Abstract

Nonstationary spatial modelling is exciting and potentially rewarding, but suffers from several problems : its computational cost, the complexity and lack of interpretability of multi-layered hierarchical models, and the difficulty of model selection. We tackle those problems by introducing a nonstationary Nearest Neighbor Gaussian Process (NNGP) model.

NNGPs are a good starting point to address the problem of the computational cost because of their accuracy and affordability. We study the behavior of NNGPs that use a nonstationary covariance function, deriving some algebraic properties and exploring the impact of ordering on the effective covariance induced by NNGPs.

To ease results analysis and model selection, we introduce a readable hierarchical model architecture. In particular, we make parameters interpretation and model selection easier by integrating stationary range, nonstationary range with circular parameters, and nonstationary range with elliptic parameters in a consistent framework.

Given the NNGP approximation and the model architecture, we propose two *ad hoc* MCMC algorithms based on Metropolis Adjusted Langevin Algorithm and Chromatic Sampling, both being improved using interweaving of parametrizations.

We carry out experiments on synthetic data sets to find empirical practical rules concerning on MCMC algorithm choice, hyperparameter tuning, and model selection. Finally, we use those guidelines to analyze a data set of lead contamination in the United States of America.

Keywords: Nearest Neighbor Gaussian Process, Nonstationary Spatial Modelling, Bayesian Statistics, Hierarchical Modelling, Environmental Statistics