Spatiotemporal Estimation of Radon Exposure for Epidemiologic Risk Assessment, Regional Case Study

Kyle J. Sorensen, Richard L. Smith, Eric A. Whitsel, Jason M. Collins
Department of Statistics and Operations Research & Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina @ Chapel Hill

Introduction

Problem
• Radon levels are rising across North America, linked to trends in climate change
• Radon exposure is associated with lung cancer, stroke, and other cardiovascular events
• Current estimates of radon exposure are limited, classified into three levels at low spatial resolution

Data
SRRS - EPAs State Residential Radon Survey
• Series of household-level short-term surveys
• 63,291 homes, 42 US states and six US territories
• Conducted between 1986 and 1992

GRP - USGS and EPAs Geologic Radon Potential
• Constructed from geologic, atmospheric and residential survey data
• Three levels:
  • “high” (estimated radon level > 4 picocuries per liter, or pCi/L) – zone 1
  • “moderate/variable” (2–4 pCi/L) – zone 2
  • “low” (< 2 pCi/L) – zone 3

Solution
• Kriging and other spatial modeling techniques
• Spatial blocking cross-validation

Methodology

Data selection for example analysis
• Region encompassing Tennessee
• 4919 homes across 500 zip-codes
• Relatively high spatial variability in GRP

SRRS analysis
• Translated, log transformed radon exposure data
• Better match to the assumption of a normally distributed response in the kriging algorithm
• Construction of heatmaps according to kriging estimates of radon exposure

SRRS+GRP analysis
• Formulation of z^2
• Background radon level conditioned on observed radon level
• Estimation of a variance parameter σ
• Analogous translation and log transformation
• Construction of heatmaps according to kriging estimates of radon exposure using z^2

30-fold spatial blocking cross-validation

Analysis and Conclusions

Figure 2: Visualization of region selection for spatial blocking cross-validation

Future Work and Recommendations

SRRS+RI analysis
• Can use US EPA’s Radon Index RI data in place of GRP data (trichotomization of RI data)
• Three levels – < 15 levels
• Likely to capture additional variability in radon level due to
  • Aerial radioactivity
  • Geology type
  • Soil permeability
  • Architecture type

Bayesian hierarchical modeling
• Allows reversal of the conditioning in the SRRS+GRP analysis
• NIMBLE, an R package aiding in the construction of Bayesian hierarchical models, optimized for spatial data

Spatial blocking cross-validation
• SRRS analysis:
  • Mean absolute error of 1.8 pCi/L
• SRRS+GRP analysis:
  • To be determined
  • Model shows signs of improved performance, especially in prediction of extreme values
  • Will be used in validation of future models

Examples

Observations from Figures (2) and (3)
• Significant increase in range of estimated radon exposure (pCi/L)
• Consistent with inclusion of variance parameter in SRRS+GRP analysis
• Improved prediction of extreme SRRS values
• Increased range of RMSPE values
• Relative error decreases in SRRS+GRP analysis

Background

Radon
• Chemical element with symbol Rn, number 86
• Radioactive, colorless, odorless, and invisible gas
• Naturally occurring product of the decay of uranium

Exposure to radon
• Second leading cause of lung cancer
• Linked to strokes and other cardiovascular events
• Evidence of recent increases in North America

Radon and climate change - climate change may indirectly influence rises in radon exposure due to...
• Increased HVAC use
• Recycling of indoor air

Kriging
• Popular spatial modeling algorithm
• Model is a Gaussian process with...
  • Mean – function of covariates
  • Covariance – function of the spatial coordinates

Spatial dependence structure
• Nearby data is more similar than distant data
• Can cause artificially optimistic estimates of model performance

Spatial blocking cross-validation
• Folds from standard k-fold CV → geographically distinct regions
• Provides more realistic measure of model performance

References