Some applications of Bayesian methods in hydrology and applied ecology

Adam Butler, Biomathematics Statistics Scotland, Edinburgh, UK

September 30, 2013

Ecological and hydrological systems are inherently complex, and often involve complex patterns of spatio-temporal dependence. Data collection is frequently also a complex process, with inhomogeneous (and, in ecology, often unknown) levels of recording effort across space and time. Modern Bayesian methods allow us to analyze ecological and hydrological data using statistical models that are sophisticated enough to capture many of these complications. In this talk we will outline four case studies in which we have used Bayesian approaches to analyze spatio-temporal ecological or environmental data, and will highlight some of the common methodological themes - and challenges - that arise from these case studies.

The first case study concerns source apportionment: hydrologists are often interested in studying the proportions of water that arise from different sources of run-off. The evidence for this is provided by regular measurements of a chemical tracer in the target water body. If tracer data are also available for the possible sources of run-off then it is possible to perform source apportionment using compositional analysis - we do this using a Bayesian random effects approach in a hierarchical framework. We also consider the case where there are no separate data available for the source components, and develop a model for source distributions via nonlinear regression on the tracer/flow relationship and nonparametric density estimation. We apply the model to a data set from two streams in central Scotland, comprised of weekly or fortnightly readings of alkalinity over seventeen years.

The second case study involves the analysis of riverine water quality time series from multiple sites. Data of this kind present some challenging methodological problems, including nonlinearity, non-normality, seasonality and missing data. We demonstrate that periodic multivariate Normal hidden Markov models provide appropriate tools for analyzing such data, and introduce a fully Bayesian inference procedure for this class of models, where the number of hidden states of the Markov process is unknown and reversible jump Markov chain Monte Carlo (RJMCMC) methods are developed. We apply the methodology to long-term dissolved inorganic nitrogen time series measured in three Scottish rivers.

The third case study involves applying Bayesian methods to a stochastic, spatio-temporal model for the spread of invasive plant species, and thereby generating predictions of future spread. Observational data consist of species distribution data (presence/absence data) from multiple time points, and the model describes the spread of the species in terms of covariates which capture landscape heterogeneity. The implementation of the model and inference algorithm are illustrated through application to British floristic atlas data for Heracleum mantegazzianum (giant hogweed), an invasive alien plant that has rapidly increased its range since 1970.

The final case study involves identifying the spatial distribution of seabird foraging areas. At least three different sources of data on seabird foraging are available: data from GPS tags fitted to individual birds, data from boat-based line transect surveys, and data on abundance, survival and productivity at individual colonies. We illustrate the use of a hidden Markov model to account for the underlying behavioural state of a bird using GPS tagging data, and the use of a hierarchical model to analyze temporal trends in colony-based counts, and conclude by discussing the potential for using Bayesian methods to draw inferences about the spatial distribution of foraging locations using data from multiple sources.