Bayesian modeling of inseparable space-time variation in disease risk

Leonhard Knorr-Held

Laina Mercer

Department of Statistics
UW

May 23, 2013

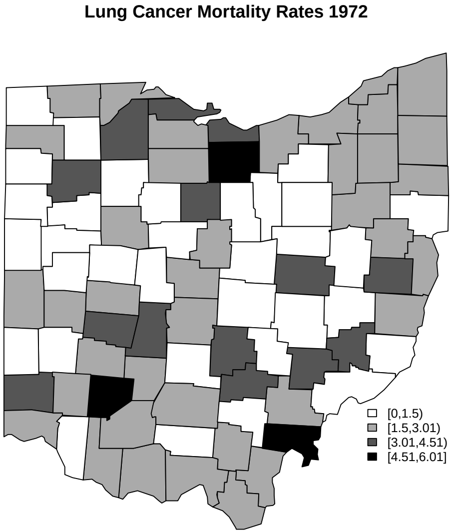
Motivation

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

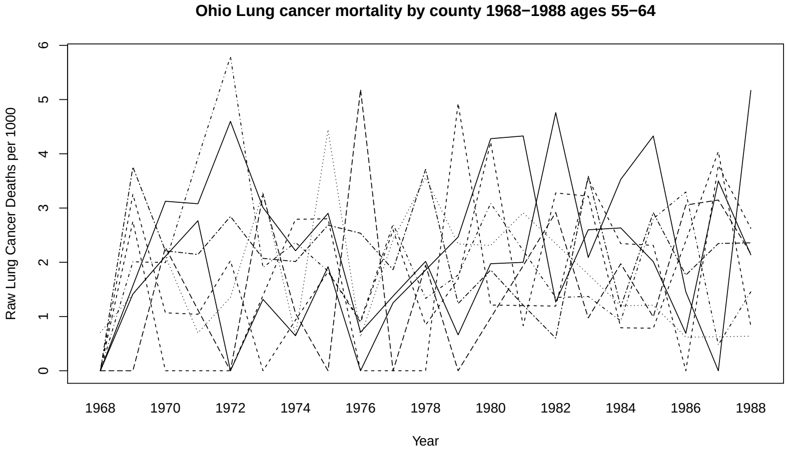
Motivation

Ohio Lung Cancer Example



Motivation

Ohio Lung Cancer Example



Background

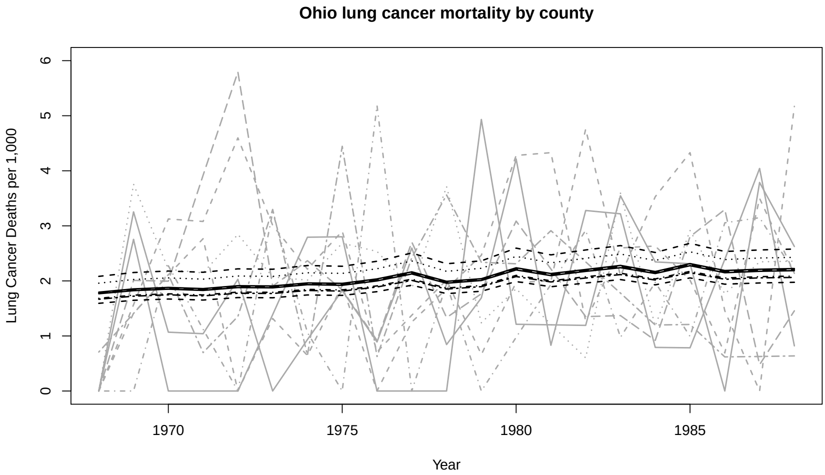
Previous approaches

Previous works have used a Hierarchical Bayesian framework to expand purely spatial models by Besag, York, and Mollie (1991) to a space $\times$ time framework.

- Bernardinelli et al. (1995)
  - Area-specific intercept and temporal trends
  - All temporal trends assumed linear
- Waller et al. (1997)
  - Spatial model for each time point
  - No spatial main effects
- Knorr-Held and Besag (1998)
  - Included spatial and temporal main effects
  - Does not allow for space $\times$ time interactions

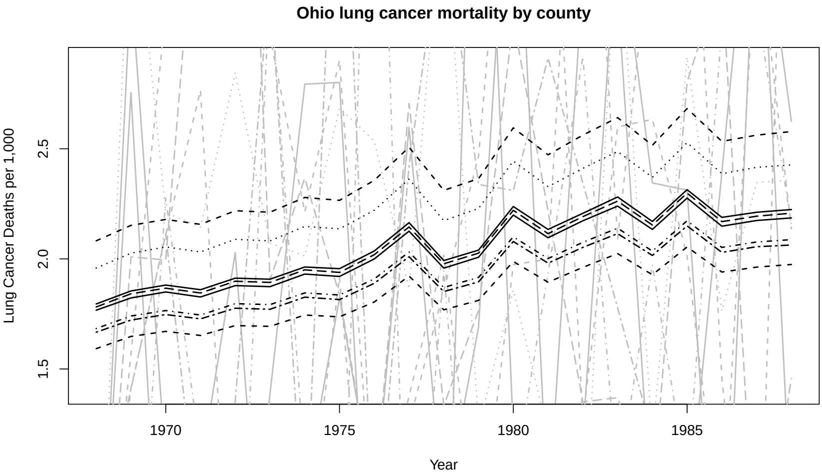
Motivation

Ohio Lung Cancer Example



Motivation

Ohio Lung Cancer Example



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

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

The Model

Stage 1

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

- 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

The exchangeable effects γ (for time) and ϕ (for space) we are assigned multivariate Gaussian priors with mean zero and precision matrix λK:

$$p(\gamma|\lambda_\gamma) \sim \mathcal{N}\left(0, \frac{1}{\lambda_\gamma} K_\gamma^{-1}\right)$$

$$p(\phi|\lambda_\phi) \sim \mathcal{N}\left(0, \frac{1}{\lambda_\phi} K_\phi^{-1}\right)$$

where $K_\gamma = I_{TxT}$ and $K_\phi = I_{kxk}$.

The Model

Stage 2 - First Order Random Walk

Temporally structured effect of time α_t is assigned a random walk.

$$\alpha_t|\alpha_{-t}, \lambda_\alpha \sim \mathcal{N}\left(\frac{1}{2}(\alpha_{t-1} + \alpha_{t+1}), 1/(2\lambda_\alpha)\right)$$

Also represented as $p(\alpha|\lambda_\alpha) \propto \exp\left(-\frac{\lambda_\alpha}{2}\alpha^T K_\alpha \alpha\right)$
where

$$K_\alpha = \begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{pmatrix}.$$

The Model

Stage 2 - Intrinsic Autoregressive (ICAR)

Spatially structured effect of area θ_i is assigned

$$\theta_i | \theta_{-i}, \lambda_\theta \sim \mathcal{N}\left(\frac{1}{m_i} \sum_{j:j\sim i} \theta_j, \frac{1}{m_i \lambda_\theta} \right)$$

where m_i is the # of neighbors. The improper joint distribution can be written as $p(\theta|\lambda_\theta) \propto \exp\left(-\frac{\lambda_\theta}{2}\theta^T K_\theta \theta\right)$, where

$$K_{\theta,ij} = \begin{cases} m_i & i = j \\ -1 & i \sim j \\ 0 & \text{otherwise} \end{cases}$$

The Model

Stage 2 - *iid* prior for δ

Type I Independent multivariate Gaussian prior

$$\delta_{it} | \delta_{-it}, \lambda_\delta \sim \mathcal{N}\left(0, 1/\lambda_\delta\right)$$

the joint distribution is

$$p(\delta | \lambda_\delta) \quad \propto \quad \exp\left(-\frac{\lambda_\delta}{2} \delta^T K_\delta \delta\right)$$

where $K_\delta = K_\phi \otimes K_\gamma = I_{kT \times kT}$.

The Model

Stage 2 - Random Walk prior for δ

Type II Temporal trends differ by area - first order random walk prior

$$\delta_{it}|\delta_{-it}, \lambda_\delta \sim \mathcal{N}\left(\frac{1}{2}(\delta_{i,t-1} + \delta_{i,t+1}), 1/(2\lambda_\delta)\right)$$

The improper joint distribution can be expressed as

$$p(\delta|\lambda_\delta) \quad \propto \quad \exp\left(-\tfrac{\lambda_\delta}{2}\delta^T K_\delta \delta\right)$$

where $K_\delta = K_\phi \otimes K_\alpha$ (rank $k(T-1)$).

Type III Spatial trends differ over time - Intrinsic autoregression prior

$$\delta_{it}|\delta_{-it}, \lambda_\delta \sim \mathcal{N}\left(\frac{1}{m_i}\sum_{j:j\sim i}\delta_{jt}, \frac{1}{m_i\lambda_\delta}\right)$$

The improper joint distribution can be expressed as

$$p(\delta|\lambda_\delta) \quad \propto \quad \exp\left(-\tfrac{\lambda_\delta}{2}\delta^T K_\delta \delta\right)$$

Where $K_\delta = K_\theta \otimes K_\gamma$ (rank $(k-1)T$).

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

$$\delta_{it}|\delta_{-it}, \lambda_\delta \sim$$
$$\mathcal{N}\left(\tfrac{1}{2}(\delta_{i,t-1} + \delta_{i,t+1}) + \tfrac{1}{m_i}\sum_{j:j\sim i}\delta_{jt} - \tfrac{1}{m_i}\sum_{j:j\sim i}(\delta_{j,t-1} + \delta_{j,t+1}), \tfrac{1}{2m_i\lambda_\delta}\right)$$

The improper joint distribution can be expressed as

$$p(\delta|\lambda_\delta) \quad \propto \quad \exp\left(-\tfrac{\lambda_\delta}{2}\delta^T K_\delta \delta\right)$$

Where $K_\delta = K_\theta \otimes K_\alpha$ (rank $(k-1)(T-1)$).

Hyperparameters were all assigned

$$\lambda \sim \mathrm{Gamma}(1, 0.01)$$

resulting in a convenient posterior distribution. For example the full conditional for λ_δ is:

$$\lambda_\delta | \delta \ \sim \mathrm{Gamma}\left(1 + 0.5 \times \mathrm{rank}(K_\delta), 0.01 + 0.5 \times \delta' k_\delta \delta\right)$$

where $\mathrm{rank}(K_\delta)$ depends on the interaction type.

To compare fit and complexity of each model the saturated deviance was calculated based on 2,500 samples from the posterior.

$$D^{(s)} = 2\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{y_{it}\log\left(\frac{y_{it}}{n_{it}\pi_{it}^{(s)}}\right) + (n_{it} - y_{it})\log\left(\frac{n_{it} - y_{it}}{n_{it}\left(1 - \pi_{it}^{(s)}\right)}\right)\right\}$$

where $\pi_{it}^{(s)} = \frac{\exp\left(\eta_{it}^{(s)}\right)}{1+\exp\left(\eta_{it}^{(s)}\right)}$.

Knorr-Held (2000) employed Markov chain Monte Carlo to sample from the implied posterior distributions.

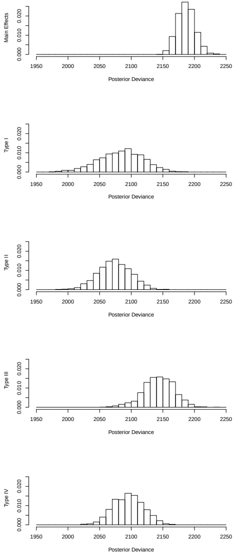
Univariate Metropolis steps were applied for each parameter and hyperparameters were updated with samples from their full conditionals.

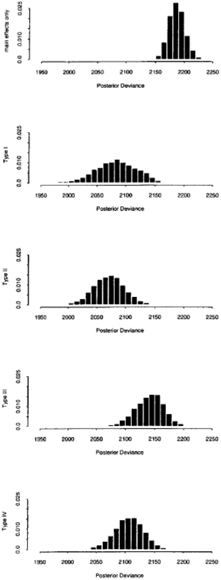
MCMC in `R`: an update for every parameter in an interaction model takes 0.1-0.2s.

Note: in `INLA` the main effects and interaction models type **I-III** all fit in less than 20min.

Ohio Lung Cancer

Posterior Distribution of the Deviance





Ohio Lung Cancer

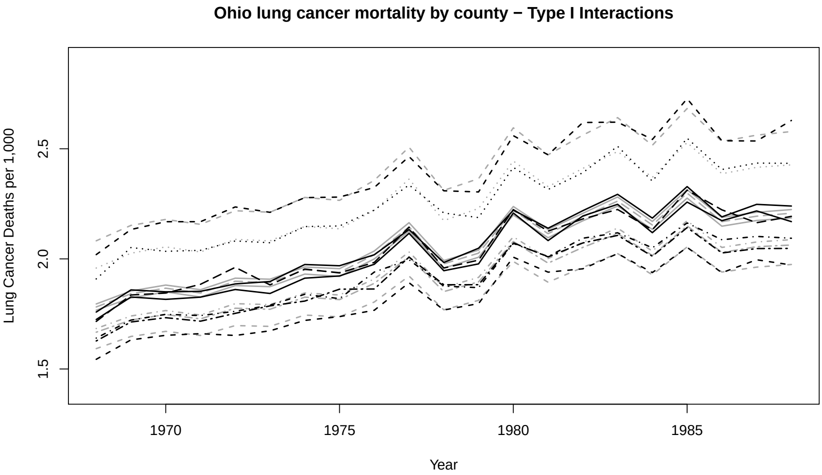
Posterior Distribution of the Deviance

Model	Median	Mean	IQR	SD
Main Effects	2187	2187	18.6	13.7
	2187	2187	18.5	13.9
Type I	2086	2084	48.4	34.9
	2083	2082	48.6	35.9
Type II	2073	2073	35.3	25.5
	2071	2071	36.6	27.0
Type III	2144	2143	32.7	24.1
	2142	2141	32.8	24.8
Type IV	2096	2096	36.2	26.0
	2106	2106	32.8	24.7

Table: Laina's values in black and Knorr-Held (2000) in gray.

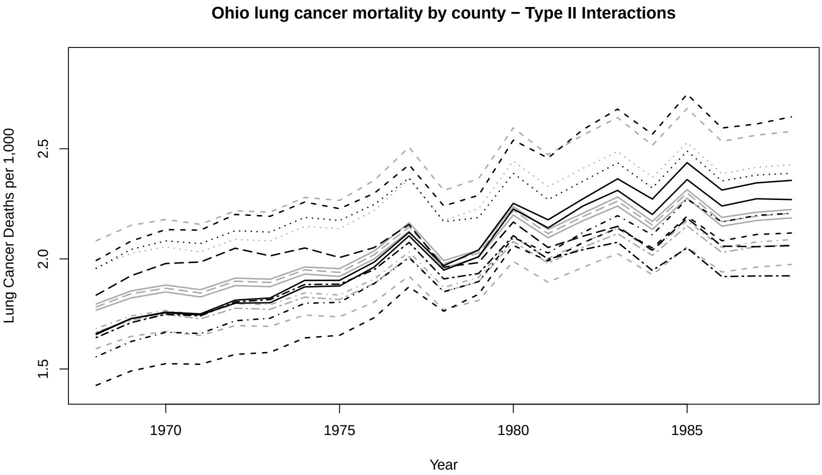
Ohio Lung Cancer

Type I interaction



Ohio Lung Cancer

Type II interaction



Knorr-Held (2000)

Conclusions & Critique

Provided and motivated a flexible approach for modeling space-time data.

Was thin on MCMC details and diagnostics.

Did not motivate use of deviance over DIC or p_D for model selection.

Focussed exclusively on non-parametric smoothing approaches.

No discussion of incorporating covariates.

Thank you!

Thank you all for your feedback and support throughout the quarter. Specifically, I would like to thank:

- William for suggesting a hair cut and shaded plots,
- Bob for suggesting enthusiasm, and
- Jon for suggesting this paper!

References

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

Besag, J., J. York, and A. Mollie (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 J. Besag (1998).
Modelling risk from a disease in time and space.
Statistics in Medicine 17, 2045–2060.

Knorr-Held, L. and H. Rue (2002).
On block updating in markov random field models for disease mapping.
Scandinavian Journal of Statistics 29, 597–614.

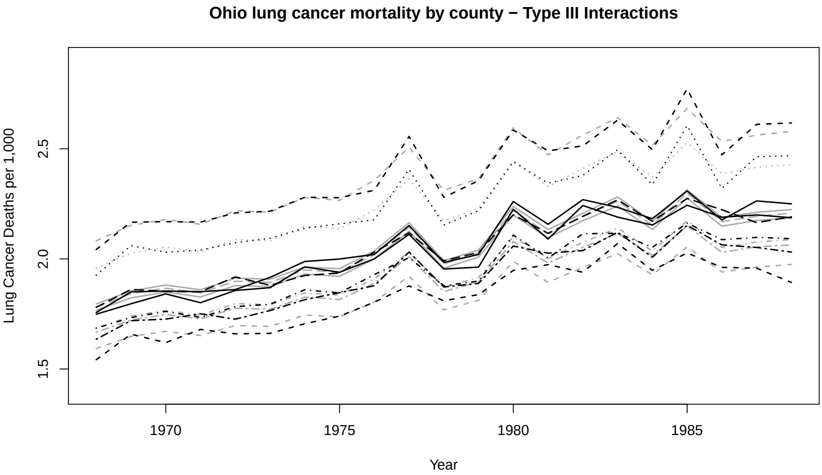
Rue, H. and L. Held (2005).
Gaussian Markov Random Fields Theory and Applications.
Number 104 in Monographs on Statistics and Applied Probability. Chapman and Hall.

Schrodle, B. and L. Held (2010).
Spatio-temporal disease mapping using inla.
Environmetrics.

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

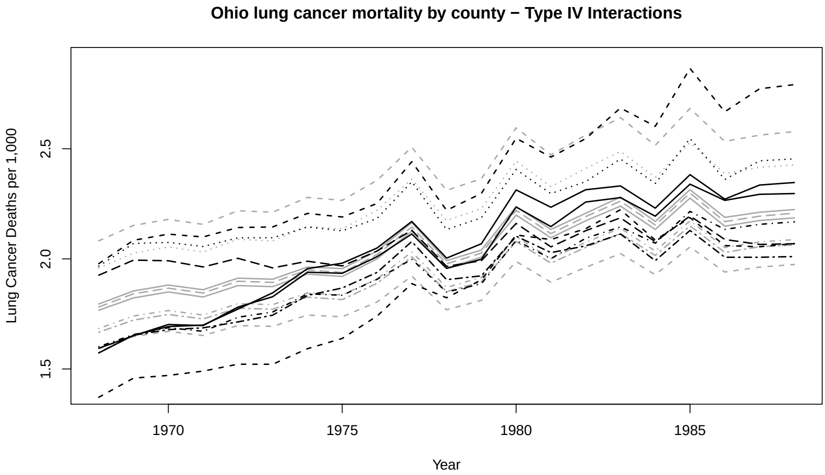
Ohio Lung Cancer

Type III interaction



Ohio Lung Cancer

Type IV interaction



A logical next step was to make fitting these models much faster.

- Knorr-Held and Rue (2002) introduced block updating
- Rue and Held (2005) great overview of Gaussian Markov Random Fields and more details on block updating
- Schrodle and Held (2010) describes (poorly) how to fit models from Knorr-Held (2000) in `INLA`.