Nonlinear Latent Process Models for Integrating Spatio-Temporal Exposure Data from Multiple Sources

ABSTRACT: Spatio-temporal prediction of levels of an environmental exposure is an important problem in environmental epidemiology. When multiple sources of exposure information are available, a joint model that pools information across sources maximizes data coverage over both space and time, thereby reducing the prediction error. We consider a Bayesian hierarchical framework where a joint model consists of a set of submodels, one for each data source, and a model for the latent process that serves to relate the submodels to one another. However, if a submodel depends on the latent process nonlinearly, inference using standard MCMC techniques can be computationally prohibitive. To make such problems tractable, we 'linearize' the nonlinear components with respect to the latent process and induce sparsity in the covariance matrix of the latent process using compactly supported covariance functions. We propose an efficient MCMC scheme that takes advantage of these approximations. We then apply our methods to data on the spatio-temporal distribution of mobile source particles in the greater Boston area.