Bayesian Variable Selection for Censored Spatial Responses with Application to PFAS Concentrations in California

Suman Majumder
Indian Statistical Institute, Kolkata Centre

Abstract

Per- and polyfluoroalkyl substances (PFAS) are persistent environmental contaminants of major public health concern due to their resistance to degradation, widespread occurrence, and potential adverse health effects. Statistical analysis of PFAS concentrations in groundwater is complicated by left-censoring from detection limits, strong spatial dependence, and high-dimensional covariates. While PFAS levels are believed to be influenced by diverse sociodemographic, industrial, and environmental factors, their relative contributions remain unclear, motivating rigorous statistical approaches that can isolate key predictors from a large candidate set. We develop a Bayesian hierarchical framework that embeds censoring within a spatial process model via approximate Gaussian processes and employs a global–local shrinkage prior for effective high-dimensional variable selection. To further refine inference, we compare three post-selection strategies, credible interval rules, shrinkage weight thresholding, and clustering-based inclusion, based on their predictive accuracy, robustness to censoring, and stability of variable inclusion. Applying this framework to PFOS concentrations in California groundwater, we identify a parsimonious yet scientifically meaningful set of drivers. Significant covariates include demographic factors (percentage of American Indian population, local gender composition), industrial sources (metal coating, textile and leather, oil and gas, electronics, chemical manufacturing and cement manufacturing facilities), proximity to airports, traffic density, and environmental features such as herbaceous cover, elevation and ozone concentration. Our findings show that the proposed Bayesian variable selection framework provides interpretable and stable inference in high-dimensional censored spatial settings, while simultaneously offering actionable insights into the factors driving PFAS contamination in groundwater.