



## On Spatial Process Models for High-Dimensional Geostatistical Datasets

Sudipto Banerjee\*

University of California, Los Angeles, USA - sudipto@ucla.edu

Abhirup Datta

University of Minnesota, Twin Cities, USA - datta013@umn.edu

Andrew O. Finley

Michigan State University, East Lansing, USA - finleya@msu.edu

Alan E. Gelfand

Duke University, Durham, USA - alan@stat.duke.edu

With the growing capabilities of Geographical Information Systems (GIS) and user-friendly software, statisticians today routinely encounter geographically referenced datasets containing observations from a large number of spatial locations. Over the last decade, hierarchical spatial process models have become widely deployed statistical tools for researchers to better understanding the complex nature of spatial variability. However, fitting hierarchical spatial models often involves expensive matrix decompositions whose computational complexity increases in cubic order with the number of spatial locations. This renders such models infeasible for large spatial data sets. In this talk, I will present two approaches for constructing well-defined spatial stochastic processes that accrue substantial computational savings. Both these processes can be used as "priors" for spatial random fields. The first approach constructs a low-rank process operating on a lower-dimensional subspace. The second approach constructs a Nearest-Neighbor Gaussian Process (NNGP) that can be exploited as a dimension-reducing prior embedded within a rich and flexible hierarchical modeling framework to deliver exact Bayesian inference. Both these approaches lead to Markov chain Monte Carlo algorithms with floating point operations (flops) that are linear in the number of spatial locations (per iteration). We compare these methods and demonstrate its use in inferring on the spatial distribution of forest biomass from the US Forest Inventory database spanning the continental US.

**Keywords:** Bayesian inference; Gaussian Processes; Low-Rank Models; Nearest-Neighbor Models.