Bayesian modelling of inseparable space-time variation in disease risk

by Leonhard Knorr-Held

Laina Mercer

Department of Statistics
UW

April 11, 2013

Motivation

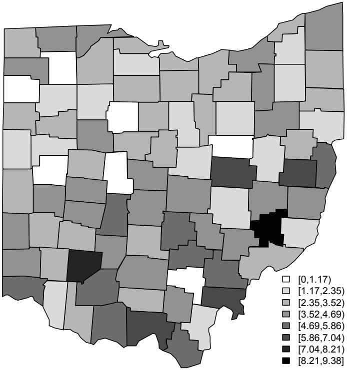
Area and time-specific disease rates

- Area and time-specific disease rates are of great interest for health care and policy purposes
- Facilitate effective allocation of resources and targeted interventions
- Sample size often too small at granular space-time scale for reliable estimates
- Bayesian approach to 'borrow strength' over space and time to improve reliability

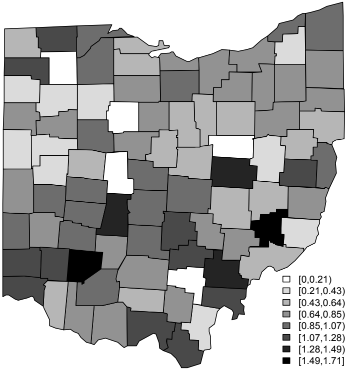
Motivation

Ohio Lung Cancer Example

Lung Cancer Mortality Rates 1972



Lung Cancer Standarized Mortality Ratio 1972



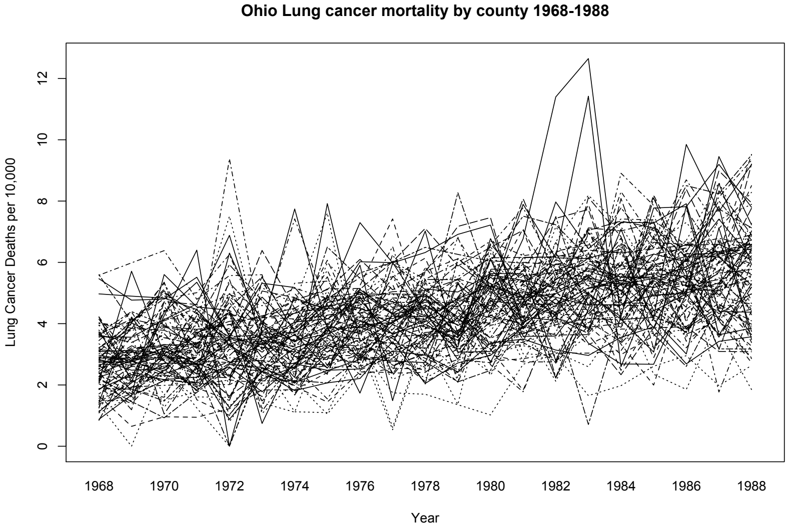
Background

Previous approaches

Previous works have used a Hierarchical Bayesian framework to expand purely spatial models by [Besag et al., 1991] to a space $\times$ time framework.

- [Waller et al., 1997] - Spatial model for each time point
  - No spatial main effects
- [Bernardinelli et al., 1995] - Area-specific intercept and temporal trends
  - All temporal trends assumed linear

Background



Ohio Lung cancer mortality by county 1968-1988
Lung Cancer Deaths per 10,000
0
2
4
6
8
10
12
1968
1970
1972
1974
1976
1978
1980
1982
1984
1986
1988
Year

Background

[Knorr-Held and Besag, 1998]

The Main Effect Model:

- n_{it} - persons at risk in county i ($i = 1, ..., n$) at time t ($t = 1, ..., T$)
- y_{it} - cases or deaths in county i at time t
  - $y_{it} \sim \mathrm{Bin}(n_{it}, \pi_{it})$
- $\eta_{it} = \log\left(\frac{\pi_{it}}{1-\pi_{it}}\right)$

Background

[Knorr-Held and Besag, 1998]

The Main Effect Model:

$$\eta_{it} = \mu + \alpha_t + \gamma_t + \theta_i + \phi_i$$

- μ - overall risk level
- α_t - temporally structured effect of time t
- γ_t - independent effect of time t
- θ_i - spatially structured effect of county i
- ϕ_i - independent effect of county i

Background

[Knorr-Held and Besag, 1998]

Prior Specifications:

$$\eta_{it} = \mu + \alpha_t + \gamma_t + \theta_i + \phi_i$$

- μ - flat non-informative
- $p(\alpha|\lambda_\alpha) \propto \exp\left(-\frac{\lambda_\alpha}{2}\sum_{t=2}^{T}(\alpha_t - \alpha_{t-1})^2\right)$
- $p(\gamma|\lambda_\gamma) \propto \exp\left(-\frac{\lambda_\gamma}{2}\sum_{t=1}^{T}\gamma_t^2\right)$
- $p(\theta|\lambda_\theta) \propto \exp\left(-\frac{\lambda_\theta}{2}\sum_{i\sim j}(\theta_i - \theta_j)^2\right)$
- $p(\phi|\lambda_\phi) \propto \exp\left(-\frac{\lambda_\phi}{2}\sum_{i=1}^{n}\phi_i^2\right)$

With precision matrix λK where the structure of K depends on the assumptions about the prior interrelationship between parameters.

Background

[Knorr-Held and Besag, 1998]

$$\eta_{it} = \mu + \alpha_t + \gamma_t + \theta_i + \phi_i$$

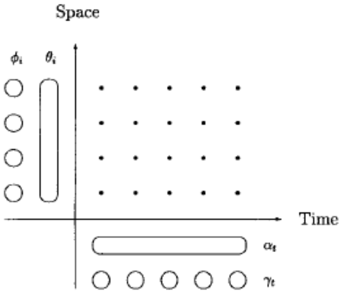


Figure 1. Symbolic representation of the main effect model. Circles represent prior independence, ovals represent prior dependence. Observations in time × space are indicated by small dots.

Background

[Knorr-Held and Besag, 1998]

Limitations

- Combines temporal and spatial main effects additively
- Does not allow for space $\times$ time interactions

[Knorr-Held, 2000]

The Model

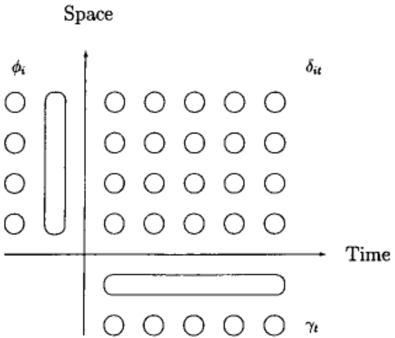
Address the situation where the disease variation cannot be separated into temporal and spatial main effects and spatio-temporal interactions become and important feature.

$$\eta_{it} = \mu + \alpha_t + \gamma_t + \theta_i + \phi_i + \delta_{it}$$

- μ - overall risk level
- α_t - temporally structured effect of time t
- γ_t - independent effect of time t
- θ_i - spatially structured effect of county i
- ϕ_i - independent effect of county i
- δ_{it} - space $\times$ time interaction

[Knorr-Held, 2000]

Four Space × Time Interactions

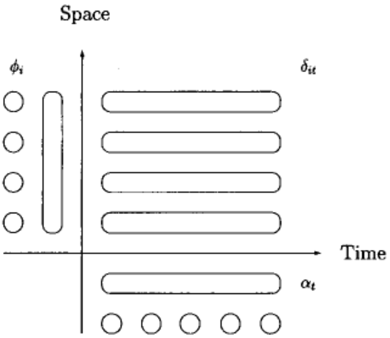


Type I Independent - Independent multivariate gaussian prior

$p(\delta|\lambda_\delta) \propto \exp\left(-\frac{\lambda_\delta}{2}\sum_{i=1}^{n}\sum_{t=1}^{T}\delta_{it}^2\right)$ with $K_\delta = K_\gamma \otimes K_\phi$ (rank nT)

[Knorr-Held, 2000]

Four Space × Time Interactions

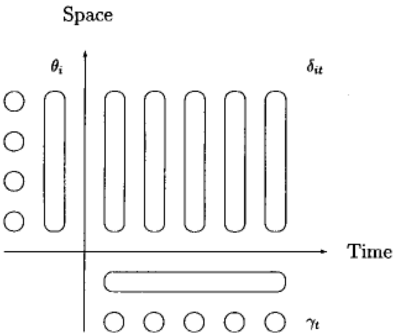


Type II Temporal trends differ by area - Random walk prior

$$p(\delta|\lambda_\delta) \propto \exp\left(-\frac{\lambda_\delta}{2}\sum_{i=1}^{n}\sum_{t=2}^{T}(\delta_{it}-\delta_{i,t-1})^2\right) \text{ with}$$
$$K_\delta = K_\alpha \otimes K_\phi \text{ (rank } n(T-1))$$

[Knorr-Held, 2000]

Four Space $\times$ Time Interactions



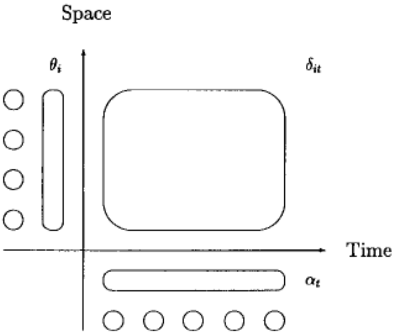
Type III Spatial trends differ over time - Intrinsic autoregression prior

$$p(\delta|\lambda_\delta) \propto \exp\left(-\frac{\lambda_\delta}{2}\sum_{i\sim j}\sum_{t=1}^{T}(\delta_{it}-\delta_{jt})^2\right) \text{ with}$$

$$K_\delta = K_\gamma \otimes K_\theta \text{ (rank } (n-1)T)$$

[Knorr-Held, 2000]

Four Space $\times$ Time Interactions



Type IV Spatio-temporal interaction - conditional depends on first and second order neighbors

$$p(\delta|\lambda_\delta) \propto \exp\left(-\frac{\lambda_\delta}{2}\sum_{i\sim j}\sum_{t=2}^{T}(\delta_{it}-\delta_{jt}-\delta_{i,t-1}+\delta_{j,t-1})^2\right)$$

with $K_\delta = K_\alpha \otimes K_\theta$ (rank $(n-1)(T-1)$)

[Knorr-Held, 2000]

Evaluation

Used Markov chain Monte Carlo (using univariate Metropolis steps) to sample from the posterior distribution of :

- $\eta_{it} = \mu + \alpha_t + \gamma_t + \theta_i + \phi_i$ (one model)
- $\eta_{it} = \mu + \alpha_t + \gamma_t + \theta_i + \phi_i + \delta_{it}$ (one model with four variations)

to the Ohio Lung Cancer data and compared posterior deviance.

The End

Questions?

References

Bernardinelli, L., Clayton, D., Pascutto, C., Montomoli, C., Ghislandi, M., and Songini, M. (1995).
Bayesian analysis of space-time variation in disease risk.
Statistics in Medicine, 14(21-22):2433–2443.

Besag, J., Green, P., Higdon, D., and Mengersen, K. (1995).
Bayesian computation and stochastic system.
Statistical Science, 10(1):3–66.

Besag, J. and Kooperberg, C. (1995).
On conditional and intrinsic autoregression.
Biometrika, 82(4):733–746.

Besag, J., York, J., and Mollie, A. (1991).
Bayesian image restoration, with two applications in spatial statistics.
Annals of the Institute of Statistical Mathematics, 43:1–59.

Knorr-Held, L. (2000).
Bayesian modelling of inseperable space-time variation in disease risk.
Statistics in Medicine, 19:2555–2567.

Knorr-Held, L. and Besag, J. (1998).
Modelling risk from a disease in time and space.
Statistics in Medicine, 17:2045–2060.

Rue, H. and Held, L. (2005).
Gaussian Markov Random Fields Theory and Applications.
Number 104 in Monographs on Statistics and Applied Probability. Chapman and Hall.

Waller, L. A., Carlin, B. P., Xia, H., and Gelfand, A. E. (1997).
Hierarchical spatio-temporal mapping of disease rates.
Journal of the American Statistical Association, 92(438):607–617.