Harmful Algal Bloom (HAB) Detection and Prediction System for the San José Lagoon

Project Overview

This project aims to develop an advanced early warning system for Harmful Algal Blooms (HABs) in the San José Lagoon, San Juan, Puerto Rico. The lagoon frequently experiences toxic cyanobacterial blooms, which deplete oxygen levels, harm aquatic life, and negatively impact local recreation and the economy. To address this, we will leverage recent advancements in remote sensing, in-situ monitoring, and Machine Learning (ML). By integrating multi-spectral satellite imagery with in-situ environmental data, we propose to create a state-of-the-art, data-driven predictive system for HAB events. This system will serve as the foundation for a decision support tool designed to detect and predict the onset, progression, and decline of HABs, providing crucial information for lagoon managers and communities.

Progress Summary

To build an accurate HAB detection and prediction model, two analyses were undertaken to understand:

- the temporal variability of environmental variables potentially influencing HAB events;
- the effectiveness of using remote sensing to detect chlorophyll-a concentrations in the San José Lagoon.

HABs in freshwater systems are primarily driven by excess anthropogenic nutrients, particularly phosphorus. Additional factors influencing HAB proliferation include lake depth, water column stability, water temperature, reduced wind speed, and landscape alterations such as urbanization and agricultural practices. Here, we consider the selection of predictor variables derived from environmental observations that are readily accessible and in close proximity to the San José Lagoon system. These variables include average wind speed, air temperature, and daily precipitation values from the San Juan, Luis Muñoz Marín International Airport, located approximately 2 km from the San José Lagoon. These data (Fig. 1) were obtained through NOAA's National Centers for Environmental Information (NCEI) application programming interface (API).

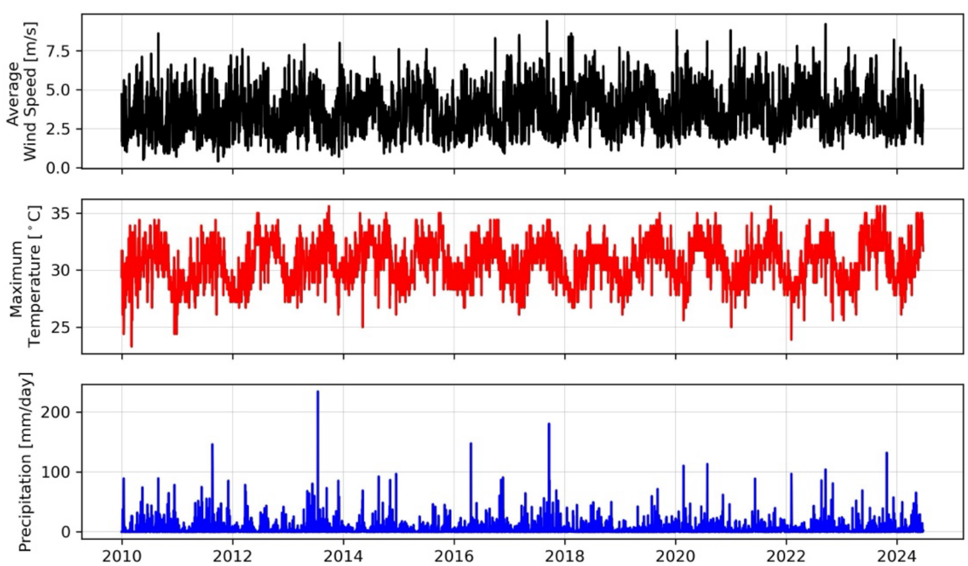

Fig. 1. Daily observations of (top panel) wind speed (mid panel) maximum temperature, and (bottom panel) precipitation at the San Juan Luis Muñoz Marín Airport, located about 2 km from the San Jose Lagoon.

Fig. 2 presents a climatology of these meteorological variables, based on approximately 14 years (2010-2024) of daily-averaged measurements. A distinct seasonal pattern is evident, characterized by elevated temperatures and precipitation during the summer and fall, with reduced wind speeds in the fall. These conditions are most conducive to HABs events, as indicated by elevated chlorophyll-a concentrations—a useful and commonly measured metric of water quality which is directly linked to cyanobacteria-dominated HABs. Ongoing work is focused on assessing whether higher chlorophyll-a concentrations are observed in response to these conditions. This will enable us to determine if the selected predictor variables generally correlate with chlorophyll-a values.

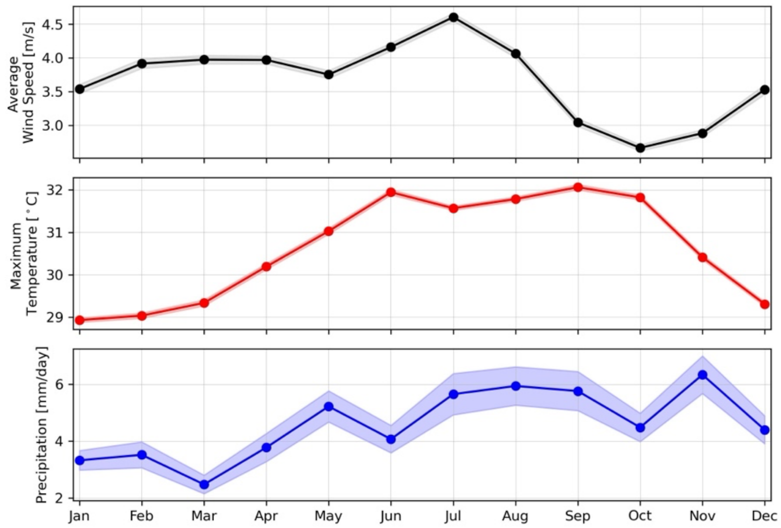

Fig. 2. Monthly-averaged meteorological observations based on approximately 14-years of data from the nearby San Juan Luis Muñoz Marín Airport weather station. The shaded regions indicate the standard error of the sample mean for each variable.

Additionally, we have conducted a preliminary evaluation of the chlorophyll-a detection skill of Sentinel-2 and Planet's PlanetScope satellites. Level 2 Sentinel-2 data were retrieved from the Sentinel data distribution system (https://scihub.copernicus.eu./). PlanetScope images were processed using ACOLITE (https://github.com/acolite/acolite/), where they underwent atmospheric correction and extraction of spectral remote sensing reflectances (Rrs). Chlorophyll concentrations are computed from the Rrs using the OCx algorithm, following the methodology described by Hu et al. (2023). This algorithm returns the near-surface concentration of the photosynthetic pigment chlorophyll-a in μg/L, calculated using empirical relationships derived from in-situ oceanic measurements. Only cloud-free images were considered in the analysis.

Fig. 3 compares remotely sensed chlorophyll-a (μg/L, left y-axis) with in-situ data (RFU, right y-axis) for the period during which in-situ data were collected (September 2023 - May 2024). Despite some negative values in the remote PlanetScope data, a positive correlation is observed with the in-situ measurements. Fig. 4 presents a scatter plot of monthly-averaged PlanetScope values versus in-situ data. A least-squares fit yields a coefficient of determination of r^2=0.65. As for the Sentinel-2 data, cloud cover and sun glint severely limited the availability of viable imagery. We are currently working on expanding the existing Sentinel-2 dataset to evaluate its accuracy.

These preliminary results indicate that the OCx algorithm is currently underestimating chlorophyll-a values. This is not entirely surprising, given that the OCx algorithm was calibrated for oligotrophic oceanic conditions. In most coastal and inland waters, where sediments or dissolved yellow substances significantly impact optical properties (Morel, 1988), the OCx algorithm may not produce accurate estimates. Therefore, some tuning of the algorithm with local in-situ measurements is likely necessary to correct the observed biases in chlorophyll-a concentrations.

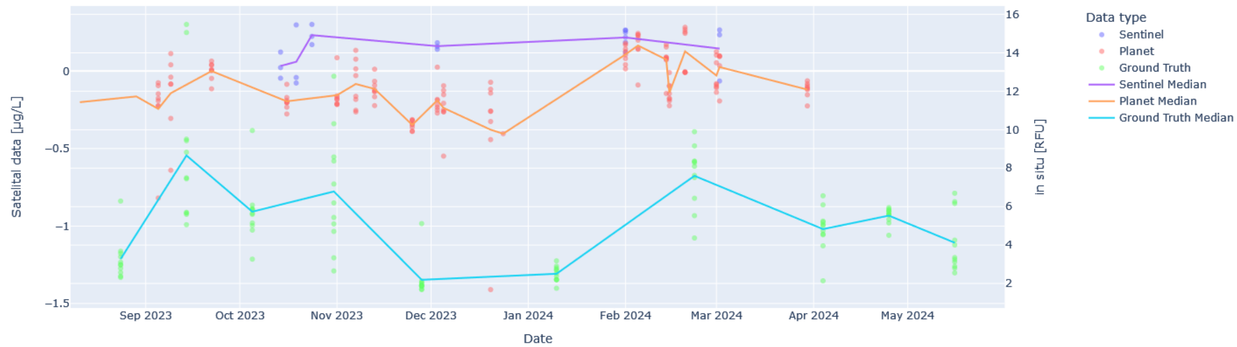

Fig. 3. Chlorophyll timeseries data derived from Sentinel-2 (purple dots), PlanetScope (red dots), and in-situ observations (green dots). Line plots illustrate median values for each sensor. Note that the remote data from Sentinel-2 and PlanetScope correspond to the left y-axis, while the in-situ field observations correspond to the right y-axis.

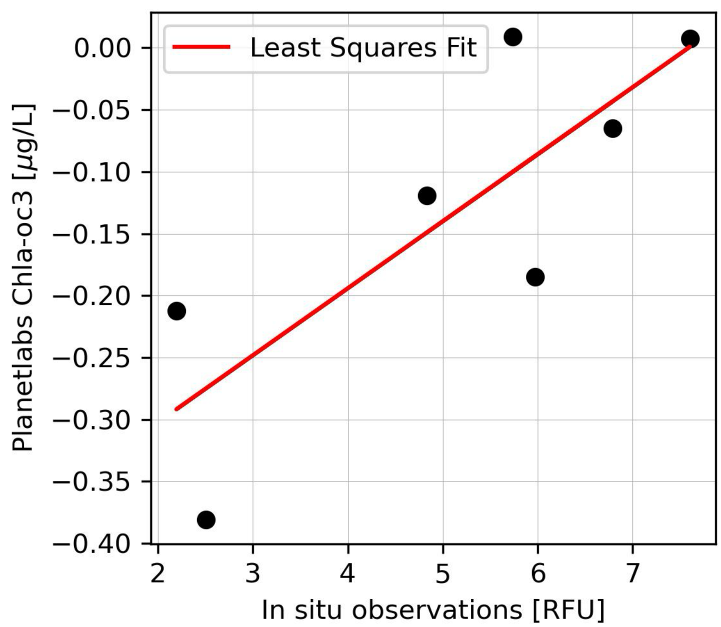

Fig. 4. Scatter plot of PlanetScope-based chlorphyll data versus in-situ field measurements. The red line shows a linear least-squares-fit.

Additional data sources and continuous data collection efforts will provide a more comprehensive understanding of the factors influencing chlorophyll levels and HAB occurrences in the San José Lagoon. It is thus crucial to obtain more data to strengthen the reliability of the HAB detection and prediction system. Future work will involve incorporating additional data from PlanetScope, Sentinel-2, and Sentinel-3 into our skill evaluation analysis. A critical aspect of this work will be testing a variety of chlorophyll algorithms for the remote sensors. Expanding the historical record will help us identify relationships with the environmental data, thereby enabling the rapid development and iteration of effective predictor variables. Finally, deploying an in-situ sensor capable of collecting continuous chlorophyll observations will allow for more robust validation of satellite data.

References

Chuanmin Hu, Werdell, J., OReilly, J., Feng, L., Zhongping Lee, Franz, B., Bailey, S., Proctor, C., 2023. Chlorophyll a, V1.0. https://doi.org/10.5067/JCQB8QALDOYD

Morel, A., 1988. Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters). J. Geophys. Res. 93, 10749–10768. https://doi.org/10.1029/JC093iC09p10749