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

Lung Cancer Mortality Rates 1972

Legend:
- [0,1.5)
- [1.5,3.01)
- [3.01,4.51)
- [4.51,6.01]
Motivation
Ohio Lung Cancer Example
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

Ohio lung cancer mortality by county

Year

Lung Cancer Deaths per 1,000

Motivation

Ohio Lung Cancer Example

Ohio lung cancer mortality by county

<table>
<thead>
<tr>
<th>Year</th>
<th>Lung Cancer Deaths per 1,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970</td>
<td>1.5</td>
</tr>
<tr>
<td>1975</td>
<td>2.0</td>
</tr>
<tr>
<td>1980</td>
<td>2.5</td>
</tr>
<tr>
<td>1985</td>
<td></td>
</tr>
</tbody>
</table>
Address the situation where the disease variation cannot be separated into temporal and spatial main effects and spatio-temporal interactions become an important feature.

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

Stage 1

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

- \( \mu \) - overall risk level
- \( \alpha_t \) - temporally structured effect of time \( t \)
- \( \gamma_t \) - independent effect of time \( t \)
- \( \theta_i \) - spatially structured effect of county \( i \)
- \( \phi_i \) - independent effect of county \( i \)
- \( \delta_{it} \) - space \( \times \) time interaction
The exchangeable effects $\gamma$ (for time) and $\phi$ (for space) we are assigned multivariate Gaussian priors with mean zero and precision matrix $\lambda K$:

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

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

where $K_\gamma = I_{T \times T}$ and $K_\phi = I_{k \times k}$. 
The Model
Stage 2 - First Order Random Walk

Temporally structured effect of time $\alpha_t$ 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

$$K_\alpha = \begin{pmatrix} 1 & -1 & \cdots & \cdots & \cdots & -1 \\ -1 & 2 & 1 & \cdots & \cdots & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \cdots & -1 & 2 & -1 \\ -1 & 2 & -1 & \cdots & \cdots & \cdots \\ -1 & 1 & \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$
The Model
Stage 2 - Intrinsic Autoregressive (ICAR)

Spatially structured effect of area \( \theta_i \) 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 - \textit{iid} prior for $\delta$

Type I Independent multivariate Gaussian prior

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

the joint distribution is

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

where $K_\delta = K_{\phi} \otimes K_{\gamma} = I_{kT \times kT}$. 
Stage 2 - Random Walk prior for $\delta$

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

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

The improper joint distribution can be expressed as

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

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

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

Where \( K_\delta = K_\theta \otimes K_\gamma \) (rank \((k - 1)T\)).
The Model
Stage 2 - Space-Time prior for $\delta$

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

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

The improper joint distribution can be expressed as

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

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

\[ \lambda \sim \text{Gamma}(1, 0.01) \]

resulting in a convenient posterior distribution. For example the full conditional for \( \lambda_\delta \) is:

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

where \( \text{rank}(K_\delta) \) 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

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

*Table: Laina’s values in black and Knorr-Held (2000) in gray.*
Ohio Lung Cancer
Type I interaction

Ohio lung cancer mortality by county – Type I Interactions

Year
Lung Cancer Deaths per 1,000
1.5 2.0 2.5
Ohio Lung Cancer

Type II interaction

Ohio lung cancer mortality by county – Type II Interactions

Lung Cancer Deaths per 1,000

Year


1.5 2.0 2.5
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 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!


Ohio Lung Cancer
Type III interaction

Ohio lung cancer mortality by county – Type III Interactions

Year
Lung Cancer Deaths per 1,000

1.5 2.0 2.5
Ohio Lung Cancer

Type IV interaction

Ohio lung cancer mortality by county – Type IV Interactions

Year

Lung Cancer Deaths per 1,000


1.5 2.0 2.5

Ohio lung cancer mortality by county – Type IV Interactions

Year

Lung Cancer Deaths per 1,000


1.5 2.0 2.5
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.