MODELOS GAUSSIANOS GEOESTATSTICOS ESPAO-TEMPORAIS E APLICAES

Alexandre Sousa da SILVA [1]
Paulo Justiniano RIBEIRO JUNIOR[2]

■ RESUMO: A especificao de funes de covarincia espao-temporais uma das possveis estratgias para modelagem de processos dos quais observaes so tomadas em diferentes posies do espao e do tempo. Tais funes podem definir processos separveis ou no separveis e na sua especificao deve-se garantir que so funes de covarincia vlidas atendendo a condio de serem positiva definidas. Entre estratgias para obteno de tais funes esto as de Cressie e Huang (1999) e Gneiting (2002). A primeira se baseia na idia de obter funes em um espao de dimenso aumentada a partir de funes vlidas no espao original e necessita de operaes no domnio da freqncia. Alternativamente a segunda proposta utiliza combinao de funes completamente montonas e estritamente crescentes, evitando inverso de representaes espectrais. H ainda poucos relatos de uso e avaliaes comparativas das diferentes propostas. Neste trabalho considerou-se a metodologia proposta por Gneiting, com diferentes valores do parmetro que indica a fora da interao entre o espao e o tempo. Diferentes modelos foram aplicados dois conjuntos de dados, um referente a estoques de peixe na costa de Portugal, e outro referente armazenagem de gua em um solo com citros. Utilizou-se a implementao no pacote *RandomFields* do programa R, revisando-se a metodologia e investigando-se a implementao computacional. Na aplicao o modelo de covarincia separvel se mostrou adequado para descrever o comportamento das observaes disponveis sendo a escolha do modelo determinada por ajustes de mxima verossimilhana.

■ PALAVRAS-CHAVE: Campos Aleatrios; Geoestatstica; Modelos Espao-Temporias; Pacote *RandomFields*

[1]Departamento de Cincias Exatas, Universidade de So Paulo - ESALQ/USP, CEP 13418-900 Piracicaba, SP, Brasil. E-mails: *assilva@esalaq.usp.br*

[2]Departamento de Estatstica, Universidade Federal do Paran, CEP 81531-990 Curitiba, PR, Brasil. E-mails: *paulojus@ufpr.br*

1 Introduo

Dados espao-temporais so caracterizados pela variabilidade no tempo e no espao, e o objetivo da anlise estatstica deste tipo de dado est em descrever a incerteza no somente sobre as estimativas das quantidades de interesse mas, tambm, estimar valores em locais e/ou tempos no amostrados. Na literatura estatstica existe um grande nmero de publicaes referentes a processos puramente espaciais ou puramente temporais, e mais recentemente comeam a ser apresentadas propostas para modelagem conjunta no espao e tempo. Esta anlise conjunta o interesse dos modelos espao-temporais.

Modelos espao-temporais tm ganho uma crescente popularidade na ltima dcada, o que pode ser explicado por um lado pela numerosa aplicabilidade, principalmente em cincias ambientais e da sade; e por outro lado pelo crescente desenvolvimento de recursos computacionais.

Gneiting e Schlather (2002), em uma reviso sobre estratgias de modelagem, consideram duas possveis abordagens para a modelagem espao-temporal: a abordagem geoestatstica e a abordagem baseada em modelos.

Na abordagem geoestatstica, geralmente considerada para campos aleatrios gaussianos, em que o processo totalmente especificado pelo vetor de mdias e a matriz de covarincia. Em geral o vetor de mdias facilmente especificado a partir de informaes contextuais, isto no acontece com a matriz de covarincia, que o componente chave desta metodologia. Entretanto, nesta abordagem, esta matriz possui elementos dados por uma funo de covarincia vlida, que assegure a condio de matriz positiva definida.

Na abordagem baseada em modelos enfatizada a adoo de modelos estocsticos e nesta metodologia a funo de covarincia no a nica estrutura utilizada para especificar o modelo, sendo assim, este mtodo pode ir de uma funo analtica simples a uma funo intratvel e somente implicitamente induzida pelo modelo adotado. Neste contexto inclu-se os modelos Bayesianos flexveis, dinmicos e/ou homogneos e mtodos baseados na tcnica do filtro de Kalman.

Neste trabalho ser considerado a abordagem geoestatstica, para campos aleatrios espao-temporais gaussianos. Uma forma intuitiva de produzir modelos vlidos para a funo de covarincia espao-temporal atravs da combinao de funes vlidas puramente espacial e puramente temporal. Uma propriedade assegura que o produto ou a soma de funes de covarincias vlidas tem como resultado uma funo de covarincia vlida. As funes de covarincia espao-temporais obtidas desta forma so ditas separveis j que no contemplam a possibilidade de interao entre o componente espacial e temporal. Por isso, em geral no so realsticas j que assumem a independncia dos processos espaciais e temporais.

Alternativamente as funes de covarincia espao-temporais no separveis consideram a interao entre o componente espacial e temporal, mas apresentam como grande desvantagem a dificuldade em se construir funes de covarincia que sejam vlidas. Mtodos matemticos sofisticados devem ser aplicados para garantir a validade destas funes e em geral se valem do teorema de Bochener (BOCHNER, 1959), que utiliza a representao espectral para garantir a validade das funes de covarincia espao-temporais. Nesta linha, Cressie e Huang (1999), Gneiting (2002) e Stein (2005), por exemplo, propem famlias de modelos de covarincia no separveis vlidas para processos espao-temporais gaussianos.

2 Campos aleatrios

Um campo aleatrio ou uma funo aleatria um processo estocstico definido no espao $G \subset \mathbf{R}^d$, ou seja, uma funo cujos valores so realizaes de variveis aleatrias em qualquer ponto do domnio (SCHIMIDT; SANS, 2006), ou em outras palavras, uma famlia ou coleo de variveis aleatrias, em que cada um dos seus membros podem ser identificados ou localizados de acordo com a mesma mtrica (SCHABENBERGER; GOTWAY, 2005). Um campo aleatrio definido como:

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

em que $Z(\mathbf{s})$ o valor do atributo Z na localizao $\mathbf{s}$ e $d \geq 1$ a dimenso do campo aleatrio.Segundo Schmidt e Sans (2006) e Le e Zidek (2006), a descrio de um campo aleatrio obtida atravs das distribuies acumuladas finito-dimensionais F se, para qualquer conjunto de pontos nas localizaes $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_n$ pertencentes regio G e qualquer inteiro n,

$F_{\mathbf{s}_1,\mathbf{s}_2,...,\mathbf{s}_n}(z_1, z_2, ..., z_n) \equiv P\{Z(\mathbf{s}_1) \leq z_1, Z(\mathbf{s}_2) \leq z_2, ..., Z(\mathbf{s}_n) \leq z_n\}.$

Um campo aleatrio gaussiano corresponde ao caso particular em que tais distribuies finito-dimensionais so gaussianas, ou seja, um campo aleatrio gaussiano se a distribuio para qualquer conjunto de ndices $i = 1, 2, ..., n$ uma gaussiana n-variada. Conseqentemente, por propriedades da distribuio gaussiana multivariada, cada $Z(\mathbf{s}_i)$ uma varivel aleatria gaussiana univariada.

O processo gaussiano particularmente importante na anlise de campos aleatrios, j que estes so completamente especificados pelo vetor de mdias e sua matriz de covarincia (GNEITING; SCHLATHER, 2002). O vetor de mdias especificado pelo conhecimento de covariveis, isto , efeitos fixos, que influenciam a varivel de interesse. A matriz de covarincia precisa ser positiva definida para ser considerada vlida, para tanto cada um de seus elementos devem ser dados por uma funo de covarincia que garanta a condio de positiva definida. Entretanto determinar se esta funo positiva definida , em geral, no trivial. Uma forma freqentemente utilizada para verificar a validade de uma funo de covarincia verificar se a representao espectral desta funo satisfaz o teorema de Bochner (BOCHNER, 1959).

2.1 Propriedades da funo de covarincia

A funo de covarincia $C(\mathbf{h})$ de um campo aleatrio estacionrio de segunda ordem deve satisfazer as seguintes propriedades :
(i) $Cov[Z(\mathbf{s}), Z(\mathbf{s} + \mathbf{0})] = Var[Z(\mathbf{s})] = C(\mathbf{0}) \geq 0$;
(ii) $C(\mathbf{h}) = C(\mathbf{-h})$;

(iii) $C(\mathbf{0}) \geq| C(\mathbf{h}) |$;
(iv) $C(\mathbf{h}) = Cov[Z(\mathbf{s}), Z(\mathbf{s}{+}\mathbf{h})] = Cov[Z(\mathbf{0}), Z(\mathbf{h})]$;
(v) Se $C_j(\mathbf{h})$ com $j = 1, 2, ..., k$, so funes de covarincia vlidas, ento $\sum_{j=1}^{k} \mathbf{b}_j C_j(\mathbf{h})$ uma funo de covarincia vlida, se $\mathbf{b}_j \geq 0 \forall j$;
(vi) Se $C_j(\mathbf{h})$ com $j = 1, 2, ..., k$, so funes de covarincia vlidas, ento $\prod_{j=1}^{k} C_j(\mathbf{h})$ uma funo de covarincia vlida;
(vii) Se $C(\mathbf{h})$ uma funo vlida no R^d, ento ela tambm ser uma funo de covarincia vlida em R^p, com $p < d$.

De acordo com Schabenberger e Gotway (2005), as propriedades (i) e (ii) so imediatas, a (iii) segue a forma da inequao de Cauchy-Schwarz, a (iv) caracteriza a falta de importncia da coordenada absoluta, (v) assegura que combinaes lineares de funes de covarincia vlidas so tambm funes de covarincia vlidas e (vi) assegura que o produto de funes de covarincia vlidas tem como resultado uma funo de covarincia vlida.

Em campos aleatrios espao-temporais a propriedade (ii) nem sempre satisfeita, o que d origem aos modelos de covarincia espao-temporais assimtricos. J as propriedades (v) e (vi) do suporte construo de modelos de covarincia espao-temporais separveis.

A matriz de covarincia para $(Z(\mathbf{s}_1), Z(\mathbf{s}_2), ..., Z(\mathbf{s}_n))$, denotada por Σ com dimenso $n \times n$, ser uma matriz simtrica, em que $\Sigma_{ij} = C(\mathbf{s}_i - \mathbf{s}_j)$, ou seja, cada valor de Σ igual a covarincia entre respostas em duas localizaes correspondentes. Para que uma funo de covarincia de um campo aleatrio estacionrio seja considerada vlida necessrio e suficiente que C satisfaa a condio de ser positiva definida, em outras palavras, para qualquer vetor no zero $\mathbf{a}$, a forma quadrtica $\mathbf{a}'\Sigma\mathbf{a}$ deve ser maior igual a zero. Sendo assim esta condio pode ser escrita nas seguintes formas:

$$\mathbf{a}'\Sigma\mathbf{a} = \sum_{i=1}^{k} \sum_{j=1}^{k} \mathbf{a}_i \mathbf{a}_j C(\mathbf{s}_i - \mathbf{s}_j) \geq 0 = \sum_{i=1}^{k} \sum_{j=1}^{k} \mathbf{a}_i \mathbf{a}_j C(\mathbf{h}_{ij}) \geq 0,$$

em que $\mathbf{a}_i$ e $\mathbf{a}_j$ so elementos do vetor $\mathbf{a}$ e $\mathbf{h}_{ij}$ denota o vetor de separao entre $\mathbf{s}_i$ e $\mathbf{s}_j$.

Uma ferramenta de uso comum para investigao da dependncia espacial em geoestatstica o semivariograma. Para campos aleatrios intrinsecamente estacionrios define-se $2\gamma(\mathbf{h}) = E\{[\mathbf{Z}(s_i) - \mathbf{Z}(s_j)]^2\}$. Portanto, uma estimativa emprica do semivariograma dada pelo estimador de momentos:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2|N(\mathbf{h})|} \sum_{|N(\mathbf{h})|} \{Z(\mathbf{s}_i) - Z(\mathbf{s}_j)\}^2,$$

em que $|N(\mathbf{h})|$ a quantidade de pontos separados pela distncia $\mathbf{h}$ de separao entre s_i e s_j. De acordo com Schabenberger e Gotway (2005), o estimador $\hat{\gamma}(\mathbf{h})$ no viesado para $\gamma(\mathbf{h})$ se $Z(\mathbf{s})$ for intrinsecamente estacionrio. No caso do processo em que assume-se a estacionariedade de segunda ordem, o variograma informa tambm sobre a funo de correlao.

Isto justifica uma prtica comum em geoestatstica de se utilizar o semivariograma emprico para inferir sobre parmetros do modelo, atravs do ajuste de uma funo vlida para o semivariograma terico.

Um outro paradigma para estimao de parmetros, sob a pressuposio de estacionariedade forte, dado pelo mtodo de mxima verossimilhana. A estimao dos parmetros de campos aleatrios pelo mtodo de mxima verossimilhana requer pressuposies sobre a distribuio do processo. Dado $\mathbf{Z} = [Z(\mathbf{s}_1), Z(\mathbf{s}_2), ..., Z(\mathbf{s}_n)]'$, que denota o vetor de observaes e assumindo $\mathbf{Z}(\mathbf{s}) \sim G(\mu\mathbf{1}, \Sigma(\theta))$, o mtodo da mxima verossimilhana consiste em maximizar o logaritmo da distribuio gaussiana n-variada:

$$L(\theta; Z_1, Z_2, ..., Z_n) = -\frac{1}{2}\{\ln(|\Sigma(\theta)|) + n\ln(2\pi) + (Z(\mathbf{s}) - \mathbf{1}\mu)'\Sigma(\theta)^{-1}(Z(\mathbf{s}) - \mathbf{1}\mu)\}. \tag{1}$$

Estimadores de mxima verossimilhana so muito utilizados em estatstica, por assegurarem uma srie de propriedades, j que sobre condio de regularidade padro estes estimadores so assintoticamente gaussianos e eficientes. Na prtica, em geoestatstica, este mtodo de estimao pode requerer muito tempo computacional, ou at mesmo ser proibitivo, devido dimenso da matriz de covarincia.

Diggle, Ribeiro e Christensen (2003) e Diggle e Ribeiro (2006) discutem em detalhes mtodos baseados na verossimilhana para estimao de parmetros do processos geoestatsticos espaciais, incluindo mtodos bayesianos.

2.2 Campos aleatrios Espao-temporais

Processos ambientais e geofsicos como concentrao de poluentes, precipitao e superfcie do vento so exemplos de fenmenos que podem ser modelados por campos aleatrios espao-temporais ou simplesmente processos espao-temporais. Este tipo de processo caracterizado pela variabilidade na dimenso do espao e do tempo. De acordo com Schabenberger e Gotway (2005), para o estudo desta variabilidade existem trs possibilidades:

- anlise espacial para cada processo temporal;
- anlise temporal para cada processo espacial;
- anlise espacial e temporal conjunta.

As duas primeiras possibilidades isolam a parte temporal ou a espacial e aplica tcnicas padres para o tipo de processo resultante. A terceira possibilidade considera o processo espacial e temporal conjuntamente e a considerada neste trabalho. No processo de construo desta anlise conjunta, a aplicao das tcnicas padres no estudo da variabilidade espacial e temporal separadamente pode ser utilizada como uma ferramenta exploratria e auxiliar na construo de modelos adequados.

Um campo aleatrio espao-temporal definido como

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

em que $Z(\mathbf{s}, t)$ valor de atributo Z no espao $\mathbf{s} \in R^d$ e no tempo $t \in R$.

Atravs desta definio, percebe-se intuitivamente que o domnio natural do processo $R^d \times R$. Neste trabalho o componente espacial ser considerado bidimensional, ou seja, $d = 2$, por ser a situao comum na prtica, entretanto ressalta-se que a dimenso do processo pode ser qualquer nmero finito positivo.

Tempo e espao podem no ser diretamente comparveis, j que as unidades das coordenadas dos dois processos apresentam grandezas diferentes. Fisicamente existe uma clara diferena entre as dimenses espaciais e temporais e os modelos precisam consider-las. Segundo Gneiting et. al (2006) funes de covarincia ou mesmo a krigagem se utilizam de espaos Euclidianos e so diretamente aplicados problemas espao-temporais. Sendo assim a simples separao vetorial do domnio espacial e temporal tem importantes implicaes.

De acordo com Gneiting e Schlather (2002), na modelagem de processos espao-temporais duas formas para especificao do modelo podem ser consideradas:

• *Especificao geoestatstica.* Nesta metodologia necessrio o conhecimento de uma distribuio, geralmente gaussiana, para funo aleatria $Z(\mathbf{s}, t)$ que definida em toda coordenada espao-temporal. A funo de covarincia definida na dimenso espao-temporal contnua, caracterizando o modelo da funo aleatria, que pode ser utilizada na predio de qualquer localizao espacial e qualquer instante de tempo. O ajuste da funo de covarincia de importncia central nesta metodologia, e portanto expresses com forma fechada para funo de covarincia so essenciais.

• *Especificao baseada em modelos.* Esta tcnica enfatiza a adoo de modelos estocsticos para soluo de problemas prticos nos quais a funo de covarincia no especificada explicitamente, mas sim induzida pelo modelo. Esta forma de especificao espao-temporal pode entretanto ir de uma funo com forma analtica simples at uma intratvel e somente implicitamente definida. Esta metodologia engloba desde modelos hierrquicos bayesianos espao-temporais at a tcnica do filtro de Kalman, com a caracterstica de apresentar predies computacionalmente eficientes.

A especificao geoestatstica conceitualmente simples, particularmente quando combinado com a adoo do modelo gaussiano, que completamente especificado por sua mdia e estrutura de covarincia. Alm disto, a krigagem, que geralmente o objetivo da anlise de campos aleatrios, requer a especificao apropriada da funo de covarincia. Entretanto esta forma de especificao assume estacionariedade espao-temporal da estrutura de covarincia, que portanto no pode assumir formas flexveis como as induzidas pela especificao baseada em modelos. Este trabalho ser direcionado especificao geoestatstica aplicada a processos espao-temporais gaussianos.

Assumindo as condies de regularidade do campo aleatrio espao-temporal, $Var(Z(\mathbf{s}, t)) < \infty$ para todo $\mathbf{s} \in \mathbf{R}^2, t \geq 0$, (CRESSIE; HUANG, 1999), define-se a mdia e a funo de covarincia por:

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

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

De acordo com Gneiting et.al (2006), na prtica o interesse inicial das anlises verificar a presena ou ausncia de trs caractersticas: estacionariedade, simetria completa e separabilidade. Aps este estudo ser possvel decidir sobre a complexidade da modelagem.

Um campo aleatrio espao-temporal estacionrio se sua esperana no depende das coordenadas espao-temporais e sua funo de covarincia s dependa de um vetor de separao dos pontos em $R^d \times R$. Portanto um campo aleatrio espao-temporal apresenta covarincia estacionria no espao se $Cov(Z(\mathbf{s}_1, t_1, \mathbf{s}_2, t_2))$ depende somente do vetor de separao $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2$, e apresenta estacionariedade temporal se $Cov(Z(\mathbf{s}_1, t_1, \mathbf{s}_2, t_2))$ depende somente do vetor de separao $u = t_1 - t_2$. Desta forma se o campo aleatrio espao-temporal possui covarincia estacionria espacial e temporalmente, este considerada uma covarincia espao-temporal estacionria dada por:

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

O campo aleatrio espao-temporal apresenta simetria completa se

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

para todas as coordenadas espao-temporais $\mathbf{s}_1, t_1$ e $\mathbf{s}_2, t_2$ em $R^2 \times R$. No caso da covarincia ser estacionria e completamente simtrica tem-se:

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

para todo $(\mathbf{h}, u) \in R^2 \times R$.

A classe mais simples de funes de covarincia espao-temporal dada pelos modelos separveis. Funes de covarincia espao-temporais separveis podem ser definidas com base nas propriedades de aditividade e multiplicabilidade (2.1.1, vi e vii). Desta forma podero ser decompostas entre uma funo de covarincia puramente espacial e outra puramente temporal, que no caso aditivo pode ser escrita como:

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

e no caso multiplicativo por:

$$Cov(Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)) = Cov(Z(\mathbf{s}_1, \mathbf{s}_2))Cov(Z(t_1, t_2)), \tag{2}$$

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

Esta classe de modelos de covarincia espao-temporal no considera a interao entre o espao e o tempo, mas muito utilizada na prtica por ser computacionalmente tratvel e uma forma intuitiva e simples de obteno de funes de covarincia vlidas.

Todos os modelos separveis so tambm completamente simtricos. Para a demonstrao desta relao considere a eq.(2), considere tambm as variveis aleatria espao-temporais $Z(\mathbf{s}_1, t_2)$ e $Z(\mathbf{s}_2, t_1)$, a funo de covarincia separvel ser dada por:

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

desta forma conclu-se que:

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

que a definio da propriedade de campos aleatrios espao-temporais completamente simtricos.

Um exemplo simples de funo de covarincia espao-temporal separvel dado pela combinao de modelos de covarincia exponenciais :

$$Cov(Z(\mathbf{s}_1,t_1),Z(\mathbf{s}_2,t_2)) = \sigma_1^2 exp(-A_1||(\mathbf{s}_1-\mathbf{s}_2)||) + \sigma_2^2 exp(-A_2||(t_1-t_2)||),$$

em que A_1 e A_2 so as denominadas *matrizes de anisotropia*.

As matrizes de anisotropia so utilizadas para permitir a representao combinada das dimenses espaciais e temporais. Neste exemplo, suponha um processo isotrpico no espao tem-se:

$$A_1 = \begin{pmatrix} \phi_1 & 0 & 0 \\ 0 & \phi_1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \phi_2 \end{pmatrix}$$

Qualquer funo de covarincia espao-temporal que no for escrita nas duas formas separveis anteriormente mencionadas e nem combinaes destas, sero ditas funes no separveis. As funes de covarincia separveis, embora facilmente obtidas, so em geral insatisfatrias para descrever processos naturais. O que gera necessidade de se especificar funes no separveis. Kryriakidis e Journel (1999), revisam e discutem abordagens para especificao de processos geoestatsticos espao-temporais. Cressie e Huang (1999), propem classes de funes de covarincia no separveis vlidas cuja obteno no se baseiam no domnio das observaes, mas sim no domnio da freqncia, uma vez que a representao espectral e a densidade espectral podem ser utilizadas para descrever as propriedades do processo. Gneiting (2002), em uma abordagem alternativa, obtm modelos vlidos de funo de covarincia atravs da definio de funes montonas e positivas.

2.2.1 Representao de Gneiting

A representao de Cressie e Huang (1999) constri uma famlia de funes de covarincia espao-temporais estacionria e no separveis atravs da inverso de Fourier na forma fechada, o que restringe esta famlia a um pequeno grupo de funes com solues de forma fechada da inverso de Fourier. Gneiting (2002) prope uma flexvel e elegante famlia de funes de covarincia espao-temporais, cuja obteno no requer operaes no domnio espectral. A construo de funes de covarincia vlidas se d atravs de componentes elementares em que sua validade facilmente verificada.

A representao de Gneiting (2002) considerara qualquer funo montona $\phi(x)$, definida em $x \geq 0$ e qualquer funo positiva $\psi(x)$, definida em $x \geq 0$ com derivadas completamente montonas , e desta forma,

$$C(\mathbf{h},u) = \frac{\sigma^2}{(\psi(u)^2)^{d/2}} \phi\Big(\frac{||\mathbf{h}||^2}{\psi(|u|^2)}\Big) \tag{3}$$

uma funo de covarincia espao-temporal vlida em $R^d \times R$, em que $\mathbf{h} = ||h||$, u representa o vetor de distncias no tempo e d deve ser maior ou igual que a genuna dimenso espacial.

De acordo com Gneiting (2002) uma funo contnua $\phi(x)$ completamente montona se possui derivadas $\phi^{(n)}$ de todos as ordens e $(-1)^n \phi^{(n)}(x) \geq 0$, em que $x > 0, n = 0, 1, 2, \ldots$.

Assim, Gneiting prope uma famlia de funes de covarincia espao-temporal muito geral que no dependem da inverso de Fourier na forma fechada e no necessita de integrabilidade. De acordo com Elmatzoglou (2006) nesta famlia os componentes $\phi(x)$ e $\psi(x)$ podem ser associados com a estrutura espacial e temporal respectivamente. Gneiting (2002), apresenta possveis escolhas para cada funo, que so reproduzidas nas Tabelas 1 e 2.

Tabela - 1: Funes completamente montonas $\phi(x)$,$x \geq 0$

$\phi(x) = exp(-cx^{\gamma}), c > 0, 0 < \gamma \leq 1$
$\phi(x) = (1 + cx^{\gamma})^{\nu}, c > 0, 0 < \gamma \leq 1, \nu > 0$
$\phi(x) = (2^{\nu-1}\Gamma(\nu))^{-1}(cx^{1/2})^{\nu}\mathbf{K}_{\nu}(cx^{1/2}), c > 0, \nu > 0$
$\phi(x) = 2^{\nu}(exp(cx^{1/2}) + exp(-cx^{1/2}))^{\nu}, c > 0, \nu > 0$

Tabela - 2: Funes positivas $\psi(x)$, $x \geq 0$

$\psi(x) = (ax^{\alpha} + 1)^{\beta}, a > 0, 0 < \alpha \leq 1, 0 \leq \beta \leq 1$
$\psi(x) = ln(ax^{\alpha} + b)/ln(b), a > 0, b > 1, 0 < \alpha \leq 1$
$\psi(x) = (ax^{\alpha} + b)/(b(ax^{\alpha} + 1)), a > 0, 0 < b \leq 1$

Um exemplo simples da famlia de covarincia espao-temporal no separvel proposta por Gneiting dada como segue:

Considere a primeira linha das Tabelas 1 e 2, em que $\phi(x) = exp(-cx^{\gamma}), c > 0, 0 < \gamma \leq 1$ e $\psi(x) = (ax^{\alpha} + 1)^{\beta}, a > 0, 0 < \alpha \leq 1, 0 \leq \beta \leq 1$ so funes completamente montona e positiva respectivamente. Substituindo na eq.(3) tem-se:

$$C(\mathbf{h}, u) = \frac{\sigma^2}{(a|t|^{2\alpha} + 1)^{\frac{\beta d}{2}}} exp\Big\{ \frac{c||\mathbf{h}||^{2\gamma}}{(a|t|^{2\alpha} + 1)^{\beta\gamma}} \Big\}. \tag{4}$$

Nesta expresso, nota-se que para $\beta = 0$ a covarincia no depende do tempo. Multiplicando-se a eq.(4) por uma funo de covarincia puramente temporal, $C(|t|) = (a|t|^{\alpha} + 1)^{-\delta}$, tem-se como resultado:

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

no caso em que $\beta = 0$, a expresso da funo de covarincia se reduz a

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

que funo de covarincia separvel na forma,

$$C(\mathbf{h},u) = \sigma^2 C(u)C(\mathbf{h})$$

em que $C(u) = (a|t|^{2\alpha}+1)^{-\delta}$ a funo puramente temporal, enquanto $C(\mathbf{h}) = exp\{c||\mathbf{h}||^{2\gamma}\}$ a funo puramente espacial. Esta representao sugere que o parmetro β pode ser utilizado para testar a hiptese de separabilidade.

3 Resultados e discusso

Muitos modelos tem sido sugeridos na literatura e o fato de terem sido adotados no significa necessariamente que estes representem melhor a realidade. A escolha de uma determinada classe de modelos uma tarefa difcil e que requer um profundo conhecimento da realidade, por parte dos pesquisadores.

Neste trabalho, o interesse no ajustar o melhor modelo e sim verificar o comportamento de alguns dos modelos discutidos no captulo anterior. Para tanto, ser considerado um conjunto referente a capacidade de armazenagem de gua em um solo com citros.

A gua um dos mais importantes fatores para o adequado desenvolvimento de qualquer cultura agrcola e o conhecimento da sua dinmica no solo indispensvel para o aprimoramento de prticas de manejo que visam otimizao da produtividade da cultura (MORETI, 2006).

Moreti (2006) fornece um estudo da capacidade de armazenagem de gua em um Latossolo Vermelho Amarelo Argisslico cultivado com citros, em uma parcela experimental na qual foram coletados dados em duas transees com 20 pontos de observao cada, com espaamento de 4,0 m entre os pontos amostrados e 7,0 m entre as transees. Para quantificar a armazenagem de gua no solo foi instalado, em cada ponto amostral, um tubo de acesso sonda de nutros at a profundidade de 1,20 m, localizado no centro da distncia entre duas plantas ao longo da linha.

A cultura de citros foi implantada em maro de 1991, e as amostras foram coletadas de 2001 a 2004, totalizando 98 coletas de dados no eqidistantes no tempo. Neste trabalho utilizou-se dados de 25 momentos de tempo tomados um em cada quatro, aos quais ajustou-se modelos espao-temporais. A restrio de tomar apenas 25 dentre os 98 perodos de tempo visa contornar problemas computacioanais. Uma vez que o objetivo aqui investigar a metodologia de modelagem e no fornecer resultados definitivos de anlises destes dados.

3.0.2 Construindo modelos de covarincia espao-temporais

Tem-se a seguinte funo de covarincia espao-temporal separvel ajustada, com os parmetros estimados pelo mtodo da mxima verossimilhana:

$$C(\mathbf{h},u) = \frac{44.12}{(0.76|u|^{1.65}+1)^{0.08}} exp\{-0.001||\mathbf{h}||^{0.85}\} + \frac{0.25}{(0.76|u|^{1.65}+1)^{0.08}}. \quad (5)$$

J para a funo de covarincia espao-temporal no separvel, com os parmetros tambm estimados pelo mtodo de mxima verossimilhana, tem-se que o modelo ajustado dado por:

$$C(\mathbf{h},u) = \frac{1}{(0.76|u|^{1.65}+1)^{0.12}}\Big(\frac{44.12}{(0.76|u|^{3.3}+1)^{0.04}}$$
$$exp\Big\{-0.001\Big[\frac{||\mathbf{h}||}{(0.76|u|^{3.3}+1)^{0.04}}\Big]^{1.65}\Big\}\Big) + \frac{0.25}{(0.76|u|^{1.65}+1)^{0.08}}. \qquad (6)$$

A distino entre os dois modelos ajustados anteriormente dada pelo parmetro β que determina se existe ou no interao entre os componentes puramente espacial e puramente temporal, ou seja, se a correlao espacial independe do tempo e vice-versa. Ento o valor do parmetro β pode ser visto como o grau de interao espao-temporal, em que valores grandes, prximos de 1, induz em um forte efeito de interao espao-temporal, enquanto que valores pequenos, prximo de 0, induz em um efeito fraco de interao espao-temporal. Isto sugere uma forma de testar a separabilidade da funo de covarincia, alm de quantificar a intensidade da interao espao-temporal.

O valor estimado para o parmetro $\beta = 0.04$ muito prximo de zero, o que indica que o modelos separvel se ajustaria bem aos dados, em outras palavras, no h evidncia clara de interao entre o espao e o tempo. Ressaltando aqui novamente que a estimao deste parmetro pode ser numericamente instveis e em uma anlise detalhada destes dados, mtodos numricos devem ser cuidadosamente calibrados.

Para verificar o efeito da suposio de separabilidade espao-temporal foi considerado tambm o erro mdio de predio, para tanto foi realizada a krigagem simples para diferentes valores de β. Os valores preditos para o tempo 26 foram comparados com os verdadeiros valores observados para este tempo. A Figura 1, ilustra estas predies e por este grfico percebe que o modelo com $\beta = 0$ o que mais se aproxima do valor real e ainda percebe-se que o aumento do valor de β, causa um diminuio na mdia e na varincia. A Tabela 3 apresenta os erros quadrticos mdios, a mdia e varincia da predio, e atravs dela possvel verificar que o menor valor do erro quadrtico mdio ocorre quando $\beta = 0$, ou seja, o modelo separvel o mais indicado para o conjunto de dados considerado.

Na Figura 2, so ilustrados os mapas de krigagem simples com parmetros $\beta = 0, 0.25, 0.95$. As predies foram realizados para o tempo 26, que no foi considerado na estimao. Estas imagens ilustram uma pequena diferena no comportamento das "manchas", mas uma diferena um pouco maior nas escalas. Atravs das legendas percebe a clara diminuio do valor mdio considerado.

Consideraes Finais

Processos ambientais e geofsicos como concentrao de poluentes, campos de precipitao e superfcie do vento so caracterizados pela variabilidade contnua no espao e no tempo. Tais campos espao-temporais so geralmente tratados como aleatrios, uma vez que as leis que governam seu comportamento no so completamente compreendidas e sua complexidade no segue uma descrio determinstica precisa.

Tabela - 3: Erro quadrtico mdio de predio para funes de covarincia espao-temporais com diferentes valores do parmetro β

	EQM	Mdia	Varincia
Modelo separvel $\beta = 0$	5.561	17.114	0.753
Modelo no separvel $\beta = 0.25$	12.674	15.913	0.494
Modelo no separvel $\beta = 0.5$	26.060	14.371	0.309
Modelo no separvel $\beta = 0.75$	45.661	12.720	0.191
Modelo no separvel $\beta = 0.95$	64.436	11.451	0.132

Modelos estocsticos, por sua vez, so baseados tipicamente em um pequeno nmero de parmetros que podem ser modelados e inferidos a partir de observaes parciais dos processos.

O objetivo da modelagem de tais processos informar sobre padres do comportamento espao-temporal. Parte-se do princpio de que no existe um modelo que se ajuste perfeitamente aos dados e o que ocorre, que alguns modelos so mais capazes, que outros, de descrever certas situaes. Atravs da modelagem possvel resumir de forma simples um conjunto de dados complexo, testar hipteses cientificamente relevantes e fazer predies dos processos em posies no observadas do espao e/ou tempo.

Fenmenos espaciais e temporais so caracterizados por suas unicidade e no reproducibilidade. Assumindo-se ento que uma amostra parcial simples da realizao de um particular mecanismo estocstico. Conseqentemente, infercias s podem ser feitas a partir da adoo de simplificaes e suposies. A princpio a estratgia de anlise espao-temporal apresentada aqui poderiam ser utilizadas simplesmente como uma extenso da anlise puramente espacial, pois simplesmente acrescenta uma dimenso referente ao tempo. Entretanto, a diferena fsica entre espao e tempo e a ordenao natural dos tempos induzem caractersticas particulares e dificuldades adicionais.

O tempo e o espao so duas noes completamente diferentes, o que dificulta a especificao de funes de covarincia espao-temporais vlidas. Alm disso estas precisam satisfazer certas condies para serem consideradas vlidas. Uma especificao muito comum para a funo de covarincia espao-temporal assegurada por propriedades que do suporte famlia de funes de covarincia separveis. Estas podem ser separadas atravs da soma e/ou produto de funes de covarincia vlidas. A grande desvantagem deste mtodo que ele no considera na modelagem a interao entre o espao e o tempo.

Uma alternativa s funes de covarincia separveis so as funes de covarincia espao-temporais no separveis e neste caso a interao entre o espao e o tempo considerada na modelagem. Na literatura existem algumas possibilidades de funes no separvel, no presente trabalho foram consideradas a especificao de Cressie e Huang e a famlia de Gneting.

A especificao de Cressie e Huang se baseia na anlise da integral de Fourier o que requer operaes utilizando representao Espectral da funo de covarincia e a inverso de

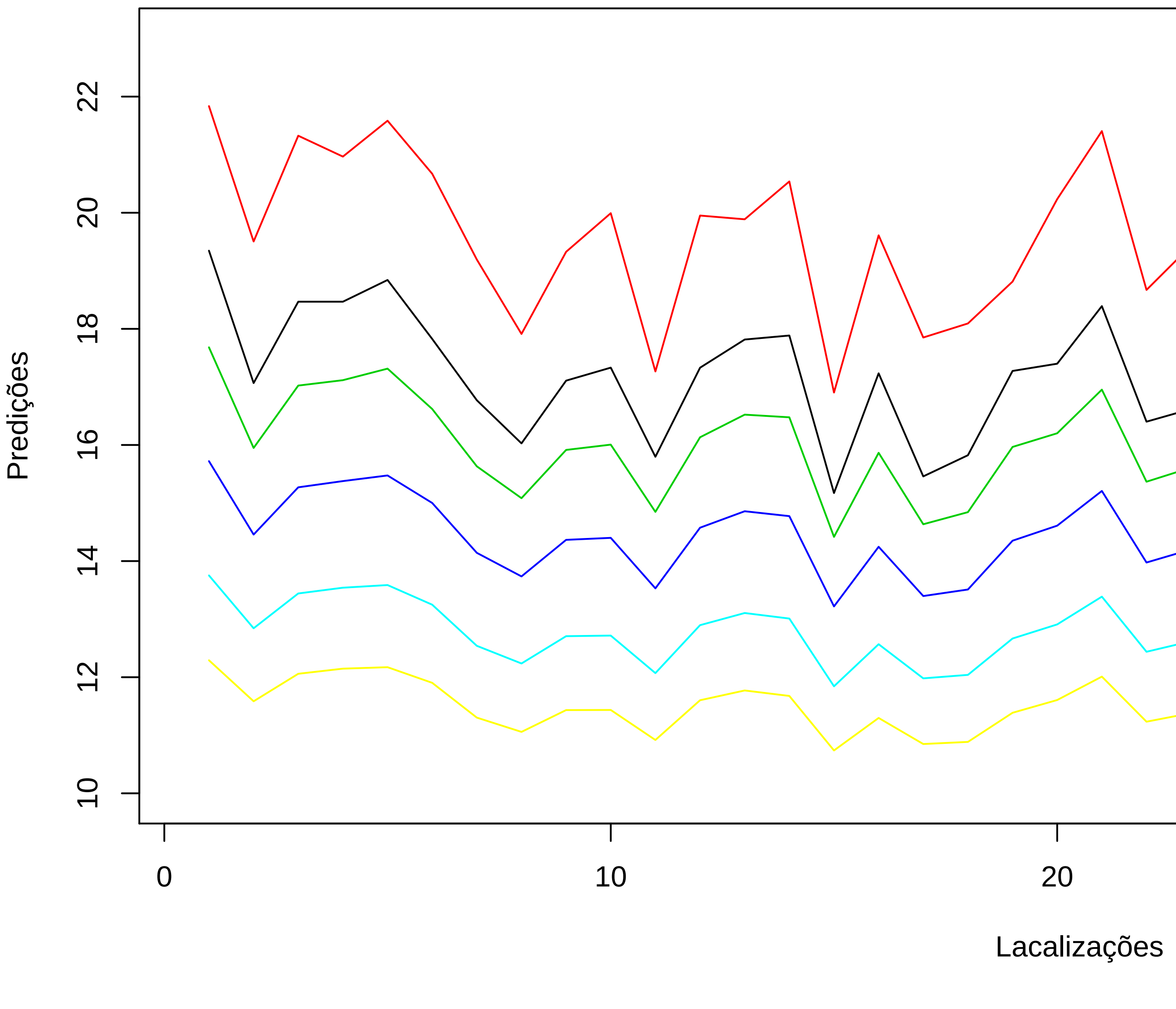

Figure - 1: Distribuio das predies para o tempo 26 e do verdadeiro valor amostrado

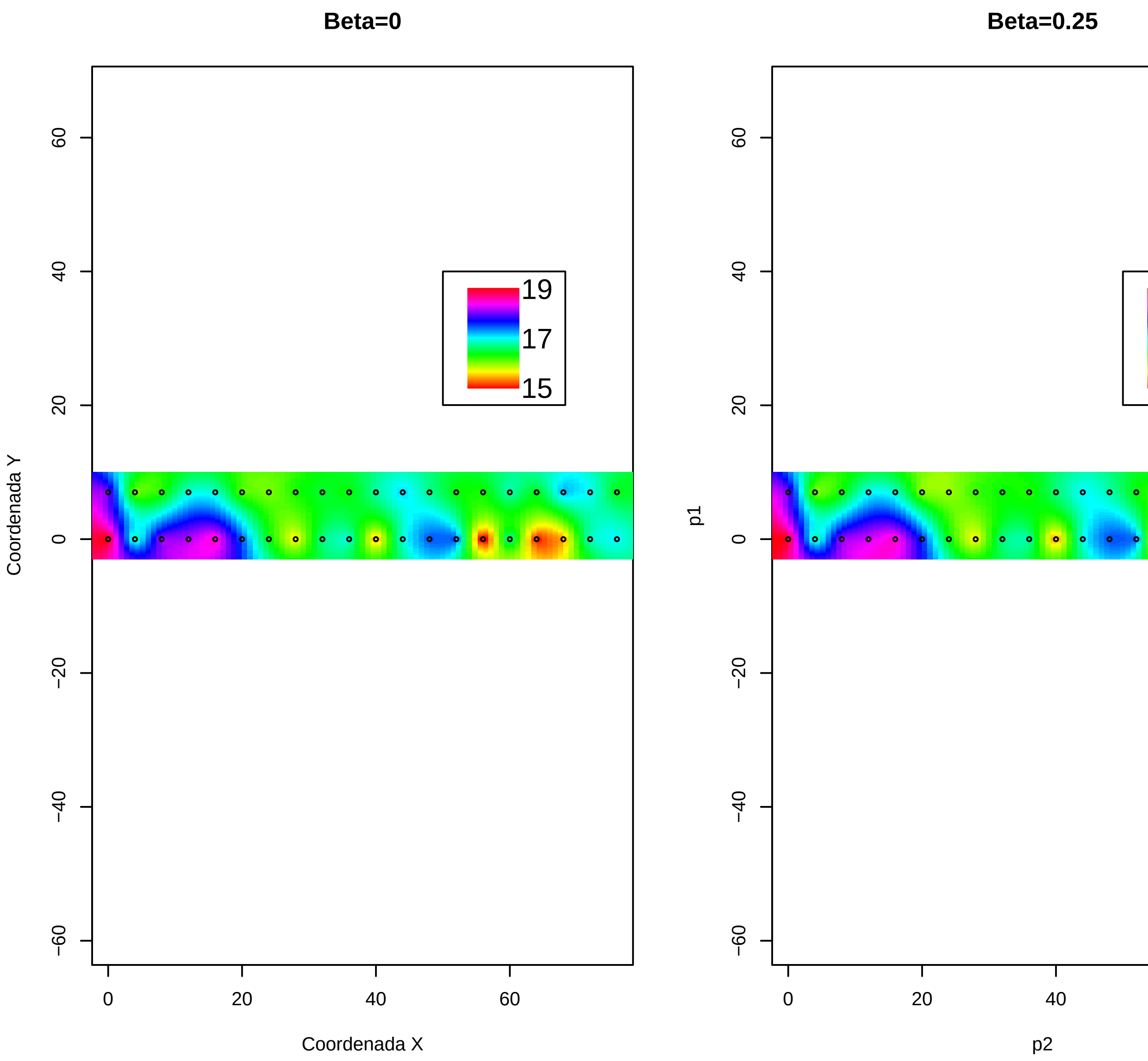

Figure - 2: Krigagem para modelo com $\beta = 0$, $\beta = 0.25$, $\beta = 0.95$,respectivamente volta do domnio da frequncia com o das observaes s obtida em casos particulares. As funes de covarincia espao-temporais propostas por Gneting consideram a combinao de funes completamente montonas e estritamente crescentes que asseguram funes de covarincia vlidas.

Nas aplicaes deste trabalho foram consideradas funes de covarincia espao-temporais pertencente famlia de Gneiting, estacionrias e completamente simtricas, que so uma escolha conveniente, visto que esta especificao no requer qualquer tipo de integrabilidade e possui implementao disponvel no pacote RandomFields.

Para os dois conjuntos de dados foi ajustado inicialmente um modelo com estrutura de covarincia separvel, composta pelo produto entre uma funo *cauchy generalizada*, para o componente puramente temporal, e uma funo *stable*, para o componente espacial. Ainda para a covarincia separvel foi adicionado um efeito pepita composto pelo produto entre uma constante e uma funo *cauchy generalizada*. Depois de estimados todos os valores para os parmetros do modelo separvel, foi ajustado um modelo de covarincia no separvel mas com a mesma estrutura de anisotropia estimada para o modelo separvel. O modelo no separvel considerado composto pelo produto entre uma funo de covarincia *cauchy generalizada*, que depende apenas do tempo, e uma funo no separvel proposta por Gneiting, que depende do tempo e do espao conjuntamente. Para o efeito pepita foi considerada a mesma estrutura estimada para o modelo separvel. importante ressaltar que o modelo separvel proposto uma caso particular do modelo no separvel, quando o parmetro $\beta = 0$.

A utilizao de conjunto de dados reais levantou dificuldades prticas na conduo das anlises. A primeira dificuldade foi a depurao das linhas de comandos do pacote "*RandomFields*", que ainda se encontra em desenvolvimento e poucas so as publicaes que se utilizao deste software, um outro grande problema relacionado parte computacional a demanda por computadores com grande capacidade de memria e processamento, j que o nmero de observaes para anlise espao-temporal cresce muito rapidamente com o aumento de obsevaes na coordenada do tempo ou no espao.

Para a anlise do conjunto de dados foram considerados os mesmos pontos em todos os tempos, sendo isto uma obrigatoriedade da implementao disponvel no pacote, j que no existe a possibilidade de se especificar coordenadas diferentes para cada tempo. Por conta destas limitaes, algumas observaes tiveram de ser descartadas da estimao e predio.

Foram encontradas dificuldades na estimao do parmetro β da funo de covarincia proposta por Gneiting, e portanto neste trabalho foram considerados outras formas de assegurar a validade dos valores estimados para este parmetro. Na anlise foi encontradas estimativa do parmetro β prximas de zero, apontando para uma fraca interao entre o espao e o tempo.

Foi considerada na anlise a predio de um tempo futuro e comparado com o verdadeiro valor observado para este tempo. O interesse neste caso foi verificar atravs do erro quadrtico mdio de predio a validade do ajuste da funo de covarincia. Na Tabela 3 existe uma grande diferena no erro quadrtico mdio, na mdia e na varincia de predio. Atravs desta tabela pode-se concluir que a estrutura de covarincia separvel a que mais se aproxima dos verdadeiros valores observados, o que pode-se perceber tambm a diminuio da mdia e da varincia de predio medida que o valor do parmetro β se aproxima de 1.

No desenvolvimento deste trabalho foram feitas ainda vrias simulaes e anlises que serviram para a melhor compreenso do pacote. Em uma destas simulaes foi estimado um campo espao-temporal com um determinado valor do parmetro β, em seguida foi realizada a estimao dos parmetros, e verificou-se que a estimativa de verossimilhana no se aproximou de forma satisfatria do valor estimado. Estes resultados so preliminares, e considera-se que estudos desta natureza devem ser efetuados de forma mais detalhada e portanto optou-se por no reportar aqui os resultados deste estudo de simulao. Uma condio que poderia ser considerada a realizao de vrias estimaes e ento seria calculada a mdia das estimativas. Espera-se que a mdia dos parmetros estimados se aproxime do valor dos parmetros simulados. Com um nmero suficientemente grande de valores estimados pode-se ter uma melhor idia da variabilidade de cada parmetro e com isto a determinao de intervalos de confiana e cobertura poderiam ser obtidos, assim como poderia se avaliar o impacto em predies.

Estas so apenas algumas entre as muitas possibilidades em se aprofundar os estudos nesta rea. Neste trabalho foi considerada apenas um resumo das novas metodologias de anlise espao-temporal e com o avano computacional e o apelo prtico desta metodologia fica cada vez mais fcil de se descrever melhor as leis da natureza. Modelos que utilizam a especificao da funo de covarincia, embora promissores e com fcil interpretao intuitiva, apresentam dificuldades e restries para uso em dados reais. Estudos complementares so ainda necessrio para auxiliar na compreenso das caractersticas de identificabilidade e aplicabilidade em situaes reais, bem como as exigncias e restries de implementao.

Agradecimentos

Este trabalho parte da dissertao de mestrado do primeiro autor do Departmento de Cincias Exatas, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de So Paulo, Piracicaba.

Agradeo a Dolorice Moreti e o Prof Dr. Paulo Leonel Libardi do Departamento de Fsica de Solos, por ter gentilmente cedido os dados da aplicao.

SILVA, A. S.; RIBEIRO JUNIOR, P. J. Space-time geostatisticals gaussian models and aplications. *Rev. Mat. Estat.,* So Paulo, v.xx, n.x, p.xx-xx, 2000. *Rev. Mat. Estat.* (São Paulo), v. 20, n.1, p. 1-10, 2000.

▪ *ABSTRACT: The specification of space-time covariance functions is one of the possible strategies to model processes observed at different locations and time points. Such functions can define separable and non-separable processes and must attend the condition of positive-definiteness. Among the strategies to obtain such valid functions are the ones suggested by Cressie and Huang (1999) and by Gneiting (2002). The former is based on the idea of obtaining valid functions in a space of increased dimension from valid functions on the primary dimension and requires operations in the frequency domain. Alternatively, the latter combines increasing monotone functions avoiding the inversion of spectral representations. There are still few reports of usage and comparisons of the strategies. This work follows Gneiting's proposals with different values for the space-time interaction parameter. Models were applied for the analysis of two real data sets, one about fish stocks in the Portuguese coast and a second on soil water storage. The implementation on the R package RandomFields was used, with methodology and computational implementation being reviewed. For both case the separable model provided a satisfactory fit, based on maximum likelihood estimation.*

▪ *KEYWORDS: Keywords: Random Fields; Geostatistic; Space-Time Models;* RandomFields *Package*

References

BOCHNER, S. *Lectures on Fourier Integrals*. Princeton, N.J.: Princeton University Press, 1959.

CRESSIE, N.; HUANG, H-C. *Classes of Non-Seperable, Spatio-temporal stationary covariance functions*. Journal of the American Statistical Association, v.94, p.1330-1340, 1999.

DIGGLE, P.J.; RIBEIRO Jr., P.J.; CHRISTENSEN, O.F. *Spatial Statistics and Computational Methods*, Springer, New York, 2003.

DIGGLE, P.J.; RIBEIRO Jr., P.J. *Model-Based Geoestatistics*, Springer, 2006.

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

GNEITING, T. Nonseperable, *Stationary Covariance Functions for Space-Time Data*. Journal of the American Statistical Association, v.97, No.458, 2002.

GNEITING, T.; SCHLATHER, M. Stochastic models wich separate fractral dimension and Hurst effect. Technical Report Series, NRCSE-TRS no. 069, 2001.

GNEITING, T.; SCHLATHER, M. *Space-time covariance models*. Encyclopedia of Environmetrics, Columbia, v.4, p.2041-2045, 2002.

GNEITING, T.; GENTON, M.G.; GUTTORP, P. Geostatistical Space-Time Models, Stationarity, Seperability and Full Symmetry. Technical Report no.475 Department of Statistics University of Washington, 2006.

KYRIAKIDIS, P.C.; JOURNEL, A.G. *Geostatistical Space-Time Models: a review.* Mathematical Geology, v.31, p.651-684, 1999.

LE, D.N.; ZIDEK, J.V. *Statistical Analysis of Environmental Space-time processes*, Springer, New York, 2006.

MORETI, D. *Avaliao espao-temporal de processos do balano de gua em solo com citros*, 2006. 138p. Tese Doutorado - Universidade de So Paulo, Piracicaba, 2006.

SCHABENBERGER, O.; GOTWAY, C.A. *Statistical Methods for Spatial Data Analysis*, Chapman & Hall / CRC, 2005.

SCHMIDT, A.M.; SANS, B. Modelagem Bayesiana da Estrutura de Covarincia de Processos Espaciais e Espao-Temporais. In: SINAPE, 17., 2006.

STEIN, M.L. Space-time covariance functions. Technical Report No. 4, University of Chicago, 2004.

STEIN, M.L. *Nonstationary spatial covariance functions.* Chicago: University of Chicago, 2005. 57 p. (Technical Report, 21)

Recebido em 01.01.2005.

Aprovado após revisão em 01.01.2005.