Modelagem de Processos Espaciais e Espaço-Temporais

Alexandra M. Schmidt
Instituto de Matemática - UFRJ
www.dme.ufrj.br/∼alex

Colóquio Inter-Institucional
Processos Estocásticos e Aplicações
IMPA - Abril de 2007

Outline Partes I e II

1. Modelos Univariados Heterogêneos
   - Introdução
   - Deformação Espacial
   - Versão Bayesiana
   - Especificação das Prioris
   - Procedimento de Inferência
   - Exemplos

2. Modelos Multivariados Heterogêneos
   - Introdução
   - Modelos Separáveis
   - Modelos de Coregionalização
   - MCL variando no espaço
   - Procedimento de Inferência
   - Exemplos

Outline Parte III

3. Modelagem de Múltiplas Séries de Vazão como função da chuva
   - Problema
   - Modelo proposto
   - Procedimento de Inferência
   - Resultados
   - Considerações Finais

Parte I

Covariância para processos espaciais univariados

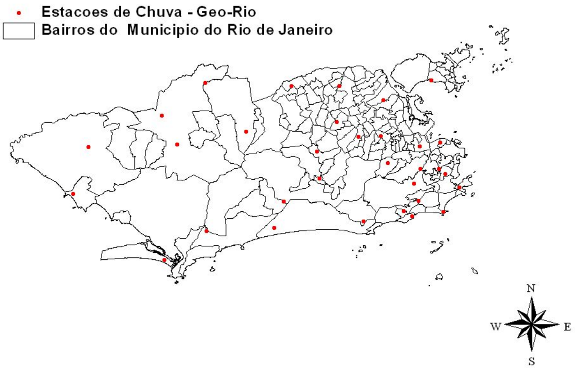


Figura: Modelagem da quantidade de chuva no Rio de Janeiro em janeiro de 2000.

Notação Básica

- Espaço Euclideano $\mathbf{S}$, i.e. $\mathbf{S} = R^d$ onde $d = 1, 2,$ ou 3. Ênfase no caso $d = 2$;
- Ponto arbitrário em $\mathbf{S}$, $\mathbf{s} \in \mathbf{S}$;
- Região de interesse, $A \subset \mathbf{S}$.
- Localizações espaciais com dados $\mathbf{s}_1, \cdots, \mathbf{s}_n$, $\mathbf{s}_i \in D \subset A \subset \mathbf{S}$
  - observado;
  - não necessariamente distintos, pode haver replicações nas localizações.
- Observações (resposta, dados) $\mathbf{Y}(\mathbf{s}_1), \cdots, \mathbf{Y}(\mathbf{s_n})$. Em geral são multivariados. Mas, na literatura, maior ênfase no caso univariado.
- Covariáveis $\mathbf{X}(\mathbf{s}_1), \cdots, \mathbf{X}(\mathbf{s}_n)$.

Principais Objetivos em Geoestatística

- Estimação - inferência sobre parâmetros de um modelo estocástico;
- Previsão (interpolação) - inferência sobre a realização do processo em localizações não-medidas de interesse;
- Planejamento de uma rede - onde colocar uma nova estação? qual retirar?
  (Ruiz, Ferreira & Schmidt (Tech. Rep. 2006))

Dois tipos importantes de estrutura espacial são estacionariedade e isotropia. Intuitivamente

(a) Estacionariedade - propriedade em que o processo é similar ao longo de $\mathbf{A}$. Isto significa que
   - a estrutura de grande escala é constante;
   - estrutura de pequena escala depende das localizações apenas através das suas posições relativas;

(b) Isotropia - o processo é estacionário E a estrutura de pequena-escala depende das localizações espaciais apenas através da distância euclideana entre elas $\Rightarrow$ invariante sob rotação e translação das localizações.

Estacionariedade Intrínseca

É definida através das primeiras diferenças:

$$
\begin{aligned}
\mathsf{E}(Y(\mathbf{s}+\mathbf{h}) - Y(\mathbf{s})) &= 0, \\
\mathsf{Var}(Y(\mathbf{s}+\mathbf{h}) - Y(\mathbf{s})) &= 2\gamma(\mathbf{h})
\end{aligned}
$$

A quantidade $2\gamma(\mathbf{h})$ é conhecida como variograma.
$\gamma(.)$ é conhecido como semi-variograma.
Em geoestatística, $2\gamma(.)$ é tratada como um *parâmetro* do processo aleatório $\{Y(\mathbf{s}) : \mathbf{s} \in D\}$ (porque descreve a estrutura de covariância).

Estacionariedade de Segunda Ordem

Um processo que satisfaça:

$$\begin{aligned} \mathsf{E}(Y(\mathbf{s})) &= \mu \quad \forall \mathbf{s} \in D \\ \mathsf{Cov}(Y(\mathbf{s}), Y(\mathbf{s}')) &= C(\mathbf{s} - \mathbf{s}') \quad \forall \mathbf{s}, \mathbf{s}' \in D \end{aligned}$$

é definido como estacionário de segunda ordem. E mais, se $C(\mathbf{s} - \mathbf{s}')$ é um função apenas de $|| \ \mathbf{s} - \mathbf{s}' \ ||$ (não é uma função das localizações) então $C(.)$ é dita isotrópica. $C(.)$ é conhecida como covariograma.

Deformação Espacial

- Paul Sampson e Peter Guttorp, da Universidade de Washington, foram pioneiros ao propor uma aproximação para o problema de heterogeneidade espacial.
- A idéia principal: transformação não-linear do espaço amostral (espaço G de geográfico) para um espaço latente D (D de dispersão), no qual a estrutura espacial é estacionária e isotrópica.
  - obtenção dos pontos observados via MDS; Em outras palavras, eles estimam as localizações medidas no espaço D de modo que as correlações observadas se ajustem à distância euclideana entre os pontos em D.
  - Interpolação das localizações não medidas via thin-plate spline.

Abordagem Bayesiana (Schmidt & O'Hagan (*JRSS, Series B*, 2003))

Assumimos que $Y_{it} = Y(s_i, t)$, $i = 1, \ldots, n$ e $t = 1, \ldots, T$. Seja $\mathbf{Y}_t = (Y_{1t}, Y_{2t}, \ldots, Y_{nt})'$ para $t = 1, \ldots, T$.

$\mathbf{Y}_1, \ldots, \mathbf{Y}_T$ são i.i.d. $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

O interesse principal está na estimação da matriz de covariâncias verdadeira, $\boldsymbol{\Sigma}$.

Abordagem Bayesiana (Schmidt & O'Hagan (*JRSS, Series B*, 2003))

Assumimos que $Y_{it} = Y(s_i, t)$, $i = 1, \ldots, n$ e $t = 1, \ldots, T$. Seja $\mathbf{Y}_t = (Y_{1t}, Y_{2t}, \ldots, Y_{nt})'$ para $t = 1, \ldots, T$.

$\mathbf{Y}_1, \ldots, \mathbf{Y}_T$ são i.i.d. $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

O interesse principal está na estimação da matriz de covariâncias verdadeira, $\boldsymbol{\Sigma}$.

A verossimilhança para $\boldsymbol{\Sigma}$ tem a forma da Wishart

$$f(\mathbf{S} \mid \boldsymbol{\Sigma}) \propto \mid \boldsymbol{\Sigma} \mid^{-\frac{T-1}{2}} \exp\left\{-\frac{T}{2} \text{ tr } \mathbf{S}\boldsymbol{\Sigma}^{-1}\right\}. \qquad (1)$$

Abordagem Bayesiana
(Schmidt & O'Hagan (*JRSS, Series B*, 2003))

Assumimos que $Y_{it} = Y(s_i, t)$, $i = 1, \ldots, n$ e $t = 1, \ldots, T$. Seja $\mathbf{Y}_t = (Y_{1t}, Y_{2t}, \ldots, Y_{nt})'$ para $t = 1, \ldots, T$.

$\mathbf{Y}_1, \ldots, \mathbf{Y}_T$ são i.i.d. $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

O interesse principal está na estimação da matriz de covariâncias verdadeira, $\boldsymbol{\Sigma}$.

A verossimilhança para $\boldsymbol{\Sigma}$ tem a forma da Wishart

$$f(\mathbf{S} \mid \boldsymbol{\Sigma}) \propto \mid \boldsymbol{\Sigma} \mid^{-\frac{T-1}{2}} \exp\left\{-\frac{T}{2} \text{ tr } \mathbf{S}\boldsymbol{\Sigma}^{-1}\right\}. \qquad (1)$$

A proposta é modelar cada elemento de $\boldsymbol{\Sigma}$ como

$$\text{Cov } (Y(s_i, t), Y(s_j, t)) = \sqrt{v(s_i)v(s_j)}c_d(s_i, s_j), \qquad (2)$$

onde para todo t

Abordagem Bayesiana (Schmidt & O'Hagan (*JRSS, Series B*, 2003))

Assumimos que $Y_{it} = Y(s_i, t)$, $i = 1, \ldots, n$ e $t = 1, \ldots, T$. Seja $\mathbf{Y}_t = (Y_{1t}, Y_{2t}, \ldots, Y_{nt})'$ para $t = 1, \ldots, T$.

$\mathbf{Y}_1, \ldots, \mathbf{Y}_T$ são i.i.d. $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

O interesse principal está na estimação da matriz de covariâncias verdadeira, $\boldsymbol{\Sigma}$.

A verossimilhança para $\boldsymbol{\Sigma}$ tem a forma da Wishart

$$f(\mathbf{S} \mid \boldsymbol{\Sigma}) \propto \mid \boldsymbol{\Sigma} \mid^{-\frac{T-1}{2}} \exp\left\{-\frac{T}{2} \text{ tr } \mathbf{S}\boldsymbol{\Sigma}^{-1}\right\}. \qquad (1)$$

A proposta é modelar cada elemento de $\boldsymbol{\Sigma}$ como

$$\text{Cov } (Y(s_i, t), Y(s_j, t)) = \sqrt{v(s_i)v(s_j)} c_d(s_i, s_j), \qquad (2)$$

onde para todo t

$$v(s) = \text{ Var } (Y(s, t)) \;\text{ e }\; c_d(s, s') = \text{ Corr } (Y(s, t), Y(s', t))$$

Especificações das prioris

A priori,

$$\begin{aligned} v(s) \mid \tau^2, f &\sim GI(\tau^2(f-2), f), \quad s \in G, \\ \pi(\tau^2) &\propto \tau^{-2}, \end{aligned} \qquad (3)$$

Mapeando a Correlação Espacial

O mapeamento das estações medidas ocorre através da função de correlação definida como

$$c_d(s_i, s_j) = g(\| d(s_i) - d(s_j) \|), \qquad (4)$$

onde $g(.)$ é uma função monótona.

Especificação da Função de Correlação $g(.)$

Similarmente a S&G,

$$g(h) = \sum_{k=1}^{K} a_k \exp(-b_k h^2). \quad (5)$$

K deve ser, de acordo com os dados, o menor possível. Os parâmetros de alcance b_k e os coeficientes a_k são desconhecidos e satisfazem a

- $\sum_{k=1}^{K} a_k = 1$;
- $b_1 > b_2 > \ldots > b_K$, $a_k > 0$ e $b_k > 0$, $k = 1, \ldots, K$.

Um efeito pepita pode ser introduzido permitindo que $b_1 \rightarrow \infty$ em (5).

O Processo Latente $d(.)$

Atribuímos como priori para função $d(.)$ um processo gaussiano:

$$d(.) \mid m(.), \boldsymbol{\sigma}_d^2, R_d(.,.) \sim PG(m(.), \boldsymbol{\sigma}_d^2 R_d(.,.)), \qquad (6)$$

- $m(.)$ é a função média a priori;
- $\boldsymbol{\sigma}_d^2$ é uma matriz 2×2;
- $R_d(.,.)$ mede a correlação a priori entre as estações, tq $R_d(s,s) = 1$.

O Processo Latente $d(.)$

Atribuímos como priori para função $d(.)$ um processo gaussiano:

$$d(.) \mid m(.), \boldsymbol{\sigma}_d^2, R_d(.,.) \sim PG(m(.), \boldsymbol{\sigma}_d^2 R_d(.,.)), \qquad (6)$$

- $m(.)$ é a função média a priori;
- $\boldsymbol{\sigma}_d^2$ é uma matriz 2×2;
- $R_d(.,.)$ mede a correlação a priori entre as estações, tq $R_d(s,s) = 1$.

Em particular,

- $\mathbf{D} = (d_1, \ldots, d_n)$, onde $d_i = d(s_i)$, é uma matriz $2 \times n$ contendo as coordenadas das estações monitoradoras no espaço D;
- $\mathbf{m} = (m(s_1), \ldots, m(s_n))$
- $\mathbf{R}_d$ uma matriz $n \times n$ com elementos $R_d(s_i, s_j)$,

$$(\mathbf{D} \mid \mathbf{m}, \boldsymbol{\sigma}_d^2, \mathbf{R}_d) \sim N_{(2\times n)}(\mathbf{m}, \boldsymbol{\sigma}_d^2, \mathbf{R}_d) \text{ OU}$$

$$vec(\mathbf{D}) \mid vec(\mathbf{m}), \boldsymbol{\sigma}_d^2, \mathbf{R}_d) \sim N_{(2n)}(vec(\mathbf{m}), \boldsymbol{\sigma}_d^2 \otimes \mathbf{R}_d)$$

Especificação de R_d e $\boldsymbol{\sigma}_d^2$

- R_d :
  Os elementos de $R_d(.,.)$ são modelados de acordo com uma função de correlação gaussiana

$$R_d(s, s') = \exp\left(-b_d \mid\mid s - s' \mid\mid^2\right),$$

  onde b_d controla, *a priori*, a *forma* da configuração das estações medidas em D.
  Sugestão : Fixar b_d igual a $\frac{1}{2a}$, onde a é o quadrado de uma distância típica entre localizações medidas em G.

Especificação de R_d e $\boldsymbol{\sigma}^2_d$

- R_d :
  Os elementos de $R_d(.,.)$ são modelados de acordo com uma função de correlação gaussiana

$$R_d(s, s') = \exp\left(-b_d \mid\mid s - s' \mid\mid^2\right),$$

  onde b_d controla, *a priori*, a *forma* da configuração das estações medidas em D.
  Sugestão : Fixar b_d igual a $\frac{1}{2a}$, onde a é o quadrado de uma distância típica entre localizações medidas em G.
- $\boldsymbol{\sigma}^2_d$:
  - Parâmetro é não identificável no sentido de Dawid (1979).
  - $\mathbf{S}$ traz informação sobre as distâncias no espaço D, trazendo informação, no máximo, sobre os autovalores de $\boldsymbol{\sigma}^2_d$;
  - $\boldsymbol{\sigma}^2_d$ é modelada como uma matriz diagonal, e $\boldsymbol{\sigma}^2_{d_{ii}} \sim IG(\beta_i, \alpha_i)$, $i = 1, 2$.

Procedimento de inferência

$$\pi(a_1,\ldots,a_K,b_1,\ldots,b_K \mid K) \propto \prod_{k=1}^{K} \pi_k(b_k) \quad \text{com} \quad \sum_{k=1}^{K} a_k = 1 \text{ e } b_1 > \ldots > b_K, \tag{7}$$

De acordo com o teorema de Bayes, a posteriori de $\boldsymbol{\theta}$ é proporcional a

$$\begin{aligned}
\pi(\boldsymbol{\theta} \mid \mathbf{S}) \propto & \mid \boldsymbol{\Sigma} \mid^{-\frac{T-1}{2}} \exp\left\{-\frac{T}{2} \operatorname{tr} \mathbf{S}\boldsymbol{\Sigma}^{-1}\right\} \left\{\prod_{i=1}^{n} v_i^{-(f+2)/2} \exp\left(-\frac{(f-2)\tau^2}{2v_i}\right)\right\} \\
& \times \tau^{\frac{(nf-2)}{2}} \left\{\prod_{k=1}^{K} \frac{1}{b_k} \exp\left\{\frac{-(\log(b_k)-\mu_b)^2}{2\sigma_b^2}\right\}\right\} \\
& \times \mid \boldsymbol{\sigma}_d^2 \mid^{-n/2} \mid \mathbf{R}_d \mid^{-1} \exp\left\{-\frac{1}{2} \operatorname{tr} (\mathbf{D}-\mathbf{m})'\boldsymbol{\sigma}_d^{-2}(\mathbf{D}-\mathbf{m})\mathbf{R}_d^{-1}\right\} \\
& \times (\sigma_{d_{11}}^2)^{-(\beta_1+2)/2} \exp\left\{-\frac{\alpha_1}{2\sigma_{d_{11}}^2}\right\} \times (\sigma_{d_{22}}^2)^{-(\beta_2+2)/2} \exp\left\{-\frac{\alpha_2}{2\sigma_{d_{22}}^2}\right\}.
\end{aligned} \tag{8}$$

Procedimento de inferência

$$\pi(a_1, \ldots, a_K, b_1, \ldots, b_K \mid K) \propto \prod_{k=1}^{K} \pi_k(b_k) \quad \text{com} \quad \sum_{k=1}^{K} a_k = 1 \text{ e } b_1 > \ldots > b_K, \tag{7}$$

De acordo com o teorema de Bayes, a posteriori de $\boldsymbol{\theta}$ é proporcional a

$$\begin{aligned}
\pi(\boldsymbol{\theta} \mid \mathbf{S}) \propto & \mid \boldsymbol{\Sigma} \mid^{-\frac{T-1}{2}} \exp\left\{-\frac{T}{2} \text{ tr } \mathbf{S}\boldsymbol{\Sigma}^{-1}\right\} \left\{\prod_{i=1}^{n} v_i^{-(f+2)/2} \exp\left(-\frac{(f-2)\tau^2}{2v_i}\right)\right\} \\
& \times \tau^{\frac{(nf-2)}{2}} \left\{\prod_{k=1}^{K} \frac{1}{b_k} \exp\left\{\frac{-(\log(b_k) - \mu_b)^2}{2\sigma_b^2}\right\}\right\} \\
& \times \mid \boldsymbol{\sigma}_d^2 \mid^{-n/2} \mid \mathbf{R}_d \mid^{-1} \exp\left\{-\frac{1}{2} \text{ tr } (\mathbf{D} - \mathbf{m})' \boldsymbol{\sigma}_d^{-2} (\mathbf{D} - \mathbf{m}) \mathbf{R}_d^{-1}\right\} \\
& \times (\sigma_{d_{11}}^2)^{-(\beta_1+2)/2} \exp\left\{-\frac{\alpha_1}{2\sigma_{d_{11}}^2}\right\} \times (\sigma_{d_{22}}^2)^{-(\beta_2+2)/2} \exp\left\{-\frac{\alpha_2}{2\sigma_{d_{22}}^2}\right\}.
\end{aligned} \tag{8}$$

Posteriori não tem forma analítica fechada $\longrightarrow$ uso de MCMC para obtenção de amostras de $\pi(\boldsymbol{\theta} \mid \mathbf{S})$. Cuidado ao sortear $\mathbf{d}$ e $\mathbf{b}$.

Exemplo - Dados Simulados

- Fixamos as coordenadas de $n = 6$ estações no espaço G;
- Fixamos $\tau^2 = 2, f = 12, b_2 = 0.25, a_1 = 0.1, a_2 = 0.9, \sigma^2_{d_{11}} = 0.25, \sigma^2_{d_{22}} = 0.375$, and $b_d = 0.5$;
- Depois de gerar as localizações no espaço latente D, a "verdadeira" matriz de covariâncias foi gerada.

Exemplo - Dados Simulados

- Fixamos as coordenadas de $n = 6$ estações no espaço G;
- Fixamos $\tau^2 = 2, f = 12, b_2 = 0.25, a_1 = 0.1, a_2 = 0.9, \sigma^2_{d_{11}} = 0.25, \sigma^2_{d_{22}} = 0.375$, and $b_d = 0.5$;
- Depois de gerar as localizações no espaço latente D, a "verdadeira" matriz de covariâncias foi gerada.

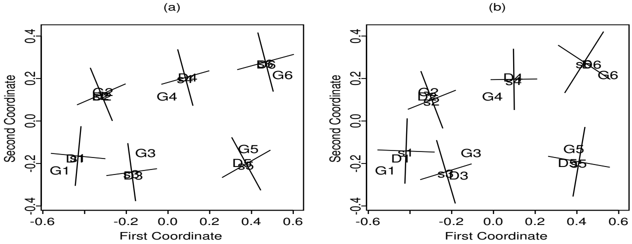


Figura: Superimposição de Procrustes da configuração média em D (s_i) sobre as localizações originais no espaço G (G_i). (a) $b_d = 0.25$ e $E(\sigma_{d_{11}}) = 0.5$, $E(\sigma_{d_{11}}) = 0.75$. (b) $b_d = 1.0$ e $E(\sigma_{d_{11}}) = 1.0$, $E(\sigma_{d_{11}}) = 1.5$.

Exemplo - Dados Simulados

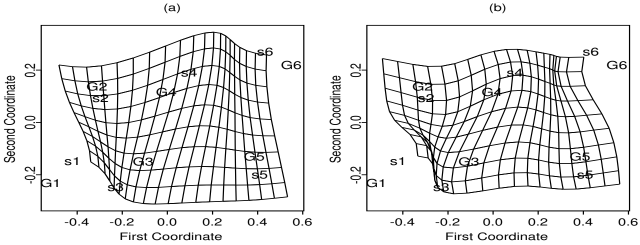


Figura: Mapeamento de uma grade regular de 200 pontos do espaço G para o espaço D. (a) $b_d = 0.25$ e $E(\sigma^2_{d_{11}}) = 0.5$, $E(\sigma^2_{d_{11}}) = 0.75$. (b) $b_d = 1.0$ e $E(\sigma^2_{d_{11}}) = 1.0$, $E(\sigma^2_{d_{11}}) = 1.5$.

Exemplo - Dados de Radiação Solar

- $n = 12$ estações monitoradoras;
- Dados referentes à primavera-verão, de 22 de março a 20 de setembro, de 1980 a 1983 $\rightarrow T = 732$.

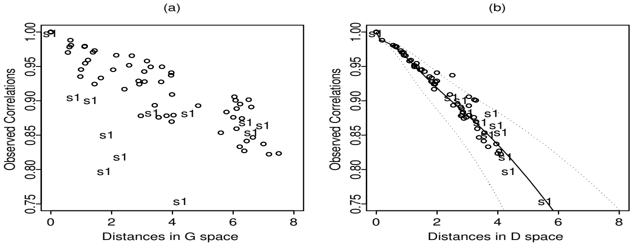


Figura: (a) Distâncias geográficas versus as correlações observadas dos dados de radiação solar. (b) Média a posteriori estimada (linha sólida) e o intervalo de 95% de credibilidade a posteriori (linha tracejada) da função de correlação para os dados de radiação solar.

Exemplo - Dados de Radiação Solar

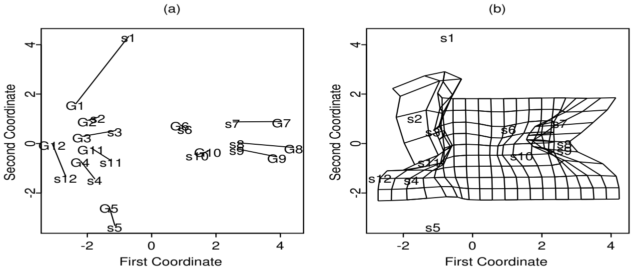
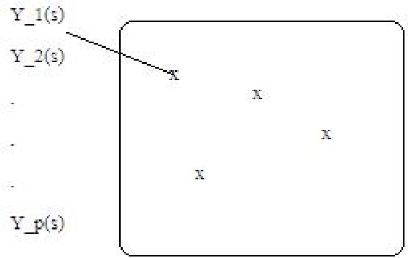


Figura: (a) Superimposição de Procrustes da média a posteriori das localizações no espaço D (s_i) sobre a as localizações da configuração original no espaço G space (G_i). (b) Mapeamento de uma grade regular de 200 pontos do espaço G para o espaço D.

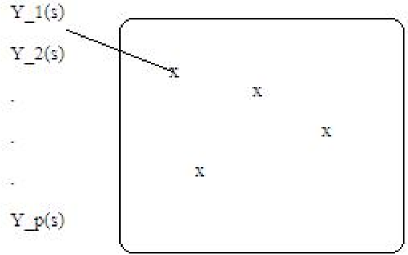
Parte II

Covariância para processos espaciais multivariados

Seja $\{\mathbf{Y}(\mathbf{s}) : \mathbf{s} \in D \subset I\!R^2; \mathbf{Y} \in \mathbb{R}^p\}$ um campo aleatório multivariado (ex. **Y(s)** pode ser $(O_3, CO_2, PM_{10})(\mathbf{s})$), $p = 3$).

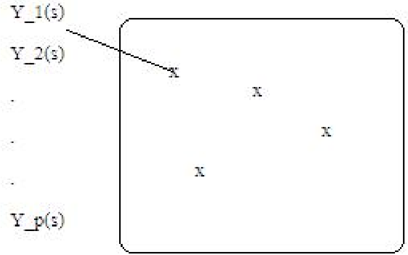


Seja $\{\mathbf{Y}(\mathbf{s}) : \mathbf{s} \in D \subset I\!R^2; \mathbf{Y} \in \mathbb{R}^p\}$ um campo aleatório multivariado (ex. **Y(s)** pode ser $(O_3, CO_2, PM_{10})(\mathbf{s})$), $p = 3$).



É preciso descrever a covariância dentro e entre as estações.

Seja $\{\mathbf{Y}(\mathbf{s}) : \mathbf{s} \in D \subset I\!R^2; \mathbf{Y} \in \mathbb{R}^p\}$ um campo aleatório multivariado (ex. **Y(s)** pode ser $(O_3, CO_2, PM_{10})(\mathbf{s}))$, $p = 3)$.



É preciso descrever a covariância dentro e entre as estações.

Objetivo: **propor uma estrutura de covariância válida e flexível**.

Modelos Separáveis

- Seja ρ uma função de correlação espacial válida;
- Seja $\mathbf{T}$ uma matriz $p \times p$, positiva definida.

Então a covariância do processo entre duas localizações quaisquer s e s', pode ser descrita por

$$C(s, s') = \rho(s, s')\mathbf{T}\,. \tag{9}$$

Considerando o empilhamento do vetor observado em n localizações, $\mathbf{Y}$, a matriz de covariâncias resultante é dada por

$$\mathbf{\Sigma} = \mathbf{R} \otimes \mathbf{T}, \tag{10}$$

Num modelo mais geral, poderíamos incluir uma componente latente, para descrever a estrutura de covariância acima, por exemplo, $\mathbf{v}(s)$, de modo que

$$\mathbf{Y}(s) = \mathbf{X}(s)\boldsymbol{\beta} + \mathbf{v}(s) + \boldsymbol{\epsilon}(s), \tag{11}$$

Considerando o modelo em (11), e empilhando as observações num vetor $\mathbf{Y}$ de dimensão np, a verossimilhança terá a seguinte forma

$$f(\mathbf{Y} \mid \mathbf{\Sigma_Y}, \boldsymbol{\beta}) \propto \mid \mathbf{\Sigma_Y} \mid^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})'\mathbf{\Sigma_Y}^{-1}(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})\right\},$$

onde $\mathbf{\Sigma_Y} = (R \otimes T) + (\mathbf{I}_n \otimes \mathbf{D})$ e $\mathbf{I}_n$ é a matriz identidade de dimensão n.

Vantagens e Desvantagens do modelo separável

- $|\mathbf{\Sigma}| = |R|^p |T|^n$ e $\mathbf{\Sigma}^{-1} = R^{-1} \otimes T^{-1}$;

Vantagens e Desvantagens do modelo separável

- $|\mathbf{\Sigma}| = |R|^p |T|^n$ e $\mathbf{\Sigma}^{-1} = R^{-1} \otimes T^{-1}$;
- $cov(Y_l(s), Y_{l'}(s')) = cov(Y_{l'}(s), Y_l(s'))$, $\forall\, l, l', s$ e s';

Vantagens e Desvantagens do modelo separável

- $|\mathbf{\Sigma}| = |R|^p |T|^n$ e $\mathbf{\Sigma}^{-1} = R^{-1} \otimes T^{-1}$;
- $cov(Y_l(s), Y_{l'}(s')) = cov(Y_{l'}(s), Y_l(s'))$, $\forall\, l, l', s$ e s';
- se ρ é simétrica e estritamente decrescente, então o alcance espacial é o mesmo para as componentes de $\mathbf{Y}(.)$;

Vantagens e Desvantagens do modelo separável

- $|\boldsymbol{\Sigma}| = |R|^p |T|^n$ e $\boldsymbol{\Sigma}^{-1} = R^{-1} \otimes T^{-1}$;
- $cov(Y_l(s), Y_{l'}(s')) = cov(Y_{l'}(s), Y_l(s'))$, $\forall\, l, l', s$ e s';
- se ρ é simétrica e estritamente decrescente, então o alcance espacial é o mesmo para as componentes de $\mathbf{Y}(.)$;
- se ρ é estacionária, a correlação generalizada, é tal que

$$\frac{cov(Y_j(s), Y_{j'}(s+h))}{\sqrt{cov(Y_j(s), Y_j(s+h)) cov(Y_{j'}(s), Y_{j'}(s+h))}} = \frac{T_{jj'}}{\sqrt{T_{jj} T_{j'j'}}},$$

independente da posição s e vetor h.

Modelos de coregionalização

Especificação Intrínseca (Mathéron, 1982)

$$\mathbf{Y}(\mathbf{s}) = \mathbf{A}\mathbf{w}(\mathbf{s})$$

$\mathbf{w}(\mathbf{s})$ são processos espaciais i.i.d.'s.

$\Rightarrow Y_l(\mathbf{s}) = \sum_{j=1}^{k} a_{lj} w_j(\mathbf{s})$ e

Cov$(\mathbf{Y}(\mathbf{s}), \mathbf{Y}(\mathbf{s}')) = \rho(\mathbf{s} - \mathbf{s}'; \phi)\mathbf{T}$
com $\mathbf{T} = \mathbf{A}\mathbf{A}^T$, $\mathbf{A}$ $p \times k$, e $k \leq p$.

$\mathbf{T}$ é conhecida como matriz de coregionalização.

Modelo Bayesiano de Coregionalização Linear
Schmidt & Gelfand (JGR, 2003) e
Gelfand, Schmidt, Banerjee & Sirmans (Test, 2004)

Seja

$$Y_l(\mathbf{s}) = \sum_{j=1}^{p} a_{lj} w_j(\mathbf{s}), \; l = 1, 2, \cdots, p.$$

mas agora assuma que $w_j(.)$ são i.i.d's com variância 1, e função de correlação$\rho(\mathbf{s} - \mathbf{s}'; \phi_j)$. Então

$$\begin{aligned} &\text{Cov}(Y_l(\mathbf{s}), Y_{l'}(\mathbf{s}')) = \\ &= \sum_{j=1}^{p} a_{lj} a_{l'j} \rho(\mathbf{s} - \mathbf{s}'; \phi_j) \end{aligned}$$

Considerando n localizações espaciais, temos que

$$\boldsymbol{\Sigma}_{\mathbf{Y}} = \begin{bmatrix} \mathbf{AA}' & \mathbf{AD}_{1,2}\mathbf{A}' & \cdots & \mathbf{AD}_{1,n}\mathbf{A}' \\ \mathbf{AD}_{2,1}\mathbf{A}' & \mathbf{AA}' & \cdots & \mathbf{AD}_{2,n}\mathbf{A}' \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{AD}_{n,1}\mathbf{A}' & \mathbf{AD}_{n,2}\mathbf{A}' & \cdots & \mathbf{AA}' \end{bmatrix}.$$

Dessa forma,

$$
\begin{aligned}
\mathbf{\Sigma_Y} &= \sum_{j=1}^{p} \mathbf{R}(\phi_j) \otimes \mathbf{a}_j \mathbf{a}_j^T \\
&= \sum_{j=1}^{p} \mathbf{R}(\phi_j) \otimes \mathbf{T}_j
\end{aligned}
$$

Calculando o alcance de $Y_j(\mathbf{s})$

Definição : distância para a qual a correlação entre $Y_j(\mathbf{s})$ e $Y_j(\mathbf{s}')$ torna-se 0.05.
O alcance, r_1, de $Y_1(\mathbf{s})$ soluciona $\rho_1(r_1;\phi_1) = 0.05$.
Mais geralmente, o alcance r_j, de $Y_j(\mathbf{s})$ soluciona

$$\frac{a_{j1}^2\rho_1(r_j;\phi_1) + \cdots + a_{jj}^2\rho(r_j;\phi_j)}{a_{j1}^2 + \cdots + a_{jj}^2} = 0.05$$

O lado esquerdo é descrescente em r.
Usualmente os ρ_j's são funções paramétricas $\rightarrow r_j$ é uma função paramétrica que não está disponível explicitamente.
Sob o enfoque bayesiano, a distribuição a posteriori de $\mathbf{A}$ e dos parâmetros em $\rho(.;.)$ fornece meios para obter a distribuição a posteriori de r_j.

Modelo Multivariado Geral

$$\mathbf{Y}(\mathbf{s}) = \boldsymbol{\mu}(\mathbf{s}) + \mathbf{v}(\mathbf{s}) + \boldsymbol{\epsilon}(\mathbf{s}) \tag{12}$$

com

- $\boldsymbol{\epsilon}(\mathbf{s}) \sim N(\mathbf{0}, \mathbf{D}_\epsilon)$, $(D_\epsilon)_{jj} = \tau_j^2$;
- $\mathbf{v}(\mathbf{s}) = \mathbf{A}\mathbf{w}(\mathbf{s})$ seguindo a especificação anterior, e $w_j(\mathbf{s})$ são PG, com média 0 e variância 1 e função de correlação $\rho_j(.)$;
- $\boldsymbol{\mu}(\mathbf{s})$ são obtidos a partir de $\mu_j(\mathbf{s}) = \mathbf{X}_j^T(\mathbf{s})\boldsymbol{\beta}_j$.

Modelo Hierárquico

1o Estágio:

$$\mathbf{Y}(\mathbf{s}_i)|\boldsymbol{\beta}_j, \{\mathbf{v}(\mathbf{s}_i)\}, \mathbf{D}_\epsilon \sim N(\boldsymbol{\mu}(\mathbf{s}_i) + \mathbf{v}(\mathbf{s}_i), \mathbf{D}_\epsilon).$$

Modelo Hierárquico

1o Estágio:

$$\mathbf{Y}(\mathbf{s}_i)|\boldsymbol{\beta}_j, \{\mathbf{v}(\mathbf{s}_i)\}, \mathbf{D}_\epsilon \sim N(\boldsymbol{\mu}(\mathbf{s}_i) + \mathbf{v}(\mathbf{s}_i), \mathbf{D}_\epsilon).$$

2o Estágio:

$$\mathbf{v} = \begin{pmatrix} \mathbf{v}(\mathbf{s}_1) \\ \vdots \\ \mathbf{v}(\mathbf{s}_n) \end{pmatrix} \sim N(\mathbf{0}, \textstyle\sum_{j=1}^{p} \mathbf{R}_j \otimes \mathbf{T}_j)$$

Modelo Hierárquico

1o Estágio:

$\mathbf{Y}(\mathbf{s}_i)|\boldsymbol{\beta}_j, \{\mathbf{v}(\mathbf{s}_i)\}, \mathbf{D}_\epsilon \sim N(\boldsymbol{\mu}(\mathbf{s}_i) + \mathbf{v}(\mathbf{s}_i), \mathbf{D}_\epsilon)$.

2o Estágio:

$$\mathbf{v} = \begin{pmatrix} \mathbf{v}(\mathbf{s}_1) \\ \vdots \\ \mathbf{v}(\mathbf{s}_n) \end{pmatrix} \sim N(\mathbf{0}, \textstyle\sum_{j=1}^p \mathbf{R}_j \otimes \mathbf{T}_j)$$

Concatenando os $\mathbf{Y}(\mathbf{s}_i)$ num vetor, $np \times 1$, $\mathbf{Y}$, similarmente $\boldsymbol{\mu}(\mathbf{s}_i)$ num vetor $\boldsymbol{\mu}$, podemos marginalizar com respeito a $\mathbf{v}$ para obter

$$f(\mathbf{Y}|\{\boldsymbol{\beta}_j\}, \mathbf{D}_\epsilon, \{\rho_j\}, \mathbf{T}) =$$
$$N\left(\boldsymbol{\mu}, \sum_{j=1}^p (\mathbf{H}_j \otimes \mathbf{T}_j) + \mathbf{I}_{n\times n} \otimes \mathbf{D}_\epsilon\right).$$

A especificação bayesiana se completa associando distribuições a priori para $\{\boldsymbol{\beta}_j\}$, $\{\tau_j^2\}$, $\mathbf{T}$ e os parâmetros em ρ_j.

Parametrização Condicional

Assuma, por exemplo, $p = 3$, sabemos que

$$p(Y_1(\mathbf{s}), Y_2(\mathbf{s}), Y_3(\mathbf{s})) = p(Y_1(\mathbf{s}))p(Y_2(\mathbf{s}) \mid Y_1(\mathbf{s}))p(Y_3(\mathbf{s}) \mid Y_2(\mathbf{s}), Y_1(\mathbf{s}))$$

Parametrização Condicional

Assuma, por exemplo, $p = 3$, sabemos que

$$p(Y_1(\mathbf{s}), Y_2(\mathbf{s}), Y_3(\mathbf{s})) = p(Y_1(\mathbf{s}))p(Y_2(\mathbf{s}) \mid Y_1(\mathbf{s}))p(Y_3(\mathbf{s}) \mid Y_2(\mathbf{s}), Y_1(\mathbf{s}))$$

Gelfand et. al (2004) propõem o uso de uma parametrização condicional para o modelo em (12)

$$\begin{aligned} Y_1(\mathbf{s}) &= \boldsymbol{\beta}_1' \mathbf{X}(\mathbf{s}) + \sigma_1 \tilde{w}_1(\mathbf{s}) \\ Y_2(\mathbf{s}) \mid Y_1(\mathbf{s}) &= \boldsymbol{\beta}_2' \mathbf{X}(\mathbf{s}) + \alpha Y_1(\mathbf{s}) + \sigma_2 \tilde{w}_2(\mathbf{s}) \\ Y_3(\mathbf{s}) \mid Y_2(\mathbf{s}), Y_1(\mathbf{s}) &= \boldsymbol{\beta}_3' \mathbf{X}(\mathbf{s}) + \alpha_1 Y_1(\mathbf{s}) + \alpha_2 Y_2(\mathbf{s}) + \sigma_3 \tilde{w}_3(\mathbf{s}) + \epsilon_3(\mathbf{s}) \end{aligned}$$

São discutidas, também, vantagens e desvantagens de uma parametrização com relação a outra.

MCL variando no espaço

Substitui-se $\mathbf{A}$ por $\mathbf{A}(\mathbf{s})$, assim

$$\mathbf{v}(s) = \mathbf{A}(s)\mathbf{w}(s) \tag{13}$$

$\rightarrow$ MCL com variação espacial (MCLVE).

MCL variando no espaço

Substitui-se $\mathbf{A}$ por $\mathbf{A}(\mathbf{s})$, assim

$$\mathbf{v}(s) = \mathbf{A}(s)\mathbf{w}(s) \tag{13}$$

$\rightarrow$ MCL com variação espacial (MCLVE).
Seja $\mathbf{T}(s) = \mathbf{A}(s)\mathbf{A}(s)^T$. Novamente, por conveniência, $\mathbf{A}(s)$ pode ser assumida triangular inferior. Agora $C(s, s')$ é tal que

$$C(s, s') = \sum_j \rho_j(s - s')\mathbf{a}_j(s)\mathbf{a}_j^T(s'). \tag{14}$$

MCL variando no espaço

Substitui-se $\mathbf{A}$ por $\mathbf{A}(\mathbf{s})$, assim

$$\mathbf{v}(s) = \mathbf{A}(s)\mathbf{w}(s) \tag{13}$$

$\rightarrow$ MCL com variação espacial (MCLVE).
Seja $\mathbf{T}(s) = \mathbf{A}(s)\mathbf{A}(s)^T$. Novamente, por conveniência, $\mathbf{A}(s)$ pode ser assumida triangular inferior. Agora $C(s, s')$ é tal que

$$C(s, s') = \sum_j \rho_j(s - s')\mathbf{a}_j(s)\mathbf{a}_j^T(s'). \tag{14}$$

Assim $\mathbf{v}(s)$ não é mais um processo estacionário nem separável.
Fazendo $s - s' \rightarrow 0$, a matriz de covariâncias de $\mathbf{v}(\mathbf{s})$ é $\mathbf{T}(\mathbf{s})$.

Possiblidades de modelagem de $\mathbf{A}(\mathbf{s})$

Modelar $\mathbf{A}(s)$ através da correspondência 1 a 1 com $\mathbf{T}(s)$.

Possiblidades de modelagem de $\mathbf{A(s)}$

Modelar $\mathbf{A}(s)$ através da correspondência 1 a 1 com $\mathbf{T}(s)$.

No caso univariado, escolhas para $\sigma^2(s)$ incluem:

- $\sigma^2(s, \theta)$, isto é, uma função paramétrica ou uma superfície de tendência em função da posição geográfica;

Possiblidades de modelagem de $\mathbf{A}(\mathbf{s})$

Modelar $\mathbf{A}(s)$ através da correspondência 1 a 1 com $\mathbf{T}(s)$.

No caso univariado, escolhas para $\sigma^2(s)$ incluem:

- $\sigma^2(s, \theta)$, isto é, uma função paramétrica ou uma superfície de tendência em função da posição geográfica;
- $\sigma^2(x(s)) = g(x(s))\sigma^2$ onde $x(s)$ é alguma covariável utilizada para explicar $\mathbf{Y}(s)$ e $g(.) > 0$ (então $g(x(s))$ é tipicamente $x(s)$ ou $x^2(s)$);

Possiblidades de modelagem de $\mathbf{A}(\mathbf{s})$

Modelar $\mathbf{A}(s)$ através da correspondência 1 a 1 com $\mathbf{T}(s)$.

No caso univariado, escolhas para $\sigma^2(s)$ incluem:

- $\sigma^2(s, \theta)$, isto é, uma função paramétrica ou uma superfície de tendência em função da posição geográfica;
- $\sigma^2(x(s)) = g(x(s))\sigma^2$ onde $x(s)$ é alguma covariável utilizada para explicar $\mathbf{Y}(s)$ e $g(.) > 0$ (então $g(x(s))$ é tipicamente $x(s)$ ou $x^2(s)$);
- ou $\sigma^2(s)$ é em si, um processo espacial (por exemplo, $\log \sigma^2(s)$ pode ser um processo gaussiano).

Possiblidades de modelagem de $\mathbf{A}(\mathbf{s})$

Modelar $\mathbf{A}(s)$ através da correspondência 1 a 1 com $\mathbf{T}(s)$.

No caso univariado, escolhas para $\sigma^2(s)$ incluem:

- $\sigma^2(s, \theta)$, isto é, uma função paramétrica ou uma superfície de tendência em função da posição geográfica;
- $\sigma^2(x(s)) = g(x(s))\sigma^2$ onde $x(s)$ é alguma covariável utilizada para explicar $\mathbf{Y}(s)$ e $g(.) > 0$ (então $g(x(s))$ é tipicamente $x(s)$ ou $x^2(s)$);
- ou $\sigma^2(s)$ é em si, um processo espacial (por exemplo, $\log \sigma^2(s)$ pode ser um processo gaussiano).

Segunda possibilidade, fazermos $\mathbf{T}(s) = g(x(s))\mathbf{T}$.

Possiblidades de modelagem de $\mathbf{A}(\mathbf{s})$

Modelar $\mathbf{A}(s)$ através da correspondência 1 a 1 com $\mathbf{T}(s)$.

No caso univariado, escolhas para $\sigma^2(s)$ incluem:

- $\sigma^2(s, \theta)$, isto é, uma função paramétrica ou uma superfície de tendência em função da posição geográfica;
- $\sigma^2(x(s)) = g(x(s))\sigma^2$ onde $x(s)$ é alguma covariável utilizada para explicar $\mathbf{Y}(s)$ e $g(.) > 0$ (então $g(x(s))$ é tipicamente $x(s)$ ou $x^2(s)$);
- ou $\sigma^2(s)$ é em si, um processo espacial (por exemplo, $\log \sigma^2(s)$ pode ser um processo gaussiano).

Segunda possibilidade, fazermos $\mathbf{T}(s) = g(x(s))\mathbf{T}$.

Terceira possibilidade, $\mathbf{T}(\mathbf{s})$ é um processo espacial matriz-variado.

Modelagem de $\mathbf{T}(\mathbf{s})$

- Definição da Wishart, $\mathbf{\Omega} = \mathbf{\Gamma}\mathbf{Z}\mathbf{Z}^T\mathbf{\Gamma}^T \sim W_p(\nu, \mathbf{\Gamma}\mathbf{\Gamma}^T)$ se $\mathbf{Z} = (\mathbf{Z}_1, \ldots, \mathbf{Z}_\nu)$ é $p \times \nu$ com Z_{lj} i.i.d. $N(0,1)$, $l = 1, \ldots, \nu$, $j = 1, \ldots, p$.

Modelagem de $\mathbf{T}(\mathbf{s})$

- Definição da Wishart, $\mathbf{\Omega} = \mathbf{\Gamma} \mathbf{Z} \mathbf{Z}^T \mathbf{\Gamma}^T \sim W_p(\nu, \mathbf{\Gamma}\mathbf{\Gamma}^T)$ se $\mathbf{Z} = (\mathbf{Z}_1, \ldots, \mathbf{Z}_\nu)$ é $p \times \nu$ com Z_{lj} i.i.d. $N(0,1)$, $l = 1, \ldots, \nu$, $j = 1, \ldots, p$.
- Suponha que temos νp processos espaciais gaussianos, estacionários, independentes, com média 0 tal que $Z_{lj}(s)$ tem função de correlação $\rho_j(s - s')$. Isto é, temos p processos espaciais independentes, diferentes e ν replicações de cada um.

Modelagem de $\mathbf{T}(\mathbf{s})$

- Definição da Wishart, $\boldsymbol{\Omega} = \boldsymbol{\Gamma}\mathbf{Z}\mathbf{Z}^T\boldsymbol{\Gamma}^T \sim W_p(\nu, \boldsymbol{\Gamma}\boldsymbol{\Gamma}^T)$ se $\mathbf{Z} = (\mathbf{Z}_1, \ldots, \mathbf{Z}_\nu)$ é $p \times \nu$ com Z_{lj} i.i.d. $N(0,1)$, $l = 1, \ldots, \nu$, $j = 1, \ldots, p$.
- Suponha que temos νp processos espaciais gaussianos, estacionários, independentes, com média 0 tal que $Z_{lj}(s)$ tem função de correlação $\rho_j(s - s')$. Isto é, temos p processos espaciais independentes, diferentes e ν replicações de cada um.
- Defina $\boldsymbol{\Omega}(s) = \boldsymbol{\Gamma}\mathbf{Z}(s)\mathbf{Z}^T(s)\boldsymbol{\Gamma}^T$, dizemos que $\boldsymbol{\Omega}(s)$ é um processo espacial estacionário Wishart matriz-variado, $\boldsymbol{\Omega}(s) \sim SW_p(\nu, \boldsymbol{\Gamma}\boldsymbol{\Gamma}^T, \rho_1, \ldots, \rho_p)$. Resulta num processo não estacionário e não-gaussiano.

Procedimento de Inferência - MCL Conjunto

- $\boldsymbol{\theta} = (\{\boldsymbol{\beta}_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{T})$, $j = 1, \cdots, p$.

Procedimento de Inferência - MCL Conjunto

- $\boldsymbol{\theta} = (\{\boldsymbol{\beta}_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{T})$, $j = 1, \cdots, p$.
- Assumindo independência ao longo de j, temos $\pi(\boldsymbol{\theta}) = \prod_j p(\boldsymbol{\beta}_j)\, p(\rho_j)\, p(\tau_j^2)\, p(\mathbf{T})$.

Procedimento de Inferência - MCL Conjunto

- $\boldsymbol{\theta} = (\{\boldsymbol{\beta}_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{T})$, $j = 1, \cdots, p$.
- Assumindo independência ao longo de j, temos $\pi(\boldsymbol{\theta}) = \prod_j p(\boldsymbol{\beta}_j)\, p(\rho_j)\, p(\tau_j^2)\, p(\mathbf{T})$.
- Portanto,

$$\pi(\boldsymbol{\theta}|\mathbf{Y}) \propto f(\mathbf{Y}|\{\boldsymbol{\beta}_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{T})\, \pi(\boldsymbol{\theta}).$$

Procedimento de Inferência - MCL Conjunto

- $\boldsymbol{\theta} = (\{\boldsymbol{\beta}_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{T})$, $j = 1, \cdots, p$.
- Assumindo independência ao longo de j, temos $\pi(\boldsymbol{\theta}) = \prod_j p(\boldsymbol{\beta}_j)\, p(\rho_j)\, p(\tau_j^2)\, p(\mathbf{T})$.
- Portanto,

$$\pi(\boldsymbol{\theta}|\mathbf{Y}) \propto f(\mathbf{Y}|\{\boldsymbol{\beta}_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{T})\, \pi(\boldsymbol{\theta}).$$

- Prioris
  - $\boldsymbol{\beta}_j \sim N(0, V)$, V matriz diagonal, com elementos grandes;

Procedimento de Inferência - MCL Conjunto

- $\boldsymbol{\theta} = (\{\boldsymbol{\beta}_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{T})$, $j = 1, \cdots, p$.
- Assumindo independência ao longo de j, temos $\pi(\boldsymbol{\theta}) = \prod_j p(\boldsymbol{\beta}_j)\, p(\rho_j)\, p(\tau_j^2)\, p(\mathbf{T})$.
- Portanto,

$$\pi(\boldsymbol{\theta}|\mathbf{Y}) \propto f(\mathbf{Y}|\{\boldsymbol{\beta}_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{T})\, \pi(\boldsymbol{\theta}).$$

- Prioris
  - $\boldsymbol{\beta}_j \sim N(0, V)$, V matriz diagonal, com elementos grandes;
  - $\tau_j^2 \sim IG(2, b)$ b, tal que $E(\tau_j^2) = b = \hat{\tau_j^2}$;

Procedimento de Inferência - MCL Conjunto

- $\boldsymbol{\theta} = (\{\boldsymbol{\beta}_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{T})$, $j = 1, \cdots, p$.
- Assumindo independência ao longo de j, temos $\pi(\boldsymbol{\theta}) = \prod_j p(\boldsymbol{\beta}_j)\, p(\rho_j)\, p(\tau_j^2)\, p(\mathbf{T})$.
- Portanto,

$$\pi(\boldsymbol{\theta}|\mathbf{Y}) \propto f(\mathbf{Y}|\{\boldsymbol{\beta}_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{T})\, \pi(\boldsymbol{\theta}).$$

- Prioris
  - $\boldsymbol{\beta}_j \sim N(0, V)$, V matriz diagonal, com elementos grandes;
  - $\tau_j^2 \sim IG(2, b)$ b, tal que $E(\tau_j^2) = b = \hat{\tau}_j^2$;
  - Se $\rho_j = \exp(-\phi_j\, d)$, $\phi \sim Ga(a_1, b_1)$, tq $\frac{a_1}{b_1} = \frac{6}{\text{dmax}}$ e variância fixa;

Procedimento de Inferência - MCL Conjunto

- $\boldsymbol{\theta} = (\{\boldsymbol{\beta}_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{T})$, $j = 1, \cdots, p$.
- Assumindo independência ao longo de j, temos $\pi(\boldsymbol{\theta}) = \prod_j p(\boldsymbol{\beta}_j)\, p(\rho_j)\, p(\tau_j^2)\, p(\mathbf{T})$.
- Portanto,

$$\pi(\boldsymbol{\theta}|\mathbf{Y}) \propto f(\mathbf{Y}|\{\boldsymbol{\beta}_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{T})\, \pi(\boldsymbol{\theta}).$$

- Prioris
  - $\boldsymbol{\beta}_j \sim N(0, V)$, V matriz diagonal, com elementos grandes;
  - $\tau_j^2 \sim IG(2, b)$ b, tal que $E(\tau_j^2) = b = \hat{\tau}_j^2$;
  - Se $\rho_j = \exp(-\phi_j\, d)$, $\phi \sim Ga(a_1, b_1)$, tq $\frac{a_1}{b_1} = \frac{6}{\text{dmax}}$ e variância fixa;
  - Relação um a um entre os elementos de $\mathbf{T}$ e a matriz triangular inferior $\mathbf{A}$; $\mathbf{T} \sim W^{-1}(\boldsymbol{\Delta}, \nu)$ tq os elementos da diagonal poderiam ser obtidos de EMQ de um modelo independente para cada elemento de $Y_j(s)$, $j = 1, \cdots, p$, e ν pequeno.

Procedimento de Inferência - MCL Conjunto

- Procedimento de inferência feito através do uso de MCMC $\Rightarrow$ Amostrador de Gibbs com passos do Metropolis-Hastings.

Procedimento de Inferência - MCL Conjunto

- Procedimento de inferência feito através do uso de MCMC $\Rightarrow$ Amostrador de Gibbs com passos do Metropolis-Hastings.
- Maior desafio é sortear eficientemente da distribuição condicional completa de $\mathbf{T}$.

$$\pi(\mathbf{T}|\{\boldsymbol{\beta}_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{Y}) \propto | \boldsymbol{\Sigma}_{\mathbf{Y}} |^{-np/2} \exp\left\{-\frac{1}{2}\left(\mathbf{Y} - \boldsymbol{\mu}\right)^T \boldsymbol{\Sigma}_{\mathbf{Y}}^{-1} \left(\mathbf{Y} - \boldsymbol{\mu}\right)\right\} \pi(\mathbf{T}) \quad (15)$$

onde $\boldsymbol{\Sigma}_{\mathbf{Y}} = \sum_{j=1}^{p} \left(\mathbf{H}_j \otimes \mathbf{T}_j\right) + \mathbf{I}_{n\times n} \otimes \mathbf{D}$.

Modelando o preço de venda e o lucro advindo de imóveis comerciais

- O preço de venda de um imóvel comercial é, teoricamente, dado pelo lucro esperado capitalizado a alguma taxa de desconto (ajustada pelo risco).
- Consideramos aqui um conjunto de dados de blocos de apartamentos em 3 mercados imobiliários bem distintos, Chicago, Dalas, e San Diego.
- Chicago é uma cidade antiga, tradicional, onde o desenvolvimento se expandiu de um centro de negócios central.
- Dalas é uma cidade mais nova, onde o desenvolvimento tende a acontecer em multi-subcentros, com o distrito central sendo menos importante no comportamento do padrão espacial.
- Já San Diego é uma cidade fisicamente mais restrita, com um desenvolvimento mais linear ao contrário do tradicional padrão "circular".

Modelando o preço de venda e o lucro advindo de imóveis comerciais

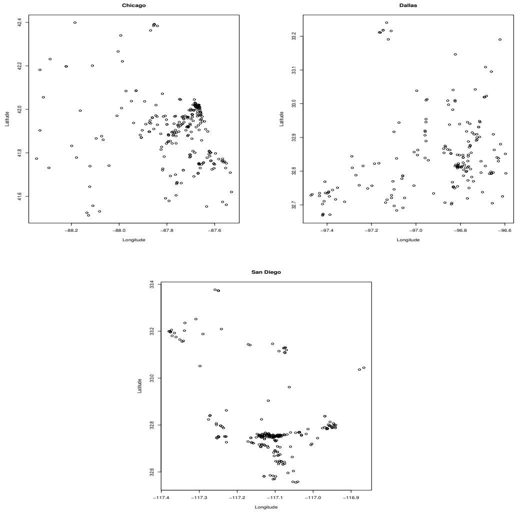


Figura: Localizações dos imóveis nos 3 mercados imobiliários, Chicago, Dalas e San Diego.

Modelando o preço de venda e o lucro advindo de imóveis comerciais

Objetivo: ajustar um modelo conjunto para o preço de venda e o lucro líquido e obter uma superfície espacial associada ao risco, $\log R = \log I - \log P$.
O modelo proposto é

$$
\begin{aligned}
I(s) &= sqft(s)\beta_{I1} + age(s)\beta_{I2} + unit(s)\beta_{I3} + v_1(s) + \epsilon_1(s) \\
P(s) &= sqft(s)\beta_{P1} + age(s)\beta_{P2} + unit(s)\beta_{P3} + v_2(s) + \epsilon_2(s).
\end{aligned}
$$

Modelando o preço de venda e o lucro advindo de imóveis comerciais

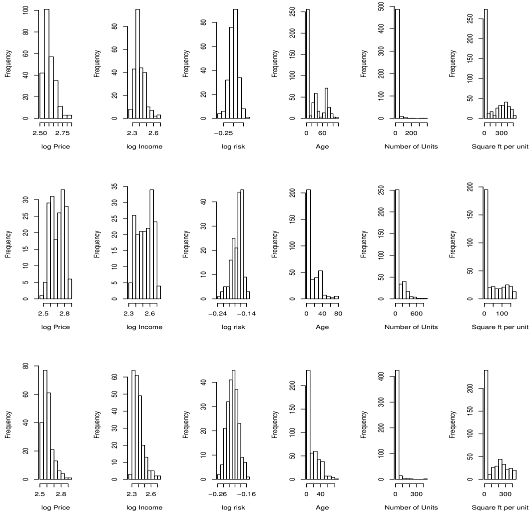
Objetivo: ajustar um modelo conjunto para o preço de venda e o lucro líquido e obter uma superfície espacial associada ao risco, $\log R = \log I - \log P$.
O modelo proposto é

$$
\begin{aligned}
I(s) &= sqft(s)\beta_{I1} + age(s)\beta_{I2} + unit(s)\beta_{I3} + v_1(s) + \epsilon_1(s) \\
P(s) &= sqft(s)\beta_{P1} + age(s)\beta_{P2} + unit(s)\beta_{P3} + v_2(s) + \epsilon_2(s).
\end{aligned}
$$

- Modelo 1 é um MCL intrínseco, isto é, ele assume uma estrutura de covariância separável para $\mathbf{v}(s)$;
- Modelo 2 assume o MCL mais geral para $\mathbf{v}(s)$;
- Modelo 3 é um MCLVE, usando a forma $\mathbf{T}(s) = (x(s))^{\psi}\mathbf{T}$ onde $x(s)$ é $unit(s)$;
- Modelo 4 utiliza um processo espacial, Wishart matriz-variado, para $\mathbf{T}(s)$.

Comparação entre diferentes modelos foi feita usando Gelfand & Ghosh (1998).

Modelando o preço de venda e o lucro advindo de imóveis comerciais



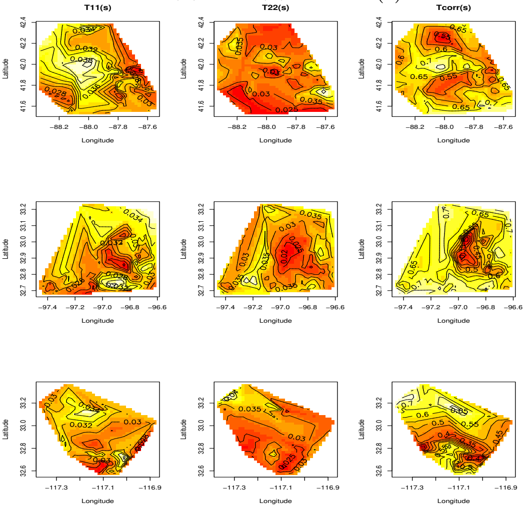
Modelando o preço de venda e o lucro advindo de imóveis comerciais

Tabela: Resultados do critério de comparação do modelo para (a) todo o conjunto de dados e (b) para a amostra retirando as 20 transações de cada mercado.

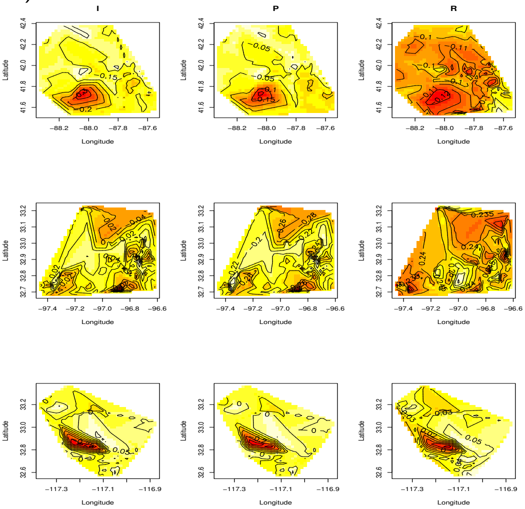
(a)	Chicago			Dalas			San Diego		
	G	P	D	G	P	D	G	P	D
1	0,1793	0,7299	0,9092	0,1126	0,5138	0,6264	0,0886	0,4842	0,5728
2	0,1772	0,6416	0,8188	0,0709	0,4767	0,5476	0,0839	0,4478	0,5317
3	0,1794	0,6368	0,8162	0,0715	0,4798	0,5513	0,0802	0,4513	0,5315
4	0,1574	0,6923	0,8497	0,0436	0,4985	0,5421	0,0713	0,4588	0,5301

(b)	Chicago			Dallas			San Diego		
	G	P	D	G	P	D	G	P	D
1	0,0219	0,0763	0,0982	0,0141	0,0631	0,0772	0,0091	0,0498	0,0589
2	0,0221	0,0755	0,0976	0,0091	0,0598	0,0689	0,0095	0,0449	0,0544
3	0,0191	0,0758	0,0949	0,0091	0,0610	0,0701	0,0087	0,0459	0,0546
4	0,0178	0,0761	0,0939	0,0059	0,0631	0,0690	0,0074	0,0469	0,0543

Superfícies espaciais associadas à variação espacial de $\mathbf{T}(s)$ para as 3 cidades, Chicago (1a. linha), Dalas (2a. linha) e San Diego (3a. linha), com as colunas correspondendo, respectivamente a, $T_{11}(s)$, $T_{22}(s)$ e $T_{corr(s)}$



Superfícies espaciais associadas à variação espacial de Renda (1a coluna), Preço (2a coluna) e Risco (3a coluna) para as 3 cidades, Chicago (1a. linha), Dalas (2a. linha) e San Diego (3a. linha).



Sequência deste trabalho

1. Modelagem de dados espacialmente desalinhados (Schmidt & Estrella (em revisão para RBE));
2. Modelagem de observações espaço-temporais (Sansó, Schmidt, & Nobre (Tech Rep, 2006));
3. Modelos espaciais multivariados para dados de contagem (Schmidt & Hoeting, em andamento).

Parte III

Modelagem de Múltiplas Séries de Vazão como função da chuva

Modelagem Bayesiana de Múltiplas Séries de Vazão

Romy R. Ravines
Alexandra M. Schmidt
Helio S. Migon

Parte da Tese de Doutorado
Defendida em Dezembro de 2006

Instituto de Matemática - UFRJ

Colocação do Problema: bacia do Rio Grande (BA)

Colocação do Problema: bacia do Rio Grande (BA)

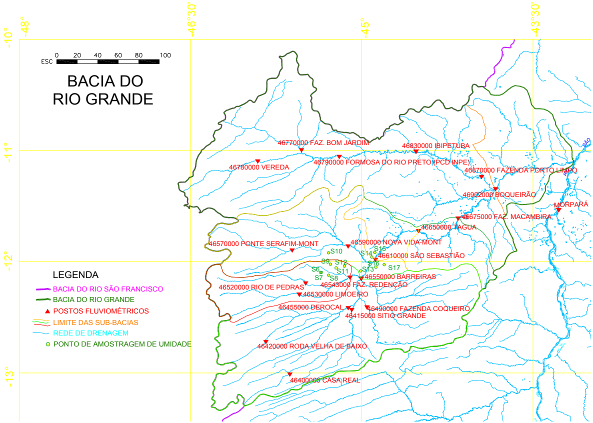
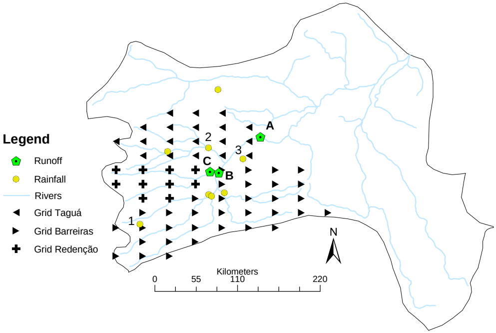
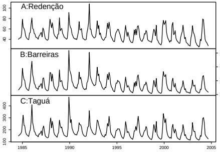


Figura: **Hidrografia e Sub-Bacias do Rio Grande.** Postos fluviométricos de interesse: A=Taguá, B=Barreiras e C=Redenção

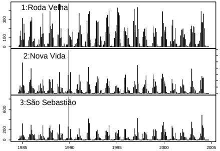
Bacia do Rio Grande (BA): dados



(a) Sub-bacias, Estações e Grade de Interpolação



(b) Vazão



(c) Chuva

Abordagem Proposta: uma série de vazão

Se Y_t é a vazão e X_t é a chuva acumulada no tempo t numa bacia, a relação chuva-vazão pode ser representada por:

$$
\begin{aligned}
Y_t &\sim p(Y_t|\mu_t, \phi_t), \quad t = 1, 2, \ldots && \text{(16a)}\\
g(\mu_t) &= f_1(\alpha_t, E_t) && \text{(16b)}\\
E_t &= f_2(E_{t-1}, \ldots, E_0, X_t) && \text{(16c)}
\end{aligned}
$$

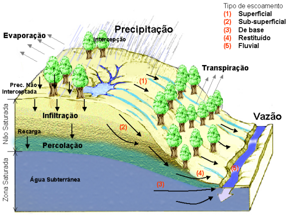


Figura: Processos físicos envolvidos na geração da vazão

- Relação não linear
- Vazão do período atual depende da vazão do período anterior e da chuva passada e corrente.
- Não existe uma retroalimentação (*feedback*) entre vazão e chuva
- Vazão é uma variável não negativa: distribuição gama, log-normal, etc.
- As séries temporais podem não ser estacionárias.

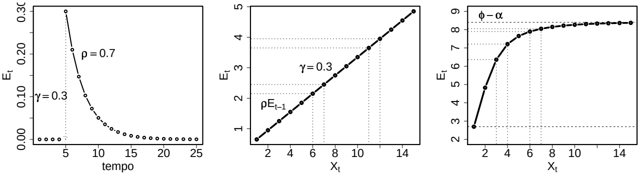
Efeito da Chuva: uma função de transferência

Com as hipóteses estabelecidas em Migon & Monteiro (1997), a relação chuva-vazão pode ser bem representada através de uma **função de transferência**.

Duas Funções de Transferência:

$$E_t = \rho_t E_{t-1} + \gamma_t X_t \tag{17a}$$

$$E_t = \rho_t E_{t-1} + [1 - \exp(-\kappa_t X_t)][\vartheta_t - (\alpha_t + \rho_t E_{t-1})] \tag{17b}$$



(a) Decaimento Exponencial (b) Retornos Proporcionais (c) Retornos Decrescentes

Figura: Hipóteses sobre as Funções de Transferência

Efeito da Chuva: uma função de transferência

Interpretação dos Parâmetros

γ_t : **Efeito Instantâneo**. Associado à velocidade de escoamento superficial, taxa de infiltração do solo e /ou interceptação da chuva pela vegetação.

ρ_t : **Fator de Recarga**. Taxa de memorização. Depende das características geológicas da bacia.

α_t : **Fluxo base**. Depende do nível do lençol freático de cada bacia.

Tipos de Efeito Instantâneo

Constante	:	$\gamma_t = \gamma$
Variando no tempo	:	$\gamma_t = \gamma_{t-1} + \delta_t$
Ganho aleatório	:	$\gamma_t = \gamma + \delta_t$

Abordagem Proposta: múltiplas séries de vazão

- Dispomos de M séries temporais de vazão observadas durante T períodos de tempo. Seja Y_t^m, a vazão observada na bacia m e tempo t, e $\mathbf{Y}_t = (Y_t^1, \ldots, Y_t^M)'$.

Abordagem Proposta: múltiplas séries de vazão

- Dispomos de M séries temporais de vazão observadas durante T períodos de tempo. Seja Y_t^m, a vazão observada na bacia m e tempo t, e $\mathbf{Y}_t = (Y_t^1, \ldots, Y_t^M)'$.
- A distribuição conjunta $p(\mathbf{Y}_t)$, condicionada a um vetor de parâmetros $\boldsymbol{\Theta}$, pode ser fatorada, por exemplo, da seguinte maneira:

$$p(\mathbf{Y}_t|\boldsymbol{\Theta}) = p(Y_t^M|Y_t^{M-1}, \boldsymbol{\Theta})p(Y_t^{M-1}|Y_t^{M-2}, \boldsymbol{\Theta}) \cdots \\ p(Y_t^2|Y_t^1, \boldsymbol{\Theta})p(Y_t^1, \boldsymbol{\Theta}),$$

onde $p(Y_t^m)$ é a distribuição de $Y_t^m, i = 1, \ldots, M$.

Abordagem Proposta: múltiplas séries de vazão

- Dispomos de M séries temporais de vazão observadas durante T períodos de tempo. Seja Y_t^m, a vazão observada na bacia m e tempo t, e $\mathbf{Y}_t = (Y_t^1, \ldots, Y_t^M)'$.
- A distribuição conjunta $p(\mathbf{Y}_t)$, condicionada a um vetor de parâmetros $\mathbf{\Theta}$, pode ser fatorada, por exemplo, da seguinte maneira:

$$p(\mathbf{Y}_t|\mathbf{\Theta}) = p(Y_t^M|Y_t^{M-1}, \mathbf{\Theta})p(Y_t^{M-1}|Y_t^{M-2}, \mathbf{\Theta}) \cdots p(Y_t^2|Y_t^1, \mathbf{\Theta})p(Y_t^1, \mathbf{\Theta}),$$

onde $p(Y_t^m)$ é a distribuição de $Y_t^m, i = 1, \ldots, M$.

- Considerando a chuva

$$p(\mathbf{Y}_t|\mathbf{X}_t, \mathbf{\Theta}) = p(Y_t^A|Y_t^B, X_t^{A|B}, \mathbf{\Theta})p(Y_t^B|X_t^B, \mathbf{\Theta})$$

onde $\mathbf{Y}_t = (Y_t^A, Y_t^B)', \mathbf{X}_t = (X_t^A, X_t^B)', X_t^{A|B} = X_t^A - X_t^B$ é a diferença entre a chuva total da sub-bacia A e da sub-bacia B.

Abordagem Proposta: múltiplas séries de vazão

Caso Particular: três séries de vazão e função de transferência em (17a).

$$
\begin{aligned}
Y_t^A|Y_t^B &\sim p(\mu_t^{A|B}, \sigma^2_{A|B}) \qquad t = 1, \ldots, T\\
Y_t^B|Y_t^C &\sim p(\mu_t^{B|C}, \sigma^2_{B|C})\\
Y_t^C &\sim p(\mu_t^C, \sigma^2_C)\\
\mu_t^{A|B} &= \alpha^{A|B} + \eta^{A|B} Y_t^B + E_t^{A|B}\\
\mu_t^{B|C} &= \alpha^{B|C} + \eta^{B|C} Y_t^C + E_t^{B|C}\\
\mu_t^C &= \alpha^C + E_t^C \quad m = A|B, B|C, C\\
E_t^m &= \rho^m E_{t-1}^m + \gamma^m X_t^m |m| + w_t^m \quad w_t^m \sim N(0, W_m);
\end{aligned}
$$

onde $|m|$ denota a área da sub-bacia m, e é multiplicada por X_t^m para representar o volume acumulado de chuva na área de drenagem correspondente. Se $m = A|B$, $X_t^{A|B}$ denota a chuva e $|A|B|$ é a área de drenagem de A sem incluir a área de drenagem de B.

Modelando a Chuva: uma bacia

- Seja $\{X_t(s), s \in B \subset \mathbb{R}^2, t = 1, 2, \cdots\}$ um campo aleatório. Aqui, $X_t(s) \geq 0$.
- A chuva numa bacia ou região B com área $|B|$, X_t, será descrita por:

$$X_t = |B|^{-1} \int_B X_t(s) ds, \quad \forall s \in B \tag{18}$$

- $X_t(s)$ pode seguir uma distribuição normal truncada e, como sugerido em Sanso & Guenni (2000), pode ser representada pelo seguinte modelo espaço-temporal:

$$X_t(s_i) = \begin{cases} w_t(s_i)^\beta & \text{se } w_t(s_i) > 0, \quad s_i \in B \\ 0 & \text{se } w_t(s_i) \leq 0 \end{cases} \tag{19a}$$

$$\mathbf{w}_t(\mathbf{s}) = \boldsymbol{\Theta}_t(\mathbf{s}) + Z_t(\mathbf{s}) + \epsilon_t(\mathbf{s}) \tag{19b}$$

$$Z_t(\mathbf{s}) \sim GP(\mathbf{0}, \sigma^2 \varrho(\|s_1, s_2\|, \lambda)) \tag{19c}$$

Modelando a Chuva: múltiplas bacias

- É necessário realizar a **troca de suporte**:

$$X_t^m = \frac{1}{|m|} \int_m X_t(s) ds \quad \forall s \in m \tag{20}$$

onde X_t^m denota a chuva no tempo t e $|m|$ denota a área da bacia m.

- Na prática, a integral em (20) é aproximada por

$$X_t^m \approx \frac{1}{N_m} \sum_{i=1}^{N_m} \hat{X}_t(s_i) \quad i = 1, \ldots, N_m \tag{21}$$

onde N_m denota o número de pontos de uma grade de interpolação construída dentro dos limites da bacia m e $\hat{X}_t(i)$ é o valor predito de chuva para a localização s_i dessa grade.

Abordagem Proposta: caso geral

Nossa proposta consiste em **realizar o ajuste do modelo para a vazão em (16) e o modelo para a chuva em (19), simultaneamente**.

- Esta proposta pode ser vista como um sistema simples mas completo e eficiente para ajustar e prever duas das mais importantes variáveis hidrológicas.
- É bastante flexível. A especificação do modelo apresentado envolve várias subclasses de modelos.
- Todos os parâmetros têm uma interpretação física clara.
- Toda a incerteza envolvida nos processos físicos é explicitamente levada em conta.

Procedimento de Inferência (uma bacia)

- Sejam $\mathbf{Y} = (Y_1, \ldots, Y_T)'$ e $\mathbf{X}(\mathbf{s}) = (X_1(\mathbf{s}), \ldots, X_T(\mathbf{s}))'$, onde $\mathbf{s} = (s_1, \ldots, s_S)$. Também, $\mathbf{X}(s_i) = (\mathbf{X}_1(s_i), \ldots, \mathbf{X}_T(s_i))', s_i \in B, i = 1, \ldots, S$, $\mathbf{X}_t(\mathbf{s}) = (X_t(s_1), \ldots, X_t(s_S))'$, e $\mathbf{X} = (X_1, \ldots, X_T)'$.
- A distribuição conjunta de $\mathbf{Y}$, $\mathbf{X}$ e $\mathbf{X}(s)$ é dada por

$$\begin{aligned} p(\mathbf{Y}, \mathbf{X}, \mathbf{X}(\mathbf{s})|\boldsymbol{\Theta}) &= p(\mathbf{Y}|\mathbf{X}, \mathbf{X}(\mathbf{s}), \boldsymbol{\Theta}_Y)p(\mathbf{X}, \mathbf{X}(\mathbf{s})|\boldsymbol{\Theta}_X) \qquad (22) \\ &= \prod_{t=1}^{T} p(Y_t|X_t, \mathbf{X}_t(\mathbf{s}), \boldsymbol{\Theta}_Y)p(X_t|\mathbf{X}_t(\mathbf{s}), \boldsymbol{\Theta}_X) \\ &\quad \prod_{i=1}^{S} p(\mathbf{X}_t(\mathbf{s}_i)|\boldsymbol{\Theta}_X) \end{aligned}$$

- A distribuição preditiva, necessária para interpolar a chuva, $X_t(s_i)$ é

$$p(\mathbf{X}(s')|\mathbf{X}(s)) = \int p(\mathbf{X}(s')|\mathbf{X}(s), \boldsymbol{\Theta}_X)p(\boldsymbol{\Theta}_X|\mathbf{X}(s))p(\boldsymbol{\Theta}_X)d\boldsymbol{\Theta}_X. \qquad (23)$$

Modelagem na Prática

Para Chuva (Sansó & Guenni (2000))

- ✓ Distribuição: Normal Truncada
- ✓ $\mathbf{F}_i = (1, \text{long}_i, 1, 0, 1, 0)'$
- ✓ $\mathbf{G} = diag(\mathbf{G}_1, \mathbf{G}_2)$. $\mathbf{G}_1 = \mathbf{I}_2$ e $\mathbf{G}_2 =$ dois harmônicos.

$$X_t(s_i) = \begin{cases} w_t(s_i)^\beta & \text{se } w_t(s_i) > 0 \\ 0 & \text{se } w_t(s_i) \leq 0 \end{cases}$$

$$\mathbf{w}_t = \mathbf{z}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim N_N(\mathbf{0}, \tau^2\mathbf{I})$$

$$\mathbf{z}_t = \mathbf{F}'\boldsymbol{\Theta}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim N_N(\mathbf{0}, \sigma^2\mathbf{V}_t)$$

$$\boldsymbol{\Theta}_t = \mathbf{G}\boldsymbol{\Theta}_{t-1} + \boldsymbol{\varepsilon}_t, \; \boldsymbol{\varepsilon}_t \sim N_k(\mathbf{0}, \mathbf{W}_t)$$

Modelagem na Prática

Para Chuva (Sansó & Guenni (2000))

- ✓ Distribuição: Normal Truncada
- ✓ $\mathbf{F}_i = (1, \text{long}_i, 1, 0, 1, 0)'$
- ✓ $\mathbf{G} = diag(\mathbf{G}_1, \mathbf{G}_2)$. $\mathbf{G}_1 = \mathbf{I}_2$ e $\mathbf{G}_2 =$ dois harmônicos.

$$X_t(s_i) = \begin{cases} w_t(s_i)^\beta & \text{se } w_t(s_i) > 0 \\ 0 & \text{se } w_t(s_i) \leq 0 \end{cases}$$

$$\mathbf{w}_t = \mathbf{z}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim N_N(\mathbf{0}, \tau^2\mathbf{I})$$

$$\mathbf{z}_t = \mathbf{F}'\boldsymbol{\Theta}_t + \boldsymbol{\epsilon}_t, \;\; \boldsymbol{\epsilon}_t \sim N_N(\mathbf{0}, \sigma^2\mathbf{V}_t)$$

$$\boldsymbol{\Theta}_t = \mathbf{G}\boldsymbol{\Theta}_{t-1} + \boldsymbol{\varepsilon}_t, \; \boldsymbol{\varepsilon}_t \sim N_k(\mathbf{0}, \mathbf{W}_t)$$

Para Vazão (Migon & Monteiro, 1997)

- ✓ Distribuição: Log-Normal ou Gama
- ✓ Séries de tempo não estacionárias
- ✓ Efeito da Chuva: FT de 1a ordem

$$Y_t|X_t \sim p(\mu_t, \phi) \; t = 1, \ldots, T$$

$$\log(\mu_t) = \alpha_t + E_t$$

$$E_t = \rho E_{t-1} + \gamma_t X_t + w_t, \; w_t \sim N(0, \sigma_E^2)$$

$$\alpha_t = \mathbf{G}_\alpha \alpha_{t-1} + w_{\alpha,t}, \; w_{\alpha,t} \sim N(0, \sigma_\alpha^2)$$

$$\gamma_t = \mathbf{G}_\gamma \gamma_{t-1} + w_{\gamma,t}, \; w_{\gamma,t} \sim N(0, \sigma_\gamma^2)$$

Obtendo Amostras da Posteriori

Maior Desafio: obter amostras da distribuição a posteriori dos parâmetros ($\boldsymbol{\Theta}_t$) do modelo de vazão.

Obtendo Amostras da Posteriori

Maior Desafio: obter amostras da distribuição a posteriori dos parâmetros ($\boldsymbol{\Theta}_t$) do modelo de vazão.
Objetivo : obter uma boa distribuição proposta para o passo de Metropolis-Hastings na amostragem de $\boldsymbol{\Theta}$.

Obtendo Amostras da Posteriori

Maior Desafio: obter amostras da distribuição a posteriori dos parâmetros ($\mathbf{\Theta}_t$) do modelo de vazão.
Objetivo : obter uma boa distribuição proposta para o passo de Metropolis-Hastings na amostragem de $\mathbf{\Theta}$.

- Um valor candidato para $\mathbf{\Theta}$, $\mathbf{\Theta}^*$, será amostrado de uma distribuição multivariada obtida combinando as idéias de West, Harrison & Migon (1985) e Frühwirth-Schnater (1994).

Obtendo Amostras da Posteriori

Maior Desafio: obter amostras da distribuição a posteriori dos parâmetros ($\mathbf{\Theta}_t$) do modelo de vazão.
Objetivo : obter uma boa distribuição proposta para o passo de Metropolis-Hastings na amostragem de $\mathbf{\Theta}$.

- Um valor candidato para $\mathbf{\Theta}$, $\mathbf{\Theta}^*$, será amostrado de uma distribuição multivariada obtida combinando as idéias de West, Harrison & Migon (1985) e Frühwirth-Schnater (1994).
- Especificamente, a distribuição normal multivariada $N(\mathbf{\Theta}^*|\mathbf{m}^s, \mathbf{C}^s)$, com média e variância *on-line* - $\mathbf{m}, \mathbf{C}$ - aproximadas pelo *Conjugate Updating*, ao invés do Filtro de Kalman.

Obtendo Amostras da Posteriori

Maior Desafio: obter amostras da distribuição a posteriori dos parâmetros ($\mathbf{\Theta}_t$) do modelo de vazão.
Objetivo : obter uma boa distribuição proposta para o passo de Metropolis-Hastings na amostragem de $\mathbf{\Theta}$.

- Um valor candidato para $\mathbf{\Theta}$, $\mathbf{\Theta}^*$, será amostrado de uma distribuição multivariada obtida combinando as idéias de West, Harrison & Migon (1985) e Frühwirth-Schnater (1994).
- Especificamente, a distribuição normal multivariada $N(\mathbf{\Theta}^*|\mathbf{m}^s, \mathbf{C}^s)$, com média e variância *on-line* - $\mathbf{m}, \mathbf{C}$ - aproximadas pelo *Conjugate Updating*, ao invés do Filtro de Kalman.
- Cada $\boldsymbol{\theta}_t^*$ é amostrado seqüencialmente de $t = T$ até $t = 1$, de suas distribuições retrospectivas, dadas pela fatorização de $N(\mathbf{\Theta}^*|\mathbf{m}^s, \mathbf{C}^s)$ em T densidades condicionais univariadas, de maneira análoga ao caso gaussiano.

Ver Ravines, Migon & Schmidt (Tech. Rep., 2007) para maiores detalhes.

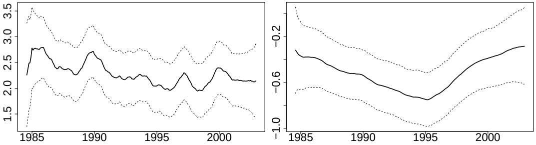
Modelo Espaço-Temporal da Chuva: Resultados (1)

O modelo espaço-temporal selecionado para a chuva tem um intercepto e um efeito linear da longitude na tendência espacial.

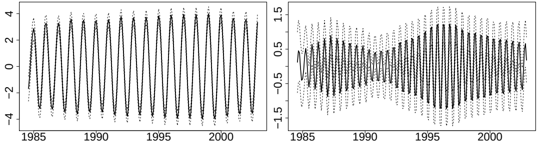
Tabela: Estatísticas a posteriori dos parâmetros estáticos

	mean	sd	2,5%	25%	50%	75%	97,5%	$\hat{R}$
β	1,732	0,016	1,701	1,722	1,732	1,743	1,764	1,001
λ	0,045	0,007	0,033	0,040	0,044	0,050	0,061	1,001
ς^2	0,719	0,040	0,644	0,691	0,718	0,746	0,798	1,001
σ^2	1,100	0,044	1,015	1,070	1,098	1,128	1,191	1,003

Modelo Espaço-Temporal da Chuva: Resultados (1)



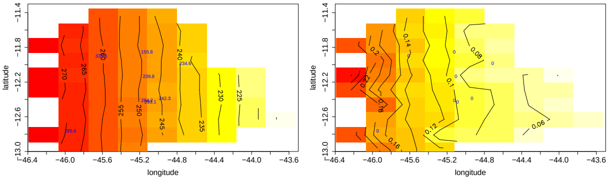
(a) Intercepto (b) Longitude



(c) 1^o Harmônico (d) 2^o Harmônico

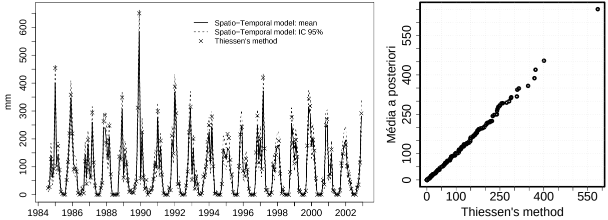
Figura: Trajetória estimada para os parâmetros dinâmicos.

Modelo Espaço-Temporal da Chuva: Resultados (2)



(a) Dezembro, 2000 (b) Junho, 2002

Figura: Médias a posteriori de chuva estimadas para dois meses diferentes.



(a) Média e Intervalo de 95% para a chuva (b) Q-Q plot

Modelo Espaço-Temporal da Chuva: Resultados (3)

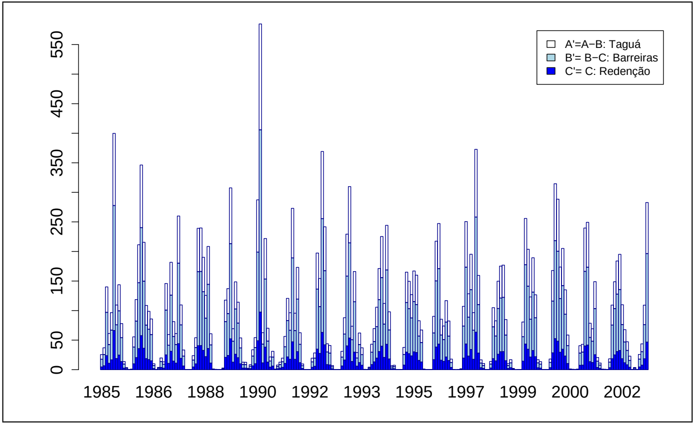


Figura: Média a posteriori estimada para a chuva acumulada por sub-bacia.

Modelando a Vazão: Resultados (1)

Assumimos que $p(Y_t^m)$ é uma distribuição $\log\text{-normal}$ com parâmetros μ_t^m e σ_m^2 e, representamos a relação entre Y_t^m e X_t^m mediante uma função de transferência de primeira ordem.

Tabela: Algumas estatísticas a posteriori

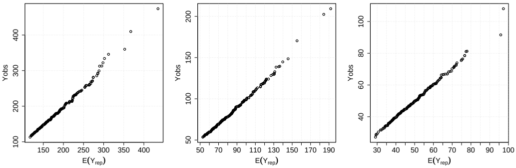
	média	25%	50%	75%	$\hat{R}$
$A\|B =$ **Taguá | Barreiras**					
α^A	1,151	1,022	1,142	1,276	1,008
η^A	0,915	0,887	0,916	0,945	1,010
ρ^A	0,510	0,387	0,533	0,658	1,001
γ^A	0,025	0,014	0,024	0,035	1,001
$\sigma^2_{E^A}$	0,003	0,002	0,003	0,004	1,015
$\sigma^2_{y^A}$	0,007	0,006	0,007	0,009	1,040
$B\|C =$ **Barreiras | Redenção**					
α^B	0,587	0,507	0,581	0,654	1,006
η^B	0,989	0,971	0,990	1,009	1,006
ρ^B	0,743	0,692	0,752	0,806	1,004
γ^B	0,005	0,003	0,005	0,008	1,001
$\sigma^2_{E^B}$	0,003	0,002	0,002	0,003	1,005
$\sigma^2_{y^B}$	0,004	0,003	0,004	0,005	1,068
$C =$ **Redenção**					
α^C	3,543	3,525	3,544	3,561	1,014
ρ^C	0,594	0,575	0,594	0,613	1,027
γ^C	1,645	1,597	1,645	1,696	1,001
$\sigma^2_{E^C}$	0,005	0,005	0,005	0,006	1,010
$\sigma^2_{y^C}$	0,002	0,002	0,002	0,003	1,107

Modelando a Vazão: Resultados (2)

Tabela: Variância e Correlação entre as três séries de vazão

Parâmetro	média	d.p.	2,5%	50%	97,5%	$\hat{R}$
Variância						
σ^2_A	0,039	0,006	0,030	0,038	0,049	1,015
σ^2_B	0,026	0,007	0,020	0,025	0,035	1,019
σ^2_C	0,010	0,001	0,008	0,010	0,013	1,000
Correlação						
$Corr[Y^A, Y^B]$	0,750	0,050	0,648	0,753	0,846	1,014
$Corr[Y^B, Y^C]$	0,630	0,054	0,526	0,630	0,732	1,014
$Corr[Y^A, Y^C]$	0,473	0,054	0,377	0,470	0,589	1,017

Modelando a Vazão: Resultados (3)



(a) $A|B$ (b) $B|C$ (c) C

Figura: Valores ajustados de vazão, usando o modelo condicional.

Tabela: Modelo Conjunto vs Modelos Univariados

	MSE	MAE
Modelo Multivariado		
$A\|B$	120,684	7,025
$B\|C$	9,060	1,820
C	2,568	1,009
Modelos Univariados Independentes		
A	162,991	7,756
B	8,047	1,652
C	2,568	1,009

Modelando a Vazão: Resultados (4)

Tabela: **Estimativas pontuais para o valor da vazão com o modelo conjunto e com os modelos univariados independentes em $t = 35, 80, 140, 200$.**

	Y_t real	média	d.p.	2,5%	50%	97,5%	$\hat{R}$
				Modelo Multivariado			
$t = 35$	149,37	154,80	17,25	125,11	153,95	194,17	1,00
$t = 80$	241,65	240,09	27,60	192,44	237,72	300,01	1,01
$t = 140$	184,90	190,56	22,03	153,09	189,39	239,06	1,00
$t = 200$	205,73	205,42	23,22	161,66	204,97	253,59	1,00
				Modelos Univariados Independentes			
$t = 35$	149,37	156,42	12,96	132,63	155,95	183,32	1,00
$t = 80$	241,65	246,58	20,49	209,10	245,61	290,81	1,00
$t = 140$	184,90	189,07	15,20	161,72	187,89	221,65	1,00
$t = 200$	205,73	202,38	17,26	172,35	201,59	237,59	1,00

Modelagem de Múltiplas Séries de Vazão

Contribuição: **Modelagem proposta: abordagem conjunta com funções de transferência bastante adequada para modelar várias séries de vazão de bacias encaixadas.**

- A idéia base é usar a representação condicional da distribuição conjunta.
- Mostramos que as estimativas pontuais (média a posteriori) das séries de vazão ficam melhores com o modelo conjunto do que com modelos individuais para cada série.
- Mostramos também que a incorporação ao modelo, da vazão de uma bacia menor, permite uma melhor interpretação e discriminação dos efeitos de cada variável.

Agradecimentos:

- Aos meus co-autores:
  A. O'Hagan, A. E. Gelfand, S. Banerjee, B. Sansó, H. S. Migon, M. A. R. Ferreira, A. Nobre, R. Ravines, R. Ruiz, R. Estrella;
- Ao CNPq pelo apoio para desenvolver meus projetos de pesquisa.

Referências:

Para maiores detalhes visite www.dme.ufrj.br/alex