

## Session 9.2: Spatio-temporal modelling for area data

Imperial College London

# Learning Objectives

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

- Explain how to extend spatial or temporal models to spatio-temporal models
- Fit Bayesian space-time models using R-INLA

The topics covered 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. PM<sub>10</sub>, 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  $y_i$  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

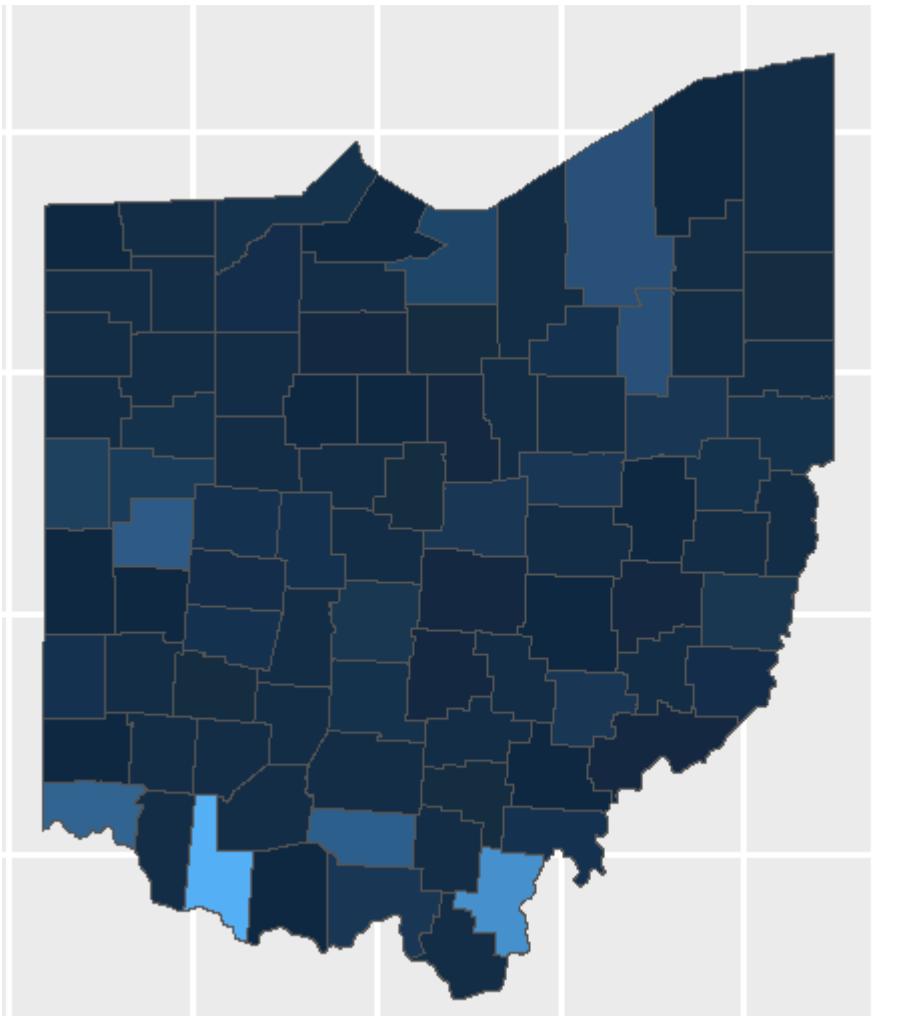
$$y_i \sim \text{Poisson}(E_i \rho_i); \quad i = 1, \dots, N$$

$$\eta_i = \log \rho_i = b_0 + b_i$$

$$\mathbf{b} = \frac{1}{\sqrt{\tau_b}} (\sqrt{1 - \phi} \mathbf{v}_* + \sqrt{\phi} \mathbf{u}_*)$$

# Ohio Lung cancer spatial pattern

- Data on lung cancer in 88 counties of Ohio, 1968-1988
- Purely spatial analysis



# 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,...)

Note that  $E_t$  makes sense if the time series covers a long period, e.g. years, otherwise it is common to assign  $E_t = \sum_t y_t / T$

## 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,...)

Note that  $E_t$  makes sense if the time series covers a long period, e.g. years, otherwise it is common to assign  $E_t = \sum_t y_t / T$

## 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,...)

Note that  $E_t$  makes sense if the time series covers a long period, e.g. years, otherwise it is common to assign  $E_t = \sum_t y_t / T$

OR

## 2. Temporal trends

$$\begin{aligned}y_t &\sim \text{Poisson}(E_t \rho_t); \quad t = 1, \dots, T \\ \log \rho_t &= b_0 + \psi_t \quad \text{global temporal smoothing} \\ \psi_t &\sim \text{Normal}(0, \sigma_\psi^2)\end{aligned}$$

# 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,...)

Note that  $E_t$  makes sense if the time series covers a long period, e.g. years, otherwise it is common to assign  $E_t = \sum_t y_t / T$

OR

## 2. Temporal trends

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

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

$$\gamma_t \sim RW(\sigma_\gamma^2)$$

# 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,...)

Note that  $E_t$  makes sense if the time series covers a long period, e.g. years, otherwise it is common to assign  $E_t = \sum_t y_t / T$

OR

## 2. Temporal trends

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

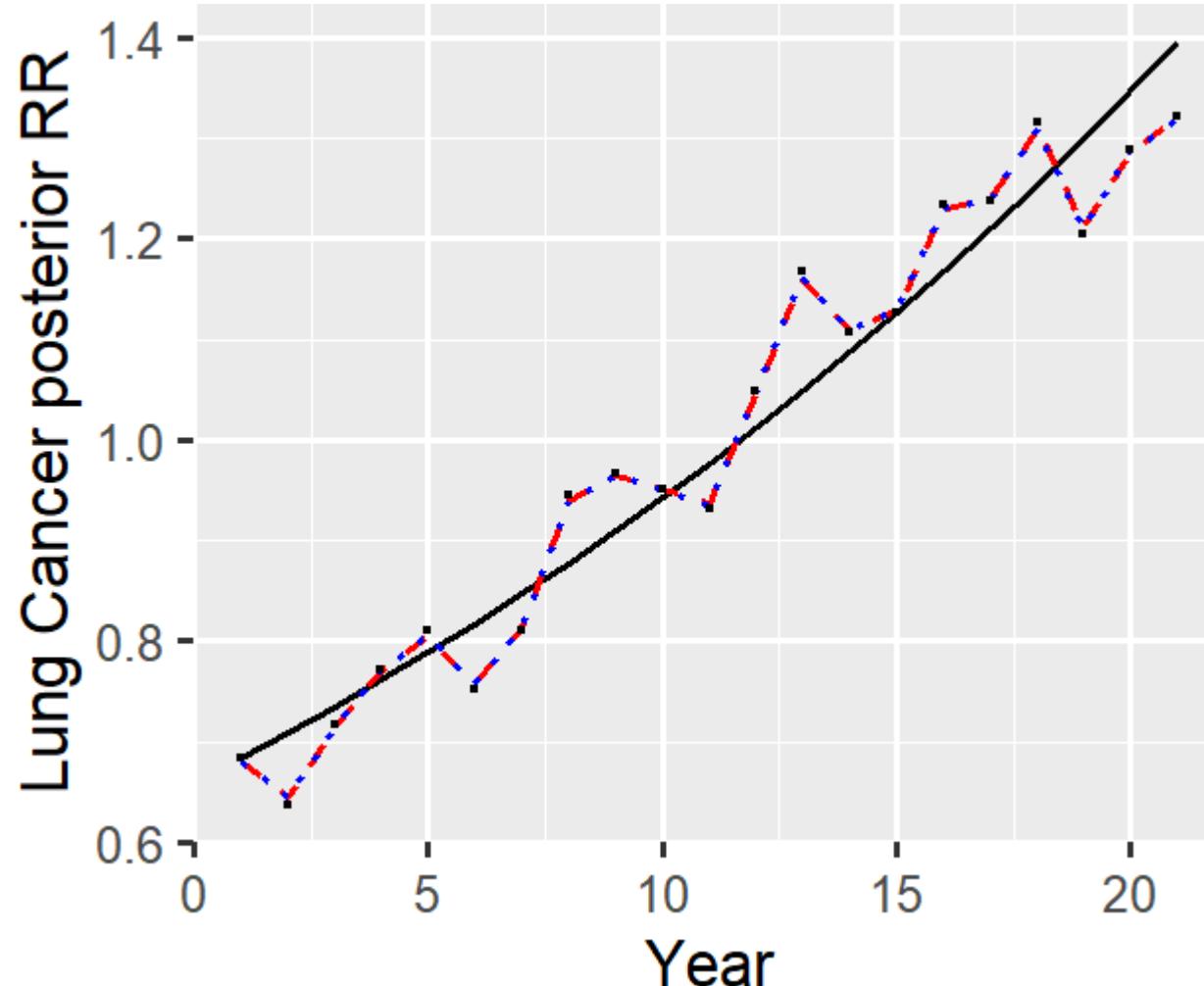
$$\log \rho_t = b_0 + \gamma_t + \psi_t \quad \text{global and local temporal smoothing}$$

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

$$\gamma_t \sim RW(\sigma_\gamma^2)$$

# Ohio Lung cancer data: modelled temporal trend

- Annual rates adjusted by gender and race
- Fitting different temporal trends (linear, local smoothing, local+global smoothing)



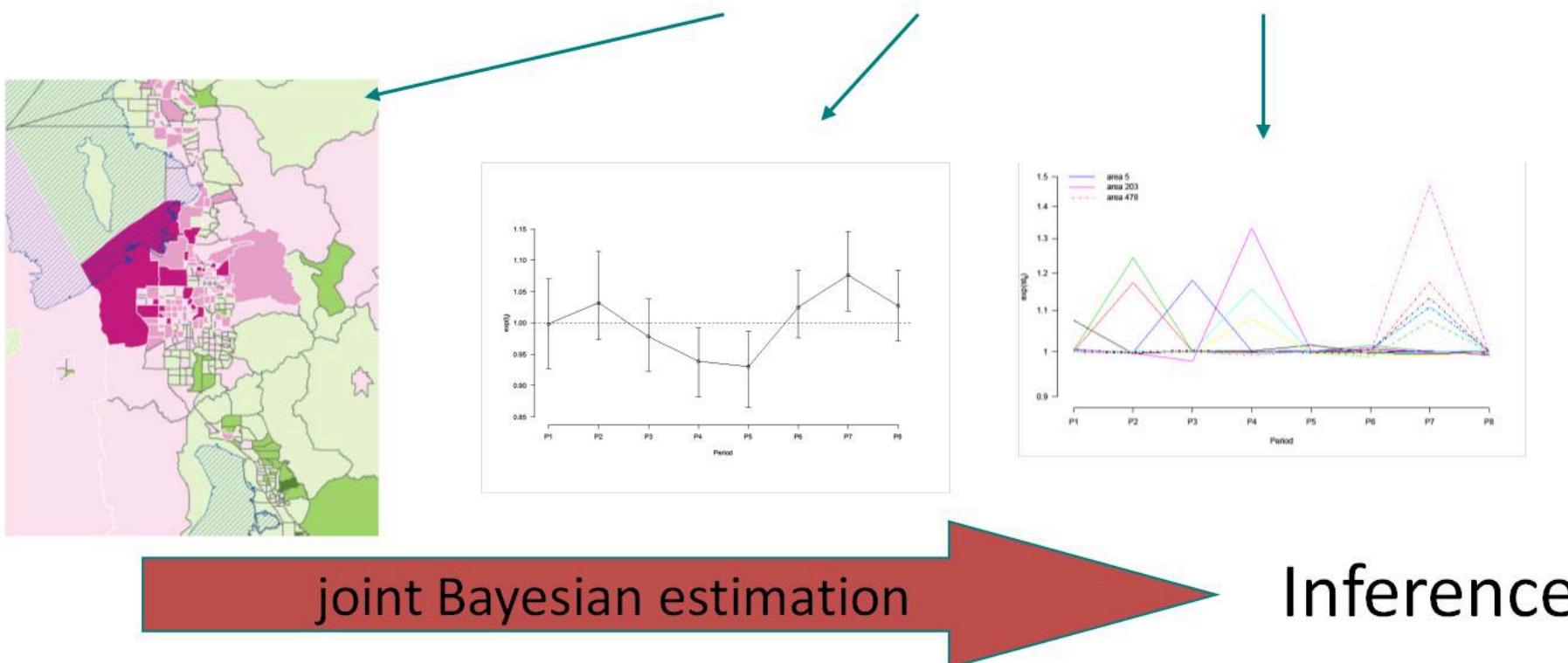
# From space to space-time

# 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:
  - $y_{it}$
  - $E_{it}$ : 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,...)

## Noise model: Poisson/Binomial

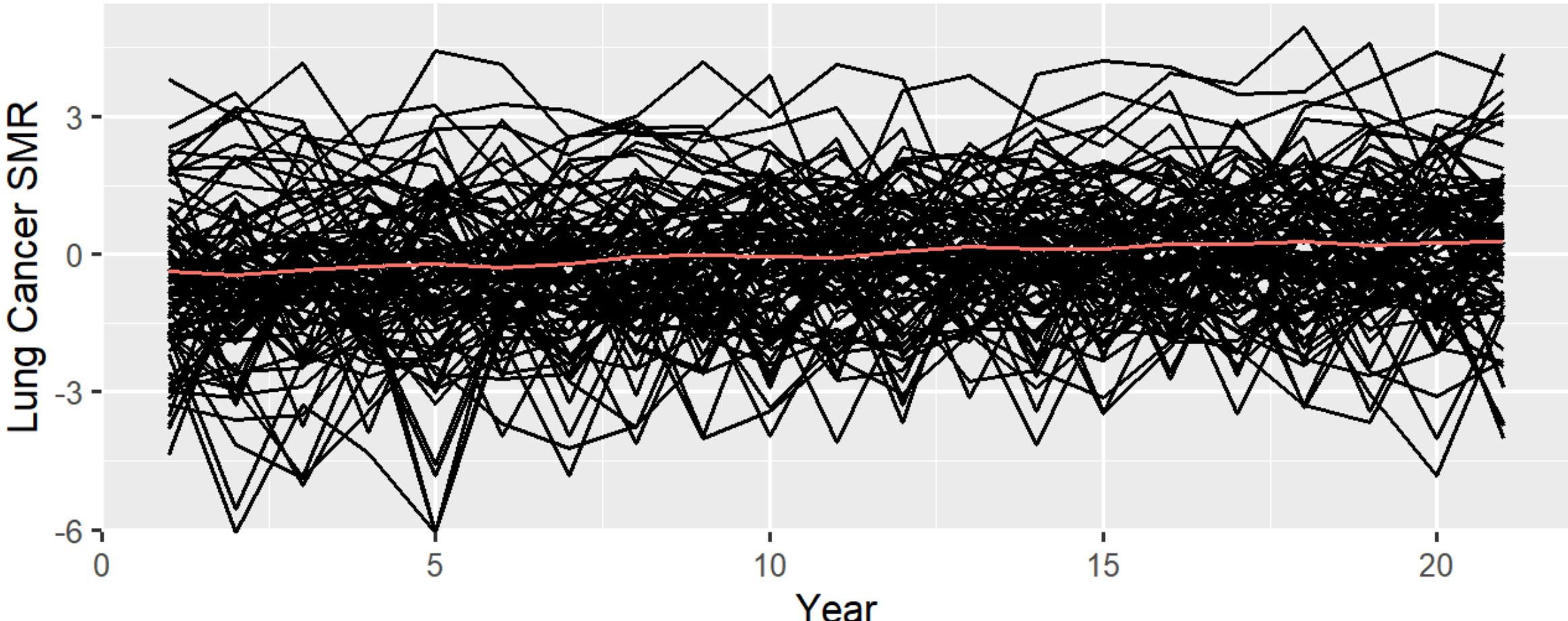
Latent structure: Space + Time + Interactions



# Ohio Lung cancer - temporal SMRs

Log SMR time trends in each county

- Slightly increasing trend but lots of variation across the counties



## Model 1: linear time + space

$$y_{it} \sim \text{Poisson}(E_{it}\rho_{it})$$
$$\log\rho_{it} = b_0 + \beta \times t + b_i$$

where

- $b_0$  overall log RR in Ohio over the 21-year period
- $b_i$  is the weighted average of spatially structured (ICAR) and unstructured random effects
- $\exp(\beta)$  is the change in the RR associated with a 1-year increase in time

```
> formula.mod1 = y ~ 1 + year + f(county, model="bym2",
+                                     graph=Ohio.adj)
```

```
> mod1 = inla(data=ohio.data, formula=formula.mod1, E=E, family="poisson", control.compute=list(waic=TRUE))
```



Additional temporal random effects can be added...

## Model 2: linear time + space + structured time

$$y_{it} \sim \text{Poisson}(E_{it}\rho_{it})$$
$$\log\rho_{it} = b_0 + \beta \times t + b_i + \gamma_t$$

where

- $b_0$  overall log RR in Ohio over the 21-year period
- $b_i$  is the weighted average of spatially structured (ICAR) and unstructured random effects
- $\exp(\beta)$  is the change in the RR associated with a 1-year increase in time
- Looking back at the purely temporal analysis it seems a random walk was able to explain the local changes in time  $\gamma_t \sim \text{Normal}(\gamma_{t-1}, \sigma_\gamma^2)$

```
> year2<-ohio.data$year  
> formula.mod2 = y ~ 1 +  
+                 year + f(county,model="bym2",  
+                 graph=Ohio.adj) +  
+                 f(year2,model="rw1")
```

```
> mod2 = inla(data=ohio.data,formula=formula.mod2, E=E, family="poisson", control.compute=list(waic=TRUE))
```



## Comparing model 1 and 2: Fixed effects

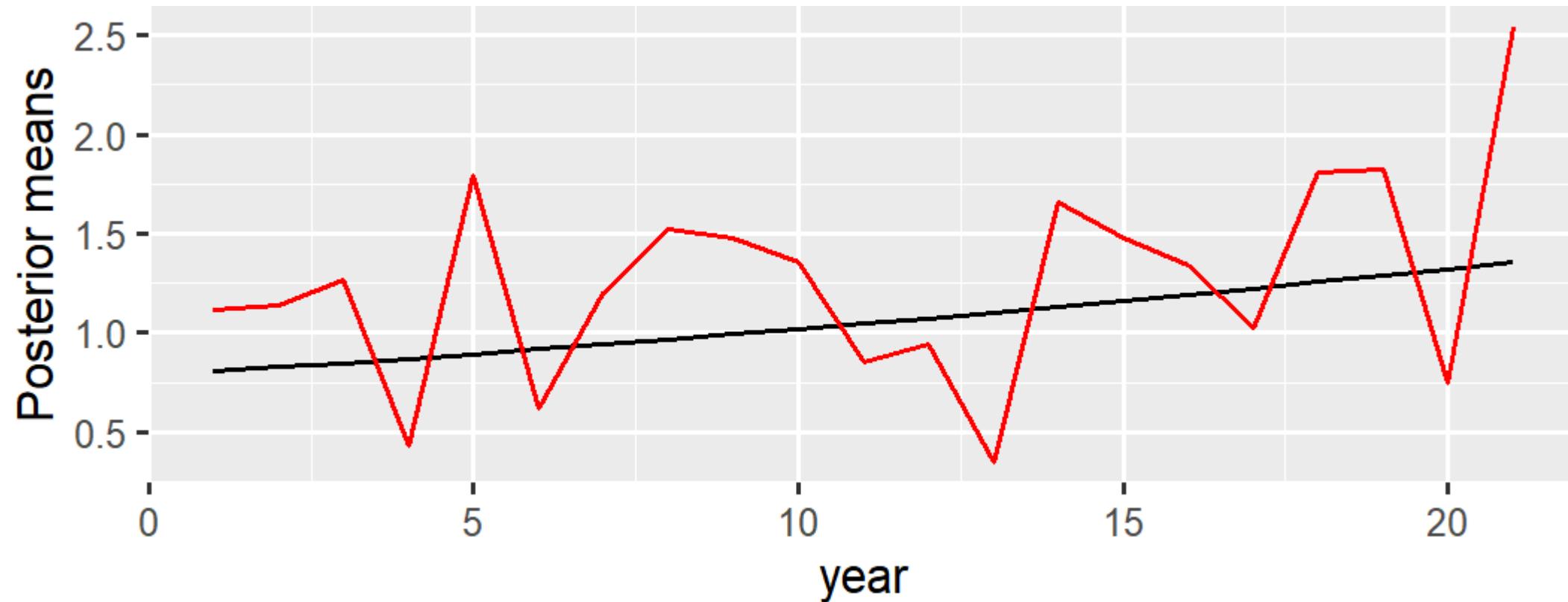
```
> mod1$summary.fixed
```

	mean	sd	0.025quant	0.5quant	0.975quant	mode	kld
(Intercept)	-0.98833298	0.11238994	-1.20942022	-0.98831790	-0.76733070	-0.98831791	1.473316e-08
year	0.02612109	0.00058628	0.02497146	0.02612109	0.02727073	0.02612109	5.527834e-11

```
> mod2$summary.fixed
```

	mean	sd	0.025quant	0.5quant	0.975quant	mode	kld
(Intercept)	-1.22479528	1.8557469	-4.8996262	-1.22480743	2.450109	-1.22480615	9.030057e-08
year	0.04083632	0.1683598	-0.2925939	0.04083756	0.374259	0.04083742	9.099081e-08

## Comparing model 1 and 2: Fitted values



## Model 3: space + structured time + unstructured time

- Exploring if a combination of random effects could be more flexible in capturing the temporal trend than a linear trend

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

where

- $b_0$  overall log RR in Ohio over the 21-year period
- $b_i$  is the weighted average of spatially structured (ICAR) and unstructured random effects
- The temporal pattern is now modelled with two random effects (global + local smoothing)

```
> formula.mod3 = y ~ 1 +  
+           f(county, model="bym2", graph=ohio.adj) +  
+           f(year, model="rw1") +  
+           f(year2, model="iid")
```

```
> mod3 = inla(data=ohio.data, formula=formula.mod3, E=E, family="poisson", control.compute=list(waic=TRUE))
```



## Comparing model 1 and 3: Fixed effects

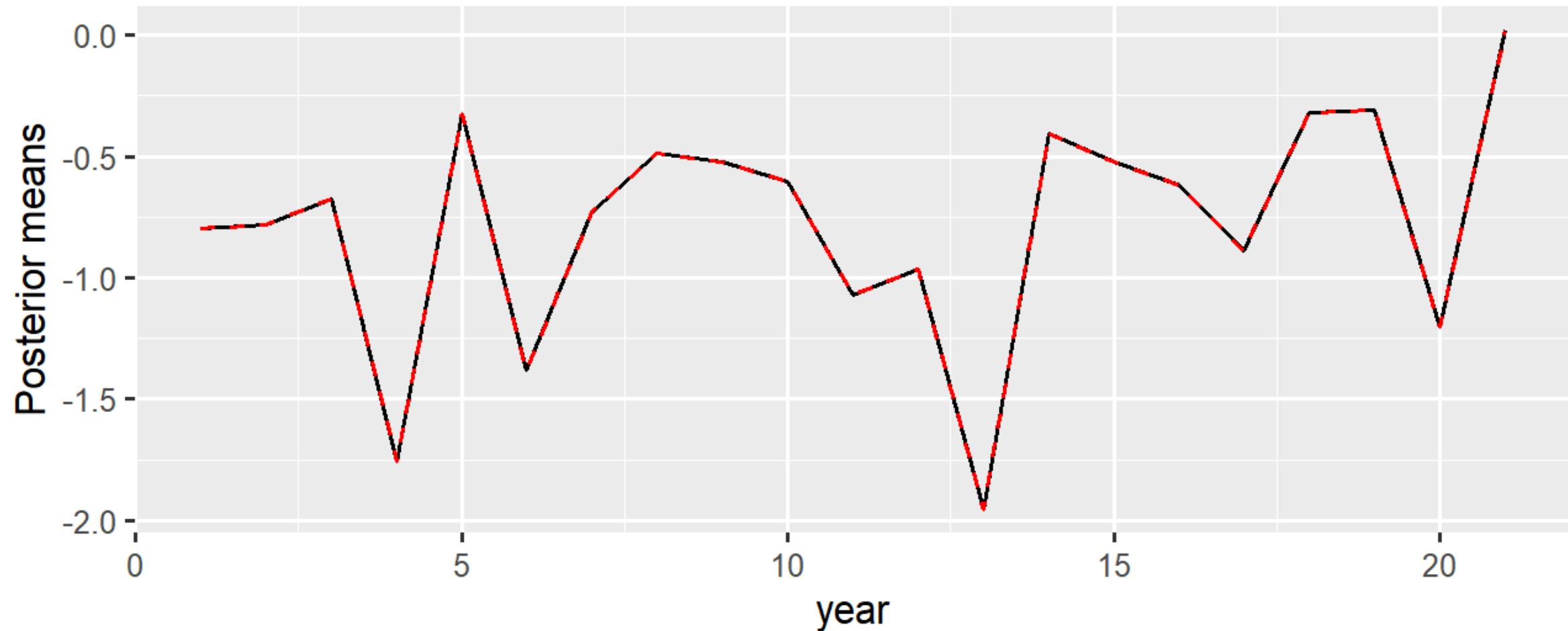
```
> mod1$summary.fixed
```

	mean	sd	0.025quant	0.5quant	0.975quant	mode	kld
(Intercept)	-0.98833298	0.11238994	-1.20942022	-0.98831790	-0.76733070	-0.98831791	1.473316e-08
year	0.02612109	0.00058628	0.02497146	0.02612109	0.02727073	0.02612109	5.527834e-11

```
> mod3$summary.fixed
```

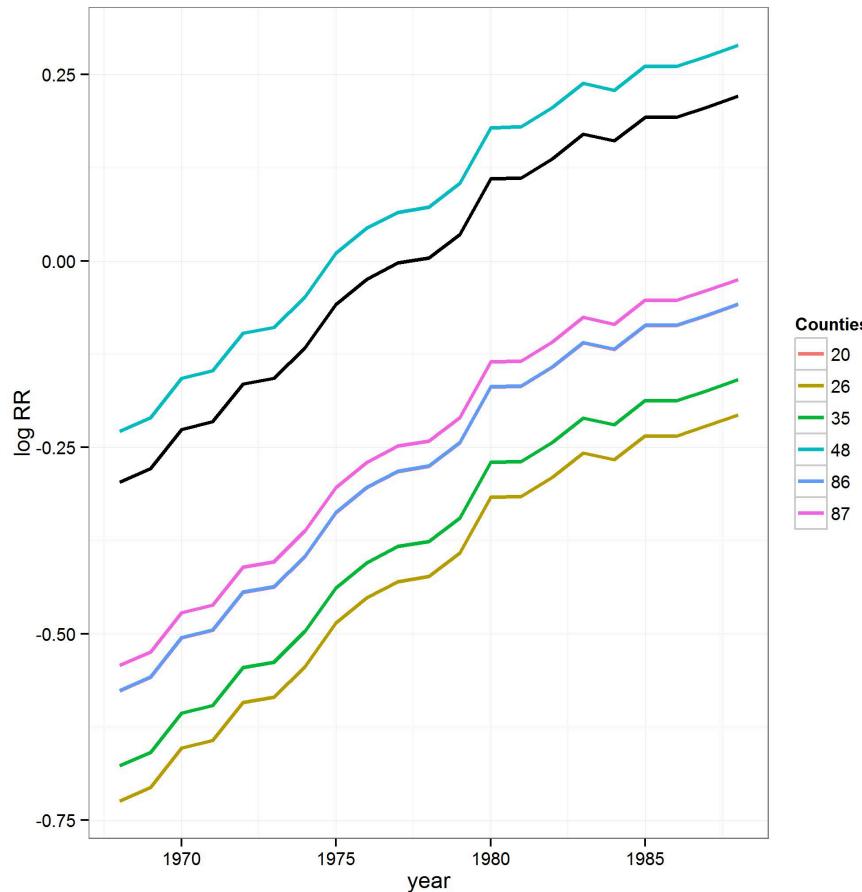
	mean	sd	0.025quant	0.5quant	0.975quant	mode	kld
(Intercept)	-0.7759713	0.1580754	-1.086885	-0.7759578	-0.4651329	-0.7759577	1.536133e-08

## Comparing model 2 and 3: Fitted values



# Do we need space-time interactions in spatio-temporal models?

- When we have temporal random effects, if we look at area specific trends this is what we would see:



Parallel lines as the temporal trend is assumed to be the same across all the spatial units

# Space-time model with exchangeable interactions

- It is easy to allow for spatio-temporal interactions, which add flexibility to the model
- A general specification of the model will be

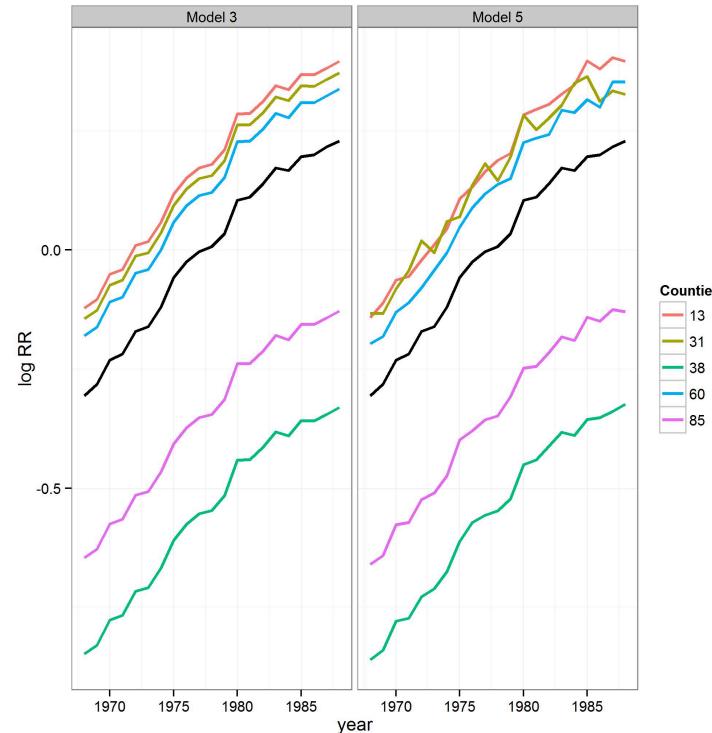
$$\begin{aligned}y_{it} &\sim \text{Poisson}(E_{it}\rho_{it}) \\ \log \rho_{it} &= b_0 + b_i + \gamma_t + \psi_t + \delta_{it} \\ b_i &= v_i + u_i \\ \gamma_t &\sim \text{RW}(1) \\ \psi_t &\sim N(0, \sigma_\psi^2) \\ \delta_{it} &\sim \text{Normal}(0, \sigma_\delta^2)\end{aligned}$$

The **space-time interaction parameter**,  $\delta_{it}$ , is modelled as exchangeable random effects, that capture departure from the additive structure (given by the BYM2 in space and RW1+iid in time).

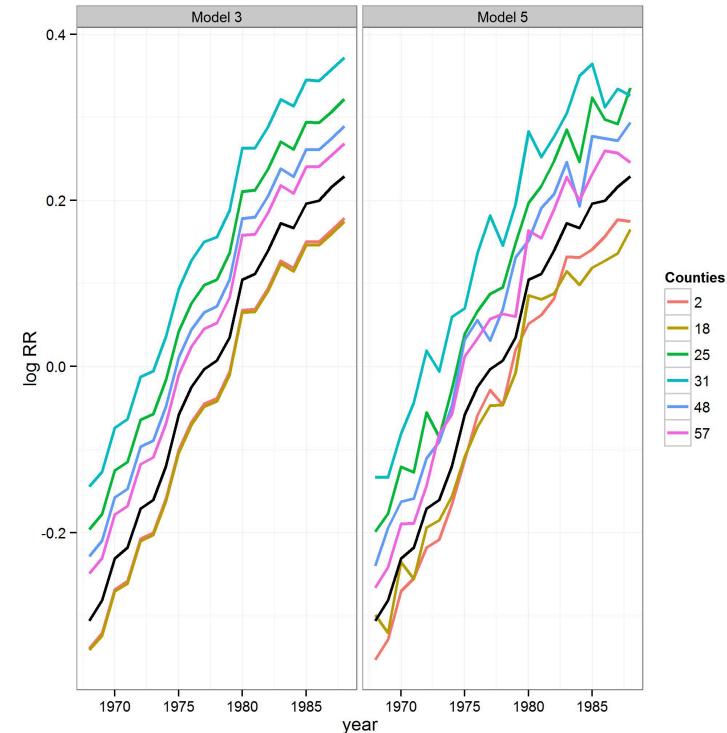
# Interpretation of the interactions

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

"usual" temporal trends



unusual temporal trends



# INLA code for model with interaction

- In INLA it 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

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

$$\gamma_t \sim RW(1)$$

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

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

```
> formula.intI = y ~ + f(county,model="bym2",
+                         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

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

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

$$b_i = v_i + u_i$$

$$\gamma_t \sim RW(1)$$

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

## Characteristics of ST interaction

Interaction	Parameters	Rank
I	Unstructured in space and Unstructured in time	$nT$
II	Unstructured in space and RW in time	$n(T-1)$ for RW1, $n(T-2)$ for RW2
III	ICAR in space and Unstructured in time	$(n-1)T$
IV	ICAR in space and RW in time	$(n-1)(T-1)$ for RW1, $(n-1)(T-2)$ for RW2

# How to model interactions

- Data in R format

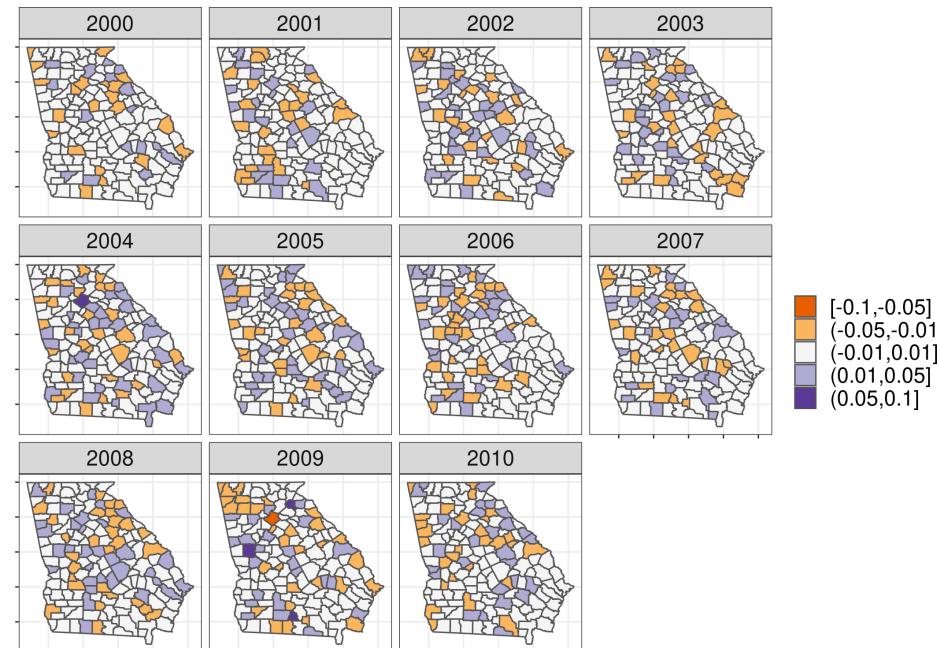
```
# A tibble: 1,749 × 8
  ID.area   Obs ID.year   Exp      y      E NAME    ID.area.year
  <int> <int> <dbl> <dbl> <int> <dbl> <chr>    <int>
1     1     20     1  25.8    20  25.8 Appling      1
2     1     24     2  25.1    24  25.1 Appling      2
3     1     25     3  22.8    25  22.8 Appling      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
# i 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 adopt Type I interaction, which assumes that the two unstructured effects  $v_i$  and  $\psi_t$  interact

```
> Georgia.adj = "Georgia.graph"
> ID.year2 = data_INLA$ID.year
> formula.intI = y ~ + f(ID.area,model="bym2",
+                         graph=Georgia.adj) +
+                         f(ID.year,model="rw1") +
+                         f(ID.year2,model="iid") +
+                         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

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

- There is a function in R which does exactly that: `kronecker(a,b)`
- Type II-IV of interactions are pretty complex and can take a **very long** time to run
- We briefly discuss, but for more information on 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  $\gamma_t$  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
```

## Type II interaction: in INLA

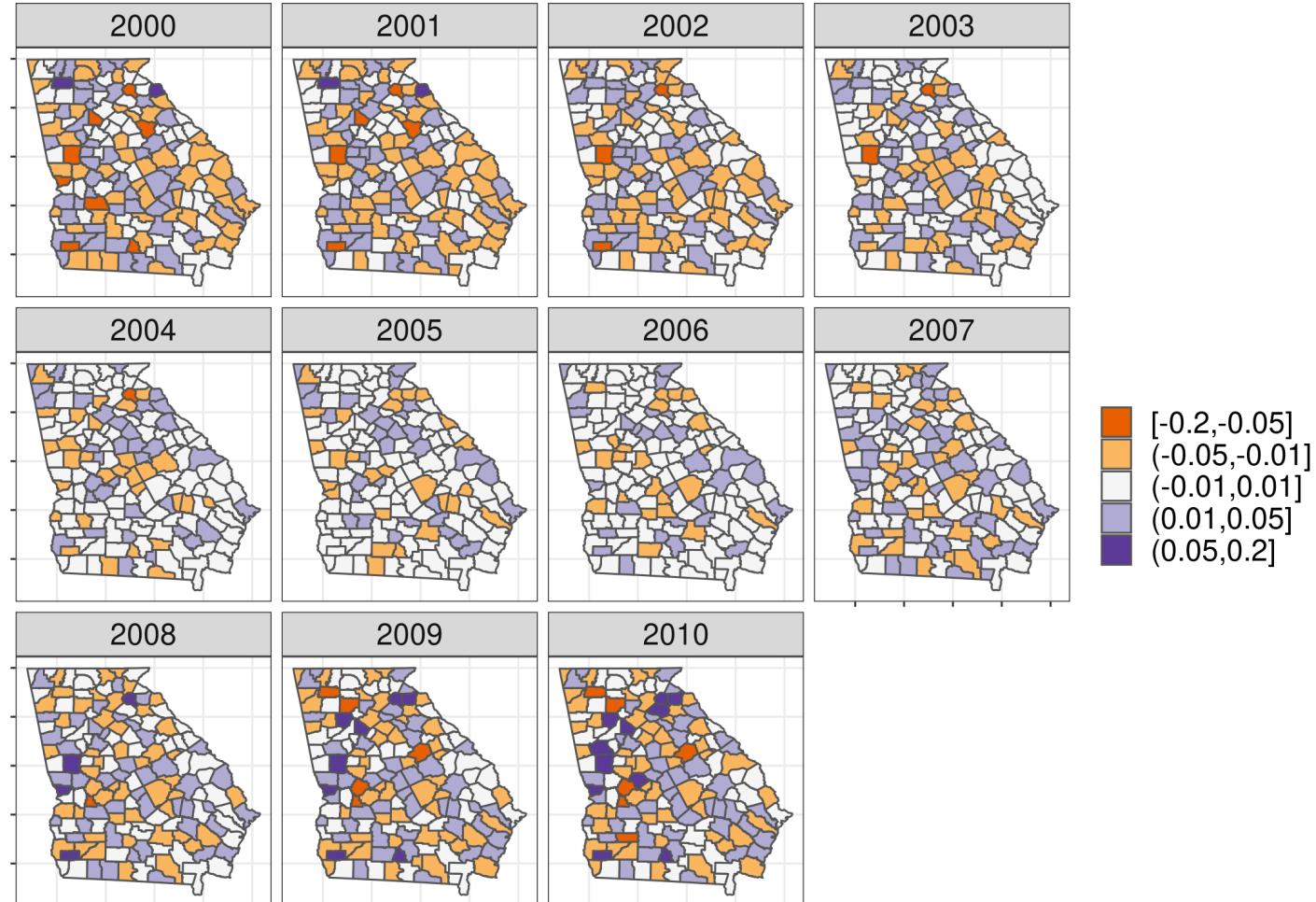
- Finally we need to impose some more constraints as the matrix have spatio-temporal interaction matrix has a rank deficiency
- We calculate the eigenvalues of the structure matrix of the interaction term and the extra constraints are those where the eigenvalues are zero

```
> eigenQ <- eigen(R)
> ids <- which(eigenQ$values < 0.00001)
> cMat <- t(eigenQ$vectors[,ids])
```

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.year2,model="iid") +
+           f(ID.area.year, model="generic0", Cmatrix=R, rankdef = nrow(cMat),
+               extraconstr = list(A = cMat, e = rep(0, nrow(cMat))))
```

# 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  $\psi_t$ .
- 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]	[,15]
[1, ]	2	0	0	0	0	0	0	0	0	0	-1	0	0	0	0
[2, ]	0	2	0	0	0	0	0	0	0	0	0	-1	0	0	0
[3, ]	0	0	2	0	0	0	0	0	0	0	0	0	-1	0	0
[4, ]	0	0	0	2	0	0	0	0	0	0	0	0	0	0	-1
[5, ]	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0

# Type III interaction in INLA

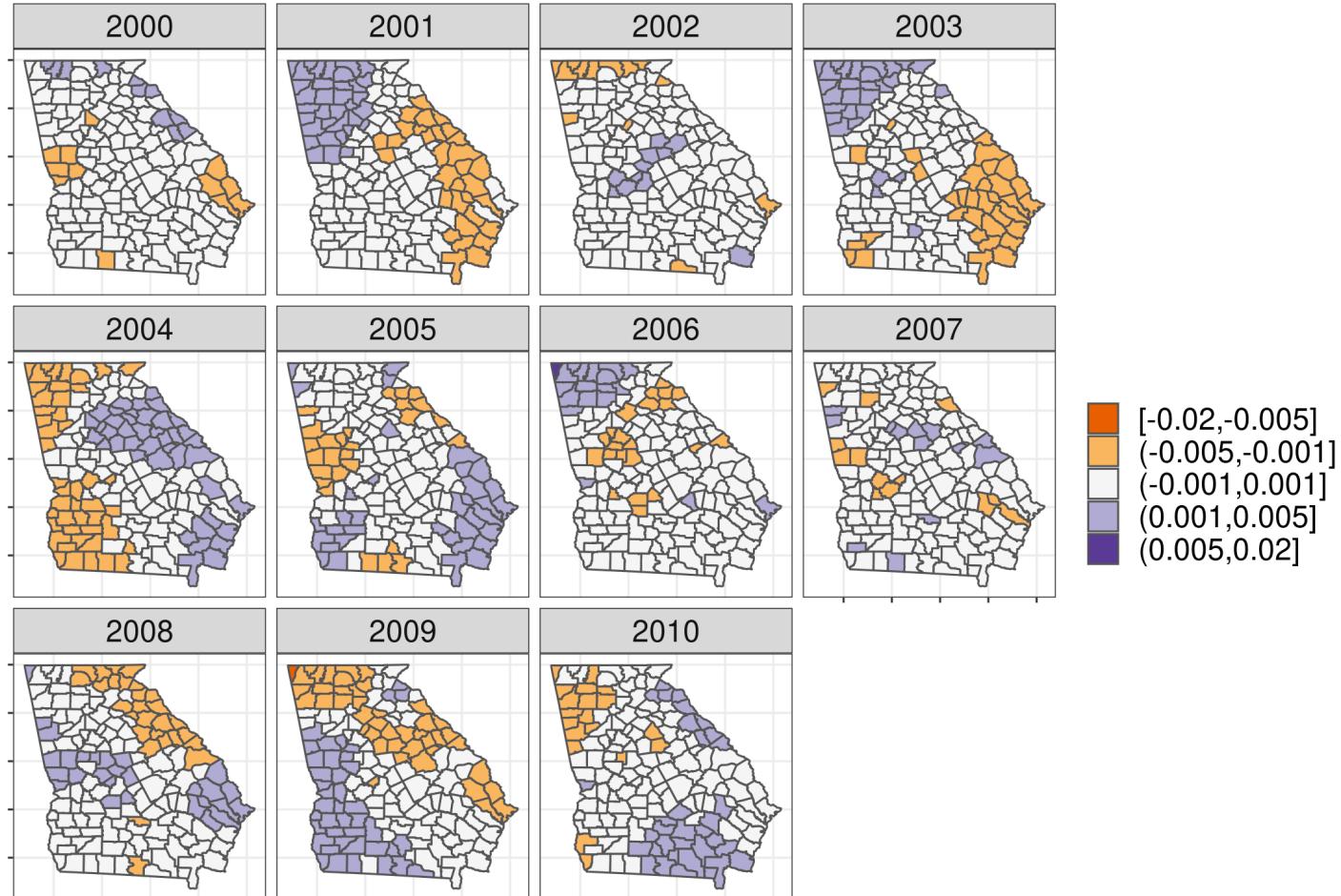
- Finally we need to impose some more constraints (on the eigenvalues equal to 0)

```
> eigenQ <- eigen(R)
> ids <- which(eigenQ$values < 0.00001)
> cMat <- t(eigenQ$vectors[,ids])
```

The formula environment becomes

```
> formula.intIII<- y ~ f(ID.area,model="bym2",graph=Georgia.adj) +
+           f(ID.year,model="rw1") +
+           f(ID.year2,model="iid") +
+           f(ID.area.year, model="generic0", Cmatrix=R, rankdef = nrow(cMat),
+               extraconstr = list(A = cMat, e = rep(0, nrow(cMat))))
```

# Type III interaction: in INLA



# Type IV interaction in INLA

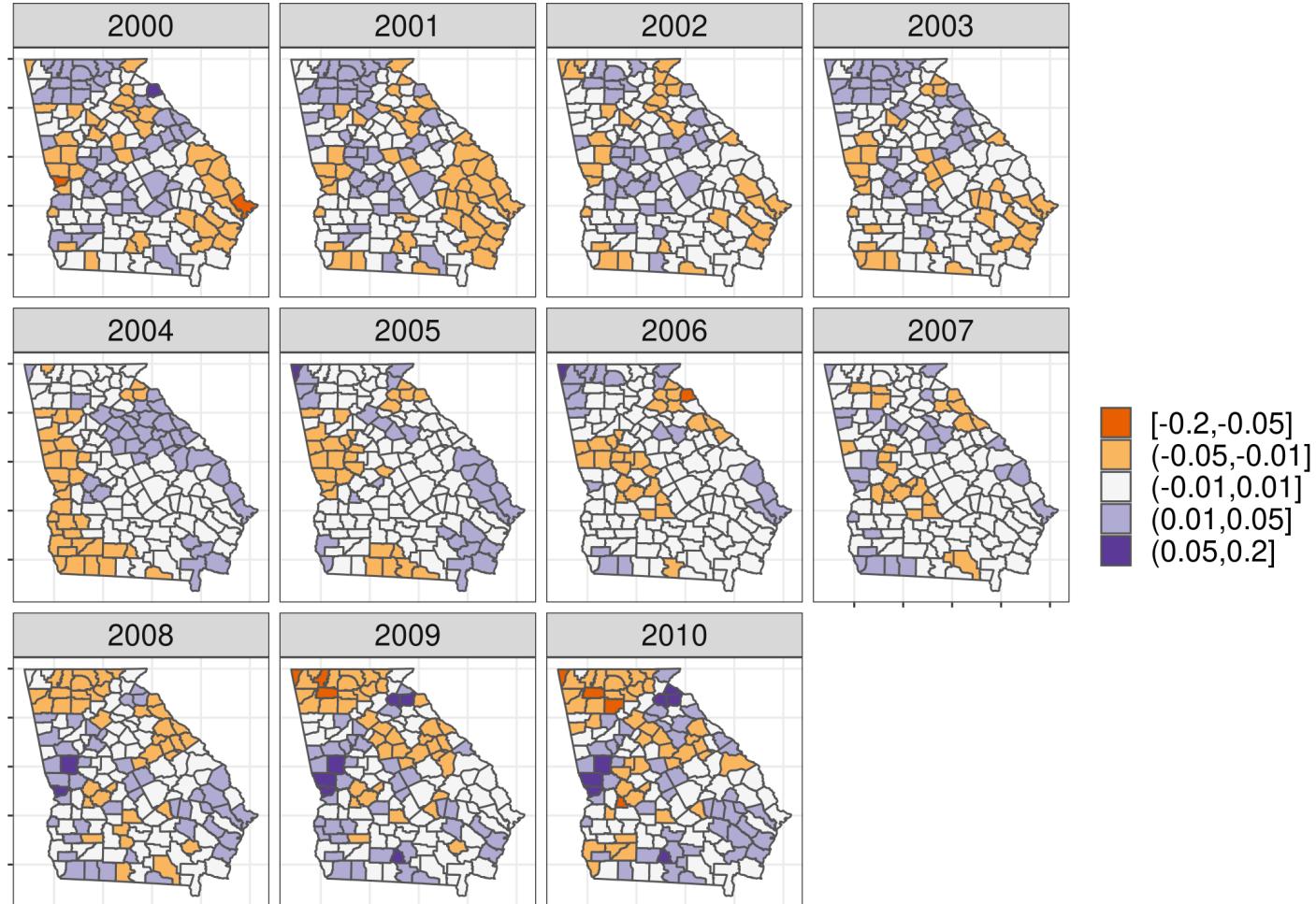
- Type IV is the most complex type of interaction, assuming that the spatially and temporally structured effects  $u_i$  and  $\gamma_t$  interact
- We assume that the temporal dependency structure for each area is not independent from all the other areas anymore, but depends on the temporal pattern 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)
> eigenQ <- eigen(R)
> ids <- which(eigenQ$values < 0.00001)
> cMat <- t(eigenQ$vectors[,ids])
```

The formula environment becomes

```
> formula.intIV<- y ~ f(ID.area,model="bym2",graph=Georgia.adj) +
+           f(ID.year,model="rw1") +
+           f(ID.year,model="iid") +
+           f(ID.area.year, model="generic0", Cmatrix=R, rankdef = nrow(cMat),
+               extraconstr = list(A = cMat, e = rep(0, nrow(cMat))))
```

# 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:

Model fitting for the different ST models

Interaction	Parameters	Rank	DIC	WAIC
I	Unstructured in space and Unstructured in time	$nT$	11570	11602
II	Unstructured in space and RW in time	$n(T-1)$ for RW1, $n(T-2)$ for RW2	11536	11569
III	ICAR in space and Unstructured in time	$(n-1)T$	11615	11659
IV	ICAR in space and RW in time	$(n-1)(T-1)$ 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.