A bayesian hierarchical model to predict spatio-temporal variability in river water quality at 102 catchments

Danlu Guo1, Anna Lintern2, Angus Webb1, Dongryeol Ryu1, Ulrike Bende-Michl3, Shuci Liu1, and Andrew Western1

1Department of Infrastructure Engineering, The University of Melbourne, Parkville, VIC Australia (danlu.guo@unimelb.edu.au)
2Department of Civil Engineering, Monash University, Clayton, VIC Australia
3Bureau of Meteorology, Parkes, ACT Australia

Our current capacity to model stream water quality is limited particularly at large spatial scales across multiple catchments. To address this, we developed a Bayesian hierarchical statistical model to simulate the spatio-temporal variability in stream water quality across the state of Victoria, Australia. The model was developed using monthly water quality monitoring data over 21 years, across 102 catchments, which span over 130,000 km². The modelling focused on six key water quality constituents: total suspended solids (TSS), total phosphorus (TP), filterable reactive phosphorus (FRP), total Kjeldahl nitrogen (TKN), nitrate-nitrite (NO₃₋), and electrical conductivity (EC). The model structure was informed by knowledge of the key factors driving water quality variation, which had been identified in two preceding studies using the same dataset. Apart from FRP, which is hardly explainable (19.9%), the model explains 38.2% (NO₃₋) to 88.6% (EC) of total spatio-temporal variability in water quality. Across constituents, the model generally captures over half of the observed spatial variability; temporal variability remains largely unexplained across all catchments, while long-term trends are well captured. The model is best used to predict proportional changes in water quality in a Box-Cox transformed scale, but can have substantial bias if used to predict absolute values for high concentrations. This model can assist catchment management by (1) identifying hot-spots and hot moments for waterway pollution; (2) predicting effects of catchment changes on water quality e.g. urbanization or forestation; and (3) identifying and explaining major water quality trends and changes. Further model improvements should focus on: (1) alternative statistical model structures to improve fitting for truncated data, for constituents where a large amount of data below the detection-limit; and (2) better representation of non-conservative constituents (e.g. FRP) by accounting for important biogeochemical processes.