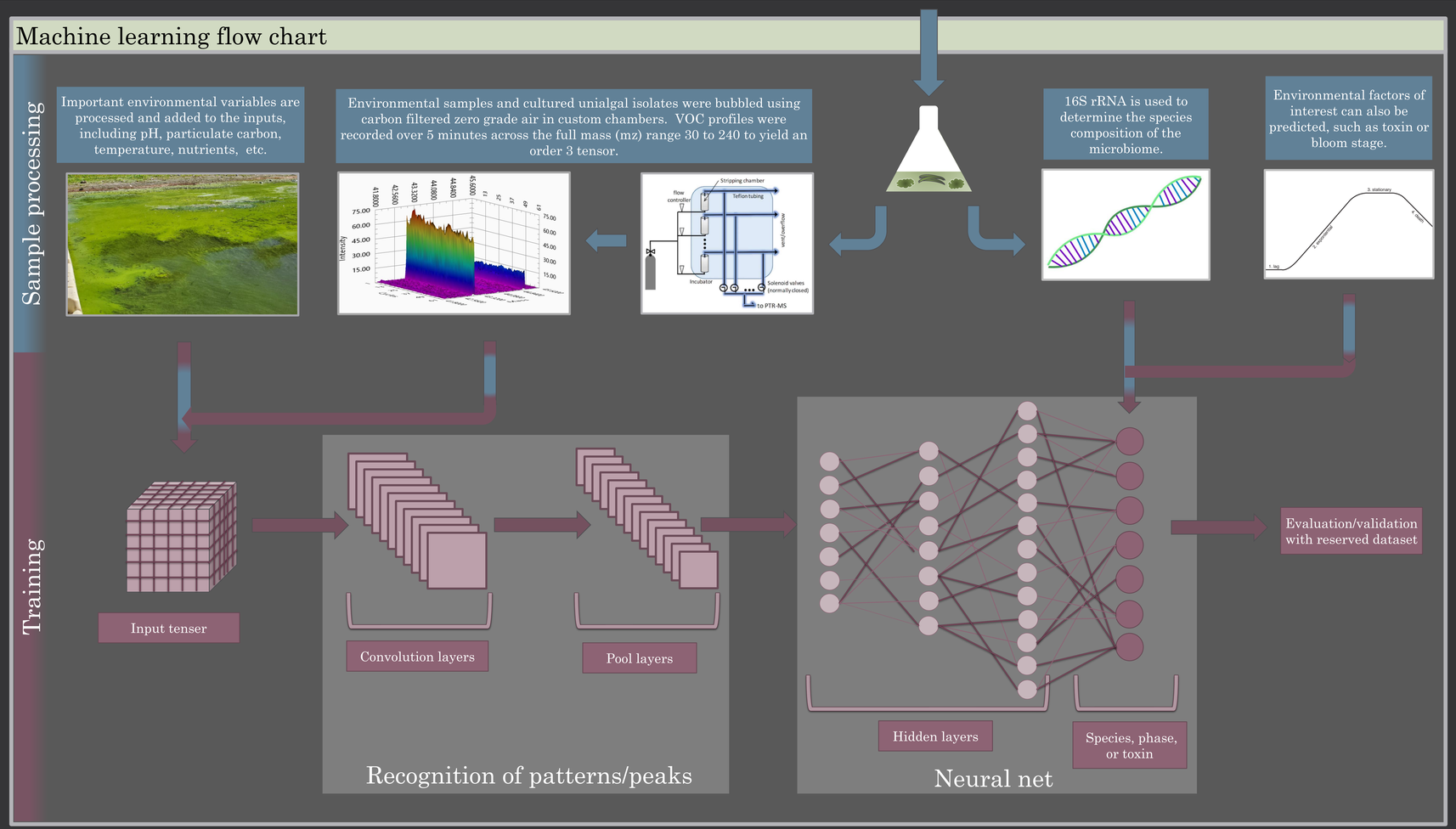
The interactions of HABs and **volatile organic compounds (VOCs)** are scarcely understood. An individual cyanobacterial strain can produce a wide range of VOCs depending on the form of nutrients that are available; some VOCs can inhibit growth of competing phytoplankton (1). However, in situ studies on the impacts of VOCs in freshwater systems and how they are influenced by biotic and physicochemical properties are rare (2). The full spectrum of these VOCs may be unique to community composition or seasons, and could pose additional risk to human health and crop production due to their contributions to ground level ozone formation(3). To address this lack of knowledge, we carried out an exploratory study to determine how VOC production, environmental variables, and microbial community are interconnected in UKL and AL. We envision that a full analysis of these interactions could be facilitated by machine learning that will yield a quick and efficient monitoring practice that may have predictive capacity for HAB dynamics.

Methods

Samples were obtained at 3 sites within UKL and AL, chosen based on their proximity to existing monitoring sites (Figure 1). Sampling was conducted May-December of 2018 and 2019. Abiotic properties were collected (chlorophyll, carbon, nitrogen, toxin concentrations, ozone, conductivity, pH, ect.) and are being processed either in house or through the Institute for Water & Watersheds' Collaboratory on Oregon State University campus. DNA samples were obtained via filtration for community analysis by 16S rRNA sequencing using MiSeq at OSU's Center for Genome Research and Biocomputing.

To obtain VOC profiles (the full spectrum of VOCs in water samples) 100 mL of sample was bubbled to "strip" the VOCs from the water and feed the VOCs into a PTR-ToF/MS (see methods flow chart for details). The samples were run in triplicate with blanks of HPLC grade water as controls for the profiles. Unialgal cultures were also isolated from each genera of interest and VOCs are being assessed in both stationary and exponential phase. The field samples and isolate data will be combined to use for unsupervised training of a convolutional neural network. Multivariant statistics and complex network theory are being used in parallel to elucidate connections between community, environmental variables, and VOCs.

Machine learning flow chart



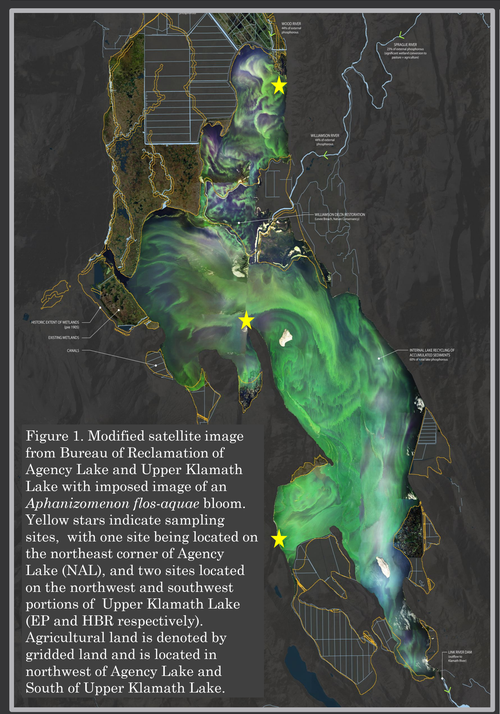
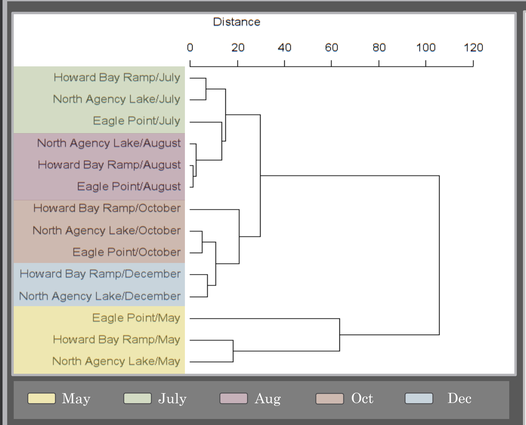


Figure 1. Modified satellite image from Bureau of Reclamation of Agency Lake and Upper Klamath Lake with imposed image of an *Aphanizomenon flos-aquae* bloom. Yellow stars indicate sampling sites, with one site being located on the northeast corner of Agency Lake (NAL), and two sites located on the northwest and southwest portions of Upper Klamath Lake (EP and HBR respectively). Agricultural land is denoted by gridded land and is located in northwest of Agency Lake and South of Upper Klamath Lake.

Current Results



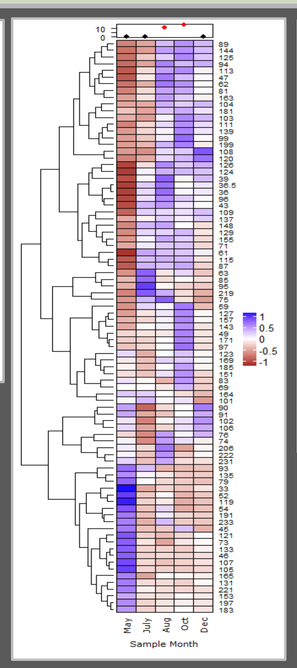


Figure 2. (Top) Cluster dendrogram of VOC profiles from samples collected at three different sites in 2018. Boxes are highlighting by date of collection. Figure 3. (Right) Correlation plot with dendrogram representing all significant VOC masses by month. VOC masses listed on y axis of chart with each row representing each mass +1 (m/z), with sampling date (x axis). Above the chart, average microcystin levels of the samples sites are represented in parts per billion (ppb), with those surpassing recreational limits highlighted in red. Color intensity is the strength of the correlation between mass number and month of year. The data shows strong correlations exist between VOC profiles and toxin trajectory.

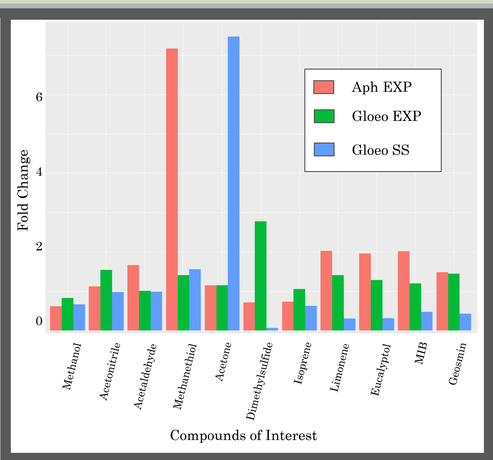


Figure 4. (Above) VOC production is strongly influenced by phase of growth. Enrichments dominated by either of 2 cyanobacterial species, *Aphanizomenon* (Aph) or *Gloeotrichia* (Gloeo) at either exponential (EXP) or stationary (SS) phases of growth (as determined by chlorophyll concentations). The values shown are the fold increase of a VOC of interest compared to the media blank. The results above will be adjusted for carbon content in the near future, but this will only exaggerate the differences seen between the exponential and stationary stages for Limonene, Eucalyptol, MIB, and Geosmin.

Discussion and Future Work

VOCs profiles from the 2018 season cluster based on season rather than site (Figure 2), and specific VOCs show seasonality with a steady decreasing/increasing in correlation with each successive sampling (Figure 3). Enriched unialgal cultures show that specific VOCs accumulate to a higher extent in exponential compared to stationary phases of growth. This result supports our idea that bloom phase is an important aspect to our analysis. We are using machine learning, specifically a convolutional neural network, to predict bloom dynamics. To circumvent our problem of a small training set, the enrichments are being supplemented with the environmental data and a pretrained model from a similar set of marine species and VOC profiles from the North Atlantic Aerosols and Marine Ecosystems Study (NAAMES) is being used as the model.

Further research and data processing is underway including 16S sequencing, processing of several environmental parameters, and incorporation of continuous monitoring data obtained from deployed buoys to account for bloom phase. Data collection is nearly competed for 2019 and focused on transitional bloom periods, as well as collection of isolates for enriched profile comparisons. A long term goal is to use metagenomic data to elucidate metabolic pathways that may reveal possible interactions with community members and their environment.

References

1. Xu, Q., Yang, L., Yang, W. *et al.* (2017) Volatile organic compounds released from Microcystis flos-aquae under nitrogen sources and their toxic effects on *Chlorella vulgaris*. *Ecotoxicol Environ Saf*, **135**, 191-200.
2. Aneja, V.P., Schlesinger, W.H., and Erisman, J.W. (2009) Effects of agriculture upon the air quality and climate: research, policy, and regulations. *Enviorn Sci and Technol*, **43**, 4234-4240.
3. McGrath, J.M., Betzelberger, A.M., Wang, S.H. *et al.* (2015) An analysis of ozone damage to historical maize and soybean yields in the United States. *PNAS*, **112**, 14390-14395.