



TÍTULO DO TRABALHO

DECODING CHLOROPHYLL-A VARIABILITY USING INTERPRETABLE MACHINE LEARNING: EVIDENCE FROM A MULTI-ARM TROPICAL RESERVOIR SYSTEM

Mathias Ahii Chia¹, Suleiman Dauda³, Rafaella de Souza Menezes², Zubeidat Precious Ahmed¹, Elizabeth Abdulmalik¹, Marcelo Luiz Martins Pompêo¹, Viviane Moschini-Carlos²

¹*Universidade de São Paulo, Departamento de Ecologia, São Paulo, São Paulo, Brasil*

²*Universidade Estadual Paulista, Instituto de Ciência e Tecnologia, Sorocaba, São Paulo, Brasil*

³*Ahmadu Bello University, Department of Botany, Zaria, Nigeria*

E-mail: machia@usp.br

Resumo: Excessive phytoplankton biomass remains a central ecological and public health concern in tropical reservoirs, yet predicting chlorophyll-a dynamics under variable nutrient and hydrological conditions remains challenging. This study applied a set of machine-learning models to a long-term monitoring dataset from the Barra Bonita reservoir system (São Paulo, Brazil), including its two main tributary arms (Rio Tietê and Rio Piracicaba). After data cleaning and transformation, models were trained using physicochemical variables and temporal features to predict chlorophyll-a concentrations. The HistGradientBoosting and Random Forest regressors produced the best performance across datasets, with log-scale R^2 values between 0.70 and 0.78, and original-scale R^2 values up to 0.50. SHAP interpretability revealed ammoniacal nitrogen, turbidity, temperature, pH, conductivity, and Kjeldahl nitrogen as the dominant predictors of phytoplankton biomass. Time-series patterns indicated strong seasonal variability and distinct bloom dynamics between arms. Correlation and feature-importance analyses confirmed nitrogen species as the most influential drivers in this highly eutrophic system. Overall, the findings demonstrate the value of explainable machine-learning models for understanding bloom dynamics in tropical reservoirs and provide a predictive foundation for proactive water-quality management.

Palavras-chave: *Unsupervised Learning; Bloom formation; Cyanobacteria; Reservoirs*