Spatial statistics short course; August, 2011
Margaret Short, Ph.D.; University of Alaska Fairbanks

The usual assumption in introductory statistics courses is that our data are independent. With many spatial data sets, we typically assume dependence: observations collected at spatially proximate locations (or regions) tend to be more similar than those at distant locations. Our models incorporate the spatial dependency, at various levels of our model hierarchy.

This short course focuses on `geostatistical data', in which data are measured at known spatial locations. We'll discuss a common initial approach in which the spatial covariance is modeled explicitly via variograms, and we'll fit these models using both a frequentist and Bayesian approach using software packages that are available for the R, a free program and environment for statistical computing and graphics.

I work primarily with Bayesian models and will illustrate several of them, including a process convolution approach. We'll discuss extensions to the basic models, e.g. for bivariate data and for space-time data, as well as nonstationary models and models with nonstandard error structures. Other topics may be included, at students' request.

The programming language R mentioned above is freely available through the web site http://www.r-project.org/. Sample R code will be provided to the extent that it is available for the various models we'll discuss. Attendees should bring a laptop on which R has been installed, along with several specialty packages (including geoR and akima; a complete list will be provided later).

An extensive list of references for further study will be provided.

Outline:

Day 1 emphasizes concepts and models for which R software is available:

- Introduction; plotting spatial data (1 hr)

- The variogram; variogram estimation (1 hr)

- Frequentist kriging (1 hr)

- Bayesian kriging (2 hr)

- Process convolution approach to spatial models, the simplest model (1-2 hr)

Day 2 emphasizes student-supplied problems & data sets, as well as more conceptual material, for which off-the-shelf may or may not be available. Selected topics (as time permits; the order is flexible):

- Non-standard error structures, e.g. mixture distribution for errors

- Nonstationary models

- Space-time models

- Handling large spatial data sets

- Spatial data on a stream network