P-spline mixed models for spatio-temporal data

María Durbán
joint work with Dae-Jin Lee

Department of Statistics
Universidad Carlos III de Madrid

June 2009

Outline

1. P-splines
   - Mixed models approach
   - Multidimensional P-splines
2. P-splines for spatial count data
   - Spatial smoothing
   - Smooth-CAR model
   - Application: Scottish Lip Cancer data
3. Spatio-temporal data Smoothing with P-splines
   - ANOVA-Type Interaction Models
   - Application Environmental spatio-temporal data
4. Spatio-temporal Disease Mapping

Outline

1. P-splines
   - Mixed models approach
   - Multidimensional P-splines
2. P-splines for spatial count data
   - Spatial smoothing
   - Smooth-CAR model
   - Application: Scottish Lip Cancer data
3. Spatio-temporal data Smoothing with P-splines
   - ANOVA-Type Interaction Models
   - Application Environmental spatio-temporal data
4. Spatio-temporal Disease Mapping

P-splines
"The Flexible Smoother"

- **Penalized Likelihood splines (Eilers & Marx, 1996):**
  - Given the data (x_i, y_i), $i = 1, ..., n$
  - Fit a **sum of local basis functions**:

$$y_i = f(x_i) + \epsilon_i, \qquad \epsilon \sim \mathcal{N}(0, \sigma^2)$$

    where $f(x_i) = B\theta$ and

    - $B = B(x)$ is a **Regression Basis**, and
    - θ is a vector of **coefficients**.
  - Control the fit through a **smoothing parameter** (λ).

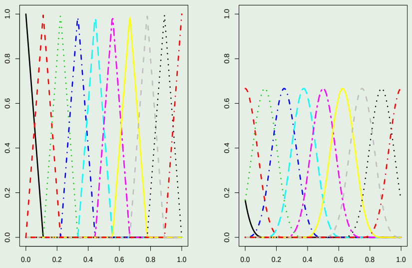
» Regression Basis

P-splines
"The Flexible Smoother"

- B-**splines Basis:**

  - $p+1$ Piece-wise polynomials of degree p.
  - Connected by **knots**.
  - In general the choice is p=3, cubic spline.

B-splines of degree p:



B-splines Basis:

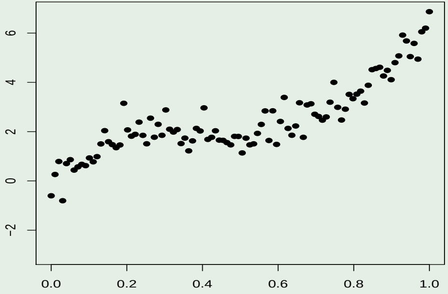
- $\widehat{y} = f(x_i) = B\widehat{\theta}$
- B-splines Regression:

$$\min S(\theta; y) = \|y - B\theta\|^2$$

$$\widehat{\theta} = (B'B)^{-1}B'y$$

  - Optimal selection of knots (Complex).
- P-Splines: add a penalty to control smoothness.

Example:



- B-**splines Basis:**

  - $\widehat{\boldsymbol{y}} = f(x_i) = \boldsymbol{B}\widehat{\boldsymbol{\theta}}$
  - B-**splines Regression:**

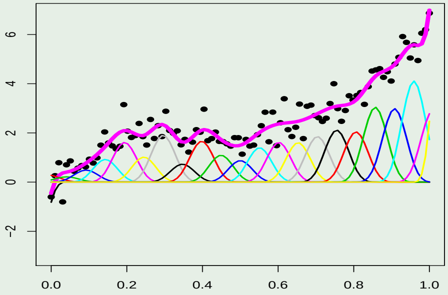
$$\min \ S(\boldsymbol{\theta}; \boldsymbol{y}) = \|\boldsymbol{y} - \boldsymbol{B}\boldsymbol{\theta}\|^2$$

$$\widehat{\boldsymbol{\theta}} = (\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'\boldsymbol{y}$$

  - Optimal selection of **knots** (**Complex**).

  - P-**Splines:** add a **penalty** to control smoothness.

Example:



P-splines
"The Flexible Smoother"

- B**-splines Basis:**

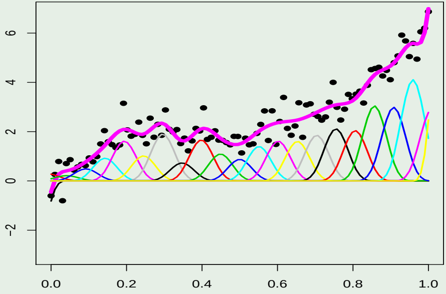
  - $\widehat{\boldsymbol{y}} = f(\boldsymbol{x}_i) = \boldsymbol{B}\widehat{\boldsymbol{\theta}}$
  - B**-splines Regression:**

$$\min \; S(\boldsymbol{\theta}; \boldsymbol{y}) = \|\boldsymbol{y} - \boldsymbol{B}\boldsymbol{\theta}\|^2$$

$$\widehat{\boldsymbol{\theta}} = (\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'\boldsymbol{y}$$

    - Optimal selection of **knots** (**Complex**).
  - P**-Splines:** add a **penalty** to control smoothness.

Example:



Methodology:

- Minimize the **penalized sum of squares** (**PSS**):

$$S(\theta; y, \lambda)_p = \|y - B\theta\|^2 + \text{PENALTY}$$

- The **PENALTY** term, controls the smoothness of the fit by λ.
  - **Eilers & Marx (1996):**
    $\Rightarrow$ (*discrete*) **Penalty** over adjacent coefficients θ.
  - **Lang & Brezger (2004):**
    $\Rightarrow$ **"Bayesian P-splines"**: random walk priors for θ, e.g.:

$$\theta|\theta_{m-1} \sim \mathcal{N}(\theta_{m-1}, \tau^2), \text{ or}$$
$$\theta|\theta_{m-1}, \theta_{m-2} \sim \mathcal{N}(2\theta_{m-1} - \theta_{m-2}, \tau^2)$$

P-splines
"The Flexible Smoother"

Methodology:

- Minimize the **penalized sum of squares** (**PSS**):

$$S(\theta; y, \lambda)_p = \|y - B\theta\|^2 + \textbf{PENALTY}$$

- The **PENALTY** term, controls the smoothness of the fit by λ.
  - **Eilers & Marx (1996):**
    $\Rightarrow$ (*discrete*) **Penalty** over adjacent coefficients θ.
  - **Lang & Brezger (2004):**
    $\Rightarrow$ **"Bayesian P-splines"**: random walk priors for θ, e.g.:

$$\theta|\theta_{m-1} \sim \mathcal{N}(\theta_{m-1}, \tau^2), \text{ or}$$

$$\theta|\theta_{m-1}, \theta_{m-2} \sim \mathcal{N}(2\theta_{m-1} - \theta_{m-2}, \tau^2)$$

P-splines
"The Flexible Smoother"

Methodology:

- Minimize the **penalized sum of squares** (**PSS**):

$$S(\boldsymbol{\theta}; \boldsymbol{y}, \lambda)_p = \|\boldsymbol{y} - \boldsymbol{B}\boldsymbol{\theta}\|^2 + \textbf{PENALTY}$$

- The **PENALTY** term, controls the smoothness of the fit by λ.

  - **Eilers & Marx (1996):**

    $\Rightarrow$ (*discrete*) **Penalty** over adjacent coefficients $\boldsymbol{\theta}$.

  - **Lang & Brezger (2004):**

    $\Rightarrow$ **"Bayesian P-splines"**: random walk priors for $\boldsymbol{\theta}$, e.g.:

$$\boldsymbol{\theta}|\boldsymbol{\theta}_{m-1} \sim \mathcal{N}(\boldsymbol{\theta}_{m-1}, \tau^2), \text{ or}$$
$$\boldsymbol{\theta}|\boldsymbol{\theta}_{m-1}, \boldsymbol{\theta}_{m-2} \sim \mathcal{N}(2\boldsymbol{\theta}_{m-1} - \boldsymbol{\theta}_{m-2}, \tau^2)$$

P-splines

"The Flexible Smoother"

- **PSS** becomes:

$$S(\theta; y, \lambda)_p = \|y - B\theta\|^2 + \theta' P\theta$$

  - $P = \lambda D'D$.
  - λ is the smoothing parameter.
  - D are difference matrices.

- For given λ , $\min S(\theta; y, \lambda)_p$

$$\widehat{\theta} = \left(B'B + \lambda D'D\right)^{-1} B'y$$

  - λ can be selected by **CV**, **GCV**, **AIC** or **BIC**.

P-splines
"The Flexible Smoother"

- **PSS** becomes:

$$S(\boldsymbol{\theta}; \boldsymbol{y}, \lambda)_p = \|\boldsymbol{y} - \boldsymbol{B\theta}\|^2 + \boldsymbol{\theta}'\boldsymbol{P\theta}$$

  - $\boldsymbol{P} = \lambda \boldsymbol{D}'\boldsymbol{D}$.
  - λ is the smoothing parameter.
  - $\boldsymbol{D}$ are difference matrices.

- For given λ , $\min S(\boldsymbol{\theta}; \boldsymbol{y}, \lambda)_p$

$$\widehat{\boldsymbol{\theta}} = \left(\boldsymbol{B}'\boldsymbol{B} + \lambda \boldsymbol{D}'\boldsymbol{D}\right)^{-1} \boldsymbol{B}'\boldsymbol{y}$$

  - λ can be selected by **CV**, **GCV**, **AIC** or **BIC**.

P-splines
"The Flexible Smoother"

- **PSS** becomes:

$$S(\boldsymbol{\theta}; \boldsymbol{y}, \lambda)_p = \|\boldsymbol{y} - \boldsymbol{B}\boldsymbol{\theta}\|^2 + \boldsymbol{\theta}'\boldsymbol{P}\boldsymbol{\theta}$$

  - $\boldsymbol{P} = \lambda \boldsymbol{D}'\boldsymbol{D}$.
  - λ is the smoothing parameter.
  - $\boldsymbol{D}$ are difference matrices.

- For given λ, $\min S(\boldsymbol{\theta}; \boldsymbol{y}, \lambda)_p$

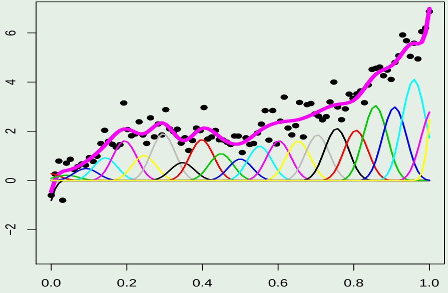
$$\widehat{\boldsymbol{\theta}} = \left(\boldsymbol{B}'\boldsymbol{B} + \lambda \boldsymbol{D}'\boldsymbol{D}\right)^{-1} \boldsymbol{B}'\boldsymbol{y}$$

  - λ can be selected by **CV**, **GCV**, **AIC** or **BIC**.

▶ $1d$ P-splines:

- No penalty over coefficients.
- Penalty over coefficients.

Example:



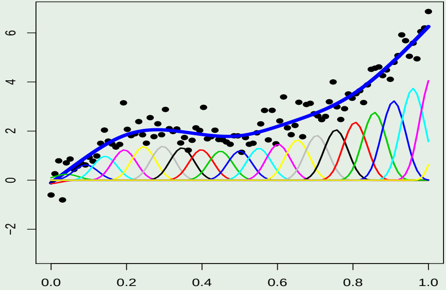
B-splines basis and $\boldsymbol{\theta}$ **without** penalty

P-splines
"The Flexible Smoother"

▶ $1d$ P-**splines:**

- No penalty over coefficients.
- Penalty over coefficients.

Example:



B-splines basis and $\boldsymbol{\theta}$ **with** penalty

Advantages over other smoothers:

- *Low-Rank*: "`dim(`**B**`)` $<$ `dim(data)`".
- Computationally efficient: "`# knots` $\leq$ `40`".
- Selection of number and Location of knots is **NOT** and issue.
- Discrete Penalties over the $\boldsymbol{\theta}$, not over the fitted curve.
- Easy extension to:
  - Mixed models,
  - non-gaussian data (GLM's) and
  - Multidimensional smoothing.
  - Spatial and Spatio-temporal smoothing.

» Mixed models

- **Reformulate:**
  - Model $y = B\theta + \epsilon$, into

$$y = X\beta + Z\alpha + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma^2 I)$$

  - where X and Z are *"fixed"* and *"random"* effects matrices.
  - with coefficients β and $\alpha \sim \mathcal{N}(0, G)$, and $G = \sigma_\alpha^2 R$
  - $\lambda = \frac{\sigma^2}{\sigma_\alpha^2}$

P-splines
A mixed model approach

- **Reformulate:**
  - Model $\boldsymbol{y} = \boldsymbol{B\theta} + \boldsymbol{\epsilon}$, into

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z\alpha} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \boldsymbol{\sigma}^2 \boldsymbol{I})$$

    - where $\boldsymbol{X}$ and $\boldsymbol{Z}$ are *"fixed"* and *"random"* effects matrices.
    - with coefficients β and $\boldsymbol{\alpha} \sim \mathcal{N}(0, \boldsymbol{G})$, and $\boldsymbol{G} = \boldsymbol{\sigma}^2_{\boldsymbol{\alpha}} \boldsymbol{R}$
    - $\lambda = \frac{\boldsymbol{\sigma}^2}{\boldsymbol{\sigma}^2_{\boldsymbol{\alpha}}}$

» Reparameterization

- **Reparameterization:**

$$B \equiv [\, X : Z \,] \Rightarrow B\theta = X\beta + Z\alpha$$

  - We use the **Singular Value Decomposition** (**SVD**) on $D'D$

» SVD

P-splines
A mixed model approach

- **Singular Value Decomposition (SVD)**

$$D'D = U\Sigma U'$$

- with $U = [U_n : U_s]$

$$D'D = [U_n : U_s] \left[\begin{array}{c|c} 0_d & \\ \hline & \widetilde{\Sigma} \end{array}\right] \left[\begin{array}{c} U_n' \\ \hline U_s' \end{array}\right]$$

  - $\widetilde{\Sigma} \equiv$ non-null eigenvalues.
  - $U_n \equiv$ eigenvectors corresponding to the null eigenvalues.
  - $U_s \equiv$ eigenvectors corresponding to the non-null eigenvalues.

- The **fix effects** (β) are **unpenalized** and
- The **Penalty** $\theta' P \theta$ becomes

$$\alpha' F \alpha$$

where $F = \lambda\widetilde{\Sigma}$ is **diagonal**.

- And the **random effects** (α) covariance matrix G:

$$G = \sigma^2 F^{-1}$$

- **Mixed Model Basis:**

$$X = [\,\mathbf{1} : x\,]$$
$$Z = BU_s$$

Advantages:

- Flexibility:
  - Easy incorporation of smoothing in complex models (*“spatial”* random effects and/or correlated errors).
- Mixed Models Theory:
  - Estimation and Inference.
- Software Implementation.
  - R, Splus, MATLAB or SAS.
- Extension to non-gaussian data:
  - Generalized Linear Mixed Models (**GLMM**)

Multidimensional P-splines

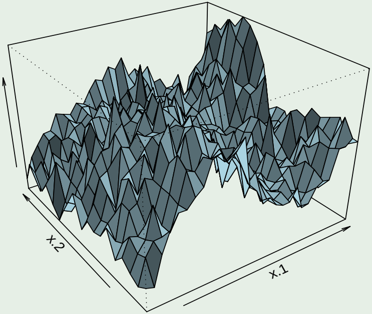
Example: 2d-array

- **Data** $\boldsymbol{Y} = \boldsymbol{y}_{ij}$, $i = 1, \ldots, n_1$ and $j = 1, \ldots, n_2$
- **Array structure**: n_1 rows and n_2 columns

$$\boldsymbol{Y} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n_2} \\ y_{21} & y_{22} & \cdots & y_{2n_2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n_1 1} & \cdots & \cdots & y_{n_1 n_2} \end{bmatrix}$$

- **Regressors:**

$$\boldsymbol{x}_1 = (x_{11}, \cdots, x_{1n_1})'$$
$$\boldsymbol{x}_2 = (x_{21}, \cdots, x_{2n_2})'$$



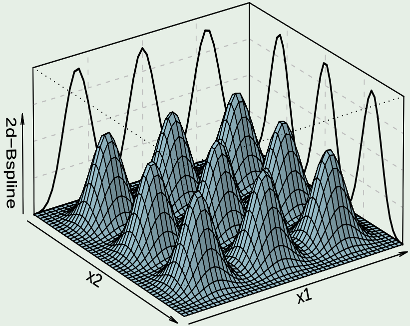
Multidimensional P-splines

- **Use of Tensor Products of B-splines (Durbán et al, 2002):**

Example: 2d-array

- **Marginal Basis:**
  - $B_1 = B_1(x_1)$, of dim. $n_1 \times c_1$.
  - $B_2 = B_2(x_2)$, of dim. $n_2 \times c_2$.
- $2d$ B**-splines Basis:**
  - Kronecker Product ($\otimes$) of marginal basis:

$$B = B_2 \otimes B_1, \quad \text{of dim. } n_1 n_2 \times c_1 c_2$$



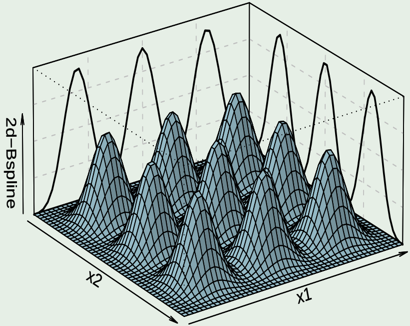
Multidimensional P-splines

- **Use of Tensor Products of B-splines (Durbán et al, 2002):**

Example: 2d-array

- **Marginal Basis:**
  - $\boldsymbol{B}_1 = \boldsymbol{B}_1(\boldsymbol{x}_1)$, of dim. $n_1 \times c_1$.
  - $\boldsymbol{B}_2 = \boldsymbol{B}_2(\boldsymbol{x}_2)$, of dim. $n_2 \times c_2$.
- $2d$ B**-splines Basis:**
  - Kronecker Product ($\otimes$) of marginal basis:

$$\boldsymbol{B} = \boldsymbol{B}_2 \otimes \boldsymbol{B}_1, \quad \text{of dim. } n_1 n_2 \times c_1 c_2$$



Multidimensional P-splines

Model:

$$y = f(x_1, x_2) + \epsilon,$$

with $y_{n_1 n_2 \times 1}$

- **In matrix form**, $\hat{y} = B\theta$ can be written as:

$$\hat{Y} = B_1 A B_2, \text{ of dim } n_1 \times n_2$$

where A is a matrix $c_1 \times c_2$ of coefficients θ of length $c_1 c_2 \times 1$.

IDEA:

- Set penalties over Θ.
  - **Row-wise** Penalty: $\theta' \left(I_{c_2} \otimes D_1' D_1\right) \theta$
  - **Column-wise** Penalty: $\theta' \left(D_2' D_2 \otimes I_{c_1}\right) \theta$

Multidimensional P-splines

Model:

$$y = f(x_1, x_2) + \epsilon,$$

with $y_{n_1 n_2 \times 1}$

- **In matrix form**, $\widehat{y} = B\theta$ can be written as:

$$\widehat{Y} = B_1 A B_2, \text{ of dim } n_1 \times n_2$$

where A is a matrix $c_1 \times c_2$ of coefficients θ of length $c_1 c_2 \times 1$.

IDEA:

- Set penalties over Θ.
  - **Row-wise** Penalty: $\theta' \left(I_{c_2} \otimes D_1' D_1 \right) \theta$
  - **Column-wise** Penalty: $\theta' \left(D_2' D_2 \otimes I_{c_1} \right) \theta$

Multidimensional P-splines

- **Penalty Matrix in $2d$:**

$$\boldsymbol{P} = \lambda_1 \underbrace{\boldsymbol{I}_{c_2} \otimes \boldsymbol{D}_1'\boldsymbol{D}_1}_{\boldsymbol{P}_1} + \lambda_2 \underbrace{\boldsymbol{D}_2'\boldsymbol{D}_2 \otimes \boldsymbol{I}_{c_1}}_{\boldsymbol{P}_2}$$

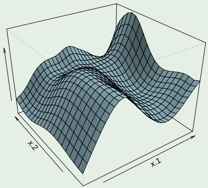
  - λ_1 and λ_2 are the smoothing parameters in each dimension.
  - **Anisotropy:** ($\lambda_1 \neq \lambda_2$)

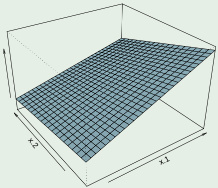
- As in $1d$ **Case**:

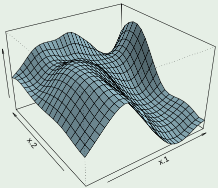
Example:

- The Mixed Model consists of:

$$\widehat{\boldsymbol{y}} \quad = \quad \underset{\text{(Linear/Fixed)}}{\boldsymbol{X}\widehat{\boldsymbol{\beta}}} \quad + \quad \underset{\text{(Non-Linear/Random)}}{\boldsymbol{Z}\widehat{\boldsymbol{\alpha}}}$$







Multidimensional P-splines

- **Mixed Models Representation:**
  - As in $1d$ **case**, the **aim** is:

$$\boldsymbol{B} \equiv [\, \boldsymbol{X} : \boldsymbol{Z} \,] \Longrightarrow \boldsymbol{B}\boldsymbol{\theta} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{\alpha}$$

  - The SVD over $\boldsymbol{P}$ allows the simultaneous diagonalization of $\boldsymbol{D}_1'\boldsymbol{D}_1$ and $\boldsymbol{D}_2'\boldsymbol{D}_2$
  - The penalty $\boldsymbol{P}$ becomes $\boldsymbol{F}$ (**block diagonal matrix**):

$$\boldsymbol{F} = \begin{pmatrix} \lambda_2 \widetilde{\boldsymbol{\Sigma}}_2 \otimes \boldsymbol{I}_2 & & \\ & \lambda_1 \boldsymbol{I}_2 \otimes \widetilde{\boldsymbol{\Sigma}}_1 & \\ & & \lambda_1 \boldsymbol{I}_{c_2-2} \otimes \widetilde{\boldsymbol{\Sigma}}_1 + \lambda_2 \widetilde{\boldsymbol{\Sigma}}_2 \otimes \boldsymbol{I}_{c_1-2} \end{pmatrix}$$

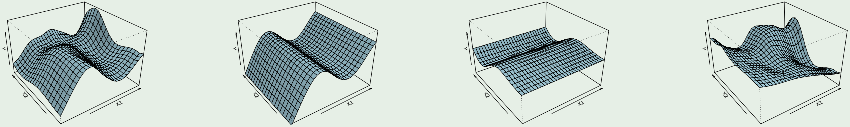
» Model

Multidimensional P-splines

ANOVA-type Decomposition of Smooth Surfaces:

$$\hat{y} = f(x_1) + f(x_2) + f(x_1, x_2)$$

(additive term for x_1) (additive term for x_2) (interaction term for x_1, x_2)



» Advantages

Multidimensional P-splines

Advantages:

- Extension to d-dimensions:

$$B = B_2 \otimes B_1 \otimes \cdots \otimes B_d$$

- **Efficient algorithms**:
  - **Currie et al (2006)**: Generalized Linear Array Models (**GLAM**)
- **Anisotropy** (different smoothing for each dimension):
- **Complex models:** spatial data smoothing

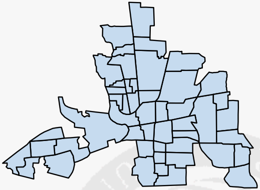
Outline

1. P-splines
   - Mixed models approach
   - Multidimensional P-splines
2. **P-splines for spatial count data**
   - Spatial smoothing
   - Smooth-CAR model
   - Application: Scottish Lip Cancer data
3. Spatio-temporal data Smoothing with P-splines
   - ANOVA-Type Interaction Models
   - Application Environmental spatio-temporal data
4. Spatio-temporal Disease Mapping

P-splines for spatial count data

P-splines for spatial smoothing

- **We propose:**



- $2d$ P**-splines:**
- **Geostatistics:** at sampling locations.
- **Regional/areal:** at the **centroids.**

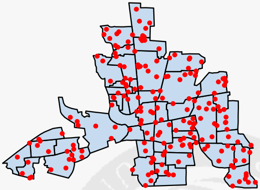
- **Models of the form:**

$$y = f(lon, lat) + \epsilon$$

where

- $f(lon, lat)$ is a **large-scale spatial smooth trend**: $X\beta + Z\alpha$.
- The **mixed model** allows the simultaneous estimation of **smoothing** and **spatial correlation**.

- **We propose:**



- $2d$ P-**splines:**
- **Geostatistics:** at sampling locations.
- **Regional/areal:** at the **centroids.**

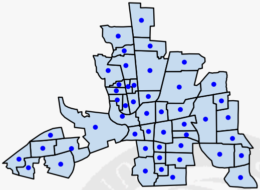
- **Models of the form:**

$$y = f(lon, lat) + \epsilon$$

where

- $f(lon, lat)$ is a **large-scale spatial smooth trend**: $X\beta + Z\alpha$.
- The **mixed model** allows the simultaneous estimation of **smoothing** and **spatial correlation**.

- **We propose:**



- $2d$ P-**splines:**
- **Geostatistics:** at sampling locations.
- **Regional/areal:** at the **centroids.**

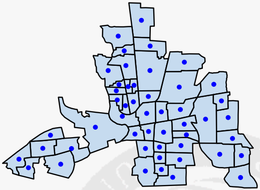
- **Models of the form:**

$$y = f(lon, lat) + \epsilon$$

where

- $f(lon, lat)$ is a **large-scale spatial smooth trend**: $X\beta + Z\alpha$.
- The **mixed model** allows the simultaneous estimation of **smoothing** and **spatial correlation**.

- **We propose:**



- $2d$ P**-splines:**
- **Geostatistics:** at sampling locations.
- **Regional/areal:** at the **centroids.**

- **Models of the form:**

$$y = f(lon, lat) + \epsilon$$

where

- $f(lon, lat)$ is a **large-scale spatial smooth trend**: $X\beta + Z\alpha$.
- The **mixed model** allows the simultaneous estimation of **smoothing** and **spatial correlation**.

- **B-spline Basis for spatial data:**
  - Given that data are **NOT** in an array

$$B = B_2 \otimes B_1 \quad \text{replace by} \quad B_2 \Box B_1$$

$\Box$ denotes the "Row-wise Kronecker" or **Box-Product**.

Def. Box-Product:

$$B_2 \Box B_1 = (B_2 \otimes \mathbf{1}_{c_1}) \odot (\mathbf{1}_{c_2} \otimes B_1)$$

$\odot$ is the "element-wise" product.

P-splines for spatial count data

- **In many applications:**
  - Collect count data **observed in regions or areas**.
    - **E.g.:** $\#$ of cases of disease or deaths
  - Counts are **Poisson** distributed.

$$y \sim \mathcal{P}(\mu)$$

Penalized-GLMM

- P-splines as **mixed models**:
  - **Linear Predictor:**

$$\eta = B\theta \Longrightarrow X\beta + Z\alpha$$

  - **Penalized log-Likelihood:**

$$\ell_p(\beta, \alpha; y) = \ell(\beta, \alpha; y) - \frac{1}{2}\alpha' F \alpha$$

- **Estimation via PQL**

P-splines for spatial count data

- **In many applications:**
  - Collect count data **observed in regions or areas**.
    - **E.g.:** $\#$ of cases of disease or deaths
  - Counts are **Poisson** distributed.

$$y \sim \mathcal{P}(\mu)$$

Penalized-GLMM

- P-splines as **mixed models**:
  - **Linear Predictor:**

$$\eta = B\theta \Longrightarrow X\beta + Z\alpha$$

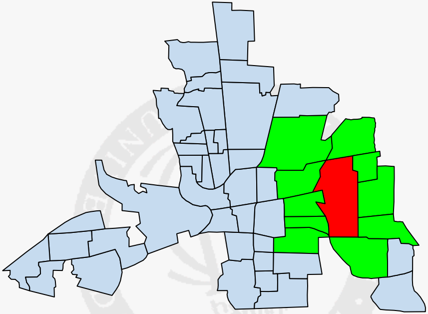
  - **Penalized log-Likelihood:**

$$\ell_p(\beta, \alpha; y) = \ell(\beta, \alpha; y) - \frac{1}{2}\alpha' F\alpha$$

- **Estimation via PQL**

Smooth-CAR model
CAR model

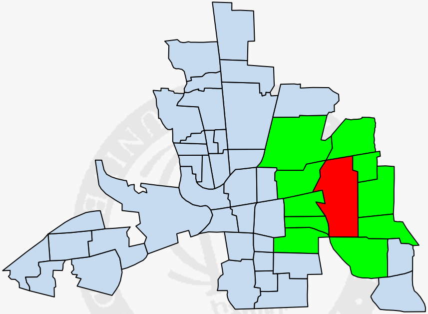
- **Most popular approach:**



- Conditional Autoregressive Models (**CAR**), Besag (1991)
- Spatial Dependence across "**neighbours**".
- Different **neighbourhood criteria**.
  - Common border.
  - Centroids distance, 4-nearest neighbours.

Smooth-CAR model
CAR model

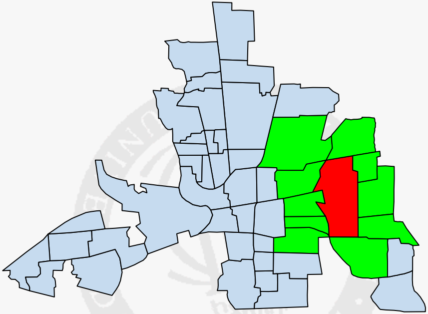
- **Most popular approach:**



- Conditional Autoregressive Models (**CAR**), Besag (1991)
- Spatial Dependence across **"neighbours"**.
- Different **neighbourhood criteria**.
  - Common border.
  - Centroids distance, 4-nearest neighbours.

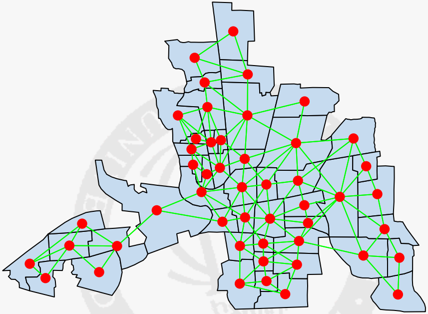
Smooth-CAR model
CAR model

- **Most popular approach:**



- Conditional Autoregressive Models (**CAR**), Besag (1991)
- Spatial Dependence across **"neighbours"**.
- Different **neighbourhood criteria**.
  - Common border.
  - Centroids distance, 4-nearest neighbours.

- **Most popular approach:**



- Conditional Autoregressive Models (**CAR**), Besag (1991)
- Spatial Dependence across **"neighbours"**.
- Different **neighbourhood criteria**.
  - Common border.
  - Centroids distance, 4-nearest neighbours.

- **Formulation:**

$$y = X\beta + b,$$

where $b = (b_1, b_2, ..., b_n)'$ is a vector for the **spatial effects**

- Impose a **spatial dependency structure** by a prior distribution for b:

$$b \sim \mathcal{N}(0, G_b)$$

where G_b depends on the "**neighbourhood structure**":

  - defined by **Contiguity matrix** (Q)

Smooth-CAR model
CAR model

- **Formulation:**

$$y = X\beta + b,$$

where $b = (b_1, b_2, ..., b_n)'$ is a vector for the **spatial effects**

- Impose a **spatial dependency structure** by a prior distribution for b:

$$b \sim \mathcal{N}(0, G_b)$$

where G_b depends on the "**neighbourhood structure**":

  - defined by **Contiguity matrix** (Q)

Smooth-CAR model
CAR model

✓ We follow an **Empirical Bayes** approach:

▶ **Intrinsic CAR:**

$$G_b = \sigma_b^2 Q^- + \kappa^{-1} I$$ **(Besag, 1991)**

- Two independent and separate variance components:
  - ▶ **Spatially-structured variation:** $\sigma_b^2 Q^-$
  - ▶ **Unstructured non-spatial correlation:** $\kappa^{-1} I$

▶ **Alternative CAR models structures:**

$$G_b = \sigma_b^2 \left(\phi Q + (1-\phi) I\right)^{-1}$$ **(Leroux et al, 1999)**

$$G_b = \sigma_b^2 \left(\phi Q^- + (1-\phi) I\right)$$ **(Dean et al, 2001)**

where

- ▶ ϕ measures the relative weight between *structured* and *unstructured* variability
- ▶ $0 \leq \phi \leq 1$

Smooth-CAR model
CAR model

✓ We follow an **Empirical Bayes** approach:

▶ **Intrinsic CAR:**

$$G_b = \sigma_b^2 Q^- + \kappa^{-1} I$$ **(Besag, 1991)**

- Two independent and separate variance components:
  - ▶ **Spatially-structured variation:** $\sigma_b^2 Q^-$
  - ▶ **Unstructured non-spatial correlation:** $\kappa^{-1} I$

▶ **Alternative CAR models structures:**

$$G_b = \sigma_b^2 \left(\phi Q + (1-\phi) I\right)^{-1}$$ **(Leroux et al, 1999)**

$$G_b = \sigma_b^2 \left(\phi Q^- + (1-\phi) I\right)$$ **(Dean et al, 2001)**

where

▶ ϕ measures the relative weight between *structured* and *unstructured* variability

▶ $0 \leq \phi \leq 1$

Smooth-CAR model
Lee and Durban (2009)

- **We propose a "hybrid" model:**
  - **Spatial P-spline** with **CAR** structure: **"Smooth-CAR" model**
  - **Model:**

$$\boldsymbol{\eta} = \boldsymbol{X\beta} + \boldsymbol{Z\alpha} + \boldsymbol{b}\ ,$$

where $\boldsymbol{b} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{G}_b)$

Our approach:

$$\eta = \underbrace{\text{Spatial Trend}}_{\substack{X\beta + Z\alpha \\ \text{(Large-scale)}}} + \underbrace{\text{Local area-level spatial correlation}}_{\substack{\text{Spatial Random Effects} \\ \text{(Small-scale)}}}$$

Smooth-CAR model

Lee and Durban (2009)

- **We propose a "hybrid" model:**
  - **Spatial P-spline** with **CAR** structure: **"Smooth-CAR" model**
  - **Model:**

$$\eta = X\beta + Z\alpha + b \; ,$$

where $b \sim \mathcal{N}(\mathbf{0}, G_b)$

Our approach:

$$\eta \;=\; \underbrace{\text{Spatial Trend}}_{\substack{X\beta + Z\alpha \\ \text{(Large-scale)}}} \;+\; \underbrace{\text{Local area-level spatial correlation}}_{\substack{\text{Spatial Random Effects} \\ \text{(Small-scale)}}}$$

Smooth-CAR model

- **Summary:**

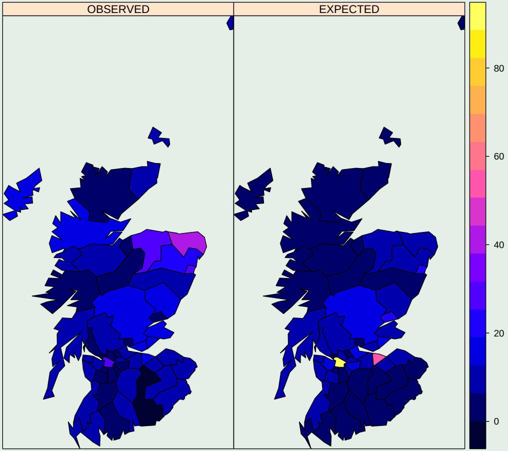
Model	Linear Predictor	Area-level var.
Poisson	$X\beta + Z\alpha$	$-$
CAR	$X\beta + b$	$b \sim \mathcal{N}(0, G_b)$
Smooth-CAR	$X\beta + Z\alpha + b$	$b \sim \mathcal{N}(0, G_b)$

- The **Smooth-CAR**:
  - Allow us model the **spatial trend** ($X\beta + Z\alpha$) along large geographical distances and
  - **Local area-level** correlation by a **CAR** component (b).

Application: Scottish Lip Cancer data

Example: Scottish Lip Cancer

- **Breslow and Clayton (1993)**
- **Observed** (y) and **Expected** (e) cases of lip cancer
- 56 **counties** in Scotland
- **Period:** $1975 - 1980$.



- **We fit several models:**

  - Smooth P-**splines** models:

$$\boldsymbol{\eta} = \log(\boldsymbol{e}) + \boldsymbol{X\beta} + \boldsymbol{Z\alpha} \qquad \textbf{(Poisson)}$$

  $\log(\boldsymbol{e})$ is the **offset** term.

  - **CAR models:**

$$\boldsymbol{\eta} = \log(\boldsymbol{e}) + \boldsymbol{X\beta} + \boldsymbol{b}\,, \quad \boldsymbol{b} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{G}_b),$$

  with:

$$\boldsymbol{G}_b = \sigma_b^2\,(\phi \boldsymbol{Q}^- + (1-\phi)\boldsymbol{I}) \qquad \textbf{(Dean)}$$

  - **Smooth-CAR model:**

$$\boldsymbol{\eta} = \log(\boldsymbol{e}) + \boldsymbol{X\beta} + \boldsymbol{Z\alpha} + \boldsymbol{b}\,, \quad \boldsymbol{b} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{G}_b)$$

Application: Scottish Lip Cancer data

Fitted Models

- **We fit several models:**
  - Smooth P-**splines** models:

$$\boldsymbol{\eta} = \log(\boldsymbol{e}) + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{\alpha} \qquad \textbf{(Poisson)}$$

    $\log(\boldsymbol{e})$ is the **offset** term.

  - **CAR** models:

$$\boldsymbol{\eta} = \log(\boldsymbol{e}) + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{b}\ , \quad \boldsymbol{b} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{G}_b),$$

    with:

$$\boldsymbol{G}_b = \sigma_b^2\ (\phi \boldsymbol{Q}^{-} + (1-\phi)\boldsymbol{I}) \qquad \textbf{(Dean)}$$

  - Smooth-CAR model:

$$\boldsymbol{\eta} = \log(\boldsymbol{e}) + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{\alpha} + \boldsymbol{b}\ , \quad \boldsymbol{b} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{G}_b)$$

Application: Scottish Lip Cancer data
Fitted Models

- **We fit several models:**
  - Smooth P**-splines** models:

$$\eta = \log(e) + X\beta + Z\alpha \qquad \textbf{(Poisson)}$$

$\log(e)$ is the **offset** term.

  - **CAR** models:

$$\eta = \log(e) + X\beta + b \ , \quad b \sim \mathcal{N}(0, G_b),$$

with:

$$G_b = \sigma_b^2 \left(\phi Q^- + (1-\phi)I\right) \qquad \textbf{(Dean)}$$

  - **Smooth-CAR model:**

$$\eta = \log(e) + X\beta + Z\alpha + b \ , \quad b \sim \mathcal{N}(0, G_b)$$

Application: Scottish Lip Cancer data
Models comparison criteria

- **In order to compare the proposed models we use:**

$$\textbf{AIC} = \textbf{Dev} + 2 \times \textbf{df}$$
$$\textbf{BIC} = \textbf{Dev} + \log(n) \times \textbf{df}$$

where:

- **df** is the effective dimension of the model (“degrees of freedom”).
  - is a measure of the complexity of the fitted model,
  - Calculated as the trace(H),

$$\hat{y} = Hy$$

Application: Scottish Lip Cancer data
Models comparison criteria

- **In order to compare the proposed models we use:**

$$\textbf{AIC} = \textbf{Dev} + 2 \times \textbf{df}$$
$$\textbf{BIC} = \textbf{Dev} + \log(n) \times \textbf{df}$$

where:

- **df** is the effective dimension of the model ("degrees of freedom").
  - is a measure of the complexity of the fitted model,
  - Calculated as the $\mathrm{trace}(H)$,

$$\widehat{y} = Hy$$

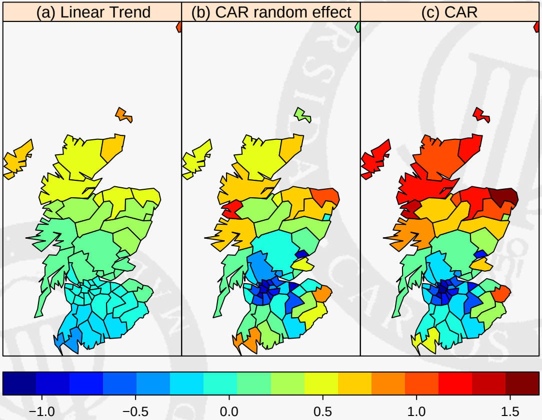
Application: Scottish Lip Cancer data

Comparisons of fitted models

Model		Parameters λ_1	λ_2	σ_s^2	κ^{-1}	ϕ	AIC	BIC	df
Smooth:	*Poisson*	11.75	3.63	-	-	-	114.04	228.46	15.90
CAR:	*Dean*	-	-	0.78	-	0.99	89.36	179.56	32.78
Smooth-CAR:	*Dean*	30.11	16.37	0.53	-	0.97	87.46	175.70	30.64

- **Observations:**
  - $\phi \approx 1 \longrightarrow$ Overdispersion is due to *"structured"* spatial correlation ($\sigma_b^2 Q^-$).
  - **Smooth-CAR** performs better in terms of the selected criteria.

▶ Dean's CAR model:

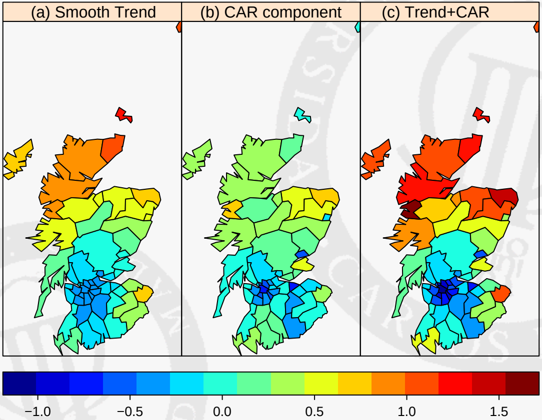


(a) Large-scale linear trend: $X\boldsymbol{\beta}$

(b) CAR structured random effects: $\boldsymbol{b} \sim \mathcal{N}(\mathbf{0}, G_b)$

(c) $X\boldsymbol{\beta} + \boldsymbol{b}$

▶ Smooth-CAR model:



(a) Smooth large-scale spatial trend: $X\beta + Z\alpha$

(b) CAR structured random effects: $b \sim \mathcal{N}(\mathbf{0}, G_b)$

(c) $X\beta + Z\alpha + b$

Outline

1. P-splines
   - Mixed models approach
   - Multidimensional P-splines
2. P-splines for spatial count data
   - Spatial smoothing
   - Smooth-CAR model
   - Application: Scottish Lip Cancer data
3. **Spatio-temporal data Smoothing with P-splines**
   - ANOVA-Type Interaction Models
   - Application Environmental spatio-temporal data
4. Spatio-temporal Disease Mapping

Spatio-temporal data

- **Response variable,** y_{ijt}
  - measured over **geographical locations**, $s = (x_i, x_j)$, with $i, j = 1, .., n$
  - and over **time periods**, x_t, for $t = 1, ..., T$
- **ISSUE:** huge amount of data available
  - **e.g. :** Environmental data, epidemiologic studies, disease mapping applications, ...
- **Smoothing techniques:**
  - Study spatial and temporal trends.
  - Space and time interactions.
  - ✓ 3**-dimensional smoothing**: P-splines and **GLAM**.

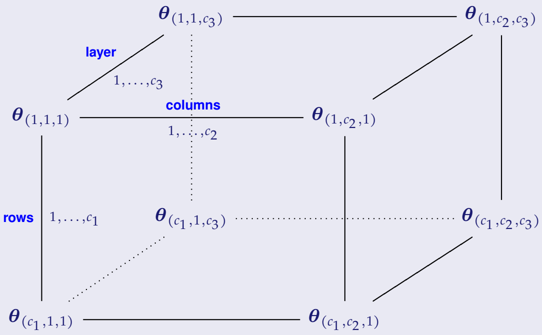
Example of GLAM in $3d$

Currie et. al (2006)

- $3d$-**case:**

$$f(x_1, x_2, x_3) = B\theta$$

- **Basis:** $B = B_1 \otimes B_2 \otimes B_3$
  - θ can be expressed as a $3d$-array $A = \{\theta\}_{ijk}$ of dim. $c_1 \times c_2 \times c_3$



- $3d$-**Penalty matrix:**
  - Set penalties over the $3d$-array A:

$$P = \lambda_1 \underbrace{D_1'D_1 \otimes I_{c_2} \otimes I_{c_3}}_{\textbf{row-wise}} + \lambda_2 \underbrace{I_{c_1} \otimes D_2'D_2 \otimes I_{c_3}}_{\textbf{column-wise}} + \lambda_t \underbrace{I_{c_1} \otimes I_{c_2} \otimes D_t'D_t}_{\textbf{layer-wise}}$$

  - For **spatio-temporal data**:

$$f(\underbrace{\text{longitude}, \text{latitude}}_{\textbf{Space}}, \textbf{time})$$

    - **Spatial anisotropy** ($\lambda_1 \neq \lambda_2$), different amount of smoothing for latitude and longitude.
    - **Temporal smoothing** (λ_t)
    - Space-time **interaction**.

- For **spatio-temporal data**, we propose:

B-splines Basis:

$$B = B_s \otimes B_t,$$

where

$B_s \equiv$ is the spatial B-spline basis $(B_1 \Box B_2)$ and

$B_t \equiv$ is the B-spline basis for time of dim. $t \times c_3$.

✓ **as GLAM:**

Given $y_{ijt} = Y_{t \times n}$, and $\theta_{ijt} = A_{c_t \times c_s}$, we have

$$\mathbb{E}[Y] = B_t A B_s'$$

✓ **as Mixed models**

$$B\theta = X\beta + Z\alpha$$

Smooth-ANOVA decomposition models

- **Chen (1993), Gu (2002):**
  - **"Smoothing-Spline ANOVA" (SS-ANOVA)**.
  - Interpretation as **"main effects"** and **"interactions"**.
  - **Models of type:**

$$
\begin{aligned}
\widehat{y} = \quad & f(x_1) + f(x_2) + f(x_t) && \text{"Main/additive effects"} \\
& + \\
& f(x_1, x_2) + f(x_1, x_t) + f(x_2, x_t) && \text{"2-way interactions"} \\
& + \\
& f(x_1, x_2, x_t) && \text{"3-way interactions"}
\end{aligned}
$$

  - **PROBLEMS:**
    - identifiability, and
    - basis dimension (*"curse of dimensionality"*)

P-spline ANOVA model

for spatio-temporal smoothing

- **Lee and Durbán (2009a)**, consider:

$$y = \gamma + f_s(x_1, x_2) + f_s(time) + f_{st}(x_1, x_2, time) + \epsilon \ ,$$

where

$$\begin{aligned} f_s(x_1, x_2) &\equiv \text{Spatial } 2d \text{ smooth surface} \\ f_t(time) &\equiv \text{Smooth time trend} \\ f_{st}(x_1, x_2, time) &\equiv \text{Space-time interaction} \end{aligned}$$

  - We need to construct an identifiable model.
  - **Our approach is based on:**
    - *low-rank* basis (P-splines)
    - the **mixed model representation** and **SVD** properties.

Basis, Coefficients and Penalty

- For each smooth term $f(\cdot)$, in spatio-temporal ANOVA model we have
  - $B-$spline **basis**:
$$B = [\mathbf{1}_{nt} : B_s \otimes \mathbf{1}_t : \mathbf{1}_n \otimes B_t : B_s \otimes B_t]$$
  - vector of **coefficients**:
$$\boldsymbol{\theta} = (\gamma, \boldsymbol{\theta}^{(s)'}, \boldsymbol{\theta}^{(t)'}, \boldsymbol{\theta}^{(st)'})'$$
  - and a blockdiagonal **Penalty**:
$$P = \begin{pmatrix} 0 & & & \\ & P_s & & \\ & & P_t & \\ & & & P_{st} \end{pmatrix},$$

  where

$$\begin{aligned} P_s &= 2d\text{-spatial penalty} \\ P_t &= 1d\text{-penalty for time} \\ P_{st} &= 3d \text{ space-time penalty} \end{aligned}$$

Basis, Coefficients and Penalty

- For each smooth term $f(\cdot)$, in spatio-temporal ANOVA model we have
  - $B-$spline **basis**:

$$\boldsymbol{B} = [\mathbf{1}_{nt} : \boldsymbol{B}_s \otimes \mathbf{1}_t : \mathbf{1}_n \otimes \boldsymbol{B}_t : \boldsymbol{B}_s \otimes \boldsymbol{B}_t]$$

  - vector of **coefficients**:

$$\boldsymbol{\theta} = (\gamma, \boldsymbol{\theta}^{(s)'}, \boldsymbol{\theta}^{(t)'}, \boldsymbol{\theta}^{(st)'})'$$

  - and a blockdiagonal **Penalty**:

$$\boldsymbol{P} = \begin{pmatrix} 0 & & & \\ & \boldsymbol{P}_s & & \\ & & \boldsymbol{P}_t & \\ & & & \boldsymbol{P}_{st} \end{pmatrix},$$

where

$$\begin{aligned} \boldsymbol{P}_s &= 2d\text{-spatial penalty} \\ \boldsymbol{P}_t &= 1d\text{-penalty for time} \\ \boldsymbol{P}_{st} &= 3d \text{ space-time penalty} \end{aligned}$$

Basis, Coefficients and Penalty

- For each smooth term $f(\cdot)$, in spatio-temporal ANOVA model we have
  - $B-$spline **basis**:
$$\boldsymbol{B} = [\mathbf{1}_{nt} : \boldsymbol{B}_s \otimes \mathbf{1}_t : \mathbf{1}_n \otimes \boldsymbol{B}_t : \boldsymbol{B}_s \otimes \boldsymbol{B}_t]$$
  - vector of **coefficients**:
$$\boldsymbol{\theta} = (\gamma, \boldsymbol{\theta}^{(s)'}, \boldsymbol{\theta}^{(t)'}, \boldsymbol{\theta}^{(st)'})'$$
  - and a blockdiagonal **Penalty**:
$$\boldsymbol{P} = \begin{pmatrix} 0 & & & \\ & \boldsymbol{P}_s & & \\ & & \boldsymbol{P}_t & \\ & & & \boldsymbol{P}_{st} \end{pmatrix},$$

where

$$\begin{aligned} \boldsymbol{P}_s &= 2d\text{-spatial penalty} \\ \boldsymbol{P}_t &= 1d\text{-penalty for time} \\ \boldsymbol{P}_{st} &= 3d \text{ space-time penalty} \end{aligned}$$

Basis, Coefficients and Penalty

- For each smooth term $f(\cdot)$, in spatio-temporal ANOVA model we have
  - $B-$spline **basis**:
$$\boldsymbol{B} = [\mathbf{1}_{nt} : \boldsymbol{B}_s \otimes \mathbf{1}_t : \mathbf{1}_n \otimes \boldsymbol{B}_t : \boldsymbol{B}_s \otimes \boldsymbol{B}_t]$$
  - vector of **coefficients**:
$$\boldsymbol{\theta} = (\gamma, \boldsymbol{\theta}^{(s)'}, \boldsymbol{\theta}^{(t)'}, \boldsymbol{\theta}^{(st)'})'$$
  - and a blockdiagonal **Penalty**:
$$\boldsymbol{P} = \begin{pmatrix} 0 & & & \\ & \boldsymbol{P}_s & & \\ & & \boldsymbol{P}_t & \\ & & & \boldsymbol{P}_{st} \end{pmatrix},$$

    where
$$\begin{aligned} \boldsymbol{P}_s &= 2d\text{-spatial penalty} \\ \boldsymbol{P}_t &= 1d\text{-penalty for time} \\ \boldsymbol{P}_{st} &= 3d \text{ space-time penalty} \end{aligned}$$

✓ However, B is **NOT full column-rank** (*"linear dependency"*)

✓ Model is **NOT identifiable**

Solution:

- Reparameterize as a mixed model (using SVD).
- For each term we have:

		Basis $[\,X : Z\,]$	
$f_s(x_1, x_2)$	$\equiv$	$x_1 : x_2$	(1)
$f_t(x_t)$	$\equiv$	x_t	(2)
$f_{st}(x_1, x_2, x_t)$	$\equiv$	$x_1 : x_2 : x_t$	(3)

- Some terms in (1) and (2) also appear in (3).

- The **mixed model representation**, allow us to **identify the columns to remove** in order to maintain the identifiability of the model.

  and obtain a blockdiagonal penalty $\boldsymbol{F}$

$$\boldsymbol{F} = \begin{pmatrix} \mathbf{0}_8 & & & \\ & \boldsymbol{F}_s & & \\ & & \boldsymbol{F}_t & \\ & & & \boldsymbol{F}_{st} \end{pmatrix}, \quad \text{with} \quad \begin{array}{l} \lambda_1, \lambda_2 \\ \lambda_t \\ \tau_1, \tau_2, \tau_t \end{array}$$

- **In P-splines context**, this is equivalent to

  ✓ apply constraints over **regression coefficients** $\boldsymbol{\theta}_{i,j,k}$

► For the ANOVA spatio-temporal model, the resultant mixed model reparameterization is equivalent to apply the next constraints:

- time effect coefficient:

$$\sum_{t=1}^{c_t} \boldsymbol{\theta}_t^{(\mathrm{t})} = 0,$$

- constraints over the spatio-temporal array of coefficients, $\boldsymbol{\Theta}^{(\mathrm{st})}$, of dimensions $c_t \times c_s$:

$$\sum_{i}^{c_1} \boldsymbol{\theta}_{t,ij}^{(\mathrm{st})} = \sum_{j}^{c_2} \boldsymbol{\theta}_{t,ij}^{(\mathrm{st})} = \sum_{i}^{c_1} \sum_{j}^{c_2} \boldsymbol{\theta}_{t,ij}^{(\mathrm{st})} = 0.$$

In practice

- We only need to construct the matrices X, Z and penalty F

			$f_s(x_1, x_2)$	$f_t(x_t)$	$f_{st}(x_1, x_2, x_t)$
X	$\equiv$	**by columns**	$x_1 : x_2$	x_t	(x_1, x_2, x_t)
Z	$\equiv$	**by blocks**	"	"	"
F	$\equiv$	**blockdiagonal**	F_s (λ_1, λ_2)	F_t λ_t	F_{st} (τ_1, τ_2, τ_t)

Ozone pollution in Europe

Lee and Durbán (2009a)

- Sample of 45 monitoring stations
- Monthly averages of O_3 levels (in $\mu g/m^3$ *units*)
- from january 1999 to december 2005 $(t = 1, \ldots, 84)$

Models:

- **Additive:**

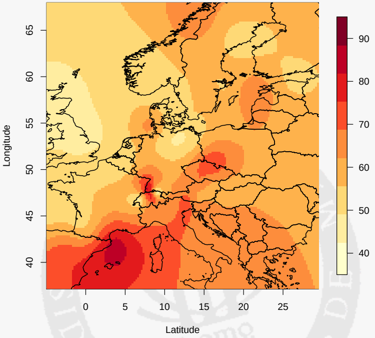
$$f_s(x_1, x_2) + f_t(x_t)$$

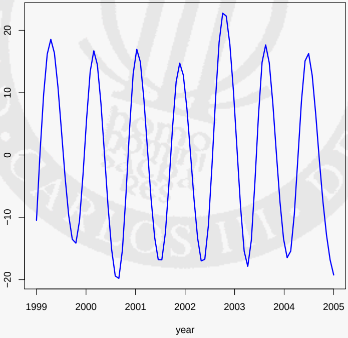
- **Spatio-temporal Interaction:**
  - ✓ **ANOVA:**

$$f_s(x_1, x_2) + f_t(x_t) + f_{st}(x_1, x_2, x_t)$$

Spatial 2d + time

$f_s(x_1, x_2)$ + $f_t(x_t)$

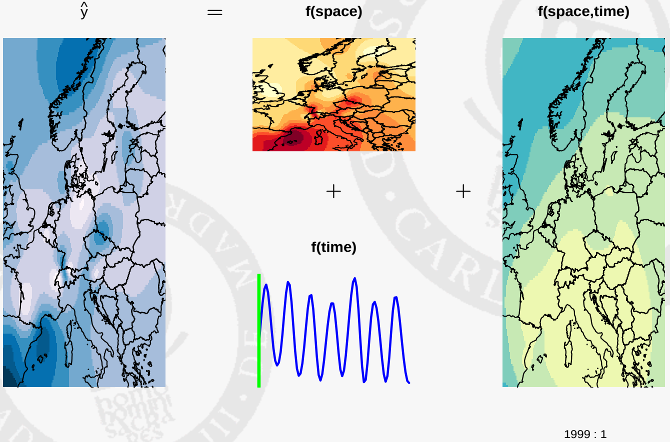




- ✓ Space-time interaction is **not** considered
- ✓ time smooth trend is **additive**

Spatio-temporal ANOVA model

Play animation

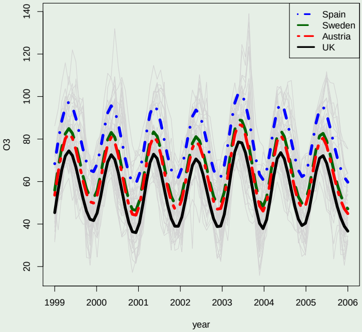


Comparison of fitted values

Additive *VS* ANOVA

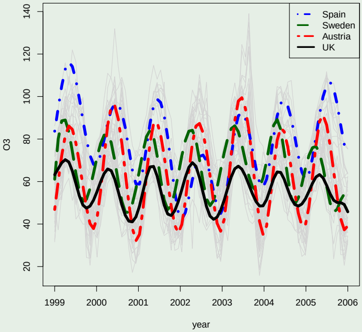
▶ Additive model fit

$$f_s(x_1,x_2)+f_s(x_t)$$



▶ ANOVA model fit

$$f_s(x_1,x_2)+f_t(x_t)+f_{st}(x_1,x_2,x_t)$$



✓ **Additive model** assumes a spatial smooth surface over all monitoring stations that remains constant over time.

✓ **ANOVA model** captures individual characteristics of the stations throughout time.

Comparison of Models

ANOVA and Additive

Model	AIC	df
ANOVA	14280.73	366.03
Additive	16506.28	65.98

- **Observations:**
  - Best overall performance of ANOVA in terms of **AIC**
  - **ANOVA model** is more realistic than Additive, and easier to decompose and interpret in terms of the fit.

Outline

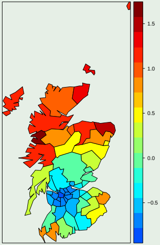
1. P-splines
   - Mixed models approach
   - Multidimensional P-splines
2. P-splines for spatial count data
   - Spatial smoothing
   - Smooth-CAR model
   - Application: Scottish Lip Cancer data
3. Spatio-temporal data Smoothing with P-splines
   - ANOVA-Type Interaction Models
   - Application Environmental spatio-temporal data
4. **Spatio-temporal Disease Mapping**

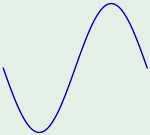
Spatio-temporal Disease Mapping

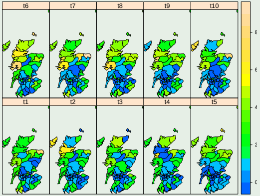
P-spline ANOVA model for disease mapping

- Y and E are $t \times n$ arrays of observed and expected cases of disease over t time periods, and $M = \log(\frac{Y}{E})$.
- Consider an ANOVA model for η

$$f_s(x_1, x_2) \quad + \quad f_t(x_t) \quad + \quad f_{st}(x_1, x_2, x_t)$$







Summary

- New flexible approach for spatial and spatio-temporal data smoothing:
  - based on P-splines as mixed models and
  - ANOVA decomposition
- Methodology also extensible for disease mapping applications.
- Computationally efficient algorithms (GLAM)

Bibliography

- Lee, D.-J. and Durbán, M. (2009)
  Smooth-CAR mixed models for spatial count data.
  Computational Statistics and Data Analysis 53(8):2968-2979.
- Lee, D.-J. and Durbán, M. (2009)
  P*-spline ANOVA-Type interaction models for spatio-temporal smoothing.*
  Submitted.
- Eilers, PHC., Currie, ID. and Durbán, M. (2006)
  Fast and compact smoothing on large multidimensional grids.
  Computational Statistics and Data Analysis, 50(1):61-76.
- Currie, ID., Durbán M. and Eilers, PHC. (2006)
  Generalized linear array models with applications to multidimensional smoothing.
  Journal of the Royal Statistical Society B, 68:1-22.