Spatio-temporal disease mapping

Birgit Schrödle

Biostatistics Unit
Institute of Social and Preventive Medicine
University of Zurich

Zürich, October 29th, 2010

Data and problem

- Data:
  Cases of BVD (bovine viral diarrhea) reported to the Swiss federal veterinary office (BVET), 2003 - 2007

Year	03	04	05	06	07
#	141	172	239	406	712

- Reporting chain:
  Livestock owner → veterinarian →Cantonal veterinarian → BVET

- Problem:
  Biased case reporting, due to e.g. low reporting motivation, veterinary political issues, ...

Spatial and temporal resolution of the data

- Time: 2003 - 2007, yearly data
- Space:
  - 184 regions
  - 26 cantons

  additionally: Principality of Liechtenstein (canton & region)
- Suspicion: federal structure of Switzerland has an influence of the case reporting

Tasks:

1. Have regions AND cantons in the model
2. Expand the BYM model for time

Spatial model - regions AND cantons

- Data Y: number Y_i of reported cases in region i
- Cases are Poisson distributed given a rate λ_i:

$$Y_i|\lambda_i \sim \text{Poisson}$$

- Linear predictor of the standard BYM model:

$$\log(\lambda_i) = \eta_i = \log(m_i) + \mu + \nu_i + \psi_i \quad i = 1, ..., 185$$

m$_i$: number of herds in region i
μ: intercept
ν_i: spatially unstructured effect, i.i.d. normally distributed
ψ_i: spatially structured effect, ICAR

Spatial model - regions AND cantons

- Linear predictor of the standard BYM model:

$$\log(\lambda_i) = \eta_i = \log(m_i) + \mu + \nu_i + \psi_i \quad i = 1, ..., 185$$

- Include an i.i.d. cantonal effect $\alpha_{j(i)}$ to account for cantonal heterogeneity:

$$\log(\lambda_i) = \eta_i = \log(m_i) + \mu + \nu_i + \alpha_{j(i)} \quad j = 1, ..., 27$$

Inclusion of a linear time trend

- Data Y: number Y_{it} of reported cases in region i at time t
- First assumption: linear time trend
  $\rightarrow$ add it to the model

$$\eta_{it} = \log(m_i) + \mu + \nu_i + \alpha_{j(i)} + \beta \cdot t$$

- Separable spatial and temporal main trends!
- But: the increase/decrease in cases could be differently steep for each canton
- Solution: include a so-called random slope for each canton

$$\eta_{it} = \log(m_i) + \mu + \nu_i + \alpha_{j(i)} + \left(\beta + \delta_{j(i)}\right) \cdot t$$

Inclusion of a nonparametric time trend

- If data are observed over a high number of years
  $\rightarrow$ assumption of a linear time trend is too strong
- Use a so-called nonparametric time trend β_t $(t = 1, ..., 5)$ and an interaction term δ_{it}

$$\eta_{it} = \log(m_i) + \mu + \nu_i + \alpha_{j(i)} + \beta_t + \delta_{it}$$

- β_t is often specified as a random walk of 1st order

Random walk of 1st order

The conditional distribution of the incidence at time t depends only on the incidence at time $t-1$ and $t+1$:

$$\beta_t | \boldsymbol{\beta}_{-t}, \sigma^2_\beta \sim \mathrm{N}(\frac{1}{2}(\beta_{t-1} + \beta_{t+1}), \sigma^2_\beta/2)$$

Bayesian model selection criterion

- Deviance information criterion (DIC):

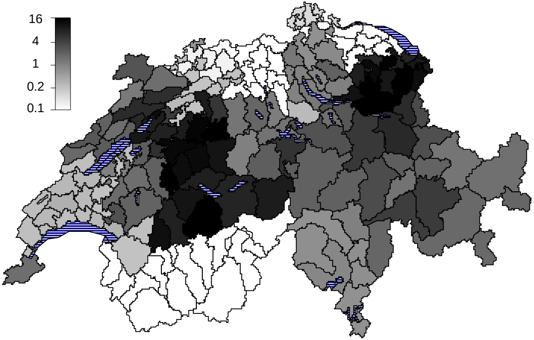
$$\text{DIC} = \bar{D} + p_D$$

  Trade-off between model fit (mean deviance: $\bar{D}$) and model complexity (effective number of parameters: p_D)

- Results:

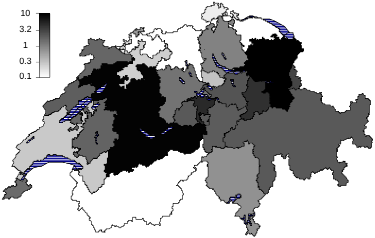
	time	$\overline{D}$	p_D	DIC
Model 1	linear	1712.1	91.3	1803.5
Model 2	RW1	1622.7	117.8	1740.5

Fitted spatial incidence of BVD - model 2

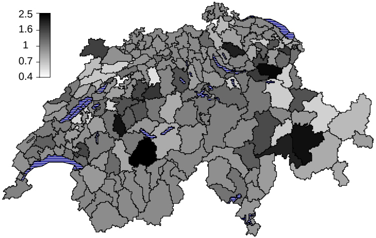


Estimated spatial effects - cantons and regions

cantons

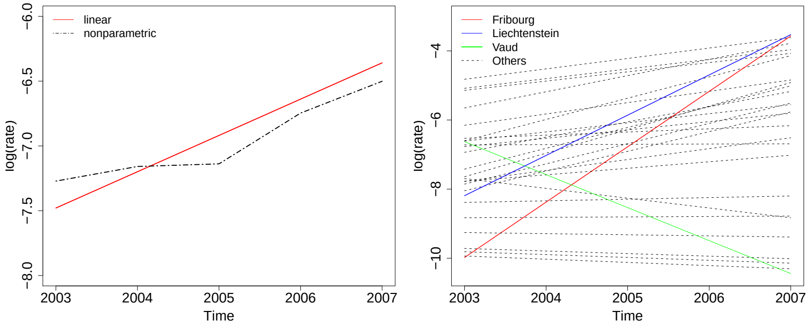


regions



→ Cantonal heterogeneity dominates spatial variation
→ Clear evidence for underreporting due to different policies of the cantons

Temporal trends - linear and nonparametric



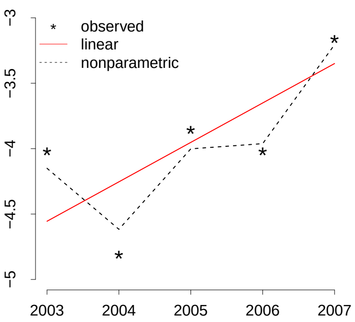
Left plot: main temporal trend for Models 1 and 2

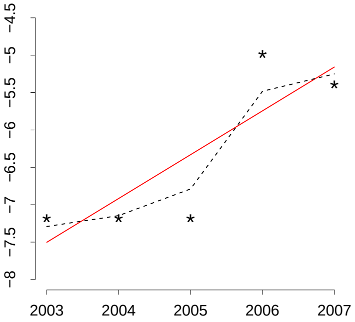
Right plot: total temporal trend for each canton, Model 1

Fitted values

1 region in Appenzell-Innerrhoden

1 region in Lucerne





THE END

- I hope you enjoyed this course!
- Please fill in the evaluation form and give it back to me...