Geography 890 - Hierarchical Bayesian Models for Environmental Spatial Data Analysis
Spring 2011

Instructor: Andrew Finley  Class: Tu Th 1:50-2:40 in 233 Geography Bldg
Office: 126 Natural Resources Building  Office Hours: By appointment
E-mail: finleya@msu.edu
Office Phone: 517-432-7219

Description
Given recent advances in spatially-enabled monitor and sensor networks along with geospatial information systems, researchers’ ability to collect, manage, and use spatial data is rapidly evolving. It is clear that the scientific community is moving into a new period where data-rich environments provide extraordinary opportunities to understand the complexity of spatially and/or temporally indexed datasets. In these settings, the focus of inference is often on specific model parameters and/or subsequent prediction at a new location or time. In these modeling exercises, rarely is it safe, or even desirable, to assume that model residuals are independent and identically distributed. The propensity to violate these assumptions is especially great in environmental datasets because the data often exhibit temporal, spatial, or hierarchical structure, or all three.

This course explores recent advancements in hierarchical random effects models using Markov chain Monte Carlo (MCMC) methods. The focus is on linear and generalized linear modeling frameworks that accommodate spatial and temporal associations. Lecture and labs offer an applied perspective on model specification, identifiability of parameters, and computational considerations for Bayesian inference from posterior distributions. The lecture series begins with a basic introduction to Bayesian hierarchical linear models and proceeds to address several common challenges in environmental data, including missing data and when the number of observations is too large to efficiently fit the desired hierarchical random effects models. The labs blend modeling, computing, and data analysis including a hands-on introduction to R statistical environment. Special attention is given to exploration and visualization of spatial-temporal data and the practical and accessible implementation of spatial-temporal models.

Grades and tentative course outline
Course grades are based solely on participation (20%) and final project (80%). To earn full credit for participation students are required to: 1) serve as a discussant for an assigned scientific paper, and; 2) present preliminary results from their final project. Guidelines for the final project will be defined later in the semester.

---

Tentative course outline:

1. **Introduction**
   
   1.1 Motivating examples from environmental sciences
   1.2 Introduction to spatial data and models
   1.3 R programming and statistical pre-requisites
   1.4 Some essentials
      
      1.4.1 Basics of matrix theory
      1.4.2 Building a high performance R computing environment

2. **Bayesian Hierarchical models**
   
   2.1 Introduction to hierarchical modeling and Bayes’ Theorem
   2.2 Bayesian linear regression
      
      2.2.1 Analysis with conjugate NIG priors
      2.2.2 Analysis with flat priors – classical analysis
      2.2.3 Prediction
      2.2.4 Assessing model fit and model comparisons
   2.3 Bayesian computation
      
      2.3.1 The Gibbs sampler
      2.3.2 The Metropolis-Hastings algorithm
      2.3.3 Adaptive Markov Chain Monte Carlo
      2.3.4 Convergence diagnosis

3. **Spatial data compilation in R**
   
   3.1 Import and Export
   3.2 Geographic distance and projection
      
      3.2.1 Distance computations
      3.2.2 Spatial data projection
   3.3 Exploratory data analysis
      
      3.3.1 Non-spatial exploratory analysis
      3.3.2 Spatial data exploratory analysis
      3.3.3 Representing continuous spatial data

4. **Model-based geostatistics**
   
   4.1 Elements of point-referenced modeling
      
      4.1.1 Stationarity
      4.1.2 Variograms
      4.1.3 Isotropy
      4.1.4 Variogram model fitting

5. **Hierarchical spatial process models**
   
   5.1 Ingredients for spatial models
5.2 Formal modeling theory for spatial processes
   5.2.1 Covariance functions
   5.2.3 Simulating spatial data
5.3 Gaussian spatial models
   5.3.1 Exact Bayesian inference
   5.3.2 Bayesian inference using MCMC
5.4 Non-Gaussian spatial models
5.5 Gaussian predictive process models
   5.5.1 Low-rank kriging and the predictive process
   5.5.2 Bias-adjusted predictive process models
   5.5.3 Knot selection

6. Multivariate spatial models
   6.1 Multivariate spatial processes
   6.2 Multivariate Gaussian spatial models
      6.2.1 Missing observations
   6.3 Multivariate non-Gaussian spatial models
   6.4 Multivariate predictive process

7. Advanced topics
   7.1 Spatially varying coefficient models
      7.1.1 Univariate models
      7.1.2 Multivariate models
   7.2 Spatiotemporal models
      7.2.1 Covariance functions
      7.2.2 Nonseparable spatiotemporal model
   7.3 Zero-inflated spatial models
   7.4 Spatial survival and frailty models
   7.5 Spatial point-process models