Lecture 7: Spatio-temporal models

May 10, 2016

Multiple ways to include interactions

- Arises for whenever there's multiple factors
- Include spatial effect (No / Yes)
- Include temporal effect (No / Independent / Smoothed)
- Include spatio-temporal effect (No / Independent / Smoothed)

Without spatial

		Temporal effect		
Spatio-temporal effect		None	Independent	Smoothed
	None	$\mathrm{gam}(c \sim 1)$	$\mathrm{gam}(c \sim 1 + factor(t))$	$\mathrm{gam}(c \sim 1 + s(t))$
	Independent	$\mathrm{gam}(c \sim 1$	$\mathrm{gam}(c \sim 1 + factor(t)$	$\mathrm{gam}(c \sim 1 + s(t)$
	Smoothed	$\mathrm{gam}(c \sim 1 + s(x, y, t))$	$\mathrm{gam}(c \sim 1 + factor(t)$	$\mathrm{gam}(c \sim 1 + s(t) + s(x, y, t))$

With spatial

		Temporal effect		
Spatio-temporal effect		None	Independent	Smoothed
	None	$\mathrm{gam}(c \sim 1 + s(x, y))$	$\mathrm{gam}(c \sim 1 + s(x, y)$	$\mathrm{gam}(c \sim 1 + s(x, y) + s(t))$
	Independent	$\mathrm{gam}(c \sim 1 + s(x, y)$	$\mathrm{gam}(c \sim 1 + factor(t)$	$\mathrm{gam}(c \sim 1 + s(x, y) + s(t)$
	Smoothed	$\mathrm{gam}(c \sim 1 + s(x, y)$	$\mathrm{gam}(c \sim 1 + s(x, y)$	$\mathrm{gam}(c \sim 1 + s(x, y) + s(t)$

Two we will focus on:

- Spatial index standardization

$$\text{gam}\left(c \sim factor(t) + s(x, y) + s(x, y, by = factor(t))\right)$$

- Spatial Gompertz model

$$\text{gam}(c \sim 1 + s(x, y) + s(x, y, t))$$

Spatial index standardization :

$$c_i \sim \mathrm{Poisson}(\exp(\lambda_i), \sigma^2)$$

$$\lambda_i = \beta_{t_i} + \omega(s_i) + \varepsilon(s_i, t_i)$$

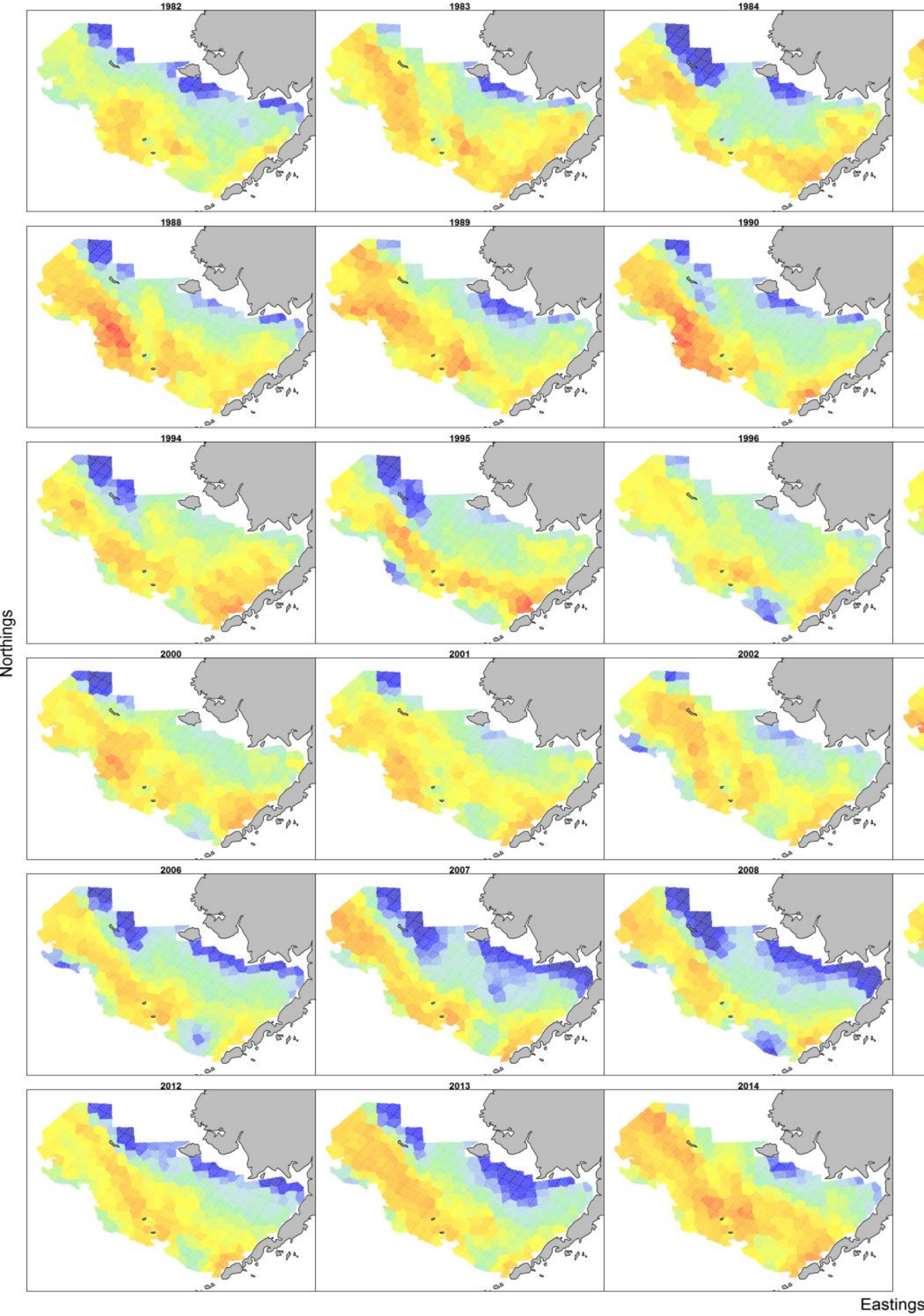
$$\boldsymbol{\omega} \sim \mathrm{MVN}(\mathbf{0}, \sigma_\omega^2 \mathbf{C}_\omega)$$

$$\boldsymbol{\varepsilon}(t) \sim \mathrm{MVN}(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{C}_\varepsilon)$$

OR [another way to write it]…

$$\mathrm{vec}(\mathbf{E}) \sim \mathrm{MVN}(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{C}_\varepsilon \otimes \mathbf{I})$$

Where $\mathbf{I}$ is an identity matrix



Alaska pollock in the Eastern Bering Sea

Spatial index standardization:

$$c_i \sim \text{Poisson}(\exp(\lambda_i), \sigma^2)$$

$$\lambda_i = \beta_{t_i} + \omega(s_i) + \varepsilon_{t_i}(s_i)$$

$$\boldsymbol{\omega} \sim MVN(0, \sigma_\omega^2 \mathbf{C}_\omega)$$

$$\boldsymbol{\varepsilon}_t \sim MVN(0, \sigma_\varepsilon^2 \mathbf{C}_\varepsilon)$$

Implications

– Expectation:

$$\mathbb{E}(\lambda_i) = \beta_{t_i}$$

Variance

$$\mathbb{V}(\lambda_i) = \sigma_\omega^2 + \sigma_\varepsilon^2$$

– Covariance

- [Show on board if anyone's interested]

Spatial index standardization:

$$c_i \sim \mathrm{Poisson}(\exp(\lambda_i), \sigma^2)$$

$$\lambda_i = \beta_{t_i} + \omega(s_i) + \varepsilon_{t_i}(s_i)$$

$$\boldsymbol{\omega} \sim MVN(0, \sigma_\omega^2 \mathbf{C}_\omega)$$

$$\boldsymbol{\varepsilon}_t \sim MVN(0, \sigma_\varepsilon^2 \mathbf{C}_\varepsilon)$$

[Show R script for model]

Derived quantities using SPDE approximation

- Estimates two parameters
  - κ: decorrelation distance
  - τ: variability
- Geostatistical range
  - Distance at which correlation is approx. 13%

$$Range = \frac{\sqrt{8}}{\kappa}$$

- Marginal standard deviation
  - Standard deviation of value at location s if re-simulating the spatial process
  - Standard deviation of value at location $s + h$ given as $h \to \infty$

$$SD = \frac{1}{\tau\kappa\sqrt{4\pi}}$$

Ways to model survey data:

$$c_i \sim g\,(\lambda_i\,;\boldsymbol{\theta})$$

$$f^{-1}(\lambda_i) = \alpha + \omega(s_i) + \varepsilon_{t_i}(s_i)$$

Where

- $g\,(\lambda_i\,;\boldsymbol{\theta})$ is the sampling distribution
- $f^{-1}(\lambda_i)$ is the (inverse) link function for the linear predictor.

1. Discrete data
   - Counts
   - Often contains more zeros than expected (true and false zeros)
2. Continuous data
   - Often arises from weighing all encountered individuals
   - Compound or delta-distribution

Discrete data

- Counts
- Often contains more zeros than expected (true and false zeros)

- Zero-inflated models

$$\Pr(c = C) = \begin{cases} p + \mathrm{Poisson}(0|\lambda) & \text{if } C = 0 \\ (1-p) \times \mathrm{Poisson}(C|\lambda) & \text{if } C > 0 \end{cases}$$

  - Where p is the probability of sampling outside occupied habitat
    - “True zero”
  - $\mathrm{Poisson}(0|\lambda)$ is the probability of sampling in occupied habitat, but encountering zero individuals
    - “False zero”

Continuous data

- Often arises from weighing all encountered individuals

- Tweedie distribution

$$d \sim Tweedie(\lambda, \mu, \sigma_w^2)$$

  - … arises as a compound distribution, i.e. counts:

$$N \sim Poisson(\lambda)$$

and individual weights

$$w_n \sim Gamma(\mu, \sigma_w^2)$$

where the total sample is a the sum of individual weights

$$d = \sum_{n=0}^{N} w_n$$

Convert from density and abundance

- Density

$$d(s,t) = \exp(\lambda(s,t))$$

- Abundance

$$B(t) = \sum_{s=1}^{n_s} d(s,t) \times a(s)$$

- $a(s)$: Area associated with each modelled location s

- Calculating areas
  1. Lay grid over domain, and count grid cells closest to location s
  2. Randomly sample from domain, and count closest samples

Derived quantities

$$d(s,t) = \exp(\lambda(s,t))$$

1. Total abundance

$$B(t) = \sum_{s=1}^{n_s} d(s,t) \times a(s)$$

2. Center of gravity

$$\bar{x}_t = \frac{1}{\sum_{s=1}^{n_s}(d(s,t) \times a(s))} \sum_{s=1}^{n_s} x(s)(d(s,t) \times a(s))$$

3. Average density

$$\bar{d}(t) = \frac{1}{\sum_{s=1}^{n_s}(d(s,t) \times a(s))} \sum_{s=1}^{n_s} d(s,t)(d(s,t) \times a(s))$$

4. Effective area occupied

$$a_t = \frac{B(t)}{\bar{d}(t)}$$

Exercise

- Divide into groups and add total abundance, center-of-gravity, average density, and effective area occupied calculations