Space-time Modeling I

BMTRY 763

Space-time (ST) Modeling (BDM13, ch 12)

- Some notation
  - Assume counts within fixed spatial and temporal periods: map evolutions
  - Both space and time are subscripts in the analysis
  - Consider separable models (with spatial and separate temporal terms)
  - Also interaction effects

Notation

$$outcome:\ y_{ij};\ \text{RRisk: } \theta_{ij}$$

$$\text{expected count: } e_{ij}$$

$$i = 1,...,m:\ \text{small areas}$$

$$j = 1,...,J:\ \text{time periods}$$

Expected Counts

- Computation (simplest - overall average):

$$e_{ij} = p_{ij} \cdot \sum_i \sum_j y_{ij} / \sum_i \sum_j p_{ij}$$

Basic retrospective model

- Infinite population; small disease probability
- Poisson assumption

$$y_{ij} \sim Pois(e_{ij}\theta_{ij})$$

$$\log(\theta_{ij}) = \alpha_0 + S_i + T_j + ST_{ij}$$

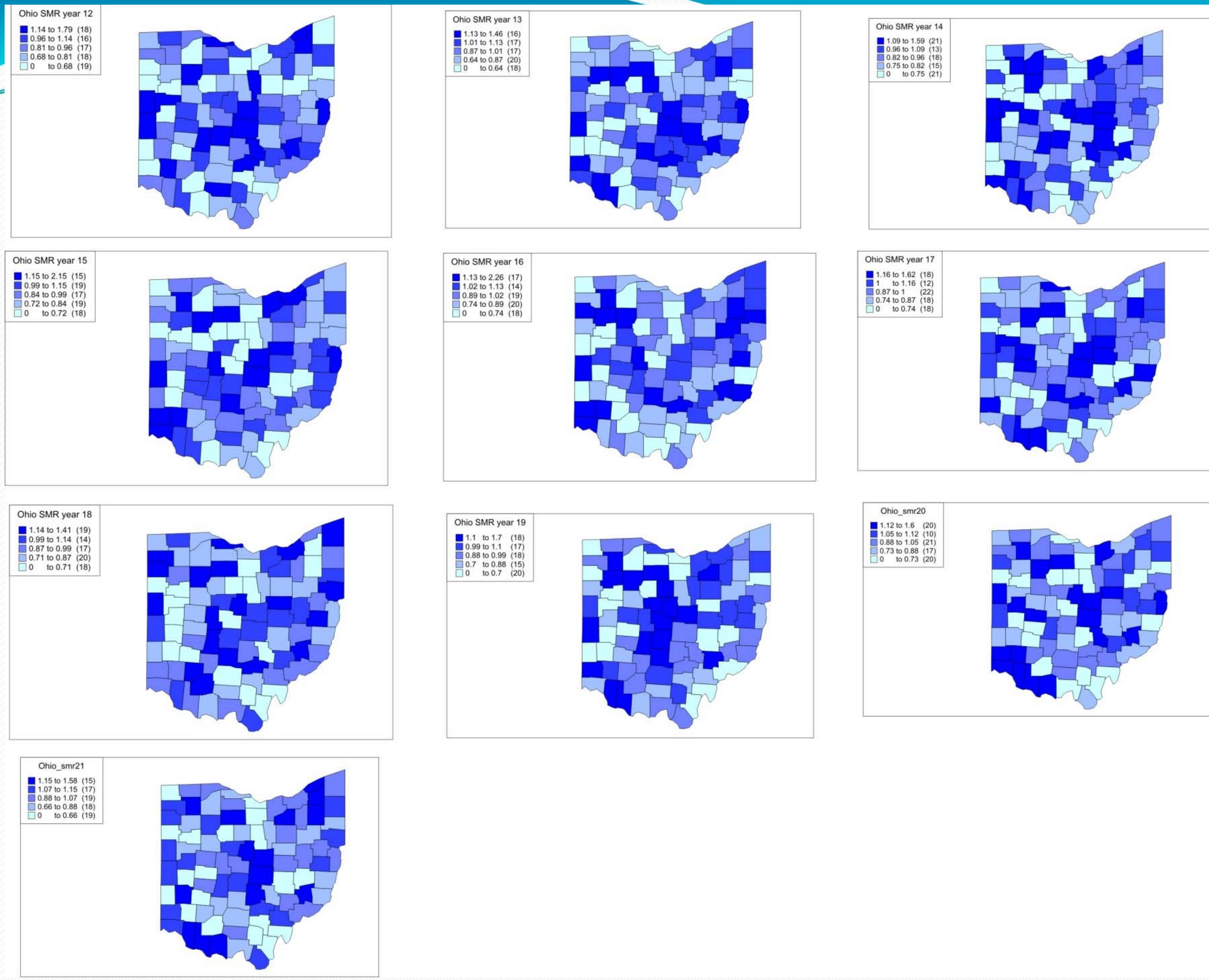
S_i : spatial terms

T_j : temporal terms

ST_{ij} : interaction

Full data set: 21 years of Ohio lung cancer

- 10 years of SMRs standardized with the statewide rate: 1979-1988
- Frequently analyzed
- Row wise from 1979


Ohio SMR year 12
1.14 to 1.79 (18)
0.96 to 1.14 (16)
0.81 to 0.96 (17)
0.68 to 0.81 (18)
0 to 0.68 (19)
Ohio SMR year 13
1.13 to 1.46 (16)
1.01 to 1.13 (17)
0.87 to 1.01 (17)
0.64 to 0.87 (20)
0 to 0.64 (18)
Ohio SMR year 14
1.09 to 1.59 (21)
0.96 to 1.09 (13)
0.82 to 0.96 (18)
0.75 to 0.82 (15)
0 to 0.75 (21)
Ohio SMR year 15
1.15 to 2.15 (15)
0.99 to 1.15 (19)
0.84 to 0.99 (17)
0.72 to 0.84 (19)
0 to 0.72 (18)
Ohio SMR year 16
1.13 to 2.26 (17)
1.02 to 1.13 (14)
0.89 to 1.02 (19)
0.74 to 0.89 (20)
0 to 0.74 (18)
Ohio SMR year 17
1.16 to 1.62 (18)
1 to 1.16 (12)
0.87 to 1 (22)
0.74 to 0.87 (18)
0 to 0.74 (18)
Ohio SMR year 18
1.14 to 1.41 (19)
0.99 to 1.14 (14)
0.87 to 0.99 (17)
0.71 to 0.87 (20)
0 to 0.71 (18)
Ohio SMR year 19
1.1 to 1.7 (18)
0.99 to 1.1 (17)
0.88 to 0.99 (18)
0.7 to 0.88 (15)
0 to 0.7 (20)
Ohio_smr20
1.12 to 1.6 (20)
1.05 to 1.12 (10)
0.88 to 1.05 (21)
0.73 to 0.88 (17)
0 to 0.73 (20)
Ohio_smr21
1.15 to 1.58 (15)
1.07 to 1.15 (17)
0.88 to 1.07 (19)
0.66 to 0.88 (18)
0 to 0.66 (19)

Some Random Effect models

model 1a:

$$\log(\theta_{ij}) = \alpha_0 + v_i + u_i + \beta t_j$$

model 1b:

$$\log(\theta_{ij}) = \alpha_0 + v_i + u_i + \gamma_j$$

model 2:

$$\log(\theta_{ij}) = \alpha_0 + v_i + u_i + \gamma_{1j} + \gamma_{2j}$$

model 3:

$$\log(\theta_{ij}) = \alpha_0 + v_i + u_i + \gamma_{2j} + \psi_{ij}$$

model 4:

$$\log(\theta_{ij}) = \alpha_0 + v_i + u_i + \gamma_{1j} + \gamma_{2j} + \psi_{ij}$$

model 5: variants of (3) with ψ_{ij}

Random Walk Prior distribution

- Model 1 b: we assume a random effect for the time element and this has a random walk prior distribution:

$$\gamma_j \sim N(\gamma_{j-1}, \tau_\gamma^{-1})$$

- More generally an AR1 prior could be used:

$$\gamma_j \sim N(\lambda\gamma_{j-1}, \tau_\gamma^{-1}); \quad 0 < \lambda \le 1$$

Interaction priors

- A variety of priors for the interaction can be assumed (both correlated and non-separable)
- Knorr-Held (2000) first suggested dependent priors (see Lawson (2013) ch12)
- Two simple separable examples of possible priors are:

$$\psi_{ij} \sim N(0, \tau_\psi) \ \text{ uncorrelated (model 3)}$$

$$\psi_{ij} \sim N(\psi_{i,j-1}, \tau_\psi) \ \text{ random walk (model 5)}$$

Model fitting Results

Model	DIC	pD
1a	5759	80
1b	5759	80
2	5759.4	79
3	5751.4	129
4	5755.3	129
5	5750.6	115

Interpretation

- The temporal trend model does not provide a better fit than the random walk (1a, 1b)
- The extra RE in model 2 is not needed
- The inclusion of the interaction in model 3 is significant but model 4 is not good
- Model 5 with the random walk interaction seems best as it has lowest DIC and smaller pD than model 3

Space-time Kalman Filter

- The Kalman Filter consists of a coupled set of equations describing the behavior of a system and also the measurement made on the system
- In a dynamic setting it is an appropriate model for an evolving system observed with error

System structure evolution

- e.g. risk evolving over time

$$\theta_{ij} \sim N(F(\theta_{i,j-1}), \Gamma)$$

e.g.

$$F(\theta_{i,j-1}) = \lambda \theta_{i,j-1} \text{ (AR1)}$$

$$\text{or } F(\theta_{i,j-1}) = \lambda_i \theta_{i,j-1}$$

Measurement model

- Poisson (independent) error

$$y_{ij} \sim Pois(\mu_{ij})$$

$$\mu_{ij} = e_{ij}\theta_{ij}$$

- In the classic Kalman filter the errors would be Gaussian and the measurement model could also have correlated errors

Gaussian approximation

- We can proceed to generalize this to allow correlation if we assume a log Gaussian form
- This leads to a hidden Markov model

$$\log(y_{ij}/e_{ij}) \sim N(\mu_{ij}, \Sigma)$$

$$\exp\{\mu_{ij}\} = \theta_{ij}$$

Full Model

- Structural model

$$\theta_{ij} \sim N(F(\theta_{i,j-1}),\Gamma),$$
$$with \;\; \Gamma_{j,k} = \text{cov}(\theta_{ij},\theta_{ik})$$

- Measurement model

$$z_{ij} = \log(y_{ij} / e_{ij}) \sim N(\mu_{ij},\Sigma)$$
$$\exp\{\mu_{ij}\} = \theta_{ij}$$
$$\Sigma_{il} = \text{cov}(z_{ij}, z_{lj})$$

WinBUGS Code

- DIC: 1657 (767)

```
for (i in 1:m){
for (j in 1:T){
Lye[i,j]<-log((y[i,j]+0.01)/(e[i,j]+0.01))
Lye[i,j]~dnorm(mu1[i,j],tauS)
theta[i,j]<-exp(Ltheta[i,j])
mu1[i,j]<-a0+Ltheta[i,j]+Struct[i]+R[j]}}
for (i in 1 :m){
Ltheta[i,1:T]~dmnorm(mu[i,],covT[,])}
for (j in 1:T){R[j]~dnorm(0,tauR)}
for(i in 1:m){
mu[i,1]<-theta[i,1]
for (j in 2:T){
mu[i,j]<-theta[i,j-1]}}
for(i in 1:T)
{for(j in 1:T)
{d[i,j]<-abs(j-i)
covT[i,j]<-sig2*pow(rho,d[i,j])}}
```

Clustering in ST data

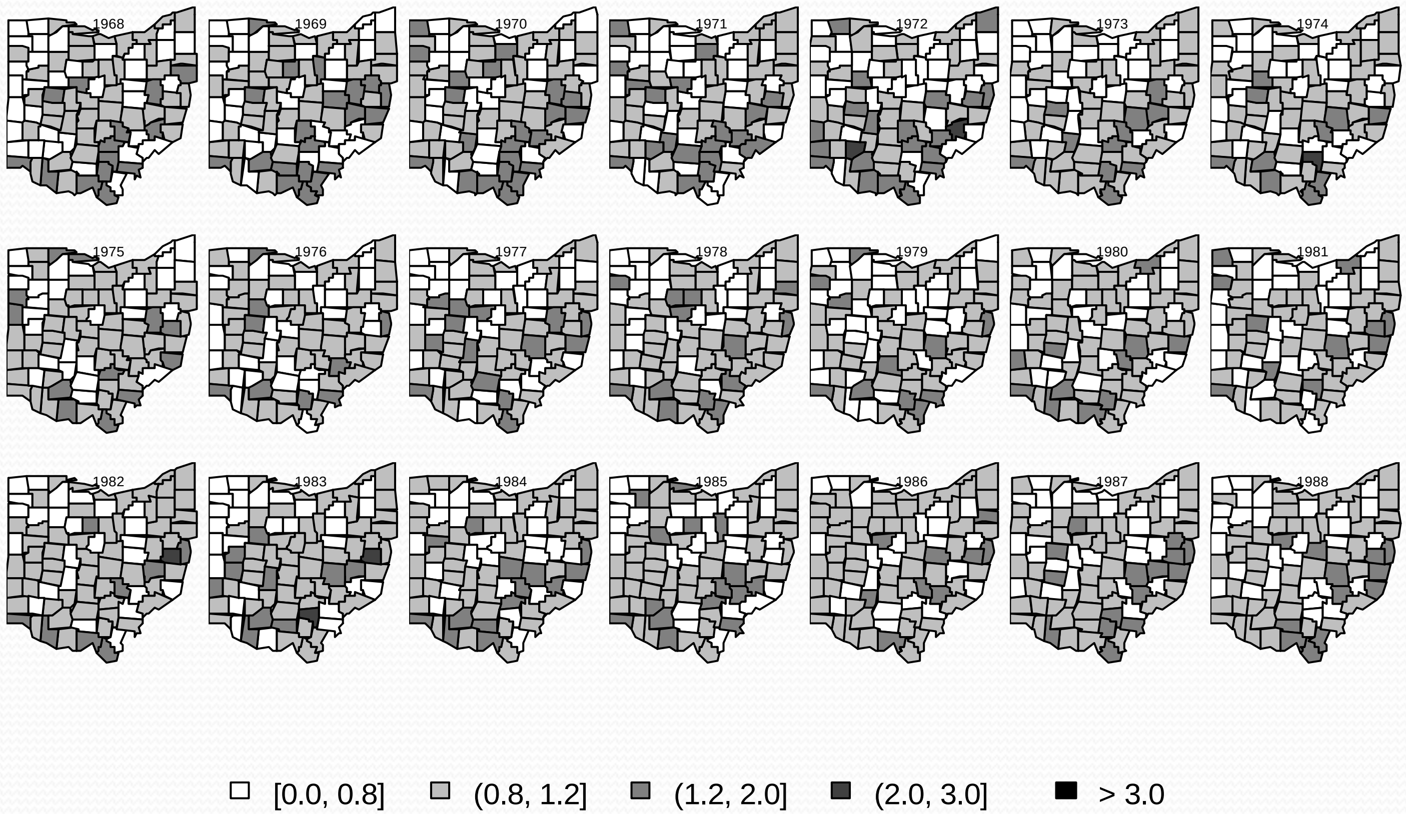
Clustering in ST data

- Clustering is a different issue.
- Earlier we examined exceedence probabilities
- These can also be used with ST data

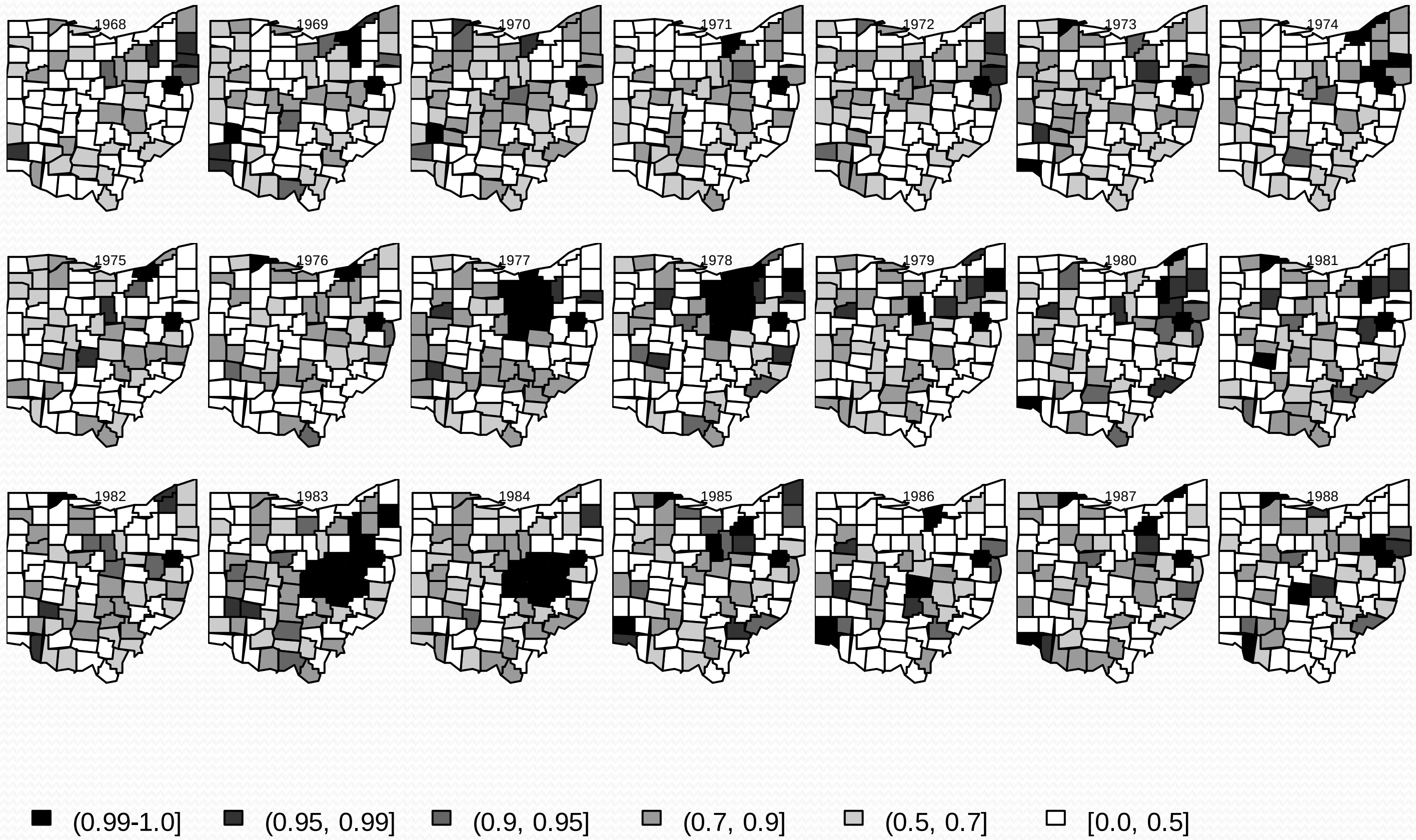
$$\Pr(\theta_{ij} > c)$$

$$c = 1 \text{ or } 2 \text{ or } 3$$

Ohio SMR 21 years



Added simulated clusters
Exceedence C=1



Exceedence C=2