A FLEXIBLE SPATIO-TEMPORAL MODEL FOR AIR POLLUTION WITH SPATIO-TEMPORAL COVARIATES

Johan Lindström, *Lund University, Sweden*
Adam Szpiro, *University of Washington, United States*
Paul D Sampson, *University of Washington, United States*
Lianne Sheppard, *University of Washington, United States*
Assaf P. Oron, *University of Washington, United States*
Mark Richards, *University of Washington, United States*
Timothy V Larson, *University of Washington, United States*

Background and Aims: We developed a spatio-temporal model for predicting ambient air pollution that combines unbalanced monitoring data with deterministic air quality model output. This work is part of the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air) study, a prospective cohort study funded by the US EPA to investigate the relationship between chronic exposure to air pollution and cardiovascular disease.

Methods: The monitoring data includes three spatial "snapshots" at 174 locations; rotating monitors giving 2-3 measurements at 84 subject homes; and 25 time series of two-week averages at fixed and regulatory monitoring (AQS) sites. The data are unbalanced because at any given time measurements are available at only a subset of measurement sites. We also have geographic covariates and a spatio-temporal covariate based on the output from a source dispersion model for traffic-related air pollution (Caline3QHC; Wilton et al 2010).

We extended our unified hierarchical spatio-temporal model (Szpiro et al, 2010) to accommodate spatio-temporal covariates. We improved the estimation approach and developed several cross-validation strategies to make the fullest use of the available data. We assessed predicted NOx concentrations in Los Angeles.

Results: The model predicted well with cross-validated spatial R^2 of approximately 0.7 at subject homes. We successfully incorporated the source dispersion model output into the prediction model, but our implementation of Caline3QHC did not improve predictions in a model with road covariates. However, we found some evidence that Caline3QHC could replace the road covariates.

Conclusions: This spatio-temporal model for air pollution concentrations provides a flexible approach to combining irregularly sampled pollution monitoring data with geographic covariates and the output from deterministic air quality models. We successfully predicted long-term average NOx concentrations while taking full advantage of the unbalanced monitoring data available to the MESA Air study. This model has been implemented in an R-package, `SpatioTemporal`.

References:
Lindström J, Szpiro AA, Sampson PD, Sheppard L, Oron A, Richards M, Larson T (2011) A Flexible Spatio-Temporal Model for Air Pollution: Allowing for Spatio-Temporal Covariates (January 19, 2011). UW Biostatistics Working Paper Series. Working Paper 370. http://www.bepress.com/uwbiostat/paper370

Szpiro A, Sampson PD, Sheppard L, Lumley T, Adar SD, Kaufman J. Predicting intra-urban variation in air pollution concentrations with complex spatio-temporal dependencies. Environmetrics, 2010, 21: 606–631.

Wilton D, Szpiro AA, Gould T, and Larson T. Improving spatial concentration estimates for nitrogen oxides using a hybrid meteorological dispersion/land use regression model in Los Angeles, CA and Seattle, WA. Science of the Total Environment, 2010, 408 :1120-1130.