ABSTRACT
A DATA-EFFICIENT APPROACH FOR MONITORING AND MODELING STREAM NUTRIENTS

Wasif Bin Mamoon

Marquette University, 2025

Streams and rivers with excessive amounts of nitrogen (N) and phosphorus (P) can compromise aquatic ecosystems, human health, and overall water resources. Precise, high-resolution measurements of nutrient concentrations are essential for pollution management programs; however, conventional monitoring methods, either grab sampling or in-situ sensors, frequently face limitations due to cost and sample restrictions. This dissertation aimed to advance the existing state-of-the-science nutrient monitoring and modeling approaches to reduce sampling requirements and model complexity.

First, this study estimated the dimensionality and predictability of nutrient signals (NOx and TP signals), that describe the nutrient dynamics. The findings showed the possibility of accurate short-term forecasting (seasonal forecasting from 1 to 5.5 months) for nutrient signals. It was also found that the majority of the signals displayed low to moderate dimensionality (5-10) and had some hidden order within their dynamics. This enabled the option of using reduced-order models to estimate concentration data.

Second, this study implemented a Data-driven Sparse Sensing (DSS) framework to estimate high-frequency nutrient data (daily NOx and TP concentrations) from minimal measurements. The simplest DSS models (training data consisting of only flow and nutrients) were able to estimate accurate daily concentrations and annual loads using only 20-80 samples per year (errors <±2% for NOx, <±9% for TP). Instead of relying on fixed-frequency sampling or “storm chasing” approach, DSS models identified the optimal sampling times that contained the maximum information across the monitoring period.

Finally, the DSS modeling framework was expanded to a broader dataset and tested under real-world monitoring scenarios. Several scenario-based analyses were performed to answer the major considerations during designing nutrient monitoring programs, such as the duration of training data, the number of gauges for training data, and the selection method for training data. Results showed that accurate reconstructions (NSE > 0.5; load errors <±4%) are possible with only 2–3 years of data from 15 to 25 sites, especially when training data includes both flow and concentration.

Overall, the initial results backed the use of reduced-order modeling approaches (i.e., DSS) that take advantage of the fact that nutrient signals are low-dimensional and sparse. Using DSS, daily nutrient concentration time-series can be generated with 80-90% reduction in the number of samples required, leading to a significant reduction in monitoring effort.