

Session 1.4: Spatio-temporal modelling for area data

VIBASS, University of Valencia

20 July 2022

Learning Objectives

At the end of this session you should be able to:

- List and describe different types of temporal dependence
- Explain how to extend spatial or temporal models to spatio-temporal models
- Distinguish the different types of structures for space time interactions and be able to search for unusual space time patterns
- Fit Bayesian space time models in INLA

The topics treated in this lecture can be found in Chapter 7 of the book **Spatial and Spatio-Temporal Bayesian Models with R-INLA**.

Outline

1. Temporal dependence
2. From space to space-time
3. Type of interactions

Temporal dependence

Temporal dependence

- Similarly to spatial dependence, it is sometimes necessary to model temporal dependence of data or of parameters:
- the weekly or monthly number of cases of many diseases exhibit often a seasonal pattern as well as short term dependence
- the underlying daily level of atmospheric pollutants, e.g. PM₁₀, will show strong correlation over consecutive days because their lifetime lasts over several days
- To the contrary of spatial models, there is a natural order to any time series data which is used in specifying the models.

Spatial patterns

1. Data

- Disease counts y_{ik} in area i and stratum k , aggregated over a time period, $i = 1, \dots, N, k = 1, \dots, K$
- Population counts n_{ik} in area i and stratum k , aggregated over a time period
- Expected numbers $E_i = \sum_k n_{ik} r_k$, where r_k reference rate for stratum (age, gender,...)

2. Spatial smoothing using BYM model

$$y_i \sim \text{Poisson}(E_i \rho_i); \quad i = 1, \dots, N$$

$$\log \rho_i = b_0 + v_i + u_i$$

$$v_i \sim \text{Normal}(0, \sigma_v^2)$$

$$\boldsymbol{u} \sim \text{ICAR}(\mathbf{W}, \sigma_u^2) \rightarrow u_i | u_j, j \neq i \sim \text{Normal} \left(\frac{\sum_j w_{ij} u_j}{\sum_j w_{ij}}, \sigma_u^2 / n_i \right)$$

with $w_{ij} = 1$ if areas i and j are neighbours, 0 otherwise

Temporal trends

1. Data

- Disease counts y_{tk} in time period t and stratum k , aggregated over space, $t = 1, \dots, T$ (equally-spaced time intervals), $k = 1, \dots, K$
- Population counts n_{tk} in time period t and stratum k , aggregated over space
- Expected numbers $E_t = \sum_k n_{tk} r_k$, where r_k reference rate for stratum (age, gender,...)

2. Temporal trends

$$y_t \sim \text{Poisson}(E_t \rho_t); \quad t = 1, \dots, T$$
$$\log \rho_t = ???$$

Temporal trends

1. Data

- Disease counts y_{tk} in time period t and stratum k , aggregated over space, $t = 1, \dots, T$ (equally-spaced time intervals), $k = 1, \dots, K$
- Population counts n_{tk} in time period t and stratum k , aggregated over space
- Expected numbers $E_t = \sum_k n_{tk} r_k$, where r_k reference rate for stratum (age, gender,...)

2. Temporal trends

$$y_t \sim \text{Poisson}(E_t \rho_t); \quad t = 1, \dots, T$$
$$\log \rho_t = b_0 + \beta t \quad \text{simple linear regression}$$

Temporal trends

1. Data

- Disease counts y_{tk} in time period t and stratum k , aggregated over space, $t = 1, \dots, T$ (equally-spaced time intervals), $k = 1, \dots, K$
- Population counts n_{tk} in time period t and stratum k , aggregated over space
- Expected numbers $E_t = \sum_k n_{tk} r_k$, where r_k reference rate for stratum (age, gender,...)

2. Temporal trends

$$y_t \sim \text{Poisson}(E_t \rho_t); \quad t = 1, \dots, T$$

$$\log \rho_t = b_0 + \beta t \quad \text{simple linear regression}$$

$$= b_0 + \psi_t \quad \text{global temporal smoothing}$$

$$\psi_t \sim \text{Normal}(0, \sigma_\psi^2)$$

Temporal trends

1. Data

- Disease counts y_{tk} in time period t and stratum k , aggregated over space, $t = 1, \dots, T$ (equally-spaced time intervals), $k = 1, \dots, K$
- Population counts n_{tk} in time period t and stratum k , aggregated over space
- Expected numbers $E_t = \sum_k n_{tk} r_k$, where r_k reference rate for stratum (age, gender,...)

2. Temporal trends

$$y_t \sim \text{Poisson}(E_t \rho_t); \quad t = 1, \dots, T$$

$$\log \rho_t = b_0 + \beta t \quad \text{simple linear regression}$$

$$= b_0 + \psi_t \quad \text{global temporal smoothing}$$

$$\psi_t \sim \text{Normal}(0, \sigma_\psi^2)$$

$$= b_0 + \gamma_t \quad \text{local temporal smoothing}$$

$$\gamma_t \sim \text{distribution ?}$$

Temporal trends

1. Data

- Disease counts y_{tk} in time period t and stratum k , aggregated over space, $t = 1, \dots, T$ (equally-spaced time intervals), $k = 1, \dots, K$
- Population counts n_{tk} in time period t and stratum k , aggregated over space
- Expected numbers $E_t = \sum_k n_{tk} r_k$, where r_k reference rate for stratum (age, gender,...)

2. Temporal trends

$$y_t \sim \text{Poisson}(E_t \rho_t); \quad t = 1, \dots, T$$

$\log \rho_t = b_0 + \beta t$ simple linear regression

$= b_0 + \psi_t$ global temporal smoothing

$$\psi_t \sim \text{Normal}(0, \sigma_\psi^2)$$

$= b_0 + \gamma_t$ local temporal smoothing

$\gamma_t \sim$ distribution ?

$= b_0 + \gamma_t + \psi_t$ global and local temporal smoothing

Autoregressive models

- Idea: predict an output based on the previous outputs
- Let $\theta = (\theta_1, \dots, \theta_T)$ be a time ordered sequence of parameters

Autoregressive models

- Idea: predict an output based on the previous outputs
- Let $\boldsymbol{\theta} = (\theta_1, \dots, \theta_T)$ be a time ordered sequence of parameters
- An autoregressive Gaussian model for $\boldsymbol{\theta}$ is defined by:
 - a time lag p
 - a set of coefficients $\{b_1, \dots, b_p\}$

so that

$$\theta_t = b_1\theta_{t-1} + b_2\theta_{t-2} + \dots + b_p\theta_{t-p} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2)$$

equivalently

$$\theta_t | \theta_{t-1}, \theta_{t-2}, \dots, \theta_{t-p} \sim N \left(\sum_{j=1}^p b_j \theta_{t-j}, \sigma_\epsilon^2 \right), \quad t = p+1, \dots, T$$

Random walks

Order 1 (RW1)

Plot of RW1

Order 2 (RW2)

Plot of RW1-RW2

$$\theta_t = \theta_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2)$$

This process is non stationary (variance doesn't exist)

$$\begin{aligned}\theta_t &= \theta_{t-1} + \epsilon_t, \quad \theta_{t-1} = \theta_{t-2} + \epsilon_{t-1} \\ \Rightarrow \theta_t &= \theta_{t-2} + \epsilon_{t-1} + \epsilon_t \\ \Rightarrow \theta_t &= \theta_0 + \epsilon_1 + \dots + \epsilon_t\end{aligned}$$

Mean $E(\theta_t) = \theta_0$

Variance $\text{Var}(\theta_t) = \text{Var}(\epsilon_1) + \dots + \text{Var}(\epsilon_t) = t\sigma_\epsilon^2 \rightarrow \infty \text{ if } t \rightarrow \infty$

- **s the difference of levels on consecutive time points** ($\theta_t - \theta_{t-1} = \epsilon_t$)

It is a stationary process if mean and variance independent of time (t)

Random walks

Order 1 (RW1)

Plot of RW1

Order 2 (RW2)

Plot of RW1-RW2



Random walks

[Order 1 \(RW1\)](#)[Plot of RW1](#)[**Order 2 \(RW2\)**](#)[Plot of RW1-RW2](#)

$$\theta_t = 2\theta_{t-1} - \theta_{t-2} + \epsilon_t$$

- This process is non stationary
- It only models a linear combination of levels on consecutive time points ($\theta_t - 2\theta_{t-1} + \theta_{t-2} = \epsilon_t$)

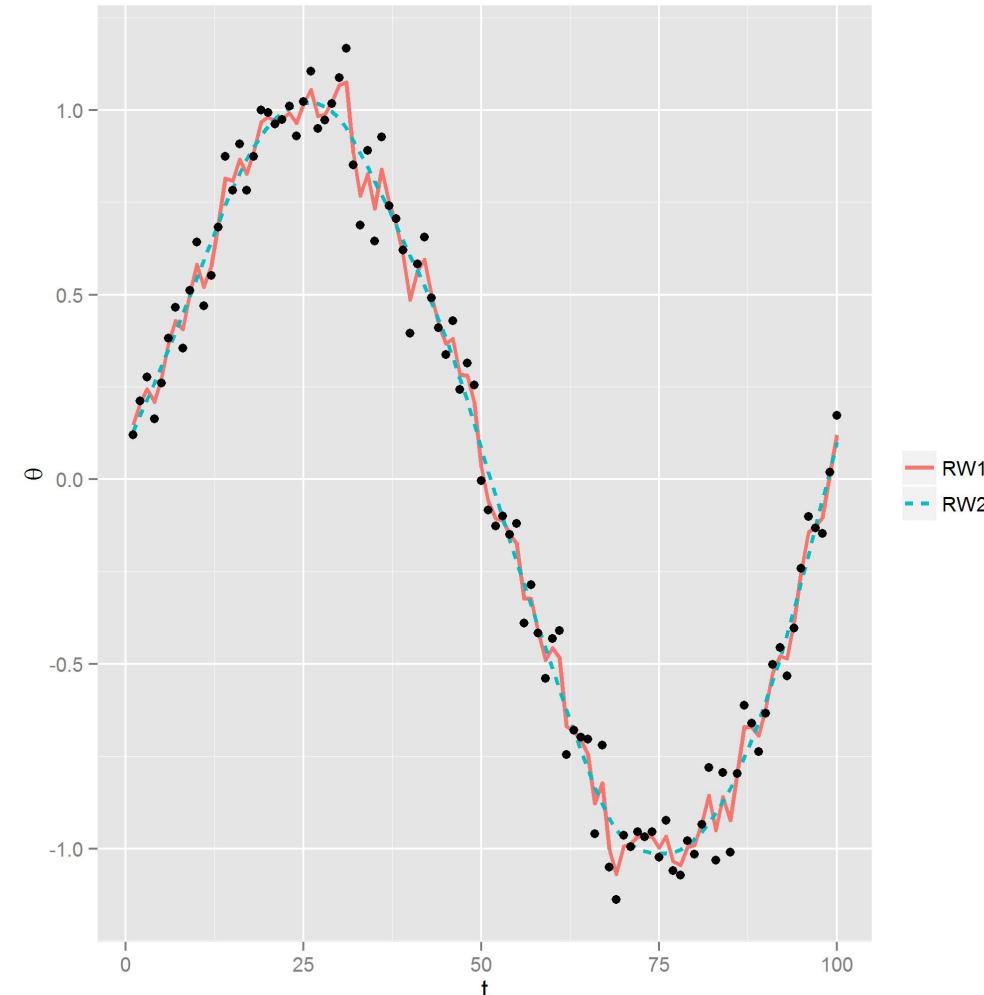
Random walks

Order 1 (RW1)

Plot of RW1

Order 2 (RW2)

Plot of RW1-RW2



Conditional distributions for a RW(1)

- The conditional distributions $p(\theta_t | \boldsymbol{\theta}_{-t})$, where $\boldsymbol{\theta}_{-t}$ represent the vector of θ s with θ_t removed, can be derived.
- RW1: The conditional distribution of θ_t involves θ_{t-1} :

$$p(\theta_t | \boldsymbol{\theta}_{-t}, \sigma_\epsilon^2) = N(\theta_{t-1}, \sigma_\epsilon^2)$$

- Similar form with the spatial autoregression case (ICAR)

$$\text{Recall: } p(X_i | X_{-i}) = \text{Normal} \left(\frac{\sum_j w_{ij} X_j}{\sum_j w_{ij}}, \frac{\sigma_u^2}{\sum_j w_{ij}} \right)$$

- time neighbours of t are $t - 1$ and $t + 1$
- if $t = 1$ or $t = T$, only 1 neighbour: t_2 and t_{T-1}
- weights: 1

In INLA

```
f(ID.time, model="rw1")
```

Conditional distributions for a RW(2)

- A different parametrisation can be specified similarly to the spatial ICAR, so that

| | | |
|---------------------|--|------------------------|
| $t = 1$ | $\theta_{t+1}, \theta_{t+2}$ | weights = 2, -1 |
| $t = 2$ | $\theta_{t-1}, \theta_{t+1}, \theta_{t+2}$ | weights = 2, 4, -1 |
| $t = 3, \dots, T-2$ | $\theta_{t-2}, \theta_{t-1}, \theta_{t+1}, \theta_{t+2}$ | weights = -1, 4, 4, -1 |
| $t = T-1$ | $\theta_{t-2}, \theta_{t-1}, \theta_T;$ | weights = -1, 4, 2 |
| $t = T$ | $\theta_{t-2}, \theta_{t-1};$ | weights = -1, 2 |

which can be re-written as

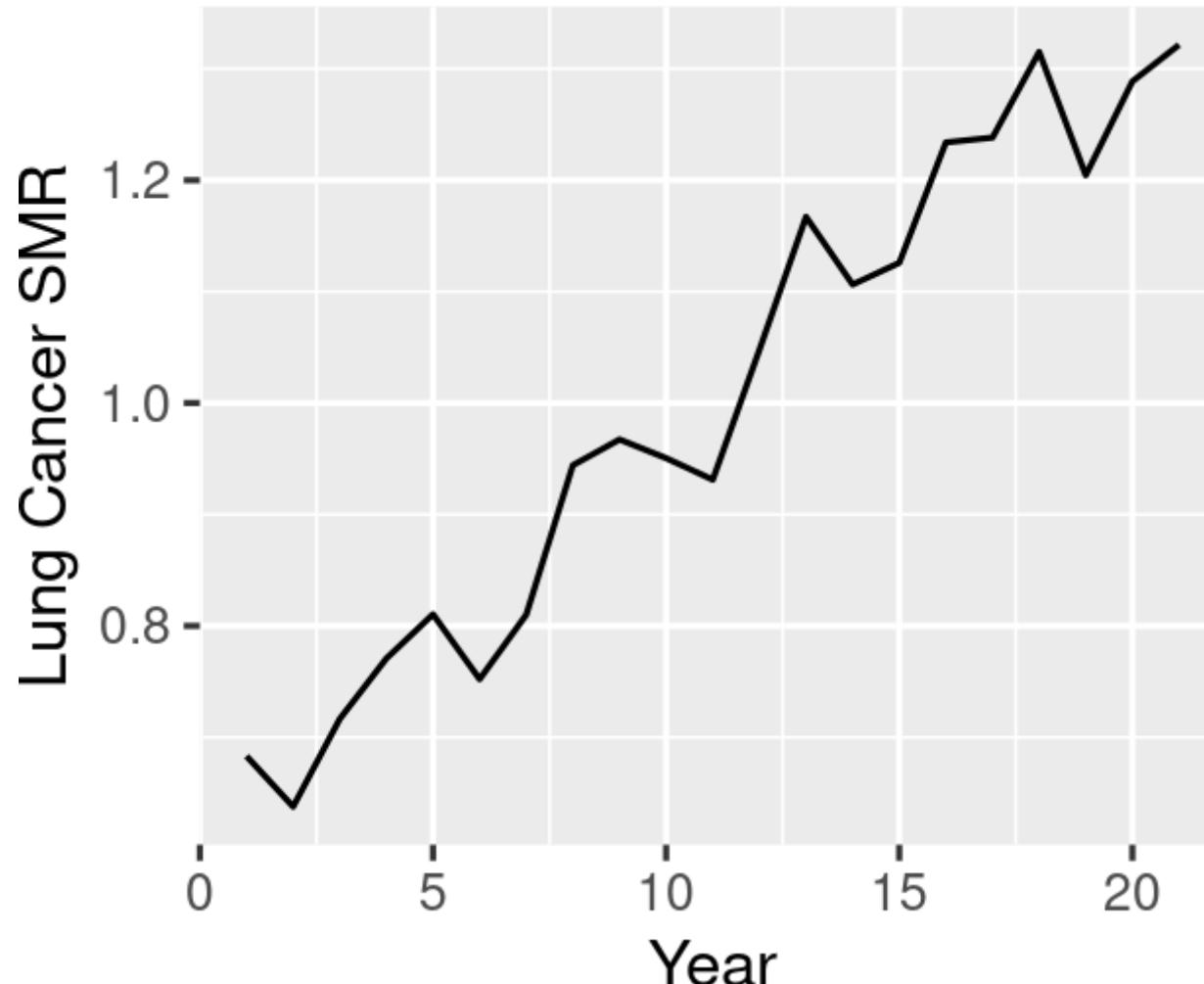
$$p(\theta_t | \boldsymbol{\theta}_{-t}, \sigma_\epsilon^2) = N(2\theta_{t-1} - \theta_{t-2}, \sigma_\epsilon^2)$$

In INLA

```
f(ID.time, model="rw2")
```

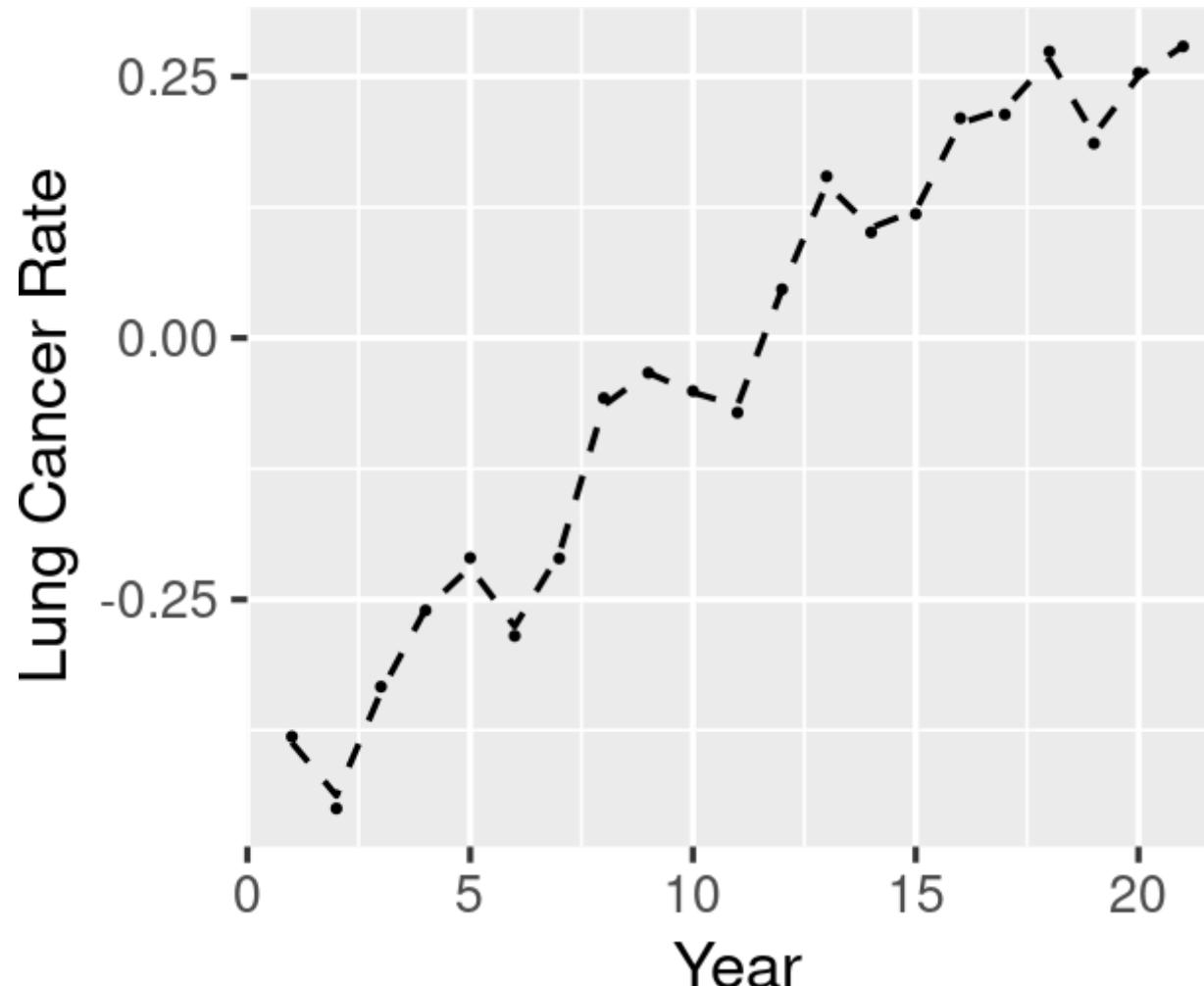
Ohio Lung cancer data

- Data on lung cancer in 88 counties of Ohio, 1968-1988
- Annual rates by gender and race

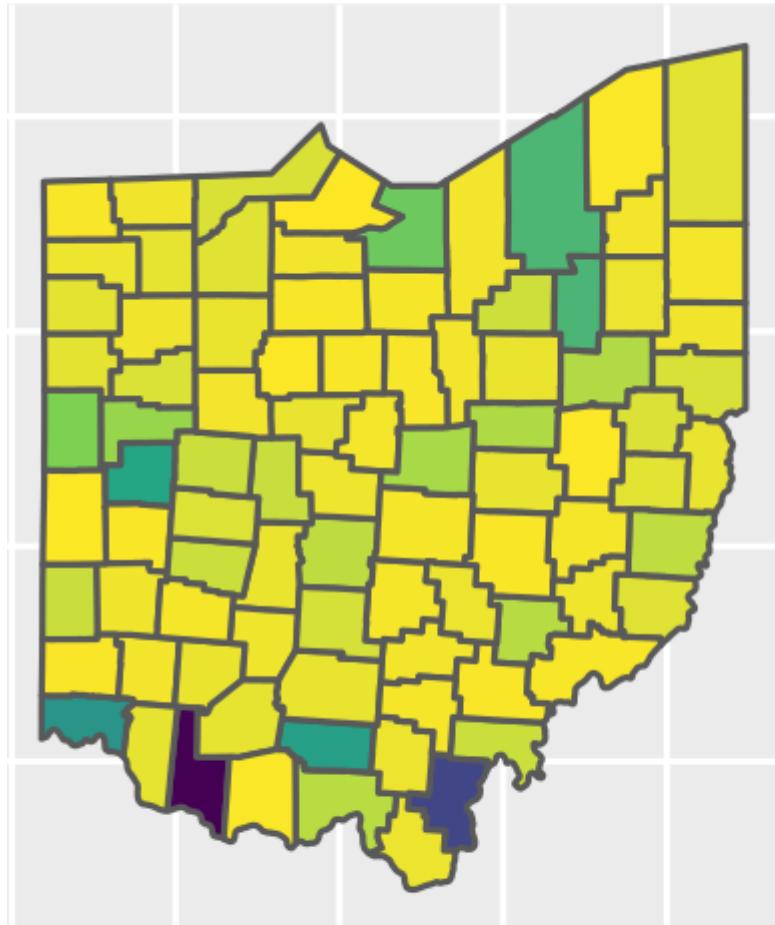


Ohio Lung cancer data: modelled temporal trend

- Smoothed temporal trend using RW1 and RW2 models (posterior mean)

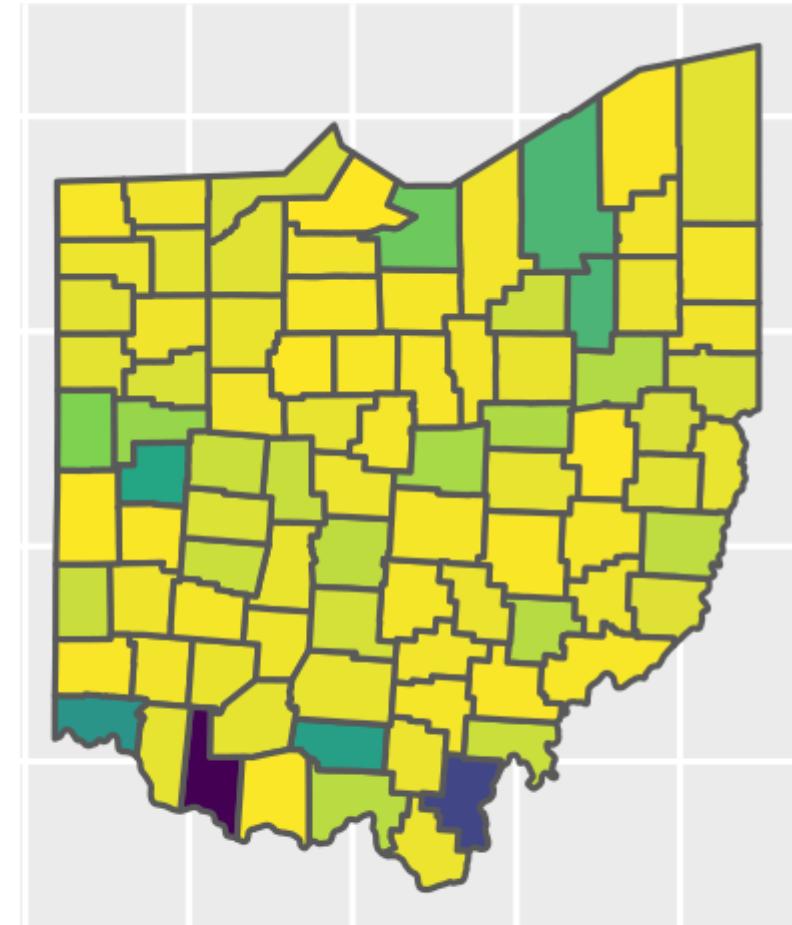


Ohio Lung cancer spatial pattern (no temporal dimension)



SMR

| | | | | | | | |
|---|---|---|---|---|---|---|----|
| | 3 | | 6 | | 9 | | 12 |
|---|---|---|---|---|---|---|----|



ρ_i

| | | | | | | | |
|---|---|---|---|---|---|---|----|
| | 3 | | 6 | | 9 | | 12 |
|---|---|---|---|---|---|---|----|

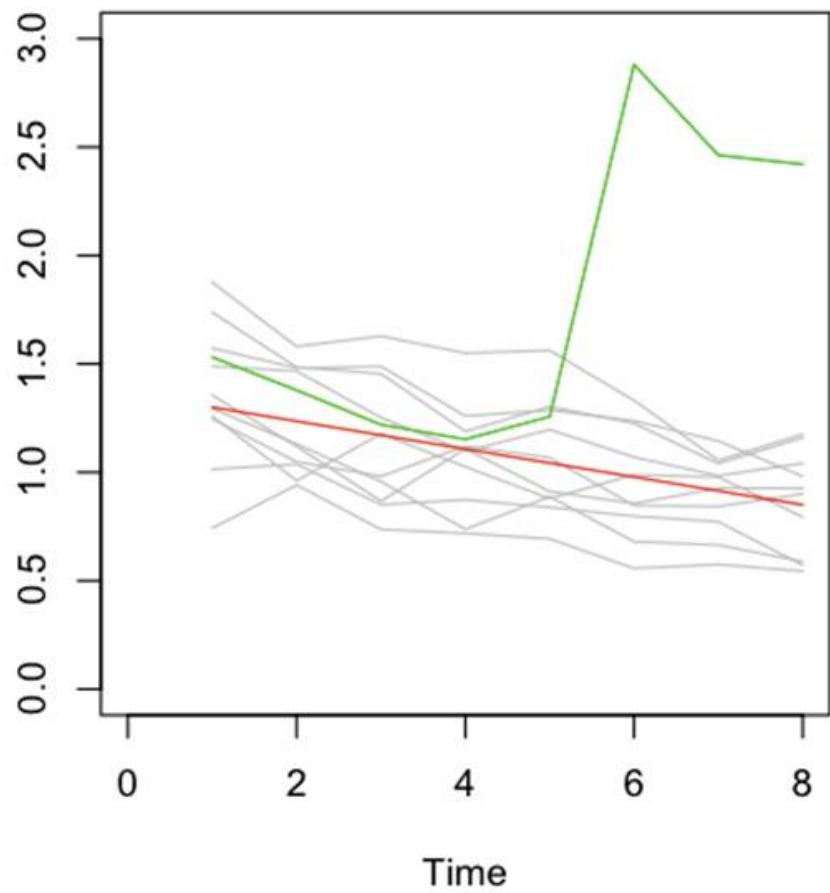
From space to space-time

Disease mapping: Extending space to space-time

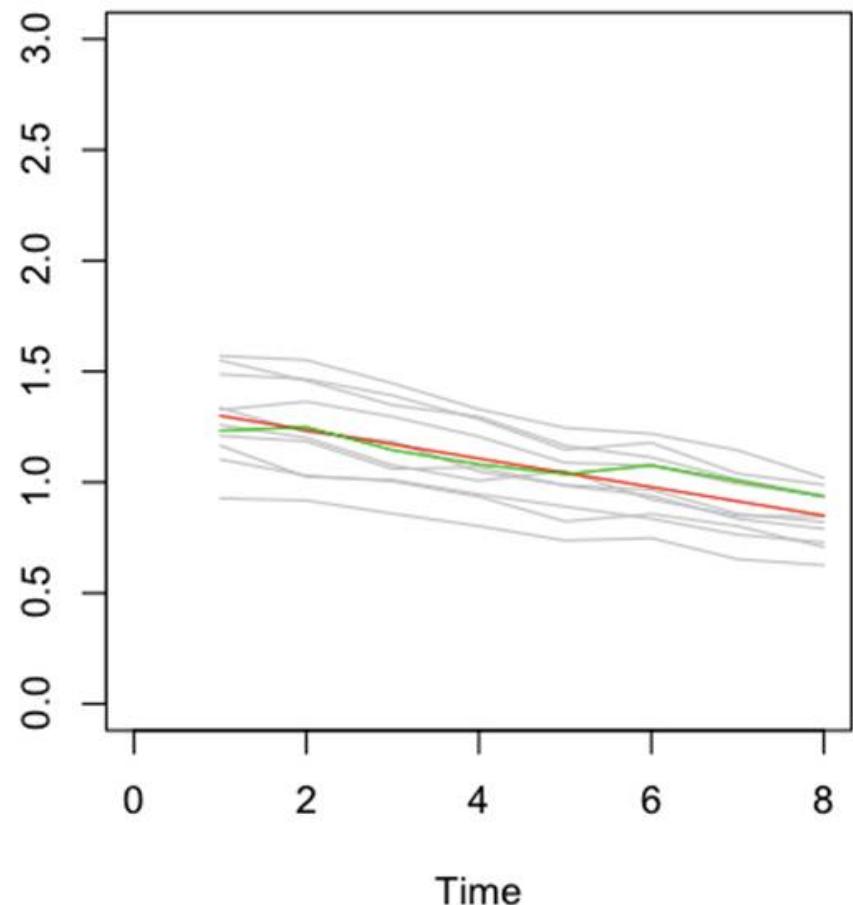
- Disease mapping is usually carried out on aggregated data over a time period
- Rather than suppressing the time dimension, it can be interesting to use models that combine the space and time dimension
- The stability (or not) of the spatial pattern can aid interpretation
- The specific space-time components of the model can potentially pinpoint unusual/emerging hazards
- Data:
 - y_{it} and
 - E_{it} : the observed and expected number of cases in area i at time t calculated as $E_{it} = \sum_k n_{itk} r_k$, where r_k reference rate for stratum (age, gender,...)

Schematic representation I

Raw SMR



Smoothed RR



Linear spatio-temporal model

- A simple parametric model assumes a linear effect of time:

$$\log \rho_{it} = b_0 + u_i + v_i + \beta \times t$$

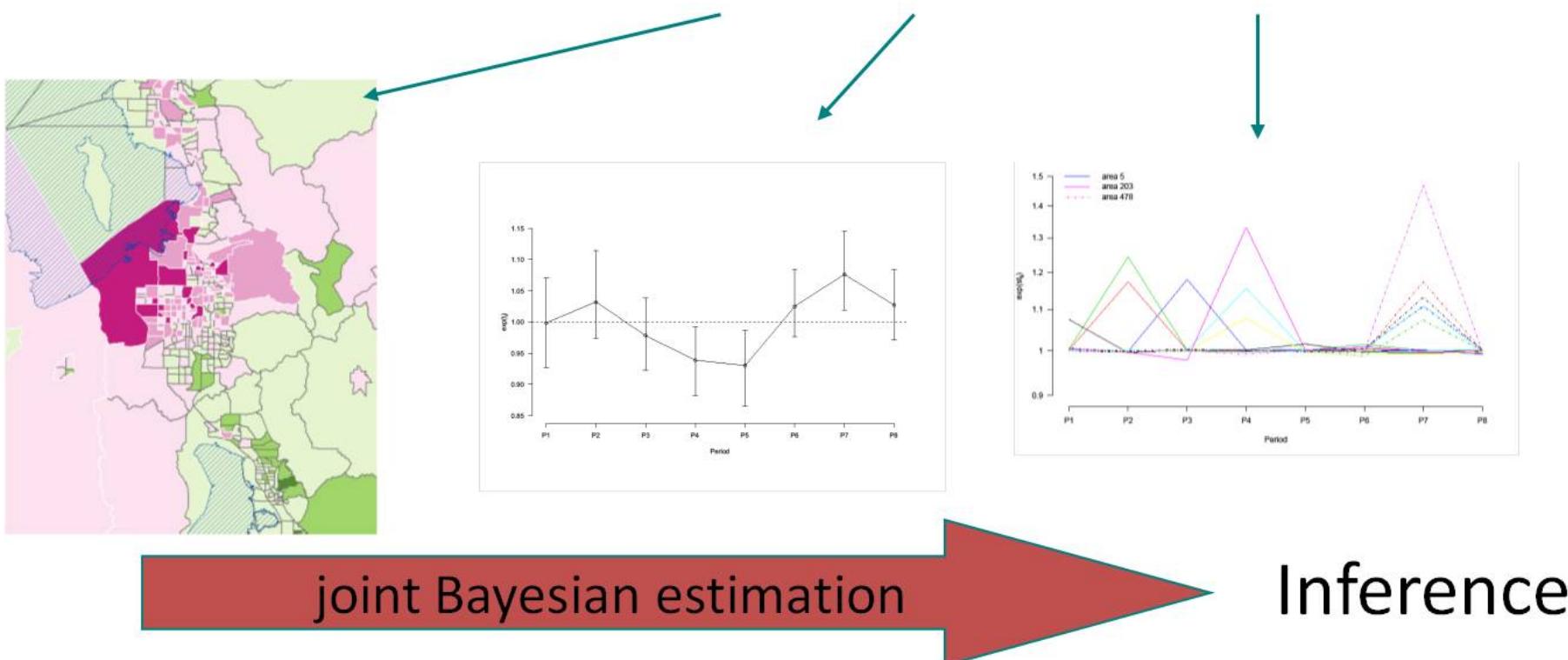
- Main spatial effect $u_i + v_i$
- Main linear trend β (global time effect)
- A differential effect δ_i can be added

$$\log \rho_{it} = b_0 + u_i + v_i + (\beta + \delta_i) \times t$$

- If $\delta_i < 0$ then the area-specific trend is less steep than the mean trend, whilst $\delta_i > 0$ implies that the area-specific trend is steeper than the mean trend.

Noise model: Poisson/Binomial

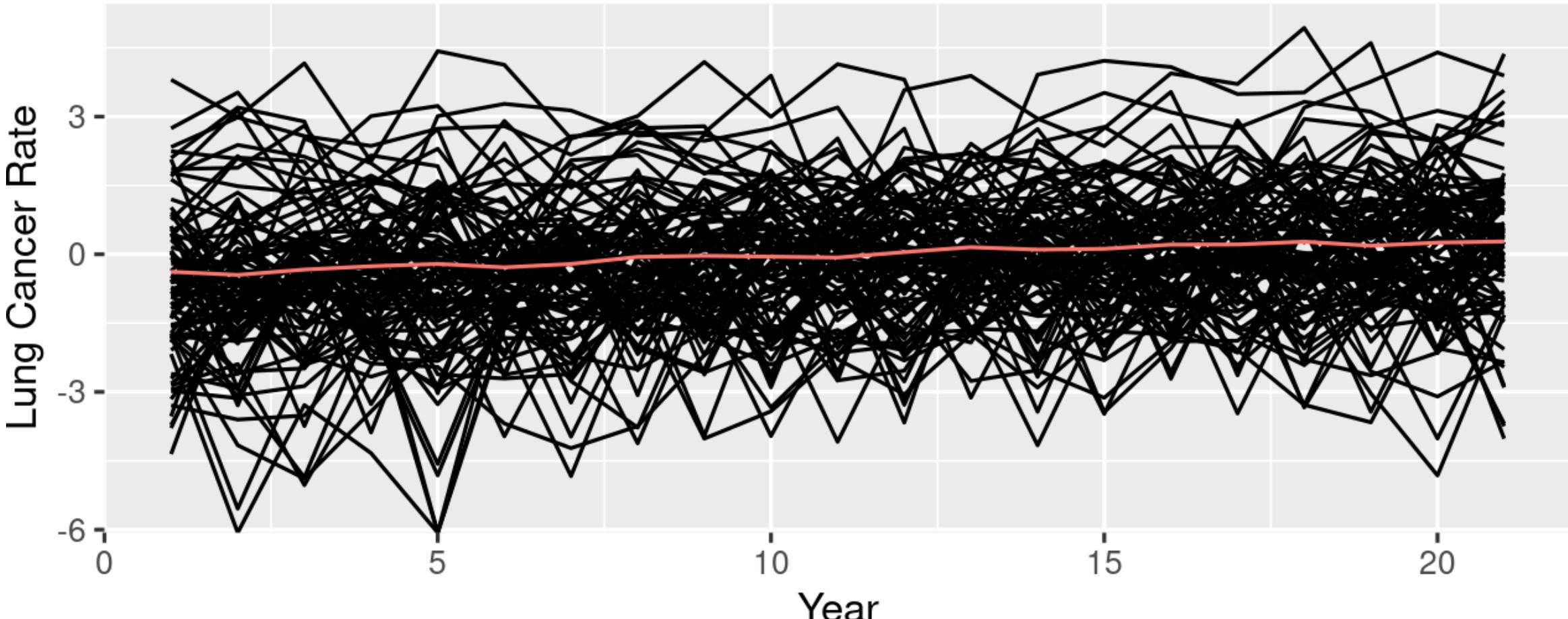
Latent structure: Space + Time + Interactions



Ohio Lung cancer - temporal SMRs

Log SMR time trends in each county

- Slightly increasing trend but lots of variation across the counties



Simple linear spatio-temporal model (Model 1)

$$y_{it} \sim \text{Poisson}(E_{it}\rho_{it})$$
$$\log\rho_{it} = b_0 + \beta * t + u_i + v_i$$

where

- b_0 overall log RR in Ohio over the 21-year period
- $v_i \sim \text{Normal}(0, \sigma_v^2)$ spatially unstructured RE
- $\mathbf{u} \sim \text{ICAR}(\mathbf{W}, \sigma_u^2)$ spatially structured RE
- $\exp(\beta)$ is the change in the RR associated with a 1-year increase in time

```
> formula.mod1 <- y ~ 1 + f(county, model="bym",
+                               graph=Ohio.adj, constr=TRUE) + year
```

Log-linear temporal model with unstructured temporal RE (Model 2)

$$y_{it} \sim \text{Poisson}(E_{it}\rho_{it})$$
$$\log \rho_{it} = b_0 + \beta * t + u_i + v_i + \psi_t$$

where

- b_0 overall log RR in Ohio over the 21-year period
- $v_i \sim \text{Normal}(0, \sigma_v^2)$ spatially unstructured RE
- $\mathbf{u} \sim \text{ICAR}(\mathbf{W}, \sigma_u^2)$ spatially structured RE
- $\exp(\beta)$ is the change in the RR associated with a 1-year increase in time
- $\psi_t \sim \text{Normal}(0, \sigma_\psi^2)$ temporally unstructured RE

```
> year2 <- ohio.data$year
> formula.mod2<- y ~ 1 + f(county,model="bym",
+                               graph=Ohio.adj, constr=TRUE) +
+                               f(year2,model="iid", constr=TRUE) +
+                               year
```

Ohio lung cancer - comparison models 1 and 2

```
> Ohio.adj=~"/Dropbox/Books/INLABook/Datasets/Ohio Lung Cancer/Ohio.graph"
> mod1<- inla(data=ohio.data,formula=formula.mod1, E=E, family="poisson")
> mod2<- inla(data=ohio.data,formula=formula.mod2, E=E, family="poisson")
>
> # Posterior mean and 95% CI for Model 1
> mod1$summary.fixed
```

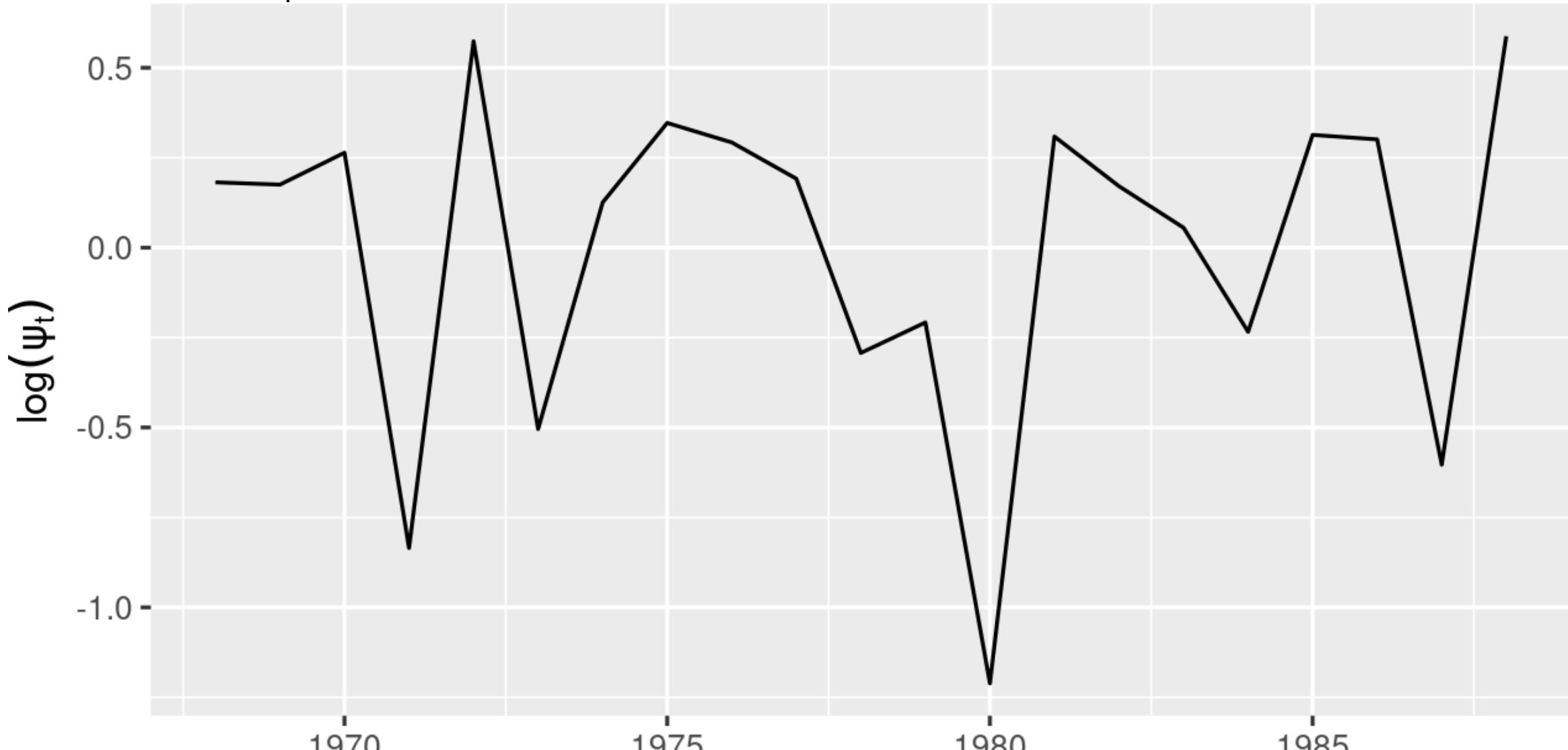
| | mean | sd | 0.025quant | 0.5quant | 0.975quant |
|-------------|-------------|--------------|-------------|-------------|-------------|
| (Intercept) | -0.98840916 | 0.1198545620 | -1.22430758 | -0.98839596 | -0.75283818 |
| year | 0.02618802 | 0.0005869893 | 0.02503579 | 0.02618792 | 0.02733977 |
| | mode | kld | | | |
| (Intercept) | -0.98835903 | 5.544091e-07 | | | |
| year | 0.02618777 | 3.834067e-07 | | | |

```
> #Posterior mean and 95% CI for Model 2
> mod2$summary.fixed
```

| | mean | sd | 0.025quant | 0.5quant | 0.975quant |
|-------------|-------------|--------------|-------------|-------------|-------------|
| (Intercept) | -1.00152385 | 0.22506963 | -1.44547777 | -1.00152150 | -0.55820692 |
| year | 0.02061841 | 0.01689569 | -0.01286484 | 0.02061805 | 0.05404213 |
| | mode | kld | | | |
| (Intercept) | -1.00149655 | 2.444035e-09 | | | |
| year | 0.02061894 | 1.060639e-09 | | | |

Ohio lung cancer - comparison models 1 and 2

Let's look at the temporal random effect in model 2



Simple additive space-time structure version 1 (Model 3)

This model assumes that the space time variations can be captured by the superimposition of a BYM spatial model and a structured time trend

$$\begin{aligned}y_{it} &\sim \text{Poisson}(E_{it}\rho_{it}) \\ \log\rho_{it} &= b_0 + b_i + \gamma_t\end{aligned}$$

- $b_i = v_i + u_i$
- $v_i \sim \text{Normal}(0, \sigma_v^2)$ spatially unstructured RE
- $\mathbf{u} \sim \text{ICAR}(\mathbf{W}, \sigma_u^2)$ spatially structured RE
- $\gamma_t \sim \text{RW}(1)$ temporally structured RE with variance parameter σ_γ^2
- Assuming a BYM2 specification, there are 3 hyperparameters
 - τ_b , the marginal precision of b_i
 - ψ , the proportion of spatial variability explained by the spatially structured component
 - σ_γ^2 , the conditional variance of the RW(1) modelling the time trend

```
> formula.mod3 <- y ~ 1 + f(county,model="bym",
+                               graph=Ohio.adj, constr=TRUE) +
+                               f(year,model="rw1", constr=TRUE)

> mod3<- inla(data=ohio.data,formula=formula.mod3, E=E, family="poisson")
```

Simple additive space-time structure version 2 (Model 4)

$$y_{it} \sim \text{Poisson}(E_{it}\rho_{it})$$
$$\log \rho_{it} = b_0 + u_i + v_i + \psi_t + \gamma_t$$

where

- b_0 overall log RR in Ohio over the 21-year period
- $v_i \sim \text{Normal}(0, \sigma_v^2)$ spatially unstructured RE
- $\mathbf{u} \sim \text{ICAR}(\mathbf{W}, \sigma_u^2)$ spatially structured RE
- $\gamma_t \sim \text{RW}(1)$ temporally structured RE with variance parameter σ_γ^2
- $\psi_t \sim \text{Normal}(0, \sigma_\psi^2)$ temporally unstructured RE}

```
> formula.mod4 <- y ~ 1 + f(county, model="bym",
+                               graph=Ohio.adj, constr=TRUE) +
+                               f(year, model="rw1", constr=TRUE) +
+                               f(year2, model="iid", constr=TRUE)
```

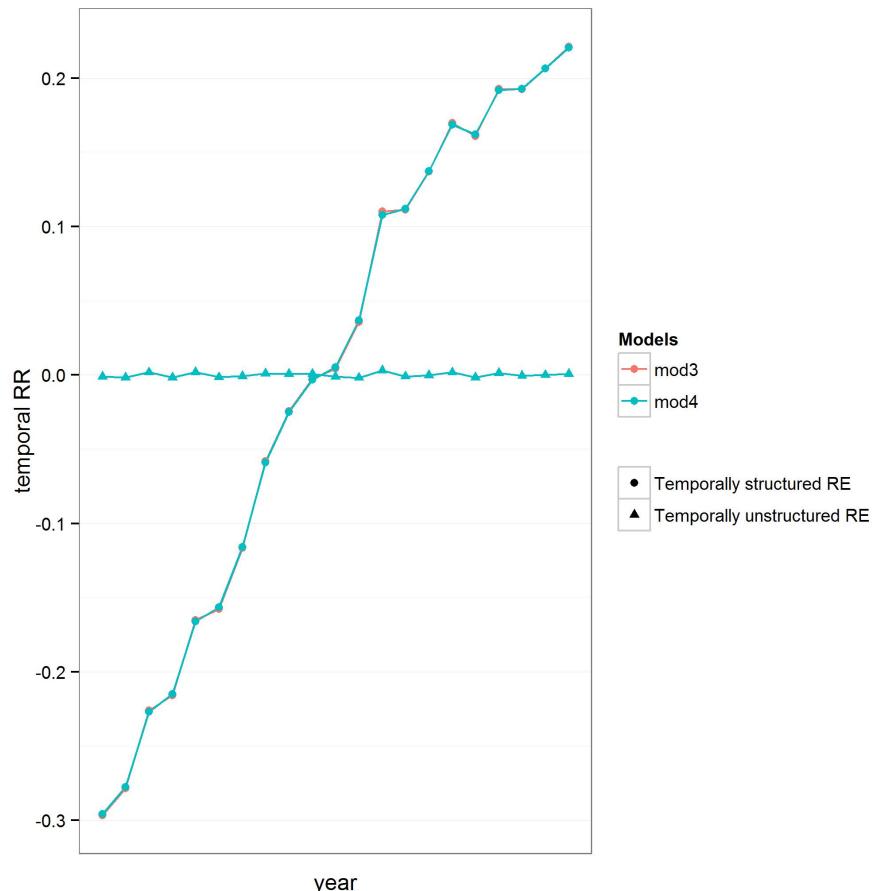
```
> mod4<- inla(data=ohio.data, formula=formula.mod4, E=E, family="poisson")
```

Ohio lung cancer - comparison models 3 and 4

Posterior means of the temporal RE

Model 3 = γ_t

Model 4 = $\psi_t + \gamma_t$

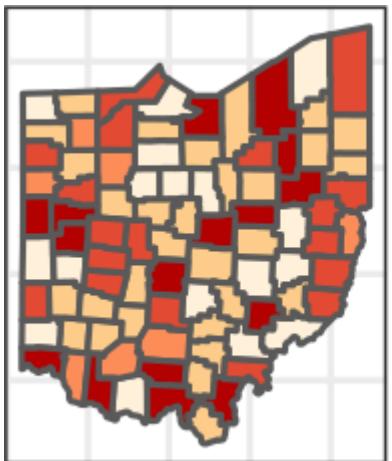


Ohio lung cancer - area-specific temporal RR

Posterior means of the temporal RE γ_t (model 3)

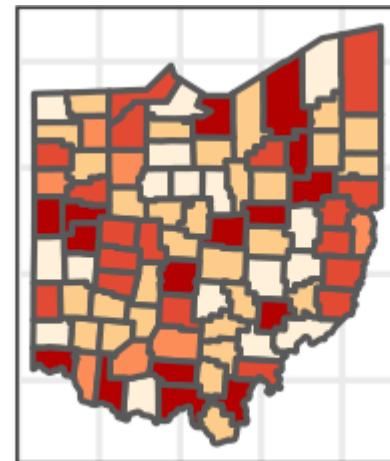
Ohio lung cancer - spatial patterns (all models)

Posterior means of the spatial RE ($u + v$) for the 4 different models



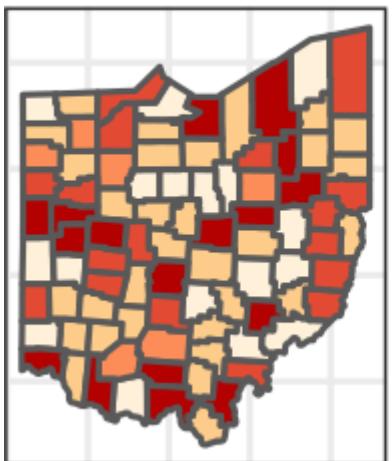
model 1

- [-4, -1]
- (-1, -0.1]
- (-0.1, 0.1]
- (0.1, 1]
- (1, 4]



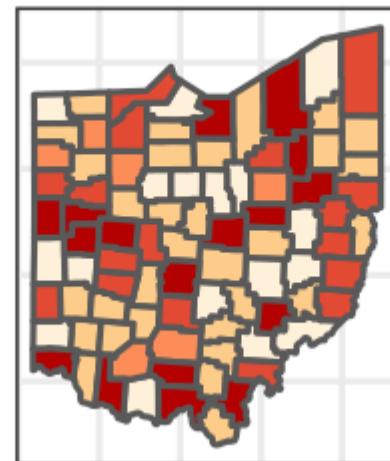
model 2

- [-4, -1]
- (-1, -0.1]
- (-0.1, 0.1]
- (0.1, 1]
- (1, 4]



model 3

- [-4, -1]
- (-1, -0.1]
- (-0.1, 0.1]
- (0.1, 1]
- (1, 4]



model 4

- [-4, -1]
- (-1, -0.1]
- (-0.1, 0.1]
- (0.1, 1]
- (1, 4]

Similar patterns with the 4 models

Space-time model with exchangeable interactions (Model 5)

$$y_{it} \sim \text{Poisson}(E_{it}\rho_{it})$$

$$\log \rho_{it} = b_0 + \theta_i + \gamma_t + \psi_t + \delta_{it}$$

$$\theta_i = v_i + u_i \text{ (BYM = ICAR + HET)}$$

$$\gamma_t \sim \text{RW}(1)$$

$$\psi_t \sim N(0, \sigma_\psi^2)$$

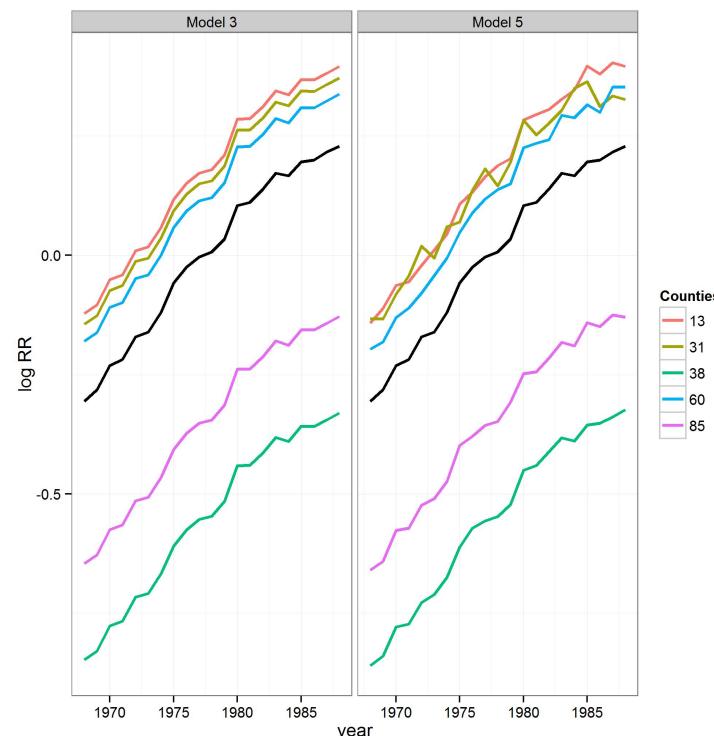
$$\delta_{it} \sim \text{Normal}(0, \sigma_\delta^2)$$

The simple additive model 4 can be extended by including **space-time interactions parameters**, δ_{it} , modelled as exchangeable random effects, that capture departure from the additive structure

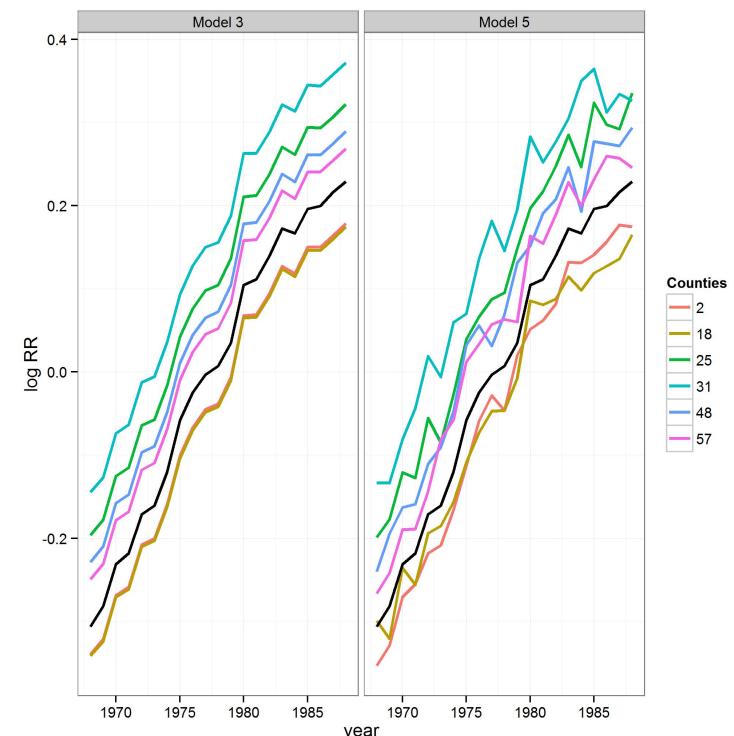
Interpretation of the interactions in model 5

- The interactions δ_{it} allow to highlight unusual temporal trends
- Rules based on the posterior probabilities $p(\delta_{it} > 0)$ for at least 1 time t
- Ohio lung cancer: 6 counties with unusual temporal trends

"usual" temporal trends



unusual temporal trends



INLA code for model with interaction (Model 5)

- In INLA is very easy to include the interaction in the formula environment
- We need to specify an index for the interaction, i.e. for each combination of area/time

$$\begin{aligned}\log \rho_{it} &= b_0 + \xi_i + \gamma_t + \delta_{it} \\ \xi_i &= v_i + u_i \text{ (BYM = ICAR + HET)} \\ \gamma_t &\sim RW(1) \\ \psi_t &\sim N(0, \sigma_\psi^2) \\ \delta_{it} &\sim \text{Normal}(0, \sigma_\nu^2)\end{aligned}$$

```
> formula.intI<- y ~ + f(county,model="bym",
+                         graph=Ohio.adj) +
+                         f(year,model="rw1") +
+                         f(year2,model="iid") +
+                         f(area.year,model="iid")
```

Types of interactions

Another example: Birth weight in Georgia

- Count of babies weighing less than 2500g in 159 counties in Georgia (US)
- Period 2000 – 2010



Different types of interactions (Knorr-Held, 2000)

$$\begin{aligned}y_{it} &\sim \text{Poisson}(E_{it}\rho_{it}) \\ \log \rho_{it} &= b_0 + v_i + u_i + \gamma_t + \psi_t + \delta_{it} \\ v_i + u_i &= (\text{BYM} = \text{HET} + \text{ICAR}) \\ \gamma_t &\sim \text{RW}(1) \\ \psi_t &\sim N(0, \sigma_\psi^2)\end{aligned}$$

Characteristics of ST interaction

Interaction Parameters Rank

| | | |
|-----|------------------|--|
| I | v and ψ | nT |
| II | v and γ | $n(T-1)$ for RW1, $n(T-2)$ for RW2 |
| III | u and ψ | $(n-1)T$ |
| IV | u and γ | $(n-1)(T-1)$ for RW1, $(n-1)(T-2)$ for RW2 |

How to model interactions

- Data in R format

```
# A tibble: 1,749 × 8
  ID.area   Obs ID.year   Exp      y      E NAME    ID.area.year
  <int> <int> <dbl> <dbl> <int> <dbl> <chr>    <int>
1     1     20     1  25.8    20  25.8 Appling      1
2     1     24     2  25.1    24  25.1 Appling      2
3     1     25     3  22.8    25  22.8 Appling      3
4     1     31     4  25.2    31  25.2 Appling      4
5     1     24     5  24.9    24  24.9 Appling      5
6     1     40     6  26.7    40  26.7 Appling      6
7     1     29     7  26.0    29  26.0 Appling      7
8     1     35     8  27.2    35  27.2 Appling      8
9     1     26     9  25.1    26  25.1 Appling      9
10    1     25    10  24.5    25  24.5 Appling     10
# ... with 1,739 more rows
```

- We need to make sure to have an index for area (ID.area), one for time (ID.year) and one for the interaction (ID.area.year)

Type I interaction in INLA

Let's start with the Type I interaction, which assumes that the two unstructured effects v_i and ψ_t interact

```
> Georgia.adj <- "~/LBW in Georgia/Georgia.graph"  
> formula.intI<- y ~ + f(ID.area,model="bym2",  
+                         graph=Georgia.adj) +  
+                         f(ID.year,model="rw1") +  
+                         f(ID.area.year,model="iid")
```



Kronecker product

- For the interactions of type II-IV we will need to use the **Kronecker product** to specify the dependencies. It is a matrix multiplication which returns a block matrix.

For instance

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \otimes \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} & 2 \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} \\ 3 \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} & 4 \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} \end{bmatrix}$$

There is a function in R which does exactly that: `kronecker(a,b)`

We will now see how to use this for obtaining the structure matrix for the different types of interactions

If you are interested in knowing more how to set interactions using the Kronecker product in INLA see Goicoa, Adin, Ugarte, and Hodges (2018)

Type II interaction: in INLA

- Type II combines the structured temporal main effect γ_t and the unstructured spatial effect v_i .
- We use first create the structure matrix for the time component assuming a random walk of order 1

```
> #RW1
> D1 <- diff(diag(11), differences=1)
> Q.gammaRW1 <- t(D1) %*% D1
> #Let's look at it
> Q.gammaRW1[1:5,1:5]
```

```
 [,1] [,2] [,3] [,4] [,5]
[1,]    1   -1    0    0    0
[2,]   -1    2   -1    0    0
[3,]    0   -1    2   -1    0
[4,]    0    0   -1    2   -1
[5,]    0    0    0   -1    2
```

Type II interaction: in INLA

- Then we create the Kronecker product $v \otimes \gamma$ and note from the table in slide 35 that the rank deficiency is n (159 in our case)

```
> #Kronecker product (for RW1)
> R <- kronecker(diag(159),Q.gammaRW1)
> R[1:5,1:5]
```

```
 [,1] [,2] [,3] [,4] [,5]
[1,]    1   -1    0    0    0
[2,]   -1    2   -1    0    0
[3,]    0   -1    2   -1    0
[4,]    0    0   -1    2   -1
[5,]    0    0    0   -1    2
```

```
> r.def <- 159
```

Type II interaction: in INLA

- Finally we need to impose some more constraints as the matrix have spatio-temporal interaction matrix has a rank deficiency
- We define the `A.constr` matrix as that where the rows are the eigenvectors of the structure matrix of the interaction term which span the null space

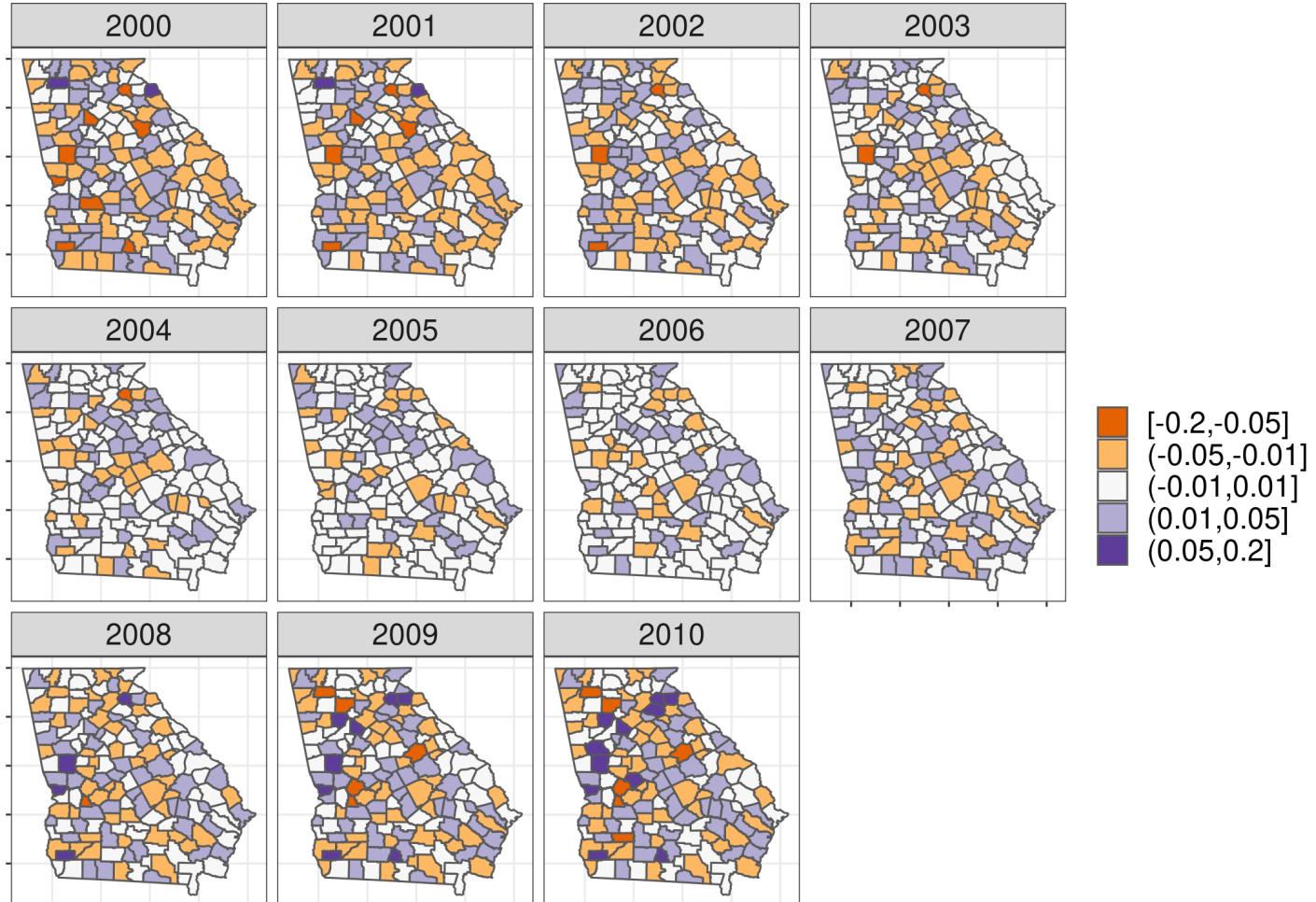
```
> A.constr <- kronecker(diag(159),matrix(1,1,11))
> A.constr <- A.constr[-1,]
> dim(A.constr)
```

```
[1] 158 1749
```

The formula environment needs to be changed to

```
> formula.intII<- y ~ f(ID.area,model="bym2",graph=Georgia.adj) +
+           f(ID.year,model="rw1") +
+           f(ID.area.year, model="generic0", Cmatrix=R, rankdef=r.def,
+               constr=TRUE,
+               extraconstr=list(A=A.constr, e=rep(0,159-1)))
```

Type II interaction: in INLA



Type III interaction in INLA

- Type III combines the spatially structured main effect u_i and the unstructured temporal effect ψ_t .
- We use first create the structure matrix for the spatial component using the neighborhood graph

```
> g <- inla.read.graph(Georgia.adj)
> Q.xi <- matrix(0, g$n, g$n)
> for (i in 1:g$n){
+   Q.xi[i,i]=g$nnbs[[i]]
+   Q.xi[i,g$nnbs[[i]]]=-1
+ }
> Q.xi[1:5,1:5]
```

| | [,1] | [,2] | [,3] | [,4] | [,5] |
|-------|------|------|------|------|------|
| [1,] | 2 | -1 | 0 | 0 | 0 |
| [2,] | -1 | 4 | 0 | 0 | 0 |
| [3,] | 0 | 0 | 5 | -1 | 0 |
| [4,] | 0 | 0 | -1 | 4 | -1 |
| [5,] | 0 | 0 | 0 | -1 | 4 |

Type III interaction in INLA

- Then we create the Kronecker product $u \otimes \psi$ and note from the table in slide 35 that the rank deficiency is T (11 in our case)

```
> #Kronecker product (for RW1)
> R <- kronecker(Q.xi,diag(11))
> R[1:5,1:15]
```

```
 [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13] [,14]
[1,] 2 0 0 0 0 0 0 0 0 0 0 -1 0 0
[2,] 0 2 0 0 0 0 0 0 0 0 0 0 -1 0
[3,] 0 0 2 0 0 0 0 0 0 0 0 0 0 -1
[4,] 0 0 0 2 0 0 0 0 0 0 0 0 0 0
[5,] 0 0 0 0 2 0 0 0 0 0 0 0 0 0
 [,15]
[1,] 0
[2,] 0
[3,] 0
[4,] -1
[5,] 0
```

```
> r.def <- 11
```

Type III interaction in INLA

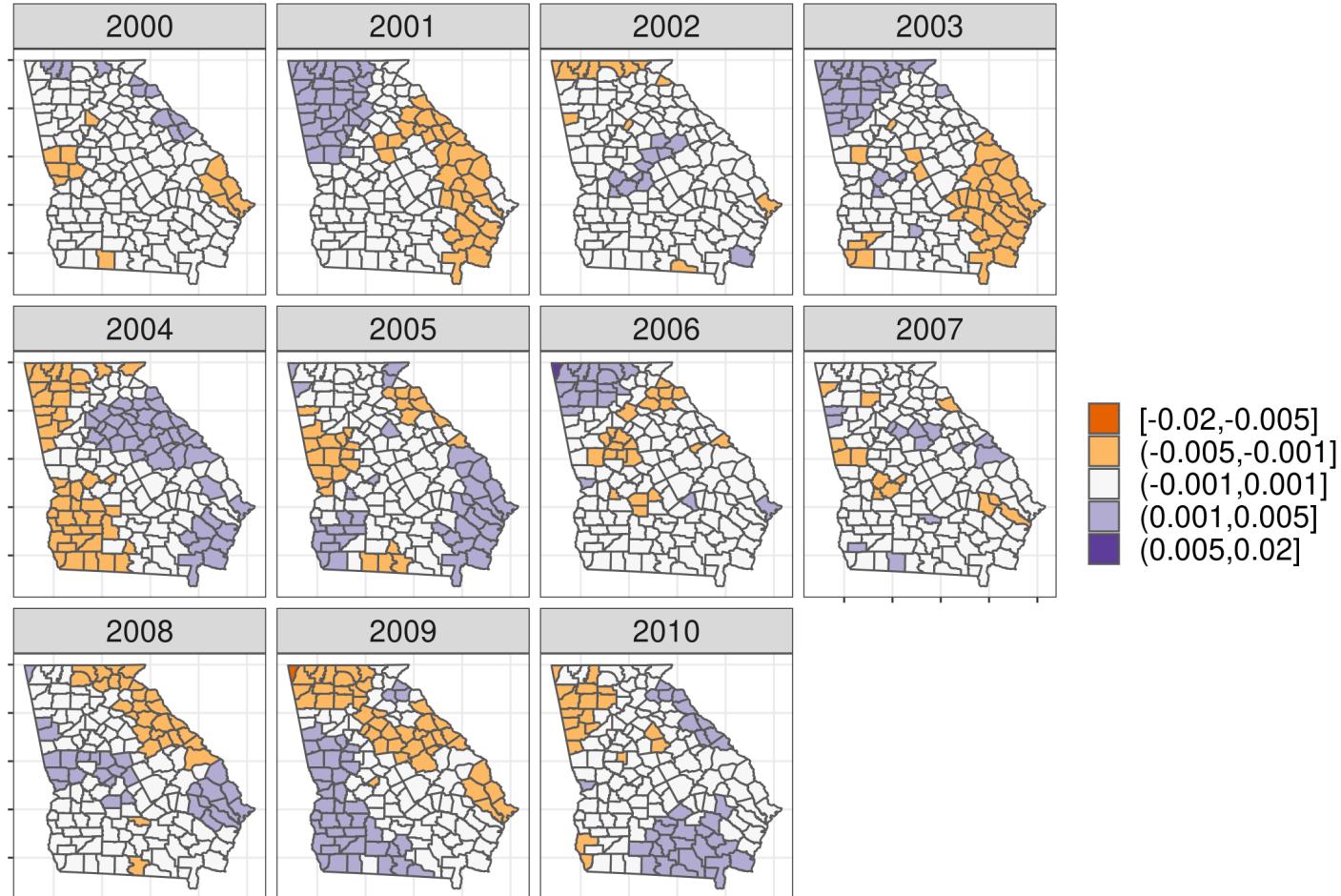
- Finally we need to impose some more constraints (sum-to-zero)

```
> A.constr <- kronecker(matrix(1,1,159),diag(11))
> A.constr <- A.constr[-1,]
```

The formula environment becomes

```
> formula.intIII<- y ~ f(ID.area,model="bym2",graph=Georgia.adj) +
+           f(ID.year,model="rw1") +
+           f(ID.area.year, model="generic0", Cmatrix=R, rankdef=r.def,
+           constr=TRUE,
+           extraconstr=list(A=A.constr, e=rep(0,11-1)))
```

Type III interaction: in INLA



Type IV interaction in INLA

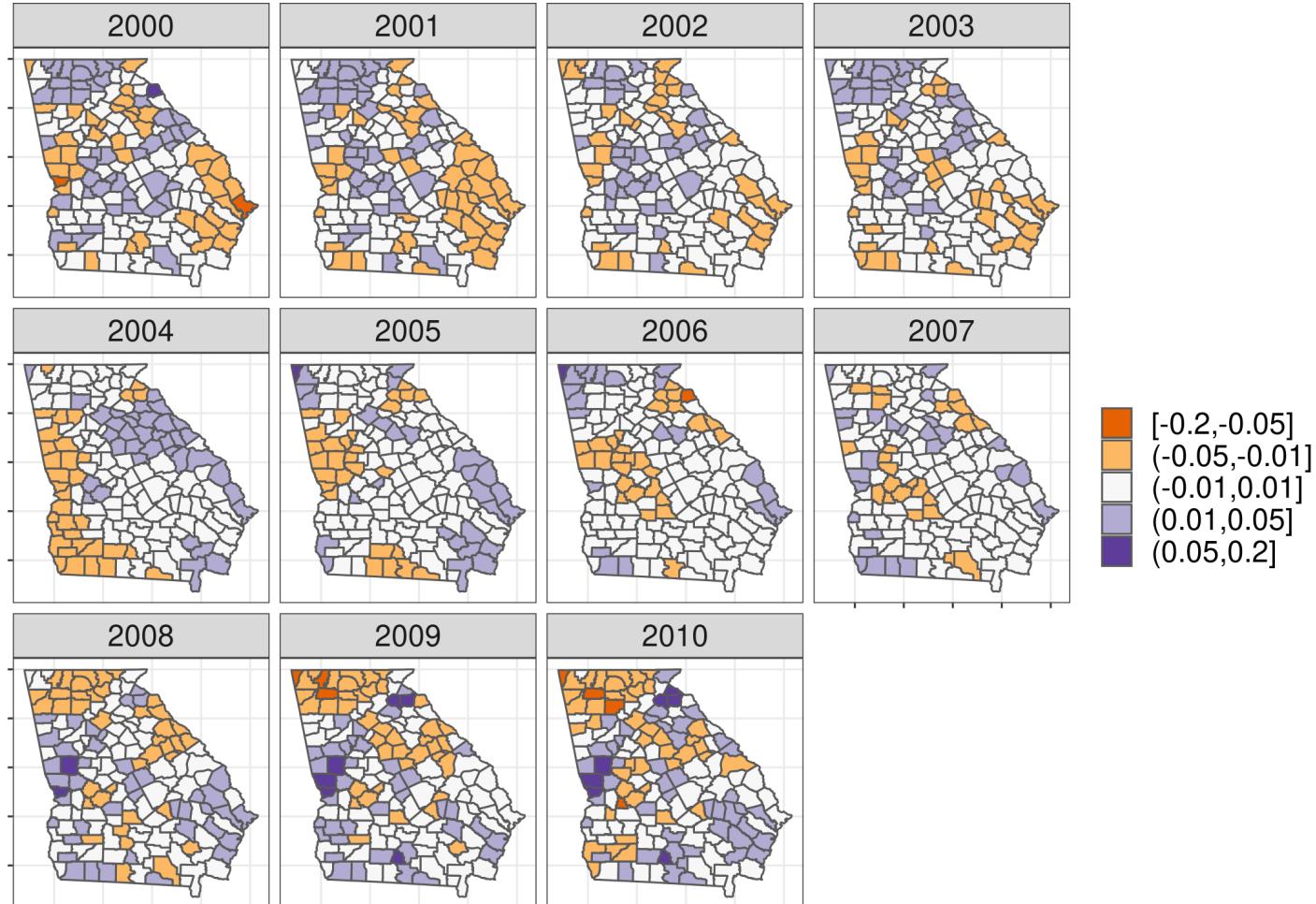
- Type IV is the most complex type of interaction, assuming that the spatially and temporally structured effects u_i and γ_t interact
- We assume that the temporal dependency structure for each area is not independent from all the other areas anymore, but depends on the temporal pattern of the neighboring areas as well
- We create the Kronecker product $u \otimes \gamma$ and note from the table in slide 35 that the rank deficiency is $n+T-1$ (169 in our case)

```
> #Kronecker product (for RW1)
> R <- kronecker(Q.xi,Q.gammaRW1)
> r.def <- 159+11-1
> A1 <- kronecker(diag(159),matrix(1,1,11))
> A2 <- kronecker(matrix(1,1,159),diag(11))
> A.constr <- rbind(A1[-1,],A2[-1,])
```

The formula environment becomes

```
> formula.intIV<- y ~ f(ID.area,model="bym2",graph=Georgia.adj) +
+           f(ID.year,model="rw1") +
+           f(ID.area.year, model="generic0", Cmatrix=R, rankdef=r.def,
+               constr=TRUE,
+               extraconstr=list(A=A.constr, e=rep(0,159+11-2)))
```

Type IV interaction: in INLA



Model selection

It might be useful to employ indexes like DIC and WAIC to decide which model is the most appropriate for the problem at hand

In this example:

```
> DIC <- c(11570,11536,11615,11570)
> WAIC <-c(11602,11569,11659,11612)
> DF$DIC<- DIC
> DF$WAIC <- WAIC
>
> kable(DF, col.names = c("Interaction", "Parameters", "Rank", "DIC", "WAIC"), escape=F, caption = "Model f
+   kable_styling(latex_options = "hold_position")
```

Model fitting for the different ST models

| Interaction | Parameters | Rank | DIC | WAIC |
|-------------|------------------|--|-------|-------|
| I | v and ψ | nT | 11570 | 11602 |
| II | v and γ | $n(T-1)$ for RW1, $n(T-2)$ for RW2 | 11536 | 11569 |
| III | u and ψ | $(n-1)T$ | 11615 | 11659 |
| IV | u and γ | $(n-1)(T-1)$ for RW1, $(n-1)(T-2)$ for RW2 | 11570 | 11612 |

So we conclude that the model with the interaction of order II is the most appropriate for this data.

Summary

- Increase quality of datasets that are both spatially and temporally indexed
- Advanced methods to deal with this type of data
- Allow to interpret the stability (or not) of the spatial patterns

Summary

- Increase quality of datasets that are both spatially and temporally indexed
- Advanced methods to deal with this type of data
- Allow to interpret the stability (or not) of the spatial patterns
- Model selection tools are useful to choose which interaction to use
- Careful with interactions of order above 1 as they can take a **very long** time to run

References

- Goicoa, T., A. Adin, M. Ugarte, et al. (2018). "In spatio-temporal disease mapping models, identifiability constraints affect PQL and INLA results". In: *Stochastic Environmental Research and Risk Assessment* 32.3, pp. 749-770.
- Knorr-Held, L. (2000). "Bayesian modelling of inseparable space-time variation in disease risk". In: *Statistics in medicine* 19.17-18, pp. 2555-2567.