American Society of Limnology and Oceanography Meeting 2022

Scientific sessions: Advancing Near-term, Iterative Ecological Forecasting in Aquatic Ecosystems / **Ecological models as tools for integrating aquatic sciences**

Title: Prediction of algal blooms in a mesotrophic lake via a two-step hybrid process-based and machine learning model architecture

Abstract: With the expanding availability of meteorological observations, and lake monitoring data, state-of-the-art data-driven machine learning (ML) models have the potential to represent the complex and nonlinear algal bloom dynamics that are not completely captured in process-based (PB) models. We applied two ML models, Gradient Boost Regressor (GBR) and Long Short-term Memory (LSTM) network, to predict algal blooms using a 17-year dataset from mesotrophic Lake Erken, in Sweden. Three prediction workflows were designed and tested using both ML models:

1. Direct use of all observed data
2. Estimating lake nutrients that have limited observations first, and then predicting chlorophyll concentrations (*Chl*) using the observed and pre-generated environmental factors
3. A hybrid PB and ML model architecture that added hydrodynamic features derived from a PB model into the above two step ML model.

The performances of the final hybrid workflow was superior to PB model in predicting nutrients and *Chl* concentrations, and in predicting the timing of algal blooms, even under the abnormal environmental conditions. This two-step hybrid model architecture has been extended to other lakes where nutrients and *Chl* observations have been collected at varying intervals, and promising predictions were achieved in most of lakes. We expect this hybrid

model architecture to not only contribute to the mechanistic explanations of algal bloom in the natural lake systems, but enable near-term and even long-term predictions.