Escola Superior de Agricultura "Luiz de Queiroz" Universidade de São Paulo

Modelos gaussianos geoestatísticos espaço-temporais e aplicações

Aluno: Alexandre Sousa da Silva
Orientador: Profº Dr. Paulo J. Ribeiro Júnior

Piracicaba
05 de Fevereiro 2007

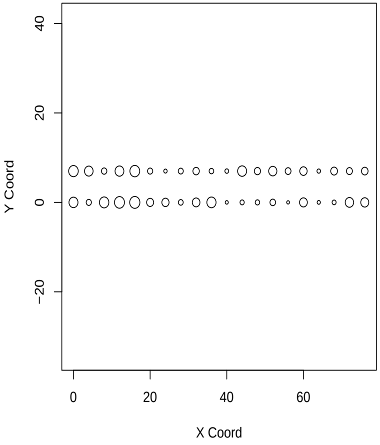
Motivação

- Falta de trabalhos sobre o tema;

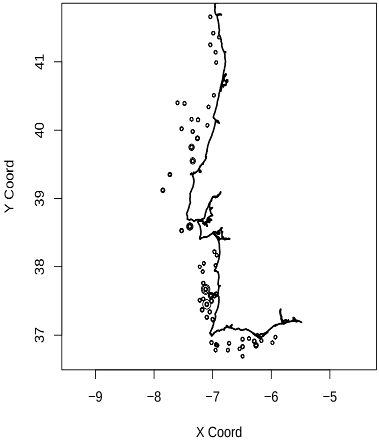
Motivação

- Falta de trabalhos sobre o tema;
- Grande aplicabilidade.

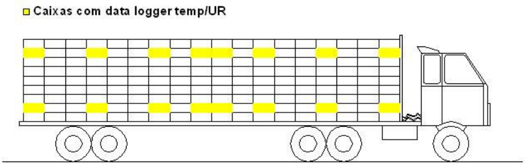
Exemplos de aplicação



Exemplos de aplicação

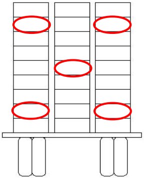


Exemplos de aplicação

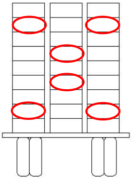


Lateral do Caminhão, com o posicionamento dos loggers. As duas laterais são iguais, porém a fileira do meio é diferente.

Corte Transversal do Caminhão (Fileira 10)



Corte Transversal do Caminhão (Demais fileiras)



Campos Aleatórios

Um campo aleatório ou uma função aleatória é um processo estocástico definido no espaço $G \subset \mathbf{R}^d$. Um campo aleatório é definido como:

$$\{Z(\mathbf{s}) : \mathbf{s} \in G \subset \mathbf{R}^d\},$$

em que $Z(\mathbf{s})$ é o valor do atributo Z na localização $\mathbf{s}$ e $d \geq 1$ é a dimensão do campo aleatório.

Campos Aleatórios Espaço-Temporais

- *Abordagem baseada em modelos*: covariância induzida pelo modelo, usualmente sem forma explícita, típica de modelos hierárquicos;

Campos Aleatórios Espaço-Temporais

- *Abordagem baseada em modelos*: covariância induzida pelo modelo, usualmente sem forma explícita, típica de modelos hierárquicos;

- *Abordagem geoestatística*: forma explícita para funções de covariância (espaço e/ou espaço temporias);

Propriedade da função de covariância estacionário de segunda ordem

(i) $Cov[Z(\mathbf{s}), Z(\mathbf{s}+\mathbf{0})] = Var[Z(\mathbf{s})] = C(\mathbf{0}) \geq 0$;

(ii) $C(\mathbf{h}) = C(\text{-}\mathbf{h})$;

(iii) $C(\mathbf{0}) \geq\mid C(\mathbf{h}) \mid$;

(iv) $C(\mathbf{h}) = Cov[Z(\mathbf{s}), Z(\mathbf{s}+\mathbf{h})] = Cov[Z(\mathbf{0}), Z(\mathbf{h})]$;

(v) If $C_j(\mathbf{h})$ com $j = 1, 2, ..., k$, são funções de covariância válidas, então $\sum_{j=1}^{k} b_j C_j(\mathbf{h})$ é uma função de covariância válida, se $b_j \geq 0 \forall j$;

(vi) If $C_j(\mathbf{h})$ com $j = 1, 2, ..., k$, são funções de covariância válidas, então $\prod_{j=1}^{k} C_j(\mathbf{h})$ é uma função de covariância válida;

(vii) If $C(\mathbf{h})$ é uma função válidas no $\mathbf{R}^d$, então ela também será uma função de covariância válida em $\mathbf{R}^p$, com $p < d$.

Para que uma função de covariância de um campo aleatório estacionário seja considerada válida é necessário e suficiente que C satisfaça a condição de ser positiva definida, ou seja:

$$\Sigma_{i=1}^{k}\Sigma_{j=1}^{k}a_i a_j C(\mathbf{s}_i - \mathbf{s}_j) \geq 0,$$

para qualquer conjunto de localizações e número real.

Campos Aleatórios Espaço-Temporais

Campo aleatório espaço-temporal é definido como:

$$\{Z(\mathbf{s}, t), \mathbf{s} \in R^d, t \in R\}.$$

Percebe-se intuitivamente que o domínio natural do processo é $R^d \times R$.
Defini-se a média e a função de covariância como:

$$\mu(\mathbf{s}, t) = E(Z(\mathbf{s}, t))$$

$$Cov(Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)) = C(\mathbf{s}_1, \mathbf{s}_2, t_1, t_2)$$

Campos Aleatórios Espaço-Temporais

- Estacionariedade;

$$Cov(Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)) = c(\mathbf{h}, u)$$

Campos Aleatórios Espaço-Temporais

- Estacionariedade;

$$Cov(Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)) = c(\mathbf{h}, u)$$

- Completa Simetria;

$$Cov(Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)) = Cov(Z(\mathbf{s}_1, t_2), Z(\mathbf{s}_2, t_1))$$

Campos Aleatórios Espaço-Temporais

- Estacionariedade;

$$Cov(Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)) = c(\mathbf{h}, u)$$

- Completa Simetria;

$$Cov(Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)) = Cov(Z(\mathbf{s}_1, t_2), Z(\mathbf{s}_2, t_1))$$

- Estacionariedade e Completa Simetria .

$$C(\mathbf{h}, u) = C(\mathbf{h}, -u) = C(-\mathbf{h}, u) = C(-\mathbf{h}, -u)$$

Separabilidade

$$Cov(Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)) = Cov(Z(\mathbf{s}_1, \mathbf{s}_2)) + Cov(Z(t_1, t_2))$$

$$Cov(Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)) = Cov(Z(\mathbf{s}_1, \mathbf{s}_2))Cov(Z(t_1, t_2)),$$

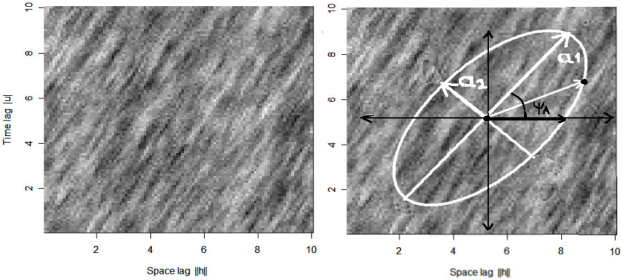
em ambos $\mathbf{s}_1, t_1$ e $\mathbf{s}_2, t_2 \in \mathbf{R}^2 \times \mathbf{R}$.

Exemplo de covariância espaço-temporal separável

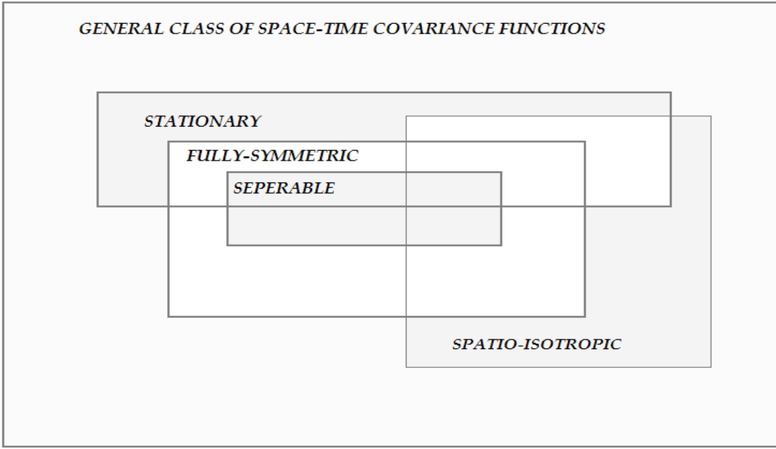
$$Cov(Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)) =$$
$$\sigma_1^2 exp(-\phi_1||(\mathbf{s}_1 - \mathbf{s}_2)||) + \sigma_2^2 exp(-\phi_2||(t_1 - t_2)||),$$

em que ϕ_1 e ϕ_2 são as matrizes de anisotropia.

Modelos e suas simplificações



Modelos e suas simplificações



Funções de covariância não separáveis, algumas alternativas:

- Cressie e Huang (1999);

Funções de covariância não separáveis, algumas alternativas:

- Cressie e Huang (1999);
- Gneiting (2002);

Campos Aleatórios Espaço-Temporais

Funções de covariância não separáveis, algumas alternativas:

- Cressie e Huang (1999);
- Gneiting (2002);
- outros...

Representação de Cressie-Huang

Considere C estacionária e contínua e que a função de distribuição espectral possua densidade espectral $g(\mathbf{w}, \tau) \geq 0$. Pelo teorema de Bochner a função de covariância espaço-temporal será considerada válida se:

$$C(\mathbf{h}, u) = \int \int exp\{i\mathbf{h}'\mathbf{w} + iu\tau\} g(\mathbf{w}, \tau) d\mathbf{w} d\tau$$

Idéia: constrir funções positivas definidas no R^d que sejam também válidas em R^{d+1} .

Representação de Cressie-Huang

Desta forma, seja $C(\mathbf{h}, u)$ integrável, temos através da tranformção inversa de Fourier que:

$$
\begin{aligned}
g(\mathbf{w}, \tau) &= (2\pi)^{-d-1} \int \int exp\{-i\mathbf{h}'\mathbf{w} - iu\tau\} C(\mathbf{h}, u) d\mathbf{h} du \\
&= (2\pi)^{-1} \int exp\{-iu\tau\} h(\mathbf{w}, u) du
\end{aligned}
$$

em que

$$
\begin{aligned}
h(\mathbf{w}, u) &= (2\pi)^{-d} \int exp\{-i\mathbf{h}'\mathbf{w}\} C(\mathbf{h}, u) d\mathbf{h} \\
&= \int exp\{iu\tau\} g(\mathbf{w}, \tau) d\tau.
\end{aligned}
$$

Representação de Cressie-Huang

Assumindo que:

$$h(\mathbf{w}, u) = \rho(\mathbf{w}, u)k(\mathbf{w})$$

as seguintes condições devem satisfeitaz:

$(C1)$ Para cada $\mathbf{w} \in R^d, \rho(\mathbf{w}, u)$, é uma função de autocorrelação contínua, $\int \rho(\mathbf{w}, u)du < \infty$, e $k(\mathbf{w}) > 0$.

$(C2)$ $\int k(\mathbf{w})d\mathbf{w} < \infty$.

tem-se que:

$$C(\mathbf{h}, u) = \int exp\{i\mathbf{h}'\mathbf{w}\}\rho(\mathbf{w}, u)k(\mathbf{w})d(\mathbf{w})$$

Desta forma o objetivo é satisfeito e para se construir uma função de covariância válida no R^{d+1}, basta que as condições $C1$ e $C2$ sejam satisfeitas na dimensão R^d.

Representação de Cressie-Huang

Um exemplo simples de covariância espaço-temporal não separável baseado nos resultados de Cressie e Huang (1999), é dado por: Considerando

$$\rho(\mathbf{w}, u) = exp\{-||\mathbf{w}||^2 u^2/4\}exp\{-\delta u^2\}; \delta > 0$$

$$k(\mathbf{w}) = exp\{-c_0||\mathbf{w}||^2/4\}; c_0 > 0,$$

que satisfizem as condições $(C1)$ e $(C2)$, tem-se:

$$C(\mathbf{h}, u) \propto \frac{1}{(u^2+c_0)^{d/2}} exp\left\{\frac{-||\mathbf{h}||^2}{u^2+c_0}\right\}exp\{-\delta u^2\}; \delta > 0,$$

a equação acima é uma função de covariância espaço-temporal contínua no $R^d \times R$.

Representação de Gneiting

A representação de Gneiting (2002) considerara qualquer função monótona $\phi(t)$, definida em $t \geq 0$ e qualquer função positiva $\psi(t)$, definida em $t \geq 0$ com derivadas completamente monótonas , então:

$$C(\mathbf{h}, u) = \frac{\sigma^2}{(\psi(u)^2)^{d/2}} \phi\left(\frac{||\mathbf{h}||^2}{\psi(|u|^2)}\right)$$

é uma função de covariância espaço-temporal válida no $R^d \times R$.

Uma função contínua $\phi(t)$ é completamente monótona se possui derivadas $\phi^{(n)}$ de todos as ordens e $(-1)^n \phi^{(n)}(t) \geq 0$, em que $t > 0, n = 0, 1, 2, \ldots$.

Representação de Gneiting

Funções completamente monótonas $\phi(t)$,$t \geq 0$
$\phi(t) = exp(-ct^{\gamma}), c > 0, 0 < \gamma \leq 1$
$\phi(t) = (1 + ct^{\gamma})^{\nu}, c > 0, 0 < \gamma \leq 1, \nu > 0$
$\phi(t) = (2^{\nu-1}\Gamma(\nu))^{-1}(ct^{1/2})^{\nu}\mathbf{K}_{\nu}(ct^{1/2}), c > 0, \nu > 0$
$\phi(t) = 2^{\nu}(exp(ct^{1/2}) + exp(-ct^{1/2}))^{\nu}, c > 0, \nu > 0$

Funções positivas $\psi(t)$, $t \geq 0$
$\psi(t) = (at^{\alpha} + 1)^{\beta}, a > 0, 0 < \alpha \leq 1, 0 \leq \beta \leq 1$
$\psi(t) = ln(at^{\alpha} + b)/ln(b), a > 0, b > 1, 0 < \alpha \leq 1$
$\psi(t) = (at^{\alpha} + b)/(b(at^{\alpha} + 1)), a > 0, o < b \leq 1$

Representação de Gneiting

Um exemplo simples da família de covariância espaço-temporal não separável proposta por Gneiting é dada como segue:

$\phi(t) = exp(-ct^{\gamma}), c > 0, 0 < \gamma \leq 1$

$\psi(t) = (at^{\alpha} + 1)^{\beta}, a > 0, 0 < \alpha \leq 1, 0 \leq \beta \leq 1$

são as funções completamente monótona e positiva respectivamente.

$$C(\mathbf{h}, u) = \frac{\sigma^2}{(\psi(u)^2)^{d/2}} \phi\left(\frac{||\mathbf{h}||^2}{\psi(|u|^2)}\right)$$

$$C(\mathbf{h}, u) = \frac{\sigma^2}{(a|t|^{2\alpha}+1)^{\frac{\beta d}{2}}} exp\{-\frac{c||\mathbf{h}||^{2\gamma}}{(a|t|^{2\alpha}+1)^{\beta\gamma}}\}.$$

Representação de Gneiting

Para $\beta = 0$ a covariância não depende do tempo, e multiplicando-se por uma função de covariância puramente temporal, $(at^{\alpha} + 1)^{-\delta}$, tem-se com resultado:

$$C(\mathbf{h}, u) = \frac{\sigma^2}{(a|t|^{2\alpha}+1)^{\delta+\frac{\beta d}{2}}} exp\{-\frac{c||\mathbf{h}||^{2\gamma}}{(a|t|^{2\alpha}+1)^{\beta\gamma}}\},$$

se $\beta = 0$ tem-se uma função de covariância separável.

Isto sugere uma forma de testar a suposição de separabilidade.

Aplicações

Monitoramento de estoques de pescada na costa portuguesa

Simplificações

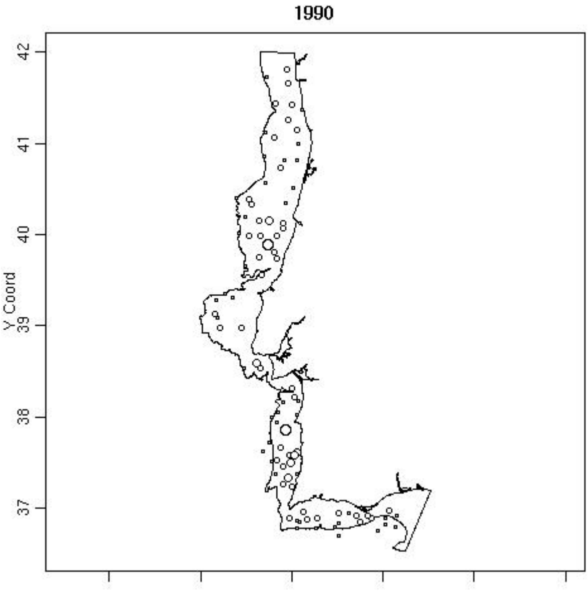
- Coletas em pontos diferentes no espaço;

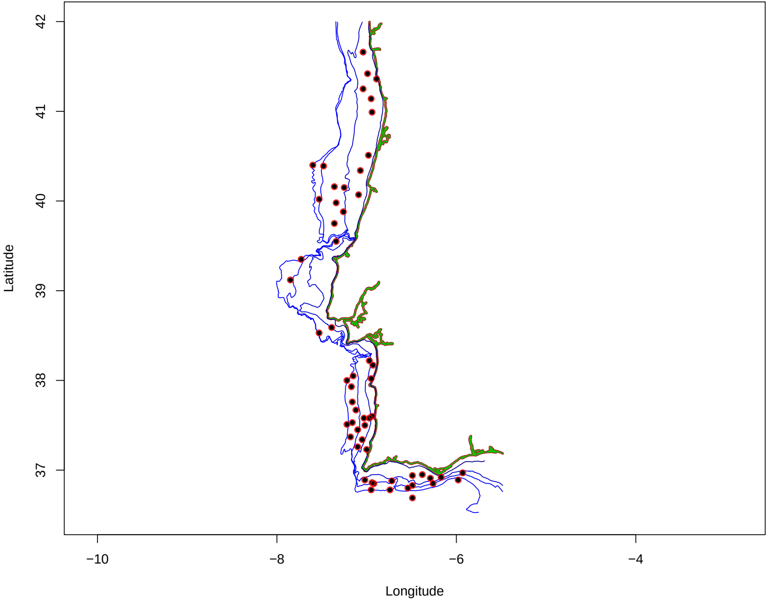
Aplicações

Monitoramento de estoques de pescada na costa portuguesa

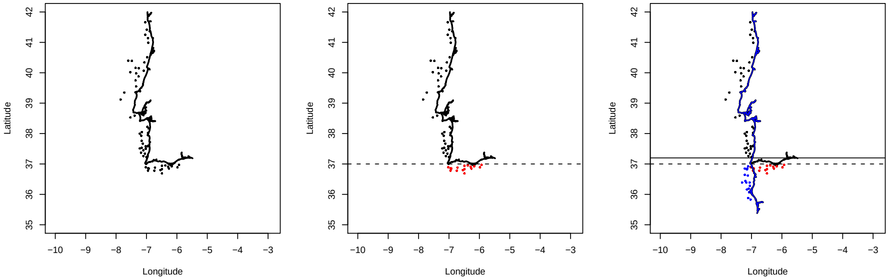
Simplificações

- Coletas em pontos diferentes no espaço;
- Rotação na coordenadas;


1990
Y Coord
42
41
40
39
38
37

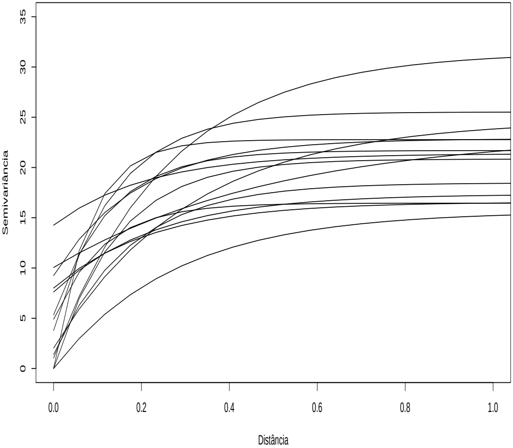


Latitude
42
41
40
39
38
37
−10
−8
−6
−4
Longitude

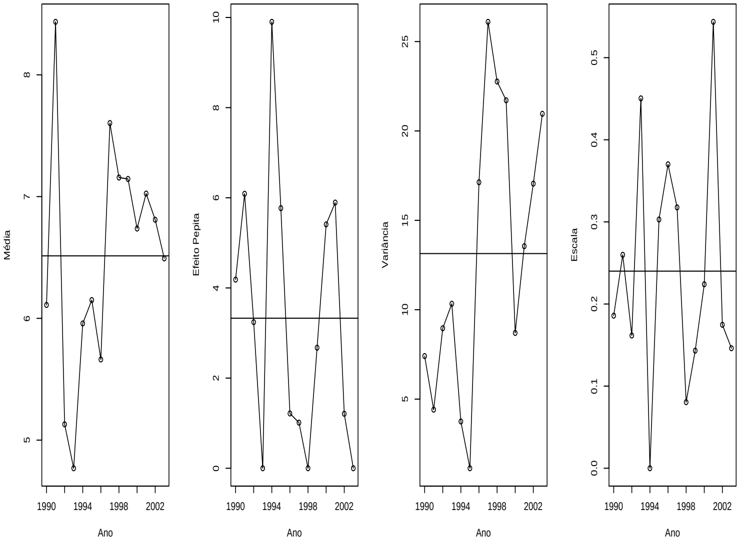


Latitude
Longitude
Latitude
Longitude
Latitude
Longitude

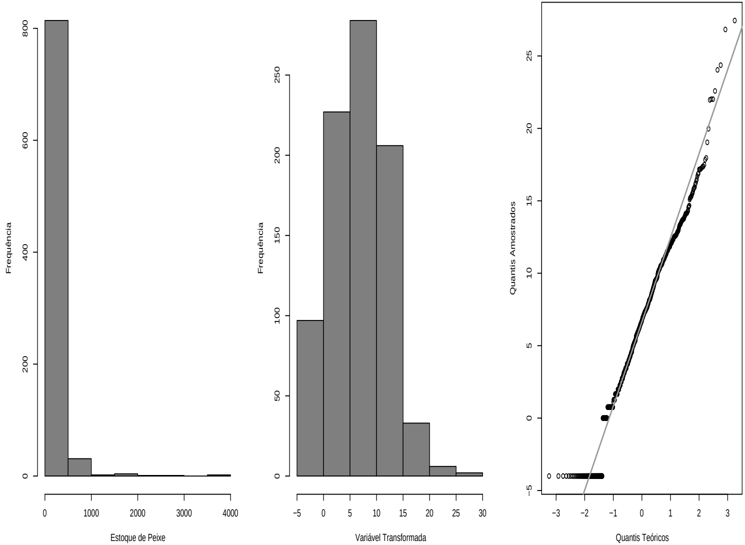
Análise descritiva



Análise descritiva



Análise descritiva



Modelagem com estrutura separável

$$C(\mathbf{h},u) = \frac{\sigma^2}{(a|u|^{\delta})} expc||h||^{\gamma} + \frac{\sigma^2_{h=0}}{(a|u|^{\delta})},$$

que é separável pois pode ser escrita da forma:

$$C(\mathbf{h},u) = C(\mathbf{h})C(u).$$

Estimando-se os parâmetros pelo método da máxima verossimilhança, obteve-se o seguinte modelo ajustado:

$$C(\mathbf{h},u) = \frac{32.26}{(0.10|u|^{0.2}+1)^{0.44}} exp\{-2.5||\mathbf{h}||^{0.83}\} + \frac{6.99}{(0.10|u|^{0.2}+1)^{0.44}}.$$

Modelagem com estrutura não separável

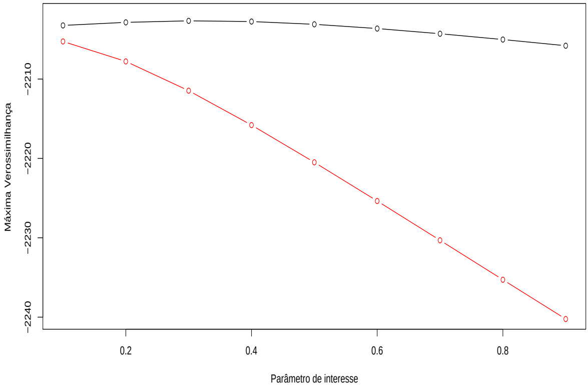
$$C(\mathbf{h}, u) = \frac{\sigma^2}{(\psi|u|^2)^{d/2}} \phi\left(\frac{||\mathbf{h}||^2}{\psi(|u|^2)}\right),$$

que para as escolhas de $\phi(.)$ e $\psi(.)$ utilizadas aqui tem como modelo ajustado pelo método da máxima verossimilhança:

$$C(\mathbf{h}, u) = \frac{1}{(0.10|u|^{0.2}+1)^{0.44}} \left(\frac{32.26}{(0.10|u|^{0.4}+1)^{0.0335}} exp\left\{ - 2.5\left[\frac{||\mathbf{h}||}{(0.10|u|^{0.4}+1)}\right]^{0.83} \right\} \right) + \frac{6.99}{(0.10|u|^{0.2}+1)^{0.44}}$$

- $\psi(|u|) = (0.10|u|^{0.4} + 1)^{0.0335}$
- $\phi(||\mathbf{h}||) = 32.26.exp\{-2.5||\mathbf{h}||^{0.83}\}$

Analisando o parâmetro β



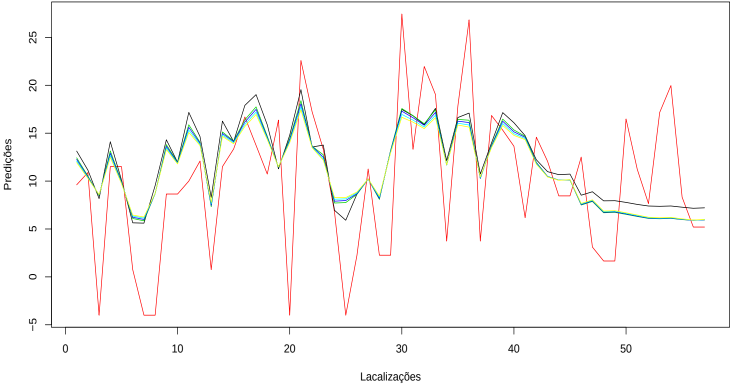
Máxima verossimilhança condicional (linha vermelha) e perfilhada

Analisando as predições

Erro quadrático médio de predição para funções de covariância espaço-temporais com diferentes valores do parâmetro β

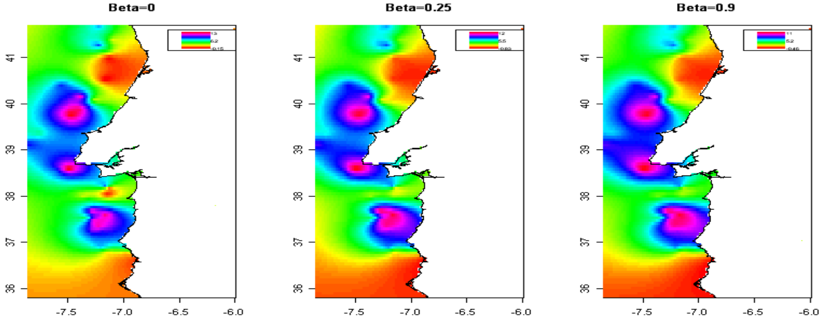
	EQM	Média	Variância
Modelo separável $\beta = 0$	41.34258	11.86858	15.73710
Modelo não separável $\beta = 0.25$	41.47026	11.33011	15.67702
Modelo não separável $\beta = 0.5$	41.61852	11.27031	14.76563
Modelo não separável $\beta = 0.75$	41.82565	11.20703	13.79720
Modelo não separável $\beta = 0.95$	42.03688	11.15387	12.98332

Analisando as predições



Distribuição das predições para 2004 e do verdadeiro valor amostrado

Analisando as predições



Krigagem para modelo com $\beta = 0$, $\beta = 0.25$ e $\beta = 0.95$, respectivamente

Conclusões

- Superioridade dos modelos espaço-temporais não separáveis;

Conclusões

- Superioridade dos modelos espaço-temporais não separáveis;
- Eficiência do RandomFields;

Conclusões

- Superioridade dos modelos espaço-temporais não separáveis;
- Eficiência do RandomFields;
- Análises futuras;

Bibbliografia

Cressie, N. and Huang, H.-C. (1999). Classes of Non-Seperable, Spatio-temporal stationary covariance functions. Journal of the American Statistical Association, 94, 1330-1340.

Elmatzoglou, I. (2006). Spatio-temporal geoestatistical models,with an aplication in fish stock. Submited for the degree of master in statistics at Lancaster University

Gneiting T.(2002), Nonseperable, Stationary Covariance Functions for Space-Time Data, American Statistical Association Journal of the American Statistical Association, June 2002, Vol.97, No.458.

Gneiting, T., and Schlather, M. (2002) Space-time covariance models. Encyclopedia of Environmetrics, Vol 4, pp 2041-2045.

Bibbliografia

Gneiting T., Genton M.G., Guttorp P.(2006), Geostatistical Space-Time Models, Stationarity, Seperability and Full Symmetry, Technical Report no.475 Department of Statistics University of Washington.

Schabenberger O., Gotway C.A. (2005) Statistical Methods for Spatial Data Analysis, Chapman Hall / CRC.

Schmidt, A.M., Sanso, B. (2006) Modelagem Bayesiana da Estrutura de Covariância de Processos Espaciais e Espaço-Temporais, 17^o SINAPE

RandomFields: Simulation and Analysis of Random Fields, Martin Schlather, R package version 1.3.28, http://www2.hsu-hh.de/schlath/index.html