Statistical modeling of water quality data on river networks

Presenter: Seoncheol Park

Department of Mathematics, Hanyang University

In this presentation, we introduce new methods for analyzing water quality data observed over river networks. These datasets contain spatio-temporal observations across stream systems, where observation times are irregular and vary across locations. Such complexities pose significant challenges for statistical modeling. We propose two new approaches. First, we develop a probabilistic forecasting method based on expectile smoothing tailored to river network data. Our approach extends flexible smoothing techniques from the Euclidean domain to the unique structure of river networks. Expectile curves of annual water quality measurements are modeled as functions of time, and a functional data forecasting method is applied to predict spatio-temporal dynamics in the network. Second, we introduce an adaptive boosting algorithm designed for data on tree-shaped linear networks. In this context, we address a classification problem and present a framework for constructing decision trees on linear networks. As a real-world application, we apply our method to water quality data on river networks. The proposed approach demonstrates improved accuracy over traditional Euclidean-based methods, particularly when observations are spatially clustered along the network. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00339064).