Space-time modelling of air pollution with array methods

Dae-Jin Lee

Universidad Carlos III de Madrid
Department of Statistics

Royal Statistical Society Conference
Edinburgh 2009
Motivation

- Increasing research on **modelling spatio-temporal** data
- Wide variety of approaches and research communities
  - *E.g.*: Environmental data, epidemiologic studies, disease mapping applications, ...

**European Environmental Agency (EEA):**

- Monitoring networks
- **EMEP project** (European Monitoring and Evaluation Programme)
- **Ozone** ($O_3$) is currently one of the air pollutants of most concern in Europe.
Monitoring stations across Europe

sample of 45 monitoring stations

Monitoring station

D.-J. Lee (Uc3m)  GLAM: Array methods in Statistics  RSS ’09 - Edinburgh
Seasonal pattern:
Motivation

- 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

Smoothing techniques:

- Study spatial and temporal trends.
- Space and time interactions.

✓ 3-dimensional smoothing: $P$-splines and GLAM.
Outline

1. $P$-splines in spatio-temporal smoothing context
2. ANOVA-Type Interaction Models
3. Application to air pollution data
4. Concluding remarks
Outline

1. $P$-splines in spatio-temporal smoothing context
2. ANOVA-Type Interaction Models
3. Application to air pollution data
4. Concluding remarks
Penalized splines

**GLAM in 3d**

- **3d-case:**
  \[ f(x_1, x_2, x_3) = B\theta \]

  - \( \theta \) can be expressed as a 3d-array \( \Theta = \{\theta\}_{ijk} \) of dim. \( c_1 \times c_2 \times c_3 \)
• **3d-Penalty matrix:**

- Set penalties over the **3d-array** \( \Theta \):

\[
P = \lambda_1 D_1' D_1 \otimes I_{c_2} \otimes I_{c_3} + \lambda_2 I_{c_1} \otimes D_2' D_2 \otimes I_{c_3} + \lambda_t I_{c_1} \otimes I_{c_2} \otimes D_t' D_t
\]

  - row-wise
  - column-wise
  - layer-wise

- For **spatio-temporal data**:

\[
f(\text{longitude, latitude, time})
\]

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

▶ Spatio-Temporal data smoothing

- For spatio-temporal data, we propose:

**Spatio-temporal \( B \)-splines Basis:**

\[
B = B_s \otimes B_t, \quad \text{of dim. } nt \times c_1 c_2 c_3
\]

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. \( t \times c_3 \).

- Note that:
  
  - GLAM framework
    
    \[
    \mathbb{E}[Y] = B_t \Theta B_s'
    \]
  
  - Mixed model
Penalized splines
▶ Spatio-Temporal data smoothing

Most of the space-time approaches consider

- Simplest **additive model** (ignores space-time interaction):
  \[ f_s(x_1, x_2) + f_t(x_t) \]

We proposed **space-time interaction models**

- 3d: \[ f_{st}(x_1, x_2, x_t) \]

Space-time ANOVA:

\[ f_s(x_1, x_2) + f_t(x_t) + f_{st}(x_1, x_2, x_t) \]
Penalized splines
▶ Spatio-Temporal data smoothing

▶ Most of the space-time approaches consider

• Simplest additive model (ignores space-time interaction):

\[ f_s(x_1, x_2) + f_t(x_t) \]

▶ We proposed space-time interaction models

• 3d:

\[ f_{st}(x_1, x_2, x_t) \]

✓ Space-time ANOVA:

\[ f_s(x_1, x_2) + f_t(x_t) + f_{st}(x_1, x_2, x_t) \]
Penalized splines

- Spatio-Temporal data smoothing

Most of the space-time approaches consider

- Simplest **additive model** (ignores space-time interaction):

\[ f_s(x_1, x_2) + f_t(x_t) \]

We proposed **space-time interaction models**

- 3d:

\[ f_{st}(x_1, x_2, x_t) \]

✓ **Space-time ANOVA**:

\[ f_s(x_1, x_2) + f_t(x_t) + f_{st}(x_1, x_2, x_t) \]
Outline

1. *P*-splines in spatio-temporal smoothing context
2. ANOVA-Type Interaction Models
3. Application to air pollution data
4. Concluding remarks
ANOVA-Type Interaction Models

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

    \[
    \hat{y} = f(x_1) + f(x_2) + f(x_t) + f(x_1, x_2) + f(x_1, x_t) + f(x_2, x_t) + f(x_1, x_2, x_t)
    \]

    “Main/additive effects”
    “2-way interactions”
    “3-way interactions”

- PROBLEMS:
  - identifiability, and
  - basis dimension (“curse of dimensionality”)

PROBLEMS:
Lee and Durbán (2009b), consider:

\[ y = \gamma + f_s(x_1, x_2) + f_t(\text{time}) + f_{st}(x_1, x_2, \text{time}) + \epsilon, \]

where

\[ f_s(x_1, x_2) \Rightarrow \text{"Spatial 2d smooth surface"} \]
\[ f_t(\text{time}) \Rightarrow \text{"Smooth time trend"} \]
\[ f_{st}(x_1, x_2, \text{time}) \Rightarrow \text{"Space-time interaction"} \]

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

for spatio-temporal smoothing

- **Model:**
  \[
  \hat{y} = \text{Intercept} + f_s(x_1, x_2) + f_t(\text{time}) + f_{st}(x_1, x_2, \text{time})
  \]

- **with Basis and Coefficients:**
  \[
  B\theta = \left( \begin{array}{c|c|c|c}
  1_{nt} & B_s \otimes 1_t & 1_n \otimes B_t & B_s \otimes B_t
  \end{array} \right)
  \begin{pmatrix}
  \gamma \\
  \theta^{(s)} \\
  \theta^{(t)} \\
  \theta^{(st)}
  \end{pmatrix}
  \]

- **and Penalty:**
  \[
P^* = \begin{pmatrix}
  0 \\
  \tau_1 D_1' D_1 \otimes I_{c_2} + \tau_2 I_{c_1} \otimes D_2' D_2 \\
  \tau_t D_t' D_t
  \end{pmatrix}
  \]
  where
  \[
P_{st} = \lambda_1 D_1' D_1 \otimes I_{c_2} \otimes I_{c_3} + \lambda_2 I_{c_1} \otimes D_2' D_2 \otimes I_{c_3} + \lambda_t I_{c_1} \otimes I_{c_2} \otimes D_t' D_t
  \]
\( \text{\textbf{P-spline ANOVA model}} \)
for spatio-temporal smoothing

- **Model:**

\[ \hat{y} = \text{Intercept} + f_s(x_1, x_2) + f_s(\text{time}) + f_{st}(x_1, x_2, \text{time}) \]

- **with Basis and Coefficients:**

\[ B\theta = \left( \begin{array}{c|c|c|c} \mathbf{1}_{nt} & B_s \otimes \mathbf{1}_t & \mathbf{1}_n \otimes B_t & B_s \otimes B_t \end{array} \right) \begin{pmatrix} \gamma \\ \theta^{(s)} \\ \theta^{(t)} \\ \theta^{(st)} \end{pmatrix} \]

- **and Penalty:**

\[ P^* = \begin{pmatrix} 0 \\ \tau_1 D'_1 D_1 \otimes I_{c_2} + \tau_2 I_{c_1} \otimes D'_2 D_2 \\ \tau_t D'_t D_t \end{pmatrix} \]

where

\[ P_{st} = \lambda_1 D'_1 D_1 \otimes I_{c_2} \otimes I_{c_3} + \lambda_2 I_{c_1} \otimes D'_2 D_2 \otimes I_{c_3} + \lambda_t I_{c_1} \otimes I_{c_2} \otimes D'_t D_t \]
**P-spline ANOVA model for spatio-temporal smoothing**

- **Model:**
  \[
  \hat{y} = \text{Intercept} + f_s(x_1, x_2) + f_s(\text{time}) + f_{st}(x_1, x_2, \text{time})
  \]

  - with **Basis** and **Coefficients**:
  \[
  B\theta = \left( \begin{array}{c} 1_{nt} \mid B_s \otimes 1_t \mid 1_n \otimes B_t \mid B_s \otimes B_t \end{array} \right) \left( \begin{array}{c} \gamma \\ \theta^{(s)} \\ \theta^{(t)} \\ \theta^{(st)} \end{array} \right)
  \]

  - and **Penalty**:
  \[
  P^* = \begin{pmatrix} 0 \\ \tau_1 D_1' D_1 \otimes I_{c_2} + \tau_2 I_{c_1} \otimes D_2' D_2 \\ \tau_t D_t' D_t \end{pmatrix}
  \]
  \[
  P_{st} = \lambda_1 D_1' D_1 \otimes I_{c_2} \otimes I_{c_3} + \lambda_2 I_{c_1} \otimes D_2' D_2 \otimes I_{c_3} + \lambda_t I_{c_1} \otimes I_{c_2} \otimes D_t' D_t
  \]
**P-spline ANOVA model**

*for spatio-temporal smoothing*

- **Model:**
  \[
  \hat{y} = \text{Intercept} + f_s(x_1, x_2) + f_s(\text{time}) + f_{st}(x_1, x_2, \text{time})
  \]

- with **Basis** and **Coefficients:**
  \[
  B\theta = \begin{pmatrix}
  1_{nt} & B_s \otimes 1_t & 1_n \otimes B_t & B_s \otimes B_t
  \end{pmatrix}
  \begin{pmatrix}
  \gamma \\
  \theta^{(s)} \\
  \theta^{(t)} \\
  \theta^{(st)}
  \end{pmatrix}
  \]

- and **Penalty:**
  \[
  P^* = \begin{pmatrix}
  0 & \tau_1 D_1' D_1 \otimes I_{c_2} + \tau_2 I_{c_1} \otimes D_2' D_2 \\
  \tau_1 D_1' D_1 & \tau_t D_t'
  \end{pmatrix}
  \]

  where

  \[
  P_{st} = \lambda_1 D_1' D_1 \otimes I_{c_2} \otimes I_{c_3} + \lambda_2 I_{c_1} \otimes D_2' D_2 \otimes I_{c_3} + \lambda_t I_{c_1} \otimes I_{c_2} \otimes D_t' D_t
  \]
**P-spline ANOVA model**
for spatio-temporal smoothing

- **Model:**
  \[ \hat{y} = \text{Intercept} + f_s(x_1, x_2) + f_s(\text{time}) + f_{st}(x_1, x_2, \text{time}) \]

  - with **Basis** and **Coefficients**:
    \[ B\theta = \begin{pmatrix} 1_{nt} & B_s \otimes 1_t & 1_n \otimes B_t & B_s \otimes B_t \end{pmatrix} \begin{pmatrix} \gamma \\ \theta^{(s)} \\ \theta^{(t)} \\ \theta^{(st)} \end{pmatrix} \]

  - and **Penalty**:
    \[ P^* = \begin{pmatrix} 0 \\ \tau_1 D'_1 D_1 \otimes I_{c_2} + \tau_2 I_{c_1} \otimes D'_2 D_2 \\ \tau_t D'_t D_t \end{pmatrix} \]

  where
  \[ P_{st} = \lambda_1 D'_1 D_1 \otimes I_{c_2} \otimes I_{c_3} + \lambda_2 I_{c_1} \otimes D'_2 D_2 \otimes I_{c_3} + \lambda_t I_{c_1} \otimes I_{c_2} \otimes D'_t D_t \]
\( p \)-spline ANOVA model
for spatio-temporal smoothing

- Model:
  \[
  \hat{y} = \text{Intercept} + f_s(x_1, x_2) + f_s(\text{time}) + f_{st}(x_1, x_2, \text{time})
  \]

  - with **Basis** and **Coefficients**:
  \[
  B\theta = \begin{pmatrix}
  1_{nt} & B_s \otimes 1_t & 1_n \otimes B_t & B_s \otimes B_t
  \end{pmatrix}
  \begin{pmatrix}
  \gamma \\
  \theta^{(s)} \\
  \theta^{(t)} \\
  \theta^{(st)}
  \end{pmatrix}
  \]

  - and **Penalty**:
  \[
  P^* = \begin{pmatrix}
  0 & \tau_1 D_1' D_1 \otimes I_{c_2} + \tau_2 I_{c_1} \otimes D_2' D_2 \\
  \tau_t D_t' D_t
  \end{pmatrix}
  \begin{pmatrix}
  P_{st}
  \end{pmatrix}
  \]

  where
  \[
  P_{st} = \lambda_1 D_1' D_1 \otimes I_{c_2} \otimes I_{c_3} + \lambda_2 I_{c_1} \otimes D_2' D_2 \otimes I_{c_3} + \lambda_t I_{c_1} \otimes I_{c_2} \otimes D_t' D_t
  \]
• In **GLAM** notation:

\[
(B_s \otimes 1_t) \theta^{(s)} \equiv 1_t \Theta_s B'_s \\
(1_n \otimes B_t) \theta^{(t)} \equiv B_t \Theta_t 1'_n \\
(B_s \otimes B_t) \theta^{(s,t)} \equiv B_t \Theta_{st} B'_s
\]
P-spline ANOVA model for spatio-temporal smoothing

- We avoid **identifiability problems** using
  - Mixed model reparameterization and
  - SVD properties

For each term we have:

\[
\begin{align*}
Basis & \quad [X : Z] \\
fs(x_1, x_2) & \equiv x_1 : x_2 \quad (1) \\
ft(x_t) & \equiv x_t \quad (2) \\
fst(x_1, x_2, x_t) & \equiv x_1 : x_2 : x_t \quad (3)
\end{align*}
\]

- Some terms in (1) and (2) also appear in (3).
- **✓ Remove linearly dependent columns** in the basis
\( P \)-spline ANOVA model
for spatio-temporal smoothing

- We avoid **identifiability problems** using
  - Mixed model reparameterization and
  - SVD properties

For each term we have:

\[
\begin{align*}
Basis \ [ \ X : Z \ ] \\
\ f_{s}(x_1, x_2) & \equiv \ x_1 : x_2 \quad (1) \\
\ f_{t}(x_t) & \equiv \ x_t \quad (2) \\
\ f_{st}(x_1, x_2, x_t) & \equiv \ x_1 : x_2 : x_t \quad (3)
\end{align*}
\]

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

✓ **Remove linearly dependent columns** in the basis
**P-spline ANOVA model**
for spatio-temporal smoothing

- Equivalent to apply **constraints** over the coefficients:

\[
\sum_{t=1}^{c_t} \theta_t^{(t)} = 0 \quad \text{(time)}
\]

\[
\sum_{i}^{c_1} \theta_{t,ij}^{(st)} = \sum_{j}^{c_2} \theta_{t,ij}^{(st)} = \sum_{i}^{c_1} \sum_{j}^{c_2} \theta_{t,ij}^{(st)} = 0 \quad \text{(space-time)}
\]

- or in **array form** over \( \Theta_{ijt} \)

» Array
\[
\sum_{i}^{c_1} \theta_{t,ij}^{(st)} = \sum_{j}^{c_2} \theta_{t,ij}^{(st)} = \sum_{i}^{c_1} \sum_{j}^{c_2} \theta_{t,ij}^{(st)} = 0
\]

✓ Centering and scaling matrix: \((I_c - 11^T / c)\)
\[ \sum_{i} c_1 \theta_{t,ij}^{(st)} = \sum_{j} c_2 \theta_{t,ij}^{(st)} = \sum_{i} \sum_{j} c_1 c_2 \theta_{t,ij}^{(st)} = 0 \]

✓ Centering and scaling matrix: \((I_c - 11'/c)\)
\[ \sum_{i}^{c_1} \theta_{t,ij}^{(st)} = \sum_{j}^{c_2} \theta_{t,ij}^{(st)} = \sum_{i}^{c_1} \sum_{j}^{c_2} \theta_{t,ij}^{(st)} = 0 \]

**Centering and scaling matrix:** \((I_c - 11'/c)\)
In practice

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

\[
f_s(x_1, x_2) \quad f_t(x_t) \quad f_{st}(x_1, x_2, x_t)
\]

\[
X \equiv \text{by columns} \quad x_1 : x_2 \quad x_t \quad (x_1, x_2, x_t)
\]

\[
Z \equiv \text{by blocks} \quad '' \quad '' \quad ''
\]

\[
F \equiv \text{blockdiagonal} \quad F_s \quad F_t \quad F_{st}
\]

\[
(\lambda_1, \lambda_2, \lambda_t) \quad \lambda_t \quad (\tau_1, \tau_2, \tau_t)
\]

- And estimate by REML
Outline

1. $P$-splines in spatio-temporal smoothing context
2. ANOVA-Type Interaction Models
3. Application to air pollution data
4. Concluding remarks
Ozone pollution in Europe
Lee and Durbán (2009b)

- 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:**

  ✓ 3d:

  $$f_{st}(x_1, x_2, x_t)$$

  ✓ **ANOVA:**

  $$f_s(x_1, x_2) + f_t(x_t) + f_{st}(x_1, x_2, x_t)$$
Space-time interaction is not considered

✓ time smooth trend is additive
Spatio-temporal ANOVA model

\[ \hat{y} = f(\text{space}) + f(\text{space, time}) + f(\text{time}) \]

1999 : 1
Comparison of fitted values

Additive VS ANOVA

✓ 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, 3d and Additive

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>14280.73</td>
<td>366.03</td>
</tr>
<tr>
<td>3d− interaction</td>
<td>14537.22</td>
<td>765.05</td>
</tr>
<tr>
<td>Additive</td>
<td>16506.28</td>
<td>65.98</td>
</tr>
</tbody>
</table>

Observations:
- Best overall performance of ANOVA in terms of AIC also with less d.f. than 3d.
- **ANOVA model** is more realistic than Additive, and easier to decompose and interpret in terms of the fit.
Outline

1. $p$-splines in spatio-temporal smoothing context
2. ANOVA-Type Interaction Models
3. Application to air pollution data
4. Concluding remarks
Concluding remarks

- Use of **Array methods** in space-time data
  - speeds up calculations
  - data and model matrices are stored efficiently

- **Spatio-temporal ANOVA** model
  - Flexible formulation and interpretability of smoothing

✓ Computational efficiency: **nested B-spline bases**
  - consider a **smaller basis** $\tilde{B}_t$ for the interaction $B_s \otimes B_t$
    ... such that
    \[ \text{rank}(\tilde{B}_t) < \text{rank}(B_t) \]
  - The size of the full basis $B$ is reduced and model is nested
    \[ f_s(x_1, x_2) + f_t(x_t) + f_{st}^*(x_1, x_2, x_t) \subset f_s(x_1, x_2) + f_t(x_t) + f_{st}(x_1, x_2, x_t) \]
Concluding remarks

- **Use of Array methods** in space-time data
  - speeds up calculations
  - data and model matrices are stored efficiently

- **Spatio-temporal ANOVA** model
  - Flexible formulation and interpretability of smoothing

✓ Computational efficiency: nested $B$-spline bases
  - consider a smaller basis $\tilde{B}_t$ for the interaction $B_s \otimes B_t$
    ... such that
    $\text{rank}(\tilde{B}_t) < \text{rank}(B_t)$

- The size of the full basis $B$ is reduced and model is nested
  $f_s(x_1, x_2) + f_t(x_t) + f_{st}^*(x_1, x_2, x_t) \subset f_s(x_1, x_2) + f_t(x_t) + f_{st}(x_1, x_2, x_t)$
Concluding remarks

- **Use of Array methods** in space-time data
  - speeds up calculations
  - data and model matrices are stored efficiently

- **Spatio-temporal ANOVA** model
  - Flexible formulation and interpretability of smoothing

✓ Computational efficiency: **nested B-spline bases**
  - consider a smaller basis $\tilde{B}_t$ for the interaction $B_s \otimes B_t$
  - such that $\text{rank}(\tilde{B}_t) < \text{rank}(B_t)$

- The size of the full basis $B$ is reduced and model is nested
  $$f_s(x_1, x_2) + f_t(x_t) + f_{st}^*(x_1, x_2, x_t) \subset f_s(x_1, x_2) + f_t(x_t) + f_{st}(x_1, x_2, x_t)$$
Concluding remarks

- Use of **Array methods** in space-time data
  - speeds up calculations
  - data and model matrices are stored efficiently

- **Spatio-temporal ANOVA** model
  - Flexible formulation and interpretability of smoothing

✓ Computational efficiency: **nested B-spline bases**
  - consider a smaller basis $\mathbf{B}_t$ for the interaction $\mathbf{B}_s \otimes \mathbf{B}_t$
  - such that

  \[
  \text{rank}(\mathbf{B}_t) < \text{rank}(\mathbf{B}_t)
  \]

  - The size of the full basis $\mathbf{B}$ is reduced and model is nested

  \[
  f_s(x_1, x_2) + f_t(x_t) + f_{st}^*(x_1, x_2, x_t) \subset f_s(x_1, x_2) + f_t(x_t) + f_{st}(x_1, x_2, x_t)
  \]
THANKS FOR YOUR ATTENTION !!!
Spatial and Spatio-temporal smoothing with $P$-splines:

Lee, D.-J. and Durbán, M. (2009a)

*Smooth-CAR mixed models for spatial count data.*
CSDA 53(8):2968-2979


$P$-spline ANOVA-Type interaction models for spatio-temporal smoothing
Submitted

$P$-splines:

Eilers, PHC. and Marx, BD.

*Flexible smoothing with B-splines and penalties*


*Generalized linear array models with applications to multidimensional smoothing*
JRSSB, 68:1-22