Hierarchical Modeling and Analysis of Spatial-Temporal Data:
Emphasis in Forestry, Ecology, and Environmental Sciences

1 Instructors

Andrew O. Finley, Department of Forestry and Geography, Michigan State University;
Sudipto Banerjee, Division of Biostatistics, School of Public Health, University of Minnesota

2 Description

Recent advances in Geographical Information Systems (GIS) and Global Positioning Systems (GPS) enable accurate geocoding of locations where scientific data are collected. This has encouraged collection of spatial-temporal datasets in many fields and has generated considerable interest in statistical modeling for time and location-referenced data. This accumulation of data and need for analysis is especially common in the broad fields of forestry and ecology. In these fields, spatially and temporally indexed data, typically consisting of one or more response variables and associated covariates, is used to estimate natural resource inventory, presence/absence, counts, and change. 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[1].

This course details recent advancements in hierarchical random effects models using Markov chain Monte Carlo (MCMC) methods. The course focus is on linear and generalized linear modeling frameworks that accommodate spatial and temporal associations. Careful attention is paid to the theoretical foundations of model specification, identifiability of parameters, and computational considerations for Bayesian inference from posterior distributions. The lecture will start with a basic introduction to Bayesian hierarchical linear models and proceed 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. Diverse settings for spatial and spatial-temporal models are considered, mostly motivated by a range of studies that employ forestry and ecological monitoring datasets. The course will blend modeling, computing, and data analysis including a hands-on introduction to the `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. In particular, participants will learn how to fit a diverse class of spatial-temporal models using the `spBayes` `R` package. Participants are encouraged to bring their own laptops with `R` and `spBayes` installed.

[1] "The first law of ecology is that everything is related to everything else." *The Closing Circle: Nature, Man, and Technology.* New York : Knopf, 1971.

The following is a broad list of topics covered. Each topic will include theory, examples and data analysis along with live interactive computing demonstrations.

Basics of the `R` statistical computing environment;

Introduction to hierarchical linear models;

Ingredients for modeling point-referenced spatial data;

Linear models for univariate point-referenced spatial data;

Bayesian "Kriging": Model-based interpolation and prediction;

Generalized linear models with spatial and temporal random effects;

Model choice and model selection;

Hierarchical models for spatial-temporal data;

Ingredients for modeling multivariate point-referenced data;

Models for multivariate point-referenced spatial data;

"Big-N problem": Predictive process models for large datasets in space and/or time;

Space varying coefficient models for spatial non-stationarity;

Ingredients for "areally-referenced" spatial and spatial-temporal data;

Simultaneous and Conditional Autoregressive models (SAR and CAR) for univariate areally-referenced data;

Multivariate extensions: A general class of multivariate CAR models;

Case studies from forestry, ecology and environmental sciences.

3 Pre-requisites

Some familiarity with classical linear models and multiple regression will be useful. A laptop with a current version of `R` and `spBayes` installed, while not required, will definitely be useful. Please visit the short course website `http://blue.for.msu.edu/JSM_09/SC` a few weeks prior to the course for software updates and `R` scripts and data that will be used for illustration.