Geostatistical Hierarchical Model for Temporally Integrated Radon Measurements

Brian J. Smith and Jacob J. Oleson

Department of Biostatistics, University of Iowa, Iowa City, Iowa, U.S.A.

The University of Iowa

UI College of Public Health

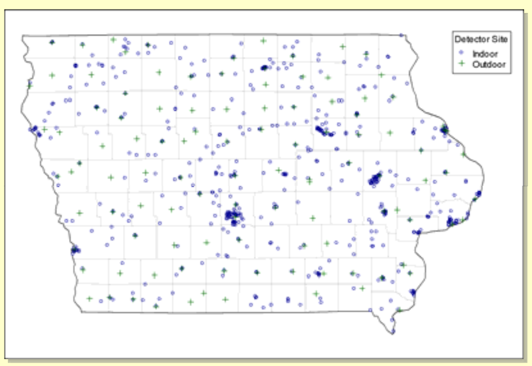
Background

Radon

- Radon is a naturally occurring radioactive gas that is potentially harmful to lung tissue.
- Ambient radon concentrations arise from uranium found in the soil, are typically affected by environmental factors, and exhibit spatial and temporal variation.
- In many epidemiologic studies, contemporary radon measurements are used as surrogates of past exposures; e.g. a measurement taken in 1995 is assumed to represent exposure over the past 15 years.

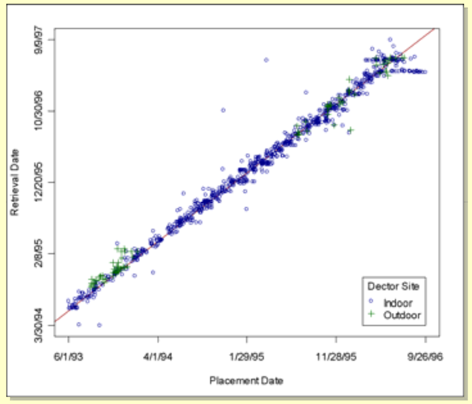
Iowa Radon Lung Cancer Study

- Epidemiologic case-control study in Iowa to estimate the effect of residential radon on lung cancer risk.
- Enrollment and data collection occurred over a four-year span beginning in 1993.
- 2,802 radon detectors were installed in 614 population-based, control subject homes.
- At least one measurement was taken on each floor of the home, for an average of 4.6 measurements per home.
- 129 radon measurements collected from an approximately uniform grid of 109 outdoor sites.



- Detectors were installed in 614 population-based, control subject homes during the first year of enrollment and in a subset of 215 homes during the second year.
- Year-long measurements – integrated averages over the installation periods – were taken to expunge seasonal trends.
- A total of 944 unique time periods are included in this analysis.

Figure: Scatter plot of placement and retrieval dates for radon detectors in the Iowa radon study. Points on the solid line represent installation periods of one year.



Methods

Analysis Goals

- Characterize the distribution of radon in Iowa.
- Allow for systematic differences due to home and outdoor environments.
- Account for important source of variability:
  - Detector measurement error,
  - Unmeasured covariates,
  - Unknown, but constant, amounts of uranium in the soil,
  - Local randomness over time.
- Prediction at any geographic site and time period.

Outdoor Radon Model

- Let $y_{OS,ik}(t)$ denote the k^{th} outdoor measurement from geographic site s_i and time point t.
- We specify the following model for this measurement:

$$\ln y_{OS,ik}(t) = \beta_{OS} + z(s_i,t) + \varepsilon_{OS}$$
$$\varepsilon_{OS} \sim N(0,\sigma^2_{OS})$$

  where β_{OS} is an overall mean parameter, ε_{OS} is an independent error term, and σ^2_{OS} is the measurement error variance.
- The $z(s_i,t)$ parameter accounts for spatio-temporal correlation among radon concentrations.

Home Radon Model

- For the k^{th} measurement from home i and time point t, we specify:

$$\ln y_{H,ik}(t) = \mathbf{x}^T_{H,ijk}\beta_H + \gamma_i + z(s_i,t) + \varepsilon_{H,ijk}$$
$$\gamma_i \sim N(0,\sigma^2_{BH})$$
$$\varepsilon_{H,ijk} \sim N(0,\sigma^2_{WH})$$

  where $x^T_{H,ijk}$ and β_H are vectors of covariates and corresponding mean parameters; γ_i an independent random effect for the home; $z(s_i,t)$ is the latent spatio-temporal parameter; and $\varepsilon_{H,ijk}$ is an independent error term.
- In the present analysis, only indicator variables for the floors on which the measurements were taken are included in the covariate vector $x_{H,ijk}$.
- The error variance σ^2_{WH} is a combination of variability due to systematic differences within the home and random detector measurement error.
- The between-home variance σ^2_{BH} includes unaccounted for differences in home environments.

Latent Spatio-Temporal Process

- The latent spatio-temporal parameters are assumed to follow a Gaussian distribution with zero mean and variance given by

$$Var\left(z(s,t), z(s',t')\right) = \sigma^2_{ST} c_S(s,s';\theta_S)\, c_T(t,t';\theta_T)$$

  where c_S and c_T are continuous parametric functions of the geographic and temporal distance between measurements, respectively.
- A two-parameter Matérn spatial correlation function is used.
- Temporal correlation is modeled with an exponential function.

Temporal Integration

- Although our geostatistical approach models correlation as a continuous function of <u>time points</u>, radon measurements are only observable for <u>time periods</u>.
- We treat each radon measurement as an integrated average over the time period P_j that the detector was installed in the home, so that

$$\ln y_{ijk} = \left|P_j\right|^{-1} \int_{P_j} \ln y_{ik}(t)\,dt$$

- Our approach is a special case of the geostatistical model for integrated areal data described by Gelfand et al. (2001).

Bayesian Methods

- A fully Bayesian approach is taken to obtain the joint posterior distribution of all model parameters and predicted radon concentrations.
- Vague priors are specified for the model parameters
- Closed-form solutions are used to integrate over time periods.
- Draws from the posterior are simulated with a Markov chain Monte Carlo computational algorithm.

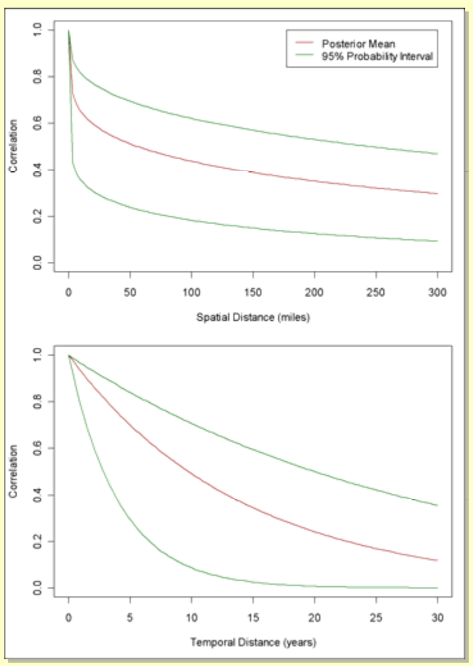
Results

State-Wide Posterior Summaries

- Outdoor geometric mean radon concentration is 0.74 pCi/L (95% HPD: 0.34-0.76).
- Mean first floor concentration is 2.41 pCi/L (1.17-4.43).
- Within-home variance is approximately 4 times the outdoor measurement error variance.
- Spatio-temporal variance is one-half times the between-home variance and 2.5 times the within-home variance.

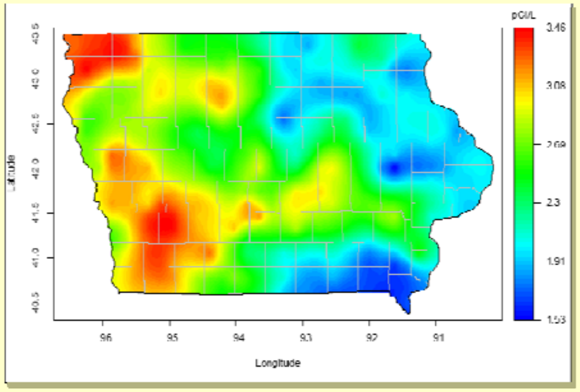
Spatio-Temporal Correlation

- Mean estimates indicate correlations of 0.65, 0.52,and 0.44 for measurements taken 10, 50, and 100 miles apart.
- Correlations of 0.70, 0.49, and 0.34 at distances of 5, 10, and 15 years are estimated for the latent temporal process.



Radon Prediction

- Posterior geometric mean of the predicted average first floor radon concentrations in Iowa for 1995.

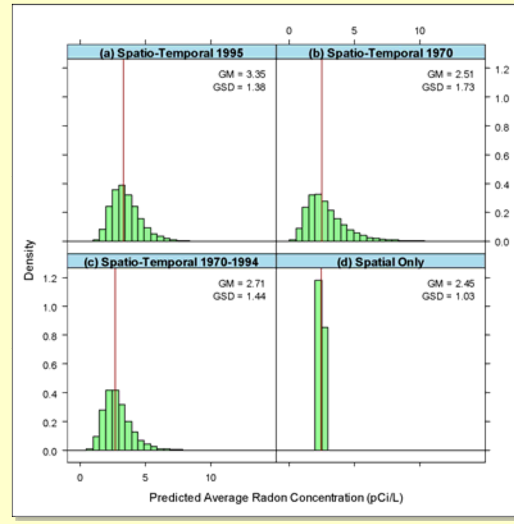


- Posterior geometric mean (geometric standard deviation) of the predicted average first floor radon concentrations (pCi/L) for select time periods and geographic sites in Iowa.

Year	SW*	Central*	NE*
1995	3.35 (1.38)	2.60 (1.37)	1.89 (1.37)
1994	3.66 (1.37)	2.83 (1.37)	1.98 (1.37)
1993	3.06 (1.39)	2.33 (1.39)	1.63 (1.40)
1992	2.96 (1.43)	2.30 (1.44)	1.66 (1.44)
1991	2.92 (1.47)	2.30 (1.47)	1.70 (1.48)
1990	2.87 (1.50)	2.31 (1.50)	1.75 (1.51)
1985	2.70 (1.61)	2.34 (1.60)	1.94 (1.61)
1980	2.61 (1.66)	2.37 (1.66)	1.94 (1.61)
1975	2.55 (1.71)	2.38 (1.70)	2.16 (1.71)
1970	2.51 (1.73)	2.39 (1.72)	2.23 (1.72)
1980-1994	2.82 (1.43)	2.37 (1.43)	1.88 (1.43)
1975-1994	2.76 (1.44)	2.37 (1.43)	1.94 (1.44)
1970-1994	2.71 (1.44)	2.37 (1.44)	1.99 (1.44)

*Longitude and latitude coordinates: SW = (-95.13, 41.06), Central = (-93.46, 42.01), and NE = (-91.80, 42.97).

- Posterior predictive distributions for radon concentrations at a SW in Iowa for select years. The results in panels (a)-(c) come from the proposed geostatistical model that accounts for both spatial and temporal correlation; those in panel (d) are from a simplified version of the model that only accounts for spatial correlation.



Select References

Field RW, Steck DJ, Smith BJ, Fisher EL, Neuberger JS, Platz CE, Robinson RA, Woolson RF, Lynch CF. Residential radon gas exposure and lung cancer: the Iowa radon lung cancer study. *American Journal of Epidemiology*. **151**:1091-1102, 2000.

Gelfand AE, Zhu L, Carlin BP. On the change of support problem for spatio-temporal data. *Biostatistics*. **2**:31-45, 2001.

Steck DJ, Field RW, Lynch CF. Exposure to atmospheric radon. *Environmental Health Perspectives*. **107**:123-127, 1999.

Conclusions

Our model allows for the prediction of radon, while accounting for spatio-temporal variation, detector measurement error, home-to-home variability, and uncertainties in estimating model parameters.

Because the model provides retrospective estimates that accurately reflect prediction errors, its use in epidemiologic radon studies could improve risk assessment.

Summary

- ✓ The proposed geostatistical model represents a new approach for characterizing residential radon.
- ✓ Accounts for important sources of uncertainty.
- ✓ Provides realistic prediction errors for concentrations at unmeasured geographic locations and time period.
- ✓ Allows for the prediction of past radon exposures based on a single contemporary measurement.
- ✓ Closed-form solutions for the temporal integration.