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

Modelos gaussianos geoestatísticos espaço-temporais e aplicações

Alexandre Sousa da Silva

Dissertação apresentada para obtenção do título de Mestre em Agronomia. Área de concentração: Estatística e Experimentação Agronômica

Piracicaba
2006

Alexandre Sousa da Silva
Estatístico

Modelos gaussianos geoestatísticos espaço-temporais e aplicações

Orientador:
Profº Dr. **PAULO JUSTINIANO RIBEIRO JUNIOR**

Dissertação apresentada para obtenção do título de Mestre em Agronomia. Área de concentração: Estatística e Experimentação Agronômica

Piracicaba
2006

Dedicatória

Aos meus pais,
Maria Teresa Sousa da Silva e
Cícero Lourenço da Silva Neto,

Aos meus irmãos: Fernando, Eduardo e Patricia;
sobrinhos: Vitória e Fernando;
aos amigos, em especial ao Gilmar,
o carinho e incentivo.

AGRADECIMENTOS

Ao Professor Dr. Paulo Justiniano Ribeiro Junior do Departamento de estatística da Universidade Federal do Paraná pela orientação, a paciência, os ensinamentos, a amizade, confiança e apoio na realização deste trabalho.

Aos professores e funcionários do Departamento de Ciências Exatas da ESALQ/USP, em especial a Clarice Demétrio, Décio Barbin, Roseli Leandro, Edwin Ortega, Cesar de Lima, Solange Sabadin e Eduardo Bonilho que me propiciaram condições para a realização deste trabalho.

Aos amigos: Ana Carolina Alexandrino, Ana Paula Soares, Cristiane Afonso, Fernanda Rizzato, Joseane Padilha, Jualiana Fachini, Pâmela Piovesan, Afrânio Vieira, Ana Alice Pilon, Ana Maria Araújo, Andréia Meyer, Angela Coelho, Cesar Taconelli, Denise Viola, Genevile Bergano, Geovana Silva, Júlio Pereira, Luciana Carvalho, Lúcio Araújo, Osmar Macedo, Sandra Cardozo, Vandely Janeiro pela ajuda, apoio, compreensão e laços de amizade.

À CNPQ pelo fundamental suporte financeiro concedido.

Ao Instituto de Investigação das Pescas e do Mar (IPIMAR) por ter gentimente cedido os dados do monitoramento de estoques de pescada da costa de Portugual.

À Dolorice Moreti e ao Profº Paulo Leonel Libardi por ter gentilmente cedido os dados de armazenagem de água no solo.

E a todos que de forma direta ou indireta contribuíram para a realização deste trabalho.

SUMÁRIO

RESUMO 6
ABSTRACT 7
1 INTRODUÇÃO 8
2 REVISÃO BIBLIOGRÁFICA 10
2.1 Campos aleatórios 10
2.1.1 Propriedades da função de covariância 14
2.1.2 Modelos de covariância 15
2.1.3 Estimação dos parâmetros 17
2.1.4 Krigagem 18
2.2 Campos aleatórios espaço-temporais 19
2.2.1 Representação de Cressie-Huang 24
2.2.2 Representação de Gneiting 25
2.2.3 Estimação 27
3 MATERIAL E MÉTODOS 29
3.1 Recursos Computacionais 29
3.2 Simulação de Campos Aleatórios Espaço-Temporais 29
4 APLICAÇÕES 36
4.1 Monitoramento de estoques de pescada na costa portuguesa 36
4.1.1 Análise descritiva 38
4.1.2 Modelagem da estrutura de covariância espaço-temporal 41
4.2 Armazenagem de água no solo 45
4.2.1 Construindo modelos de covariância espaço-temporais 49
5 CONCLUSÕES 53
REFERÊNCIAS 58
APÊNDICE 63

RESUMO

Modelos gaussianos geoestatísticos espaço-temporais e aplicações

A especificação de funções de covariância espaço-temporais é uma das possíveis estratégias para modelagem de processos dos quais observações são tomadas em diferentes posições do espaço e do tempo. Tais funções podem definir processos separáveis ou não separáveis e na sua especificação deve-se garantir que são funções de covariância válidas atendendo a condição de serem positiva definidas. Entre estratégias para obtenção de tais funções estão as de Cressie e Huang (1999) e Gneiting (2002). A primeira se baseia na idéia de obter funções em um espaço de dimensão aumentada a partir de funções válidas no espaço original e necessita de operações no domínio da freqüência. Alternativamente a segunda proposta utiliza combinação de funções completamente monótonas e estritamente crescentes, evitando inversão de representações espectrais. Há ainda poucos relatos de uso e avaliações comparativas das diferentes propostas. Neste trabalho considerou-se a metodologia proposta por Gneiting, com diferentes valores do parâmetro que indica a força da interação entre o espaço e o tempo. Diferentes modelos foram aplicados à dois conjuntos de dados, um referente a estoques de peixe na costa de Portugal, e outro referente à armazenagem de água em um solo com citros. Utilizou-se a implementação no pacote *RandomFields* do programa R, revisando-se a metodologia e investigando-se a implementação computacional. Para os dois conjuntos de dados o modelo de covariância separável se mostrou adequado para descrever o comportamento das observações disponíveis sendo a escolha do modelo determinada por ajustes de máxima verossimilhança.

Palavras-chave: Campos Aleatórios; Geoestatística; Modelos Espaço-Temporias; Pacote *RandomFields*

ABSTRACT

Space-time geostatisticals guassian models and aplications

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

1 INTRODUÇÃO

Dados espaço-temporais são caracterizados pela variabilidade no tempo e no espaço, e o objetivo da análise estatística deste tipo de dado está em descrever a incerteza não somente sobre as estimativas das quantidades de interesse mas, também, estimar valores em locais e/ou tempos não amostrados. Na literatura estatística existe um grande número de publicações referentes a processos puramente espaciais ou puramente temporais, e mais recentemente começam a ser apresentadas propostas para modelagem conjunta no espaço e tempo. Esta análise conjunta é o interesse dos modelos espaço-temporais.

Modelos espaço-temporais têm ganho uma crescente popularidade na última década, o que pode ser explicado por um lado pela numerosa aplicabilidade, principalmente em ciências ambientais e da saúde; e por outro lado pelo crescente desenvolvimento de recursos computacionais.

Gneiting e Schlather (2002), em uma revisão sobre estratégias de modelagem, consideram duas possíveis especificações para a modelagem espaço-temporal: a especificação geoestatística e a especificação baseada em modelos.

Na especificação geoestatística, geralmente considerada para campos aleatórios gaussianos, em que o processo é totalmente especificado pelo vetor de médias e a matriz de covariância. Em geral o vetor de médias é facilmente especificado a partir de informações contextuais, isto não acontece com a matriz de covariância, que é o componente chave desta metodologia. Entretanto, nesta abordagem, esta matriz possui elementos dados por uma função de covariância válida, que assegure a condição de matriz positiva definida.

Na especificação baseada em modelos é enfatizada a adoção de modelos estocásticos e nesta metodologia a função de covariância não é a única estrutura utilizada para especificar o modelo, sendo assim, este método pode ir de uma função analítica simples a uma função intratável e somente implicitamente induzida pelo modelo adotado. Neste contexto incluí-se os modelos Bayesianos flexíveis, dinâmicos e/ou homogêneos e métodos baseados na técnica do filtro de Kalman.

Neste trabalho será considerado a especificação geoestatística, para campos

aleatórios espaço-temporais gaussianos. Uma forma intuitiva de produzir modelos válidos para a função de covariância espaço-temporal é através da combinação de funções válidas puramente espacial e puramente temporal. Uma propriedade assegura que o produto ou a soma de funções de covariâncias válidas tem como resultado uma função de covariância válida. As funções de covariância espaço-temporais obtidas desta forma são ditas separáveis já que não contemplam a possibilidade de interação entre o componente espacial e temporal. Por isso, em geral não são realísticas já que assumem a independência dos processos espaciais e temporais.

Alternativamente as funções de covariância espaço-temporais não separáveis consideram a interação entre o componente espacial e temporal, mas apresentam como grande desvantagem a dificuldade em se construir funções de covariância que sejam válidas. Métodos matemáticos sofisticados devem ser aplicados para garantir a validade destas funções e em geral se valem do teorema de Bochener (BOCHNER, 1959), que utiliza a representação espectral para garantir a validade das funções de covariância espaço-temporais. Nesta linha, Cressie e Huang (1999), Gneiting (2002a) e Stein (2005), por exemplo, propõem famílias de modelos de covariância não separáveis válidas para processos espaço-temporais gaussianos.

Diante da importância de se estudar o comportamento no espaço e no tempo conjuntamente, o presente trabalho teve como objetivos gerais:

• fazer uma revisão de literatura sobre campos aleatórios espaço-temporais e metodologias para a sua caracterização;

• realizar um estudo de simulação, com o intuito de facilitar a compreensão dos conceitos definidos na Revisão de Literatura, além de possibilitar um melhor entendimento do funcionamento e dos modelos implementados no pacote "RandomFields".

• aplicar diferentes modelos à dois conjuntos de dados, um referente a estoques de peixe na costa de Portugal, e outro referente à armazenagem de água em um solo com citros.

2 REVISÃO BIBLIOGRÁFICA

2.1 Campos aleatórios

Um campo aleatório ou uma função aleatória é um processo estocástico definido no espaço $G \subset \mathbf{R}^d$, ou seja, uma função cujos valores são realizações de variáveis aleatórias em qualquer ponto do domínio (SCHIMIDT; SANSÓ, 2006), ou em outras palavras, uma família ou coleção de variáveis aleatórias, em que cada um dos seus membros podem ser identificados ou localizados de acordo com a mesma métrica (SCHABENBERGER; GOTWAY, 2005). 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.

Segundo Schmidt e Sansó (2006) e Le e Zidek (2006), a descrição de um campo aleatório é obtida através das distribuições acumuladas finito-dimensionais F se, para qualquer conjunto de pontos nas localizações $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_n$ pertencentes à região 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 aleatório gaussiano corresponde ao caso particular em que tais distribuições finito-dimensionais são gaussianas, ou seja, um campo aleatório é gaussiano se a distribuição para qualquer conjunto de índices $i = 1, 2, ..., n$ é uma gaussiana n-variada. Conseqüentemente, por propriedades da distribuição gaussiana multivariada, cada $Z(\mathbf{s}_i)$ é uma variável aleatória gaussiana univariada.

O processo gaussiano é particularmente importante na análise de campos aleatórios, já que estes são completamente especificados pelo vetor de médias e sua matriz de covariância (GNEITING; SCHLATHER, 2002). O vetor de médias é especificado pelo conhecimento de covariáveis, isto é, efeitos fixos, que influenciam a variável de interesse. A matriz de covariância precisa ser positiva definida para ser considerada válida, para tanto cada um de seus elementos devem ser dados por uma função de covariância que garanta a condição de positiva definida. Entretanto determinar se esta função é positiva definida é, em geral, não trivial. Uma forma freqüentemente utilizada para verificar a validade de uma

função de covariância é verificar se a representação espectral desta função satisfaz o teorema de Bochner (BOCHNER, 1959).

Campo aleatório no plano, $d = 2$, é o caso em que, comumente se aplica a metodologia para análise de dados geoestatísticos, neste caso considera-se geralmente as coordenadas latitude e longitude ou no caso de $d = 3$ a profundidade ou a altitude pode ser considerada a terceira dimensão. Para uma melhor compreensão deste texto, será feita à princípio uma breve revisão da análise geoestatística e em seguida uma discussão mais aprofundada sobre campos aleatórios espaço-temporais e suas particularidades.

Tipicamente dispõe-se de apenas uma realização do processo e portanto suposições adicionais são necessárias para possibilitar inferências. A mais comum é a de estacionariedade que implica na falta de importância dos valores absolutos das coordenadas. Assim um campo aleatório é estritamente estacionário se para qualquer conjunto finito de observações $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_n$ pertencentes ao espaço G, a distribuição conjunta de $(Z(\mathbf{s}_1), Z(\mathbf{s}_2), ..., Z(\mathbf{s}_n))$ é a mesma do que a de $(Z(\mathbf{s}_1+\mathbf{h}), Z(\mathbf{s}_2+\mathbf{h}), ..., Z(\mathbf{s}_n+\mathbf{h}))$, para qualquer translação determinada por $\mathbf{h}$.

Uma definição menos restritiva de estacionariedade exige que o vetor de médias e a função de covariância sejam invariantes sobre translação, desta forma tem-se:

$$\boldsymbol{E}[Z(\mathbf{s})] = \mu \qquad Cov[Z(\mathbf{s}), Z(\mathbf{s}+\mathbf{h})] = C(\mathbf{h}),$$

assim a média será constante em toda região e a covariância entre atributos em diferentes localizações dependerá apenas do vetor de distâncias $\mathbf{h}$. Usualmente refere-se a este tipo de estacionariedade como estacionariedade fraca ou de segunda ordem. Note que estacionariedade no senso estrito implica em estacionariedade fraca, mas em geral o contrário não é verdade. Para processos gaussianos as duas condições são equivalentes (SCHIMIDT; SANSÓ, 2006).

No caso de $Z(\mathbf{s})$ não apresentar estacionariedade de segunda ordem, pode-se adotar uma suposição menos restritiva de que os incrementos $[Z(\mathbf{s})-Z(\mathbf{s}+\mathbf{h})]$ devem ser estacionários. De acordo com Schabenberger e Gotway (2006), esta característica é denominada estacionariedade intrínseca. Sendo assim, um processo é dito intrinsecamente estacionário se

$E[Z(\mathbf{s})] = \mu$ e $Var[Z(\mathbf{s}) - Z(\mathbf{s}+\mathbf{h})] = 2\gamma(\mathbf{h})$, em que $\gamma(\mathbf{h})$ é denominado semivariograma. Verifica-se então a seguinte relação:

$$\begin{aligned} \boldsymbol{V}ar[Z(\mathbf{s}) - Z(\mathbf{s}+\mathbf{h})] &= Var[Z(\mathbf{s})] + Var[Z(\mathbf{s}+\mathbf{h})] - 2Cov[Z(\mathbf{s}) - Z(\mathbf{s}+\mathbf{h})] \\ &= 2\{Var[Z(\mathbf{s})] - C(\mathbf{h})\} \\ &= 2\{C(\mathbf{0}) - C(\mathbf{h})\} = 2\gamma(\mathbf{h}). \end{aligned} \quad (1)$$

Por conta desta relação, $\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h})$, a variabilidade de campos aleatórios estacionários de segunda ordem podem ser estudados através do semivariograma ou da função de covariância.

Caso o processo não apresente nenhuma forma de estacionariedade, tem-se o caso mais geral que são os campos aleatórios não-estacionários. A não-estacionariedade do campo aleatório pode vir de uma média não constante, ou mesmo da variação na estrutura de covariância. Este tipo de comportamento é uma característica muito comum principalmente em dados provenientes de processos ambientais. Um caso comum na prática é a falta de estacionariedade na média, situação na qual, pode-se tentar contornar assumindo algum modelo determinístico, como por exemplo um modelo linear. Variância e covariância não constantes podem, em algumas situações, ser contornadas por transformação na variável resposta (CHRISTENSEN; DIGGLE; RIBEIRO Jr., 2001). Outras abordagens mais gerais, para processos não estacionários, porém mais complexas são dadas por modelos de deformações espaciais (SAMPSON; GUTTORP, 1992, SHIMIDT; O'HAGAN, 2003) ou por convoluções espaciais (HIGDON, 2002, FUENTES; SMITH, 2001).

Um campo aleatório pode apresentar algum tipo de estacionariedade e ainda a função de covariância depender da direção, desta forma o campo aleatório será considerado anisotrópico. A forma mais simples de anisotropia é dada pela anisotropia geométrica. Um campo aleatório apresenta anisotropia geométrica quando a estrutura de covariância possui diferentes comportamentos em diferentes direções, conforme o traçado de uma elipse. Por outro lado, se a associação dentro do campo aleatório é a mesma em todas as direções, este é dito isotrópico. A existência de dependência direcional na estrutura de covariância adiciona parâmetros na estrutura de correlação, e embora não adicione dificuldades na modelagem, a

identificação de tais parâmetros a partir dos dados pode ser difícil.

A forma de lidar com diferentes correlações em diferentes direções, conforme a anisotropia geométrica, é através de uma transformação no sistema de coordenadas. A Figura 1 apresenta um campo aleatório gaussiano espacial com anisotropia geométrica, além da relação dos parâmetros que descrevem a anisotropia.

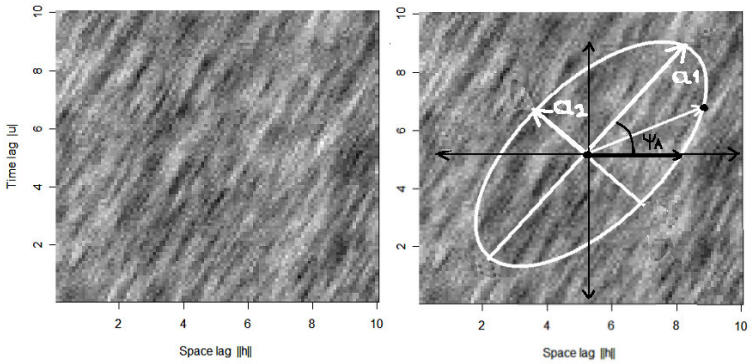

Figura 1 - Análise da anisotropia geométrica pela forma elíptica

Algebricamente, um campo aleatório com anisotropia geométrica nas coordenadas $x = (x_1, x_2)$ pode ser convertido para um campo aleatório estacionário nas coordenadas $x' = (x_1', x_2')$ através da seguinte matriz "A", conhecida como matriz de anisotropia:

$$(x_1', x_2') = (x_1, x_2)A = (x_1, x_2) \begin{pmatrix} \cos(\psi_A) & -\sin(\psi_A) \\ \sin(\psi_A) & \cos(\psi_A) \end{pmatrix} \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix},$$

em que os vetores α_1 e α_2 representam a escala de correlação nas diferentes direções e ψ_A é o ângulo de anisotropia, conforme ilustrado pela Figura 1.

De acordo com Schmidt e Sansó (2006), na literatura geoestatística, processos

espaciais estacionários e isotrópicos são chamados de homogêneos, caso contrário são conhecidos como heterogêneos. Processos homogêneos são bastante restritivos, mas suas propriedades são claramente compreendidas, podendo capturar uma grande variedade de fenômenos naturais.

2.1.1 Propriedades da função de covariância

A função de covariância $C(\mathbf{h})$ de um campo aleatório estacionário 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) $C(\mathbf{0}) \geq \mid C(\mathbf{h}) \mid$;

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

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

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

De acordo com Schabenberger e Gotway (2005), as propriedades (i) e (ii) são imediatas, a (iii) segue a forma da inequação de Cauchy-Schwarz, a (iv) caracteriza a falta de importância da coordenada absoluta, (v) assegura que combinações lineares de funções de covariância válidas são também funções de covariância válidas e (vi) assegura que o produto de funções de covariância válidas tem como resultado uma função de covariância válida.

Em campos aleatórios espaço-temporais a propriedade (ii) nem sempre é satisfeita, o que dá origem aos modelos de covariância espaço-temporais assimétricos. Já as propriedades (v) e (vi) dão suporte à construção de modelos de covariância espaço-temporais separáveis.

A matriz de covariância para $(Z(\mathbf{s}_1), Z(\mathbf{s}_2), ..., Z(\mathbf{s}_n))$, denotada por Σ com dimensão $n \times n$, será uma matriz simétrica, em que $\Sigma_{ij} = C(\mathbf{s}_i - \mathbf{s}_j)$, ou seja, cada valor de

Σ é igual a covariância entre respostas em duas localizações correspondentes. 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, em outras palavras, para qualquer vetor não zero $\mathbf{a}$, a forma quadrática $\mathbf{a}'\Sigma\mathbf{a}$ deve ser maior igual a zero. Sendo assim esta condição pode ser escrita nas seguintes formas:

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

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

2.1.2 Modelos de covariância

A função ser positiva definida é condição necessária e suficiente para ser considerada uma função de covariância válida. A dificuldade na confirmação desta condição leva à utilização, em estudos geoestatísticos, de famílias de covariância que são conhecidamente positiva-definida. A forma mais comum para o comportamento empírico da estrutura de covariância estacionária é que a correlação entre dois pontos diminua à medida que a distância entre eles aumente. Desta forma é conveniente considerar funções de covariância com este comportamento. Entretanto há modelos não monótonos, utilizados para modelar processos com periodicidade, que ainda obedecem à condição de ser positivos definidos. Um exemplo são os modelos chamados de onda ("wave"). De forma geral, uma família de covariância válida no R^d não é necessariamente válida em dimensões maiores que d, mas é certamente válida em dimensões menores que d. As duas funções descritas a seguir são muito utilizadas em estudos geoestatísticos:

A família Matérn

A função de correlação pertencente à chamada família Matérn, proposta por Berfil Matérn (1986) é definida como:

$$\boldsymbol{C}(\mathbf{h}) = \{2^{k-1}\Gamma(k)\}^{-1}(\mathbf{h}/\phi)^k\mathbf{K}_k(\mathbf{h}/\phi),$$

sendo esta uma função com dois parâmetros em que $\phi > 0$ é o parâmetro de escala com a dimensão de distância, $k > 0$ é o parâmetro de forma responsável pela suavidade do processo e $\mathbf{K}_k(.)$ é a função Bessel de ordem k.

- $k = 1/2$, Modelo Exponencial: $C(\mathbf{h}) = exp(-\mathbf{h}/\phi)$
- $k \longrightarrow \infty$, Modelo Gaussiano: $C(\mathbf{h}) = exp\{(-\mathbf{h}/\phi)^2\}$.

Alguns autores como Stein (1999), Diggle e Ribeiro (2006) advogam o uso desta função por combinar a descrição da extensão da dependência espacial através do parâmetro ϕ, com a descrição da suavidade do processo determinado pelo parâmetro k.

A família, exponencial potência

A função de correlação da família exponencial potência é dada por:

$$\boldsymbol{C}(\mathbf{h}) = exp\{-(\mathbf{h}/\phi)^k\},$$

esta é uma função, assim como a família Matérn, com dois parâmetros sendo $\phi > 0$ parâmetro de escala e $k > 0$ parâmetro de forma neste caso limitado por $0 < k \leq 2$. E possui os modelos chamados de Exponencial e Gaussianos como casos particulares para $k = 1$ e 2, respectivamente. É importante notar que a função de covariância pertencente à família Mátern com $k = 1/2$ é igual a função exponencial com $k = 1$.

A seguir serão descritas funções utilizadas na construção dos modelos de covariância espaço-temporais:

Cauchy

$$\boldsymbol{C}(\mathbf{x}) = (1 + \mathbf{x}^2)^{-k},$$

o parâmetro k é um valor positivo. O modelo possui duas generalizações, o modelo "cauchy generalizado"e o modelo "hiperbólico", este não será utilizado neste trabalho.

Cauchy generalizada "gencauchy"

$$\boldsymbol{C}(\mathbf{x}) = (1 + \mathbf{x}^{k_1})^{-\frac{k_2}{k_1}},$$

o parâmetro $k_1 \in (0, 2]$ e k_2 é positivo.

Estável "stable"

$$\boldsymbol{C}(\mathbf{x}) = exp(-\mathbf{x}^k),$$

o parâmetro $k \in (0, 2]$.

Modelo espaço-temporal não-separável "nsst"

$$\boldsymbol{C}(\mathbf{x}, y) = \psi(y)^{-k}\phi\Big(\frac{\mathbf{x}}{\psi(y)}\Big),$$

k deve ser maior ou igual a genuina dimensão espacial.

2.1.3 Estimação dos parâmetros

Uma ferramenta de uso comum para investigação da dependência espacial em geoestatística é o semivariograma. Como visto na eq.(1), para campos aleatórios intrinsecamente estacionários define-se $2\gamma(\mathbf{h}) = E\{[\mathbf{Z}(s_i) - \mathbf{Z}(s_j)]^2\}$. Portanto, uma estimativa empírica do semivariograma é dada pelo estimador de momentos:

$$\hat{\boldsymbol{\gamma}}(\mathbf{h}) = \frac{\mathbf{1}}{\mathbf{2}|\boldsymbol{N}(\mathbf{h})|}\sum_{|\boldsymbol{N}(\mathbf{h})|}\{\boldsymbol{Z}(\mathbf{s_i}) - \boldsymbol{Z}(\mathbf{s_j})\}^{\mathbf{2}},$$

em que $|N(\mathbf{h})|$ é a quantidade de pontos separados pela distância $\mathbf{h}$ de separação entre s_i e s_j. De acordo com Schabenberger e Gotway (2005), o estimador $\hat{\gamma}(\mathbf{h})$ é não viesado para $\gamma(\mathbf{h})$ se $Z(\mathbf{s})$ for intrinsecamente estacionário. No caso do processo em que assume-se a estacionariedade de segunda ordem, o variograma informa também sobre a função de correlação.

Para exemplificar a relação entre o semivariograma e a função de correlação, denotando que $C(\mathbf{h}) = C(\mathbf{0})exp(-\phi\mathbf{h})$ e $C(\mathbf{0}) = \sigma^2$, tem-se que:

$$\begin{aligned}
\boldsymbol{\gamma}(\mathbf{h}) &= C(\mathbf{0}) - C(\mathbf{h}) \\
&= C(\mathbf{0}) - C(\mathbf{0})exp(-\phi\mathbf{h}) \\
&= C(\mathbf{0})(1 - exp(-\phi\mathbf{h})) \\
&= \sigma^2(1 - exp(-\phi\mathbf{h})).
\end{aligned}$$

.

Isto justifica uma prática comum em geoestatística de se utilizar o semivariograma empírico para inferir sobre parâmetros do modelo, através do ajuste de uma função válida para o semivariograma teórico.

Um outro paradigma para estimação de parâmetros, sob a pressuposição de estacionariedade forte, é dado pelo método de máxima verossimilhança. A estimação dos parâmetros de campos aleatórios pelo método de máxima verossimilhança requer pressuposições sobre a distribuição do processo . Dado $\mathbf{Z} = [Z(\mathbf{s}_1), Z(\mathbf{s}_2), ..., Z(\mathbf{s}_n)]'$, que denota o vetor de observações e assumindo $\mathbf{Z}(\mathbf{s}) \sim G(\mu\mathbf{1}, \Sigma(\theta))$, o método da máxima verossimilhança consiste em maximizar o logaritmo da distribuição gaussiana n-variada:

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

Estimadores de máxima verossimilhança são muito utilizados em estatística, por assegurarem uma série de propriedades estatísticas, já que sobre condição de regularidade padrão estes estimadores são assintoticamente gaussianos e eficientes. Na prática, em geoestatística, este método de estimação pode requerer muito tempo computacional, ou até mesmo ser proibitivo, devido à dimensão da matriz de covariância.

Diggle, Ribeiro e Christensen (2001) e Diggle e Ribeiro (2006) discutem em detalhes métodos baseados na verossimilhança para estimação de parâmetros do processos geoestatísticos espaciais, incluindo métodos bayesianos.

2.1.4 Krigagem

Em geral a amostra observada é composta por um conjunto esparso de pontos no espaço, mas o objetivo consiste em descrever o processo continuamente em toda região de estudo. A partir dos dados amostrados e do modelo adotado, pode-se predizer valores do atributo em pontos não amostrados e este é o objetivo da predição espacial. O tamanho, a forma, a orientação e o arranjo espacial dos dados amostrados dão suporte e influenciam a capacidade em se predizer valores desconhecidos.

Em geoestatística métodos de predição são tipicamente conhecidos como métodos de "krigagem". A krigagem é muitas vezes o objetivo final na análise de campos aleatórios

e a validade da predição é influenciada pela especificação da função de covariância, bem como, a estimação de seus parâmetros. Já sua acurácia é pode ser verificada pelo erro quadrado de predição:

$$\{Z_{\mathbf{s}} - p(z;\mathbf{s})\}^2,$$

em que $Z_{\mathbf{s}}$ é o verdadeiro valor do atributo Z na localização $\mathbf{s}$ enquanto $p(z;\mathbf{s})$ é a predição na mesma localização $\mathbf{s}$, que é função do dado observado z.

Um método de predição muito utilizado é o preditor de krigagem simples, que é um preditor linear dado pela seguinte relação,

$$\boldsymbol{p}(z;\mathbf{s}) = \mu_z + \sigma^2 \mathbf{r}' \Sigma^{-1}(z - \mu_z). \tag{3}$$

De acordo com Elmatzoglou (2006), o preditor de krigagem simples é usualmente referido como o preditor linear de mínimos quadrados ótimo. No caso do campo aleatório ser gaussiano, a esperança condicional do processo nas localizações não observadas, condicionada a todos os valores observados, é dada também pela eq.(3), e assim o preditor de krigagem simples será o melhor preditor de mínimos quadrados.

2.2 Campos aleatórios espaço-temporais

Processos ambientais e geofísicos como concentração de poluentes, precipitação e superfície do vento são exemplos de fenômenos que podem ser modelados por campos aleatórios espaço-temporais ou simplesmente processos espaço-temporais. Este tipo de processo é caracterizado pela variabilidade na dimensão do espaço e do tempo. De acordo com Schabenberger e Gotway (2005), para o estudo desta variabilidade existem três possibilidades:

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

As duas primeiras possibilidades isolam a parte temporal ou a espacial e aplica técnicas padrões para o tipo de processo resultante. A terceira possibilidade considera o processo espacial e temporal conjuntamente e é a considerada neste trabalho. No processo de

construção desta análise conjunta, a aplicação das técnicas padrões no estudo da variabilidade espacial e temporal separadamente pode ser utilizada como uma ferramenta exploratória e auxiliar na construção de modelos adequados.

Um campo aleatório espaço-temporal é definido como

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

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

Através desta definição, percebe-se intuitivamente que o domínio natural do processo é $R^d \times R$. Neste trabalho o componente espacial será considerado bidimensional, ou seja, $d = 2$, por ser a situação comum na prática, entretanto ressalta-se que a dimensão do processo pode ser qualquer número finito positivo.

Tempo e espaço podem não ser diretamente comparáveis, já que as unidades das coordenadas dos dois processos apresentam grandezas diferentes. Fisicamente existe uma clara diferença entre as dimensões espaciais e temporais e os modelos precisam considerá-las. Segundo Gneiting et. al (2006) funções de covariância ou mesmo a krigagem se utilizam de espaços Euclidianos e são diretamente aplicados à problemas espaço-temporais. Sendo assim a simples separação vetorial do domínio espacial e temporal tem importantes implicações.

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

• *Especificação geoestatística.* Nesta metodologia é necessário o conhecimento de uma distribuição, geralmente gaussiana, para função aleatória $Z(\mathbf{s},t)$ que é definida em toda coordenada espaço-temporal. A função de covariância é definida na dimensão espaço-temporal contínua, caracterizando o modelo da função aleatória, que pode ser utilizada na predição de qualquer localização espacial e qualquer instante de tempo. O ajuste da função de covariância é de importância central nesta metodologia, e portanto expressões com forma fechada para função de covariância são essenciais.

• *Especificação baseada em modelos.* Esta técnica enfatiza a adoção de modelos estocásticos para solução de problemas práticos nos quais a função de covariância não é especificada explicitamente, mas sim induzida pelo modelo. Esta forma de especificação

espaço-temporal pode entretanto ir de uma função com forma analítica simples até uma intratável e somente implicitamente definida. Esta metodologia engloba desde modelos hierárquicos bayesianos espaço-temporais até a técnica do filtro de Kalman, com a característica de apresentar predições computacionalmente eficientes .

A especificação geoestatística é conceitualmente simples, particularmente quando combinado com a adoção do modelo gaussiano, que é completamente especificado por sua média e estrutura de covariância. Além disto, a krigagem, que geralmente é o objetivo da análise de campos aleatórios, requer a especificação apropriada da função de covariância. Entretanto esta forma de especificação assume estacionariedade espaço-temporal da estrutura de covariância, que portanto não pode assumir formas flexíveis como as induzidas pela especificação baseada em modelos. Este trabalho será direcionado à especificação geoestatística aplicada a processos espaço-temporais gaussianos.

Assumindo as condições de regularidade do campo aleatório espaço-temporal, $Var(Z(\mathbf{s},t)) < \infty$ para todo $\mathbf{s} \in \mathbf{R}^2, t \geq 0$, (CRESSIE; HUANG, 1999), define-se a média e a função de covariância por:

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

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

De acordo com Gneiting et.al (2006), na prática o interesse inicial das análises é verificar a presença ou ausência de três características: estacionariedade, simetria completa e separabilidade. Após este estudo será possível decidir sobre a complexidade da modelagem.

Um campo aleatório espaço-temporal é estacionário se sua esperança não depende das coordenadas espaço-temporais e sua função de covariância só dependa de um vetor de separação dos pontos em $R^d \times R$. Portanto um campo aleatório espaço-temporal apresenta covariância estacionária no espaço se $Cov(Z(\mathbf{s}_1,t_1,\mathbf{s}_2,t_2))$ depende somente do vetor de separação $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2$, e apresenta estacionariedade temporal se $Cov(Z(\mathbf{s}_1,t_1,\mathbf{s}_2,t_2))$ depende somente do vetor de separação $u = t_1 - t_2$. Desta forma se o campo aleatório espaço-temporal possui covariância estacionária espacial e temporalmente, este é considerada uma covariância

espaço-temporal estacionária dada por:

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

O campo aleatório espaço-temporal apresenta simetria completa se

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

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

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

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

A classe mais simples de funções de covariância espaço-temporal é dada pelos modelos separáveis. Funções de covariância espaço-temporais separáveis podem ser definidas com base nas propriedades de aditividade e multiplicabilidade (2.1.1, vi e vii). Desta forma poderão ser decompostas entre uma função de covariância puramente espacial e outra puramente temporal, que no caso aditivo pode ser escrita como:

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

e no caso multiplicativo por:

$$\boldsymbol{C}ov(Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)) = Cov(Z(\mathbf{s}_1, \mathbf{s}_2))Cov(Z(t_1, t_2)), \tag{4}$$

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 covariância espaço-temporal não considera a interação entre o espaço e o tempo, mas é muito utilizada na prática por ser computacionalmente tratável e uma forma intuitiva e simples de obtenção de funções de covariância válidas.

Todos os modelos separáveis são também completamente simétricos. Para a demonstração desta relação considere a eq.(4), considere também as variáveis aleatória espaço-temporais $Z(\mathbf{s}_1, t_2)$ e $Z(\mathbf{s}_2, t_1)$, a função de covariância separável será dada por:

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

desta forma concluí-se que:

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

que é a definição da propriedade de campos aleatórios espaço-temporais completamente simétricos.

Um exemplo simples de função de covariância espaço-temporal separável dado pela combinação de modelos de covariância exponenciais é:

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

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

As matrizes de anisotropia são utilizadas para permitira representação combinada das dimensões espaciais e temporais. Neste exemplo, suponha um processo isotrópico no espaço tem-se:

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

formas mais gerais de especificação destas matrizes serão discutidas em Material e Métodos.

Qualquer função de covariância espaço-temporal que não for escrita nas duas formas separáveis anteriormente mencionadas e nem combinações destas, serão ditas funções não separáveis. As funções de covariância separáveis, embora facilmente obtidas, são em geral insatisfatórias para descrever processos naturais. O que gera à necessidade de se especificar funções não separáveis. Kryriakidis e Journel (1999), revisam e discutem abordagens para especificação de processos geoestatísticos espaço-temporais. Cressie e Huang (1999), propõem classes de funções de covariância não separáveis válidas cuja obtenção não se baseiam no domínio das observações, mas sim no domínio da freqüência, uma vez que a representação espectral e a densidade espectral podem ser utilizadas para descrever as propriedades do processo. Gneiting (2002a), em uma abordagem alternativa, obtém modelos válidos de função de covariância através da definição de funções monótonas e positivas. Descreve-se a seguir as formas de obtenção propostas por estes autores.

2.2.1 Representação de Cressie-Huang

A proposta de Cressie e Huang (1999) tem como objetivo desenvolver classes de funções de covariância espaço-temporais estacionárias válidas, não separáveis e com forma fechada. Para tanto, o resultado teórico mostra como funções positivas definidas no R^{d+1} podem ser obtidas a partir de funções positivas definidas no R^d.

Assuma que C é uma função contínua e que a representação espectral possua densidade espectral $g(\mathbf{w}, \tau) \geq 0$. Lembre-se que para uma função de covariância espaço-temporal ser válida ela precisa ser positiva definida. Pelo teorema de Bochner a função de covariância espaço-temporal será considerada válida se poder ser escrita da seguinte forma:

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

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

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

em que

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

assumindo que:

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

e assumindo também que as seguintes condições são satisfeitas:

$(C1)$ 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:

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

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.

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

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

e

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

válidas em R^d satisfazendo as condições $(C1)$ e $(C2)$, tem-se que:

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

é uma função de covariância espaço-temporal contínua válida em $R^d \times R$.

2.2.2 Representação de Gneiting

A representação de Cressie e Huang (1999) constrói uma família de funções de covariância espaço-temporais estacionária e não separáveis através da inversão de Fourier na forma fechada, o que restringe esta família a um pequeno grupo de funções com soluções de forma fechada da inversão de Fourier. Gneiting (2002a) propõe uma flexível e elegante família de funções de covariância espaço-temporais, cuja obtenção não requer operações no domínio espectral. A construção de funções de covariância válidas se dá através de componentes elementares em que sua validade é facilmente verificada.

A representação de Gneiting (2002a) considerara qualquer função monótona $\phi(x)$, definida em $x \geq 0$ e qualquer função positiva $\psi(x)$, definida em $x \geq 0$ com derivadas completamente monótonas , e desta forma,

$$\boldsymbol{C}(\mathbf{h}, u) = \frac{\sigma^2}{(\psi(u)^2)^{d/2}} \phi\Big(\frac{||\mathbf{h}||^2}{\psi(|u|^2)}\Big) \tag{5}$$

é uma função de covariância espaço-temporal válida em $R^d \times R$, em que $\mathbf{h} = ||h||$, u representa o vetor de distâncias no tempo e d deve ser maior ou igual que a genuína dimensão espacial.

De acordo com Gneiting (2002a) uma função contínua $\phi(x)$ é completamente monótona 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 propõe uma família de funções de covariância espaço-temporal muito geral que não dependem da inversão de Fourier na forma fechada e não necessita de integrabilidade. De acordo com Elmatzoglou (2006) nesta família os componentes $\phi(x)$ e $\psi(x)$ podem ser associados com a estrutura espacial e temporal respectivamente. Gneiting (2002a), apresenta possíveis escolhas para cada função, que são reproduzidas nas Tabelas 1 e 2.

Tabela 1 - Funções completamente monótonas $\phi(x)$,$x \geq 0$

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

Tabela 2 - Funções 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 família de covariância espaço-temporal não separável proposta por Gneiting é 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 $\psi(x) = (ax^{\alpha} + 1)^{\beta}, a > 0, 0 < \alpha \leq 1, 0 \leq \beta \leq 1$ são funções completamente monótona e positiva respectivamente. Substituindo na eq.(5) tem-se:

$$\boldsymbol{C}(\mathbf{h}, u) = \frac{\sigma^2}{(a|t|^{2\alpha} + 1)^{\frac{\beta d}{2}}} exp\Big\{\frac{c||\mathbf{h}||^{2\gamma}}{(a|t|^{2\alpha} + 1)^{\beta\gamma}}\Big\}. \tag{6}$$

Nesta expressão, nota-se que para $\beta = 0$ a covariância não depende do tempo. Multiplicando-se a eq.(6) por uma função de covariância puramente temporal, $C(|t|) = (a|t|^{\alpha}+$

$1)^{-\delta}$, tem-se como resultado:

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

no caso em que $\beta=0$, a expressão da função de covariância se reduz a

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

que é função de covariância separável na forma,

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

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

2.2.3 Estimação

De acordo com Schabenberger e Gotway (2005), para estimação dos parâmetros da função de covariância espaço-temporal, pode-se utilizar os mesmos métodos básicos apresentados para campos aleatórios no R^2, ou seja, métodos usados em variograma ou na verossimilhança.

O semivariograma para processos espaço-temporais estacionários é definido por:

$$\begin{aligned}\boldsymbol{V}ar[Z(\mathbf{s},t)-Z(\mathbf{s}+\mathbf{h},t+u)] &= 2\{Var[Z(\mathbf{s},t)]-C(\mathbf{h},u)\}\\ &= 2\{C(\mathbf{0},0)-C(\mathbf{h},u)\}=2\gamma(\mathbf{h},u).\end{aligned}$$

Sob estacionariedade de segunda ordem, por conta da relação $\gamma(\mathbf{h},u)=C(\mathbf{0},0)-C(\mathbf{h},u)$, a variabilidade dos campos aleatórios espaço-temporais estacionários pode ser estudada através do semivariograma ou da função de covariância. Na análise geoestatística puramente espacial o semivariograma é muito utilizado, entretanto na análise espaço-temporal isto não acontece. O estimador empírico do semivariograma espaço-temporal é dado por:

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

em que $|N(\mathbf{h}, u)|$ é a quantidade de pontos que estão dentro da distância determinada por $\mathbf{h}$ em cada distância de tempo determinada por u. Quando os dados são espaçados irregularmente no tempo e/ou espaço, o estimador empírico do semivariograma deve considerar uma distância de tolerância. Para o semivariograma espaço-temporal com um componente espacial isotrópico, o gráfico será tridimensional em que os eixos serão a distância no tempo, no espaço e no terceiro eixo o valor estimado para semivariância.

Uma outra forma de se estimar os parâmetros é o método de máxima verossimilhança ou máxima verossimilhança restrita. Sob a suposição de gaussianidade, as estimativas são obtidas através do máximo da função de verossimilhança, que por sua vez é dada pela função gaussiana multivariada, conforme eq.(2).

3 MATERIAL E MÉTODOS

3.1 Recursos Computacionais

O "RandomFields"(2006) é um pacote do software estatístico R (2006) que requer versões 2.2.0 ou maiores, desenvolvido por Martin Schlather, que tem como principal funcionalidade rotinas para simulação e análise de campos aleatórios. Através deste pacote pode-se simular vários tipos de campos aleatórios espaciais ou espaço-temporais, incluindo processos anisotrópicos e simulações condicionais. Ele também possui implementado ferramentas para análise de campos aleatórios, tais como: variograma empírico, ajuste interativo dos parâmetros, estimação dos parâmetros por máxima verossimilhança e mínimos quadrados dentre outras. Além de também possuir implementado os procedimentos de interpolação espacial (krigagem).

As simulações são realizadas através da função "GaussRF()"que possibilita a simulação de campos aleatórios espaciais ou espaço-temporais estacionários gaussianos. Nesta função é necessário, dentre outras coisas, a especificação das dimensões espaciais e as temporais e também a especificação do modelo e dos parâmetros com os quais o campo aleatório deve ser gerado. O "RandomFields"possui implementado vários modelos válidos de covariância e também vários métodos de simulação, que devem ser especificados pelo usuário.

Todas as simulações e análises de campos aleatórios espaço-temporais, consideradas neste trabalho, foram realizadas através do pacote "RandomFields". Os comando estão disponíveis no Apêndice. Além disto algumas funcionalidades do pacote geoR (2001) foram também utilizados na análise dos dados.

3.2 Simulação de Campos Aleatórios Espaço-Temporais

No desenvolvimento deste trabalho foram efetuadas diversas simulações de campos aleatórios espaço-temporais com o objetivo de facilitar o entendimento dos conceitos definidos anteriormente. O pacote apresenta uma proposta bastante geral de implementação de modelos e análise de processos através da especificação de estruturas de covariância espaço-

temporais. Entretanto a habilidade em gerar simulações dos processos possibilita ainda o estudo sobre propriedades do modelo, estimabilidade, entre outros. A modelagem espaço-temporal no pacote "RandomFields"está em desenvolvimento e por conta disso seus procedimentos ainda não são exaustivos e claramente documentados e existem poucos trabalhos que se utilizam deste pacote. Desta forma a geração de simulações cumpre aqui, também o papel de possibilitar um melhor entendimento do funcionamento e dos modelos implementados no "RandomFields".

Neste trabalho será considerado apenas simulações de campos aleatórios espaço-temporais gaussianos, como dito anteriormente, estes são totalmente especificados pelo vetor de médias e a matriz de covariância e através da especificação deste dois componentes pode-se simular o processo adequadamente.

Para simulação de campos aleatórios espaço-temporais gaussianos o "RandomFields"possui implementado alguns métodos como: *circular embedding*, *turning bands* e *direct matrix decomposition*, dentre outros. Neste trabalho as simulações serão realizadas, inicialmente através do método da decomposição direta da matriz de covariância, uma vez que o número de pontos a serem simulados é baixo o suficiente para possibilitar o uso desta decomposição. Este método é utilizado para simulação de qualquer distribuição gaussiana e a decomposição pode ser obtida pela raiz quadrada da matriz de covariância, a decomposição de Cholesky ou a decomposição do valor singular.

O tempo computacional da simulação de campos aleatórios espaço-temporais aumenta em alguns casos especiais, como por exemplo no caso de simulações em grade não regular com muitas localizações ou quando a estrutura de covariância é muito complexa. Por simplificação na ilustração, as simulações serão inicialmente geradas em espaço e tempo unidimensionais, isto é em $R \times R$, sobre a suposição de estacionariedade de segunda ordem e com função de covariância exponencial.

Na simulação de campos aleatórios espaço-temporais pode-se usar matrizes de anisotropia como exemplificado em (2.1), para a especificação de processos com diferentes características quanto a simetria, isotropia espacial, separabilidade. Para isto especificam-se

matrizes de anisotropia do tipo:

$$A_i = \begin{pmatrix} \alpha_{ixx} & \alpha_{ixT} \\ \alpha_{iTx} & \alpha_{iTT} \end{pmatrix}$$

em que α_{ixx} e α_{iTT} são os parâmetros de escala para o espaço e tempo respectivamente, enquanto α_{ixT} e α_{iTx} são parâmetros de forma do campo aleatório. No caso de processos em que a dimensão espacial é em R^2 α_{ixx} é uma matriz de dimensão 2×2. Como as dimensões espacial e temporal possuem grandezas diferentes a matriz de anisotropia se faz necessária. No que se segue são fornecidas alguns exemplos de especificação de modelos a partir de definições destas matrizes.

Todos os exemplos apresentados são casos particulares do seguinte esquema de covariância:

$$\boldsymbol{C}(\mathbf{d}) = \sum_{i=1}^{N_1} C_i(||\mathbf{d}A_i||) + \prod_{j=1}^{N_2} C_j(||\mathbf{d}A_j||)$$

em que $\mathbf{d} = [\mathbf{h}, u]$, com $(\mathbf{h}, u) \in R^1 \times R_1$.

Modelos de covariância espaço-temporais separáveis são uma forma intuitiva, fácil e conveniente de se produzir funções de covariância válidas, através da multiplicação ou adição de funções puramente espaciais e puramente temporais. A conveniência computacional também é uma grande vantagem destes modelos.

A principal desvantagem destes modelos de covariância espaço-temporais é a falta de realismo, pois eles assumem que o campo aleatório espaço-temporal é composto por dois processos independentes, um no domínio do tempo e outro no domínio do espaço.

Para a simulação do campo aleatório espaço-temporal com estrutura de covariância dada pela soma das funções de covariância puramente espacial e puramente temporal, considere as seguintes matrizes de anisotropia:

$$A_1 = \begin{pmatrix} 8 & 0 \\ 0 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}$$

$$\boldsymbol{C}(\mathbf{d}) = exp(-||\mathbf{d}A_1||) + exp(||\mathbf{d}A_2||) = exp(|8\mathbf{h}|) + exp(|4u|) \tag{7}$$

A Figura 2a ilustra a simulação de campos aleatórios espaço-temporais considerando as matrizes de anisotropia acima e a eq.(7), enquanto a Figura 2b ilustra uma simulação utilizando as mesmas matrizes de anisotropia, mas com estrutura de covariância dada pelo produto das funções de covariância puramente espacial e puramente temporal dada por:

$$\boldsymbol{C}(\mathbf{d}) = exp(-||\mathbf{d}A_1||)exp(||\mathbf{d}A_2||) = exp(|8\mathbf{h}|)exp(|4u|)$$

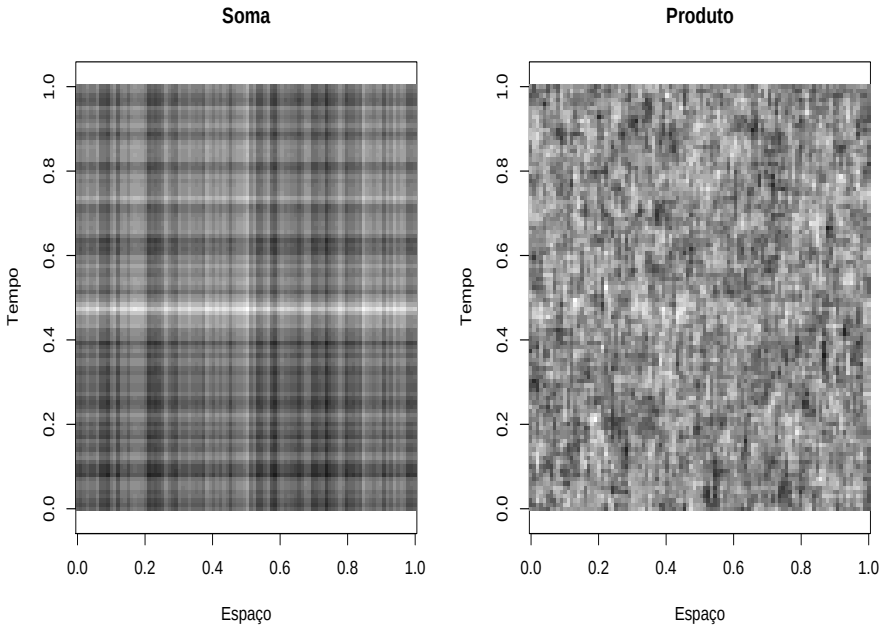

Figura 2 - a e b - Simulações utilizando estruturas de covariância separável em soma e produto, respectivamente

Uma alternativa aos modelos de covariância separáveis são os modelos de covariância espaço-temporais não separáveis, que tem como principal vantagem a interação entre o componente temporal e o espacial. A dificuldade em se construir funções de covariância válidas é a grande desvantagem destes modelos, e métodos matemáticos sofisticados são aplicados para garantia da validade das funções.

Para a simulação do campo aleatório espaço-temporal com estrutura de covariância dada pelas funções de covariância não separáveis, considere a seguinte matriz de

anisotropia completamente simétrica:

$$A_1 = \begin{pmatrix} 8 & 0 \\ 0 & 4 \end{pmatrix}$$

$$\boldsymbol{C}(\mathbf{d}) = exp(-||\mathbf{d}A_1||) = exp\Big(\sqrt{(8\mathbf{h})^2 + (4u)^2}\Big) \tag{8}$$

A Figura 3a ilustrada a simulação de um campo aleatório espaço-temporal simétrico considerando a matriz de anisotropia acima e a eq.(8). Já a Figura 3b ilustra a simulação de um campo aleatório não completamente simétrico utilizando a estrutura de covariância dada pela eq.(9) e a matriz de anisotropia assimétrica dada por:

$$A_1 = \begin{pmatrix} 8 & -2 \\ -10 & 4 \end{pmatrix}$$

e portanto a função de covariância corresponde a

$$\boldsymbol{C}(\mathbf{d}) = exp(-||\mathbf{d}A_1||) = exp\Big(\sqrt{(8\mathbf{h} - 10u)^2 + (-2\mathbf{h} + 4u)^2}\Big) \tag{9}$$

Considere agora a família de funções de covariância espaço-temporais gaussianas não separáveis proposta por Gneiting e simulada em $(\mathbf{h}, u) \in R^2 \times R$, em uma grade de 100x100 pontos no espaço e em 3 tempos, com a matriz de anisotropia,

$$A_1 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

e a seguinte função de covariância:

$$\boldsymbol{C}(\mathbf{h}, u) = \frac{\sigma^2}{(|u|^\delta + 1)^{\beta d/2}} exp\Big(- \Big(\frac{||h||^{2\gamma}}{(|u|^{2\delta} + 1)^{\beta\gamma}} \Big) \Big), \tag{10}$$

em que segundo Gneiting (2002a), tem-se que: $\phi(x) = exp(-x^\gamma)$ é chamada função estável (Stable), $\psi(x) = (x^\delta + 1)^\beta$ é a cauchy generalizada (gencauchy) e d a dimensão do espaço.

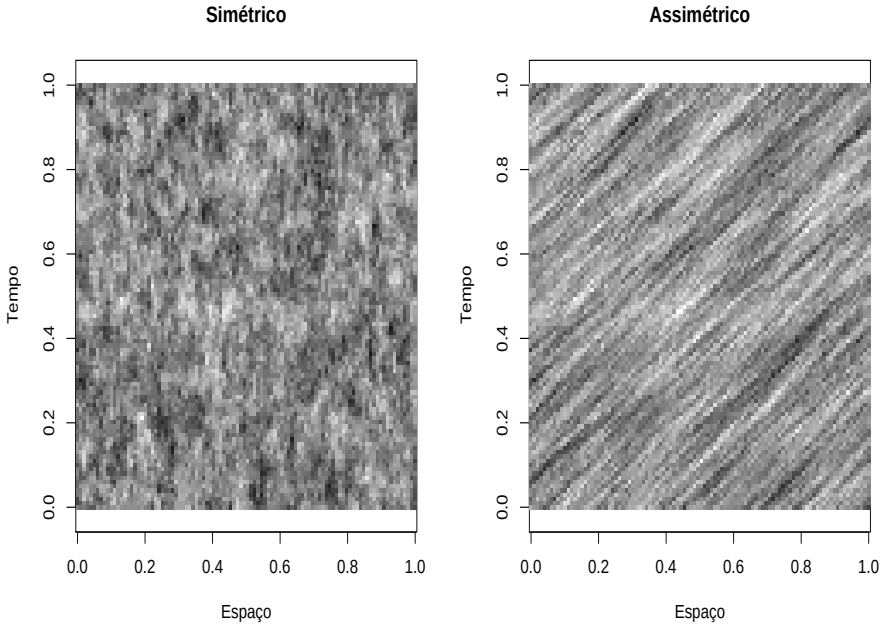

Figura 3 - a e b - Estrutura de covariância não separável simétrica e assimétrica, respectivamente

As simulações apresentadas na Figura 4 foram obtidas com os valores, $\sigma^2 = 1$, $\delta = 1.5$, $\beta = 0.5$, $d = 2$ e $\gamma = 1.2$. Desta forma tem-se que as simulações foram geradas pelo seguinte modelo de covariância:

$$\begin{aligned} \boldsymbol{C}(\mathbf{h}, u) &= \frac{1}{(|u|^{1.5}+1)^{0.5*2/2}} exp\Big(- \Big(\frac{||h||^{2*1.2}}{(|u|^{2*1.5}+1)^{0.5*1.2}} \Big) \Big) \\ &= \frac{1}{(|u|^{1.5}+1)^{0.5}} exp\Big(- \Big(\frac{||h||^{2.4}}{(|u|^{3}+1)^{0.6}} \Big) \Big) \end{aligned} \tag{11}$$

No Apêndice são incluídos os códigos em R para especificação destes e outros modelos e obtenção das simulações.

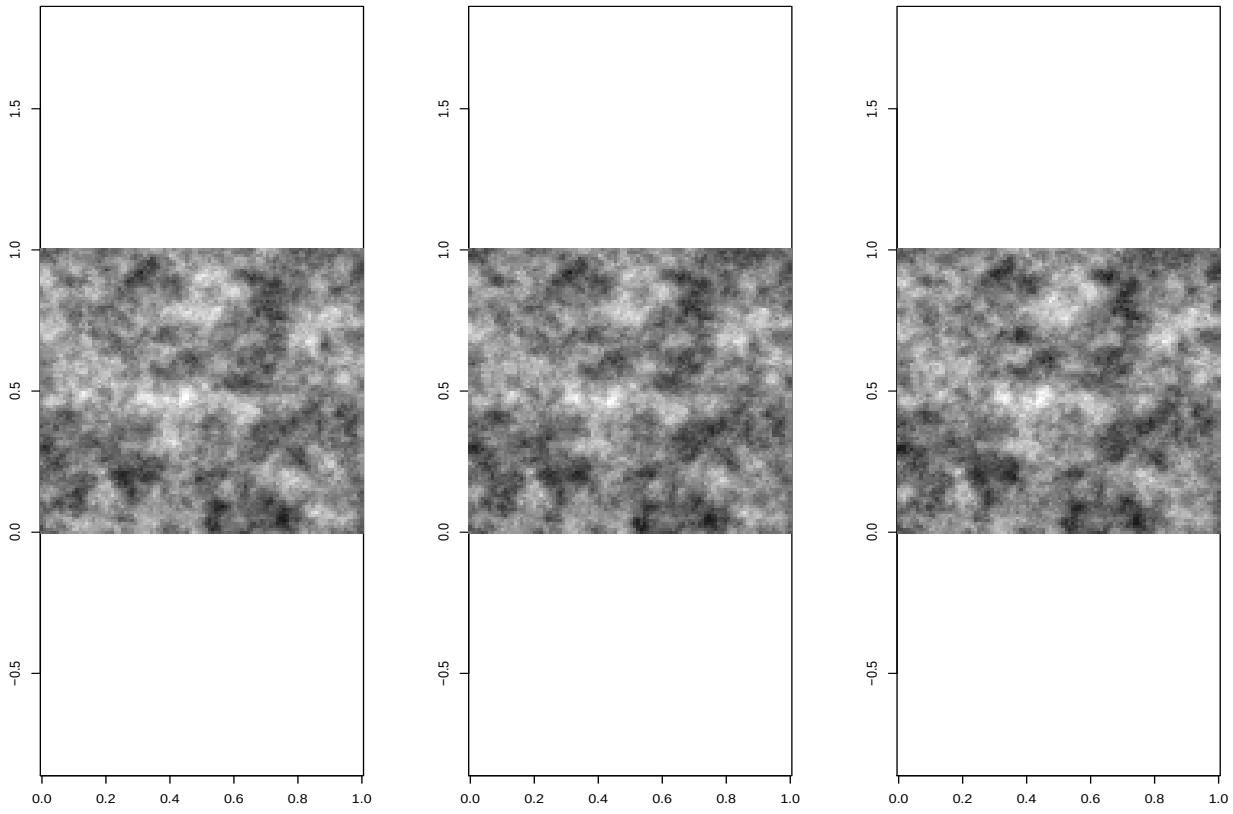

Figura 4 - Campo aleatório espaço-temporal com função de covariância proposta por Gneiting

.

4 APLICAÇÕES

Nos capítulos anteriores foram apresentadas particularidades na modelagem de processos espaço-temporais e algumas idéias de como tratá-las. Muitos modelos tem sido sugeridos na literatura e o fato de terem sido adotados não significa necessariamente que estes representem melhor a realidade. A escolha de uma determinada classe de modelos é uma tarefa difícil e que requer um profundo conhecimento da realidade, por parte dos pesquisadores.

Neste trabalho, o interesse não é ajustar o melhor modelo e sim verificar o comportamento de alguns dos modelos discutidos nos capítulos anteriores. Para tanto, dois conjuntos de dados serão considerados, o primeiro é referente ao estoque de uma determinada espécie de peixe na costa portuguesa, e o segundo é sobre a capacidade de armazenagem de água em um solo com citros.

4.1 Monitoramento de estoques de pescada na costa portuguesa

Os dados discutidos aqui foram gentilmente cedidos pelo Instituto de Investigação das Pescas e do Mar (IPIMAR) e discutidos em detalhe em Jardim e Ribeiro Jr. (2006 a e b).

Conhecer o comportamento do estoque de peixes e sua evolução no tempo é de grande importância pois pode colaborar na determinação do valor das quotas de pesca que garanta a perpetuação das espécies e ao mesmo tempo um desenvolvimento sustentável da região pesquisada. Os dados utilizados aqui não são suficientes para a determinação das quotas de pesca, e nem este é o interesse. O interesse aqui é conhecer melhor o processo e poder estimar a abundância com mais precisão e de forma que possa haver uma melhor comparação entre os anos.

Os dados referem-se a amostragem para avaliação do estoque de vários tipos de peixes e foram coletados por arrastro de fundo, no mar da costa portuguesa. Através deste método de amostragem vários tipos de peixe foram coletados, mas neste trabalho considerou-se para as análises os dados da espécie *Merluciosmerlucios* (pescada ou "hake").

O "hake"é uma espécie de pescada que é encontrada em quase toda a extensão da

costa portuguesa, o que o torna mais interessante para a análise espacial. A coleta foi realizada anualmente de 1990 à 2004, para as quais considerou-se 57 posições no espaço, o que resulta em um conjunto contendo 855 obervações obtidas espaço-temporalmente, sendo que 798 serão utilizadas na modelagem e as do último ano são utilizadas apenas como forma de verificar a qualidade das predições obtidas com diferentes escolhas para a função de covariância. A Figura 5 ilustra as localizações amostradas.

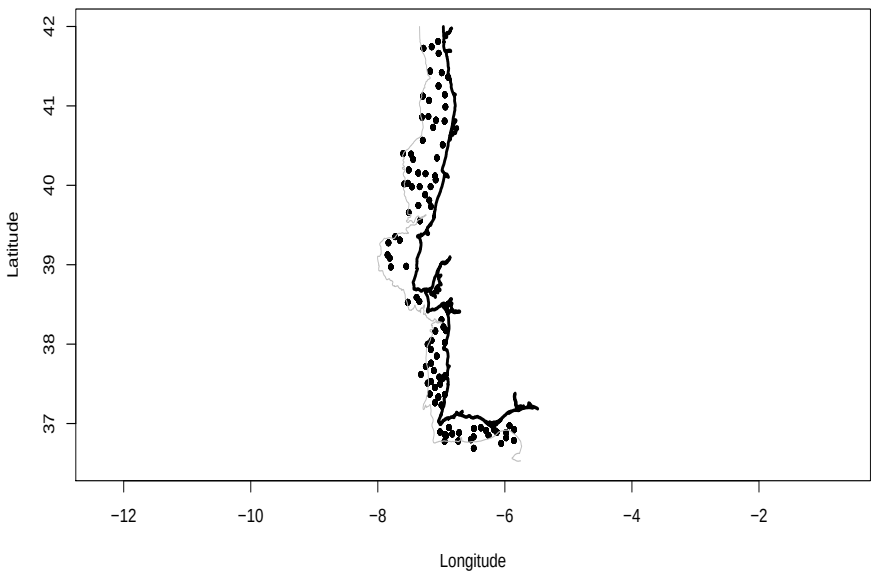

Figura 5 - Mapa das localizações amostradas

Neste tipo de análise o interesse de estudo não se restringe aos valores nas 57 localizações, mas sim em se fazer uma extensão para qualquer lugar na área e/ou qualquer tempo. O procedimento de coleta de dados envolve altos custos o que impede um grande número de pontos de coleta na área. Entretanto sabe-se por conhecimento do comportamento dos cardumes que estes se distribuem continuamente ao longo da costa, e em geral pontos próximos possuem estoque de peixes mais parecidos que pontos mais distantes. Jardim e Ribeiro Jr. (2006 a e b) apresentam análises geoestatísticas e estudos de delineamentos amostrais desta espécie em uma fração da área total da costa portuguesa.

A estrutura de dependência contínua do espaço é a característica central da análise geoestatística usual, e caso fosse fixado um determinado ponto no tempo, a distribuição do estoque de peixe poderia ser tratada como processo estocástico espacial a partir do qual predições de pontos não observados podem ser obtidas. Entretanto neste tipo de análise todas as observações obtidas em outros tempos não seriam utilizadas.

Uma outra possibilidade poderia ser considerar cada ponto observado como uma série temporal, desta forma o objetivo de se predizer um ano futuro poderia ser alcançado utilizando-se técnicas de análise de séries temporais e/ou de dados longitudinais. Entretanto as observações obtidas em pontos de coleta seriam ignoradas. Além disto, neste caso específico dispõe-se de apenas 15 observações no tempo o que restringe as possibilidades de modelagem temporal.

O objetivo aqui é portanto descrever o processo através de uma análise que trata espaço e tempo conjuntamente objetivando que os mecanismos propostos por esta metodologia sejam capazes de descrever o comportamento do sistema espaço-temporal.

Para auxiliar na construção de modelos espaço-temporais é sempre importante a realização prévia de uma análises exploratórias detalhadas, que foram efetuadas para os dados em questão e aqui serão reportadas algumas delas.

4.1.1 Análise descritiva

A distribuição marginal dos dados amostrados se apresenta bastante assimétrica conforme Figura 6a o que invalida a suposição de distribuição gaussiana. Optou-se aqui por adotar a transformação *Box-Cox* (BOX; COX, 1964) com parâmetro $\lambda = 0.25$ que resulta em uma transformação satisfatória conforme Figura 6b e 6c. O uso de transformações Box-Cox em modelos geoestatísticos é discutido em Chistensen, Diggle e Ribeiro (1998) e De Oliveira, Kaden e Shart (1997).

O formato da costa portuguesa e a localização dos pontos de coleta torna inviável a utilização das coordenadas originais nas análises. Desta forma optou-se aqui por se fazer uma rotação nas coordenadas da região ao sul de Portugal (região do Algarve), como pode ser visto na Figura 7. Toda a análise foi baseada nas coordenadas transformadas. Através da

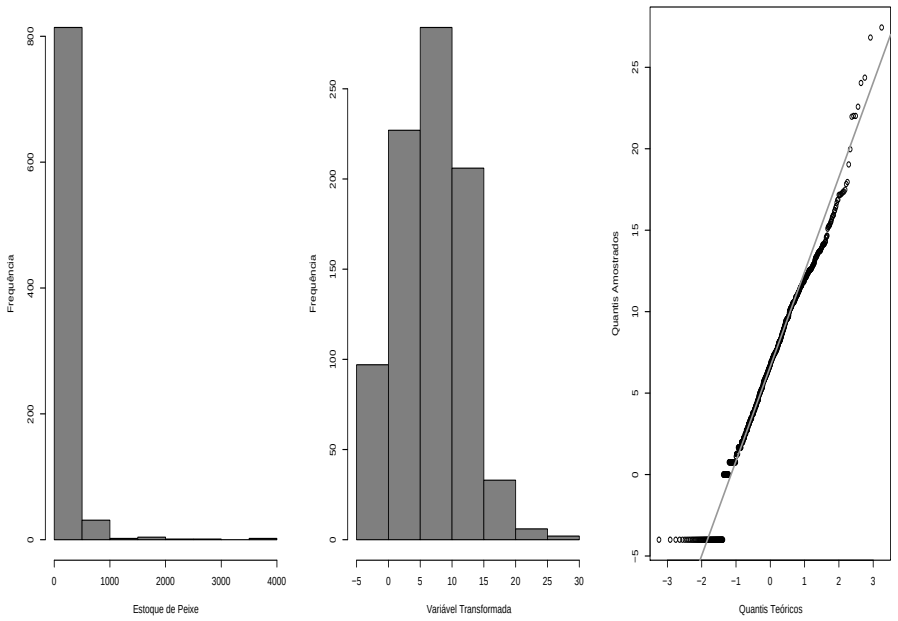

Figura 6 - a, b e c - Histograma dos dados, histograma da transformação e QQ-plot, respectivamente

transformação foi possível considerar as distâncias ao longo da costa, em uma mesma métrica possibilitando o cálculo de distâncias Euclidianas entre as localizações amostradas.

Ainda como análise descritiva considerou-se os 14 tempos separadamente e através da análise geoestatística usual, puramente espacial, estimou-se os parâmetros com cada um dos campos aleatórios resultantes do fatiamento dos dados no tempo. A Figura 8 mostra os 14 variogramas ajustados cuja a estimação de parâmetros foi obtido pelo método de máxima verossimilhança, considerando-se uma estrutura de covariância comum dada pelo modelo exponencial. Os ajustes mostram uma dependência espacial de pequena escala e com parâmetros que parecem diferir entre os diferentes tempos.

A Figura 9 apresenta os valores dos parâmetros estimados para cada tempo em que, os parâmetros do modelo são: média, efeito pepita (variância do erro de medida), variância do efeito espacial e escala da função de correlação. Os gráficos da Figura 9 não parecem apresentar nenhuma forma de tendência, ao longo do tempo. Entretanto as seqüencia

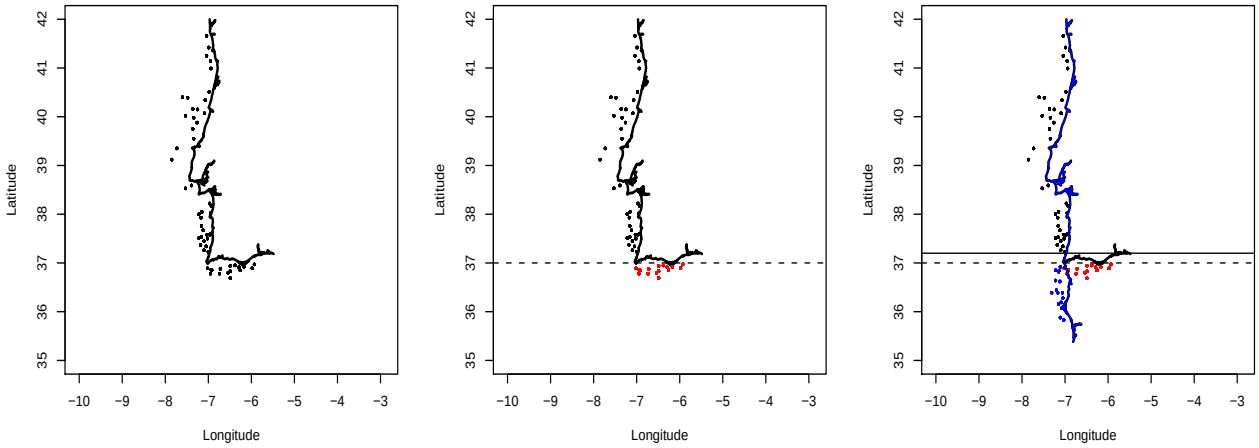

Figura 7 - Correção das coordenadas

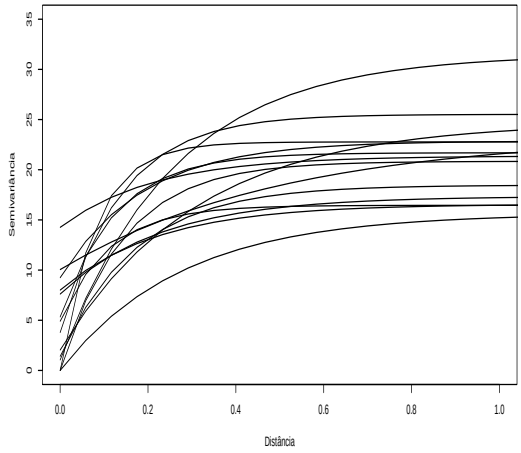

Figura 8 - Modelos de variograma ajustados para cada um dos tempos individualmente

de valores estimados para a média e variância do efeito espacial parecem ser correlacionadas no tempo. É importante ressaltar que estas observações devem ser avaliadas com parcimônia tendo em vista o pequeno número de observações no tempo.

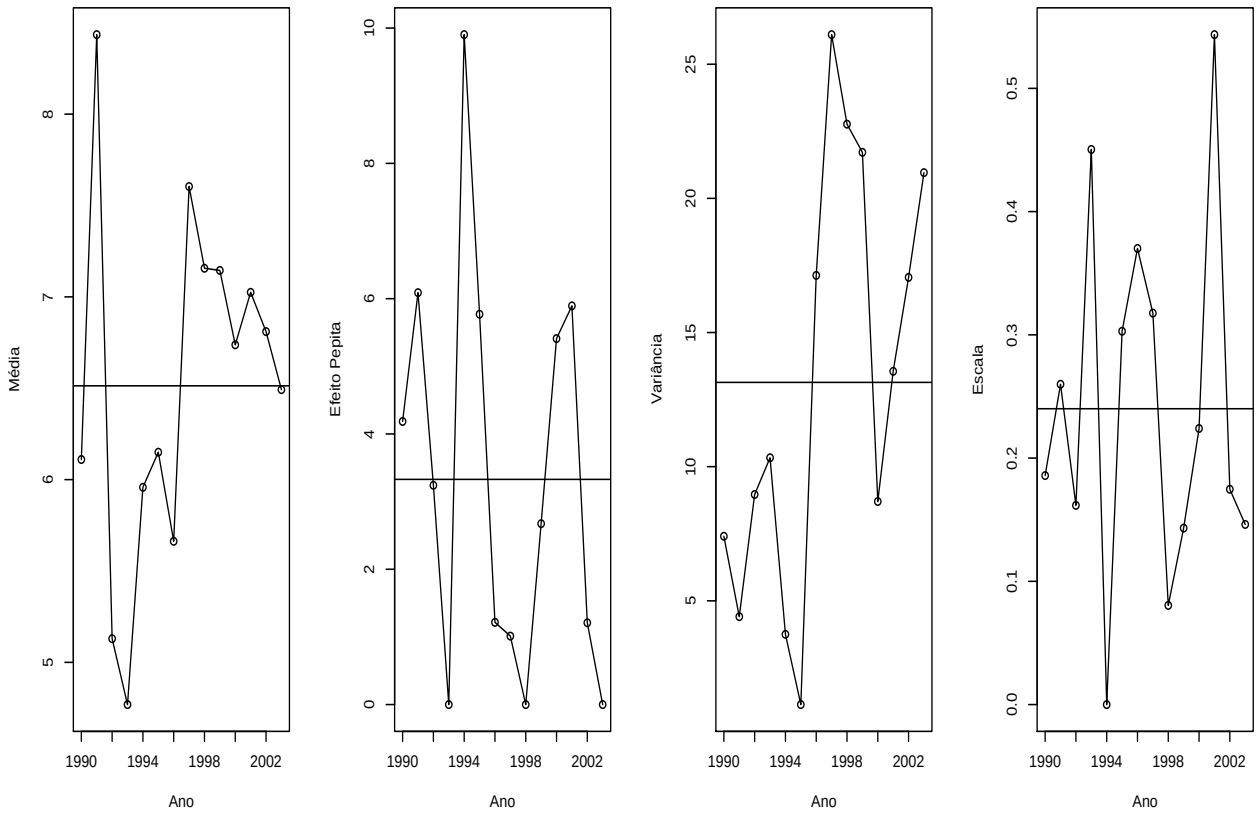

Figura 9 - Estimativas dos parâmetros média, efeito pepita, variância e escala, para cada um dos 14 anos. As linhas horizontais correspondem à média das estimativas

4.1.2 Modelagem da estrutura de covariância espaço-temporal

Para a análise espaço-temporal conjunta será considerada a princípio uma função de covariância com estrutura separável e em seguida serão utilizadas funções de covariância espaço-temporais não separáveis propostas por Gneiting.

Assumir funções de covariância separáveis é a forma mais simples de modelagem da estrutura de covariância de um campo aleatório espaço-temporal. As funções separáveis são escritas na forma de soma ou produto de funções puramente temporais ou puramente espaciais. Considerou-se inicialmente uma função:

$$\boldsymbol{C}(\mathbf{h}, u) = \frac{\sigma^2}{(a|u|^{\delta})} expc||h||^{\gamma} + \frac{\sigma^2_{h=0}}{(a|u|^{\delta})}, \tag{12}$$

que é separável pois pode ser escrita da forma:

$$\boldsymbol{C}(\mathbf{h}, u) = C(\mathbf{h})C(u). \tag{13}$$

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

$$\boldsymbol{C}(\mathbf{h}, u) = \frac{32.26}{(0.10|u|^{0.2}+1)^{0.44}} exp\{-2.5||\mathbf{h}||^{0.83}\} + \frac{6.99}{(0.10|u|^{0.2}+1)^{0.44}}. \tag{14}$$

Nota-se que este modelo é o produto de três modelos básicos como foi descrito em (2.1.2), em que o primeiro corresponde à função de covariância *cauchy generalizada*, que é ajustada ao componente puramente temporal, enquanto que o segundo membro corresponde à função de covariância *stable*, ajustada ao componente puramente espacial. O terceiro membro corresponde ao efeito pepita que depende do tempo e é ajustado pelo produto de uma constante e uma função *cauchy generalizada*.

O modelo de covariância espaço-temporal separável proposto acima é um caso particular da família de funções de covariância espaço-temporais não separável propostas por Gneiting, com o valor do parâmetro β igual a zero.

Como visto anteriormente a família de covariância não separável proposta por Gneiting possui a seguinte forma geral:

$$\boldsymbol{C}(\mathbf{h}, u) = \frac{\sigma^2}{(\psi|u|^2)^{d/2}} \phi\Big(\frac{||\mathbf{h}||^2}{\psi(|u|^2)}\Big), \tag{15}$$

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

$$\boldsymbol{C}(\mathbf{h}, u) = \frac{1}{(0.10|u|^{0.2}+1)^{0.44}} \Big(\frac{32.26}{(0.10|u|^{0.4}+1)^{0.0335}} exp\Big\{ -2.5\Big[\frac{||\mathbf{h}||}{(0.10|u|^{0.4}+1)}\Big]^{0.83}\Big\}\Big) + \frac{6.99}{(0.10|u|^{0.2}+1)^{0.44}} \tag{16}$$

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

Esta função de covariância pode ser dividida em quatro partes como segue:

$$\boldsymbol{C}(\mathbf{h}, u) = C_1(u).C_2(\mathbf{h}, u) + C_{3\ h=0}.C_4(u),$$

em que $C_1 = (0.10|u|^{0.2} + 1)^{-0.44}$ é a função de covariância *cauchy generalizada*, $C_2 = \left(\frac{32.26}{(0.10|u|^{0.4}+1)^{0.0335}} exp\left\{ -2.5\left[\frac{||\mathbf{h}||}{(0.10|u|^{0.4}+1)}\right]^{0.83}\right\}\right)$, uma função não separável proposta por Gneiting e os últimos dois termos $C_3 = C_1$ e $C_4 = 6.99$, correspondem a uma função de covariância *cauchy generalizada*, que depende do tempo, e um efeito pepita puramente espacial. Cada um dos componentes separadamente é uma função de covariância válida e por conseqüência a combinação destes componentes tem como resultado uma função de covariância válida.

Portanto, a distinção entre os dois modelos ajustados anteriormente é dada pelo parâmetro β que determina se existe ou não interação entre os componentes puramente espacial e puramente temporal, ou seja, se a correlação espacial independe do tempo e vice-versa. Então o valor do parâmetro β pode ser visto como o grau de interação espaço-temporal, em que valores grandes, próximos de 1, induz em um forte efeito de interação espaço-temporal, enquanto que valores pequenos, próximo de 0, induz em um efeito fraco de interação espaço-temporal. Isto sugere uma forma de testar a separabilidade da função de covariância, além de quantificar a intensidade da interação espaço-temporal.

No caso do estoque de peixe o valor estimado pelo método de máxima verossimilhança, para o parâmetro β foi 0.0335, sendo este valor muito próximo de zero o que sugere a independência entre as estruturas temporal e espacial.

Na prática podem ocorrer dificuldades nos algorítmos numéricos para identificar este parâmetro. Uma possível forma de contornar uma eventual dificuldade numérica é "isolar"a estimação deste parâmetro. A Figura 10 mostra duas diferentes estratégias dadas pela verossimilhança condicional e verossimilhança perfilhada.

Denotando-se o vetor de parâmetros por (θ, β) em que, θ é o vetor contendo todos os demais parâmetros do modelo e tomando-se uma seqüência de valores para β, tem-se que na primeira obtemos a função de verossimilhança para valores de β fixando os valores de θ nas suas estimativas de verossimilhança, isto é $L(\beta, \theta) = \hat{\theta}$. Para a verossimilhança perfilhada obtemos para cada valor de β a verossimilhança maximizada em relação aos demais parâmetros em θ, isto é $L(\beta, \theta) = \hat{\theta}_\beta$.

A diferença entre os gráficos mostra a não ortogonalidade de β em relação aos demais parâmetros do modelo, e ressalta o fato de que a verossimilhança perfilhada é mais

apropriada para investigar o parâmetro de interesse.

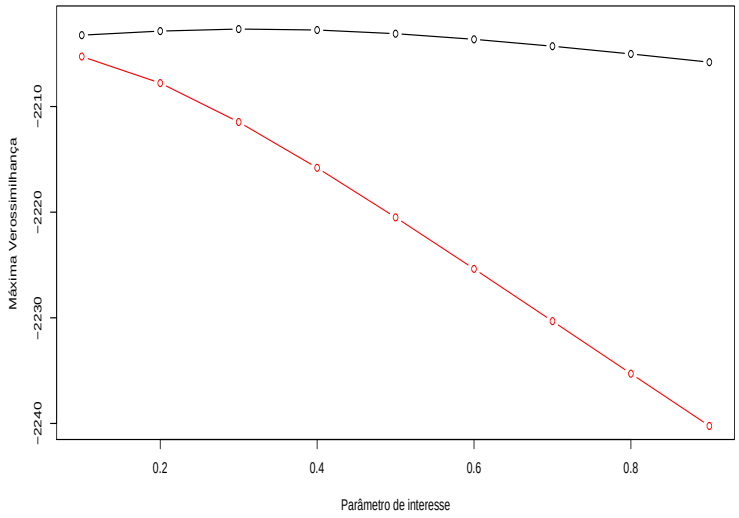

Figura 10 - Máxima verossimilhança condicional (linha vermelha) e perfilhada

O gráfico da verossimilhança condicional está condizente com a estimativa de $\beta = 0.0335$, já que o máximo desta função deva ocorrer quando $\beta = 0$ o que sugere um modelo de covariância espaço-temporal separável. Já para a verossimilhança perfilhada o máximo desta função ocorre quando $\beta = 0.3$, mas a pequena diferença no valor do máximo da verossimilhança leva também à conclusão de que o modelo de covariância espaço-temporal separável é mais indicado. Pelo estudo destas verossimilhanças conclui-se que este parâmetro é muito instável e difícil de ser estimado, e por sua variância não ser conhecida testes estatísticos mais sofisticados não são viáveis.

Afim de comparar o impacto da escolha do valor de β nas predições, a krigagem simples foi realizada para diferentes valores de β. Os valores preditos para o ano de 2004 foram comparados com os verdadeiros valores observados para o mesmo ano.

A Tabela 3 apresenta o erro quadrático médio de predição para os modelos de covariância espaço-temporal propostos por Gneiting com parâmetro $d = 0, 0.25, 0.5, 0.75, 0.95$, além da média e da variância da predição para cada um dos modelos. Pode-se perceber uma

diferença mínima entre os erros de predição, além de uma pequena diferença na média e na variância da predição. Estes resultados sugerem a escolha de um modelo de covariância espaço-temporal separável, já que o ganho no erro de predição é muito pequeno, e o modelo separável possui um parâmetro a menos e é mais facilmente interpretado.

Tabela 3 - 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

A Figura 11 ilustra a pequena diferença entre as predições dos 5 modelos ajustados, sendo a linha vermelha o verdadeiro valor para o estoque de peixe amostrado em 2004 e as demais linhas correspondem aos modelos de covariância.

Os pontos referem-se a uma ordenação das coordenadas ordenadas em relação aos valores da latitude, o que indica que na região da do sul de Portugal as predições foram ainda mais distantes dos valores observados.

As Figuras 12 e 13, ilustram a krigagem para os modelos de covariância espaço-temporais pertencente à família de Gneiting com o parâmetro β com valores 0, 0.25, 0.95. Não é possível verificar grandes diferenças entre os mapas de krigagem.

4.2 Armazenagem de água no solo

A água é um dos mais importantes fatores para o adequado desenvolvimento de qualquer cultura agrícola e o conhecimento da sua dinâmica no solo é indispensável para o aprimoramento de práticas de manejo que visam à otimização da produtividade da cultura (MORETI, 2006).

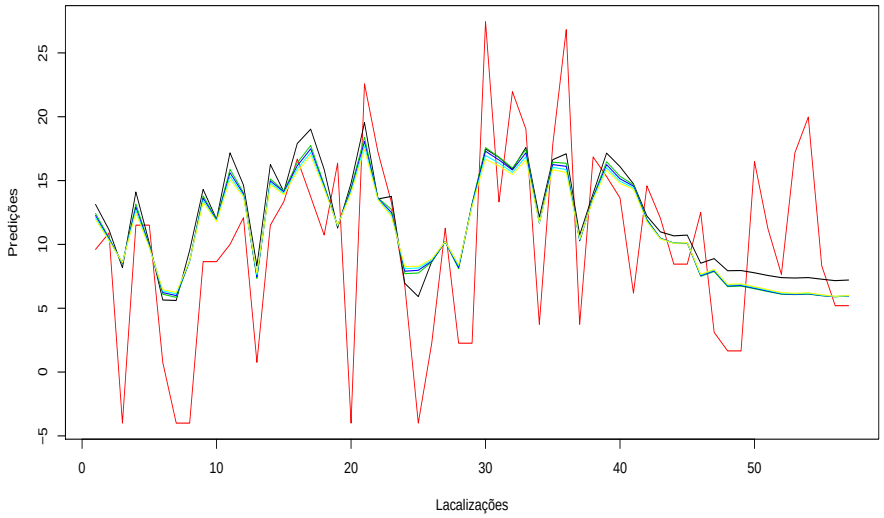

Figura 11 - Distribuição das predições para 2004 e do verdadeiro valor amostrado

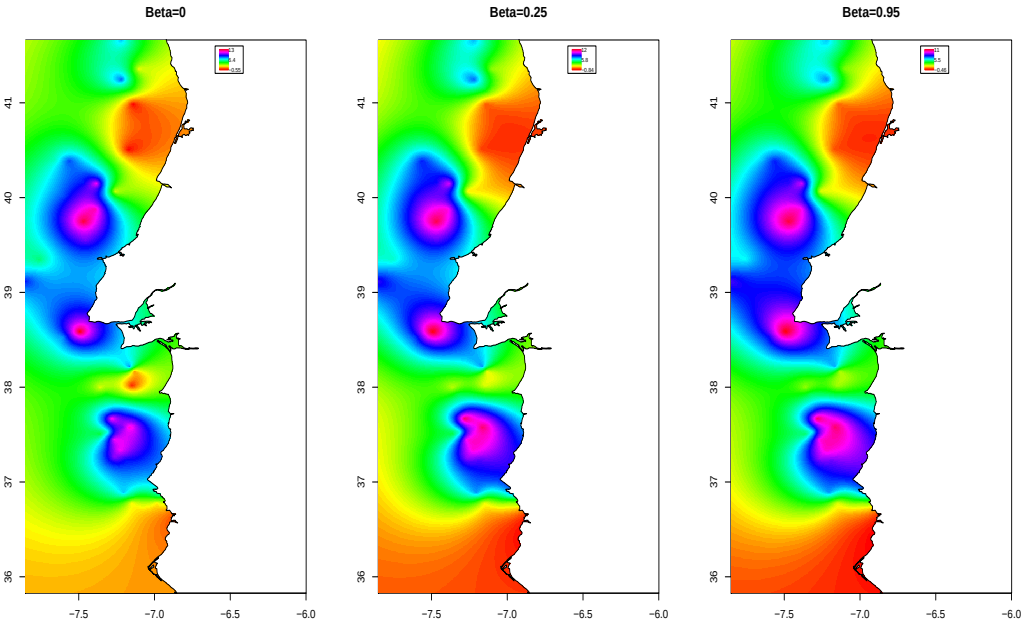

Figura 12 - Krigagem para modelo com $\beta = 0$, $\beta = 0.25$ e $\beta = 0.95$, respectivamente

Moreti (2006) fornece um estudo da capacidade de armazenagem de água em um Latossolo Vermelho Amarelo Argissólico cultivado com citros, em uma parcela experimental na qual foram coletados dados em duas transeções com 20 pontos de observação cada, com espaçamento de 4,0 m entre os pontos amostrados e 7,0 m entre as transeções, conforme ilustrado na Figura 13. Para quantificar a armazenagem de água no solo foi instalado, em cada ponto amostral, um tubo de acesso à sonda de nêutros até a profundidade de 1,20 m, localizado no centro da distância entre duas plantas ao longo da linha.

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

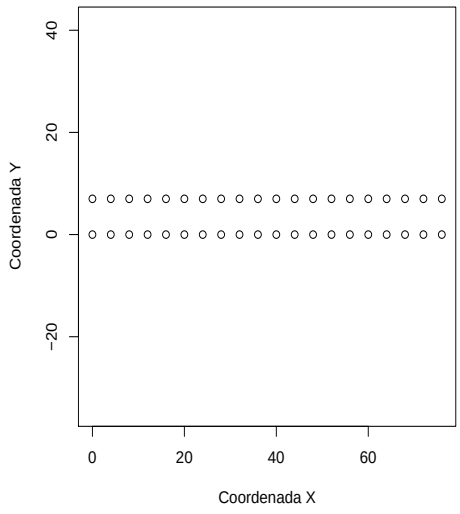

Figura 13 - Parcelas amostradas

Como primeira técnica de análise descritiva a metodologia de Box-Cox foi aplicada e não sugeriu nenhuma transformação na distribuição marginal dos dados amostrados.

Em seguida foram isolados os componentes puramente espacial e puramente temporal. A Figura 14a mostra o comportamento da série nos 25 tempos considerados, em que cada linha representa um ponto amostrado. Pode-se perceber a existência de uma estabilidade e uma baixa variabilidade da armazenagem de água no solo considerando o período de tempo. Já o item b da figura ilustra os 25 variograma ajustados cuja estimação dos parâmetros foi obtido pelo método da máxima verossimilhança, considerando-se uma estrutura de covariância comum dada pelo modelo exponencial. Os ajustes mostram uma dependência espacial de pequena escala e com parâmetros que parecem diferir entre os diferentes tempos.

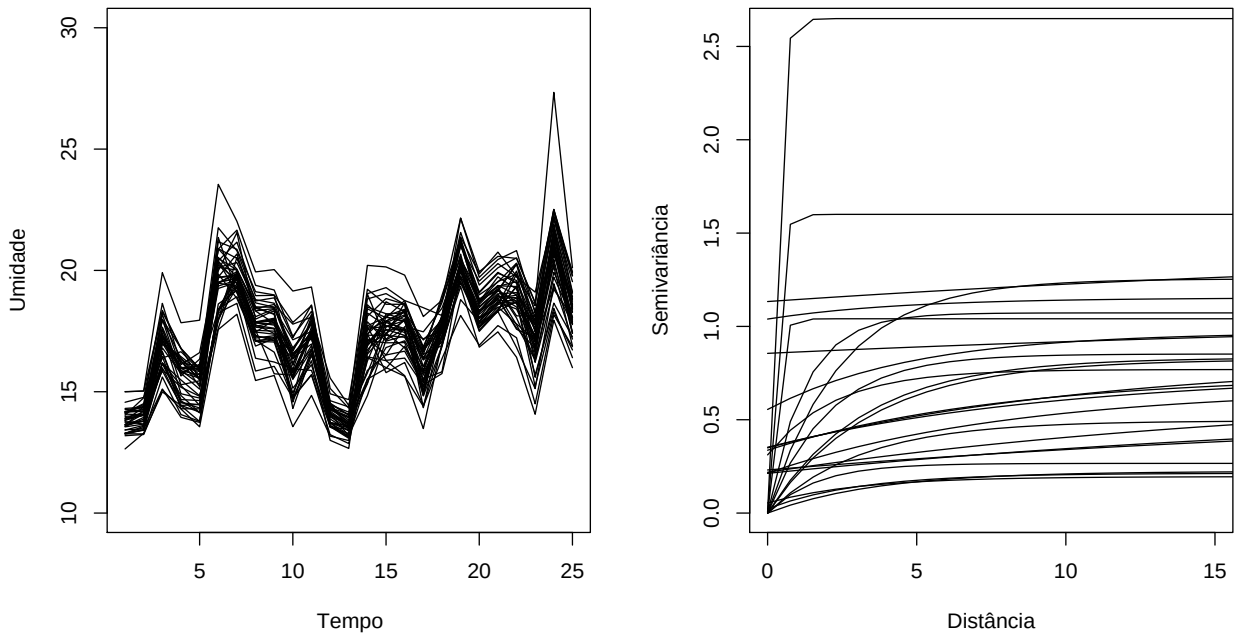

Figura 14 - a e b - Séries temporais para os dados nos diferentes pontos de amostragem e semivariogramas ajustados para cada um dos tempos, respectivamente

Na Figura 15 são apresentados estimativas dos parâmetros do modelo obtido pelo método da máxima verossimilhança para cada tempo individualmente. Nestes gráficos não é possível perceber nenhum tipo de tendência nítida para os parâmetros considerados ao longo do tempo e assim como para os dados de pescada estas séries dos parâmetros de média

e variância do efeito espacial parecem apresentar correlação temporal. Estes efeitos poderiam ser melhor estudados considerando-se dados para série de tempo mais longos.

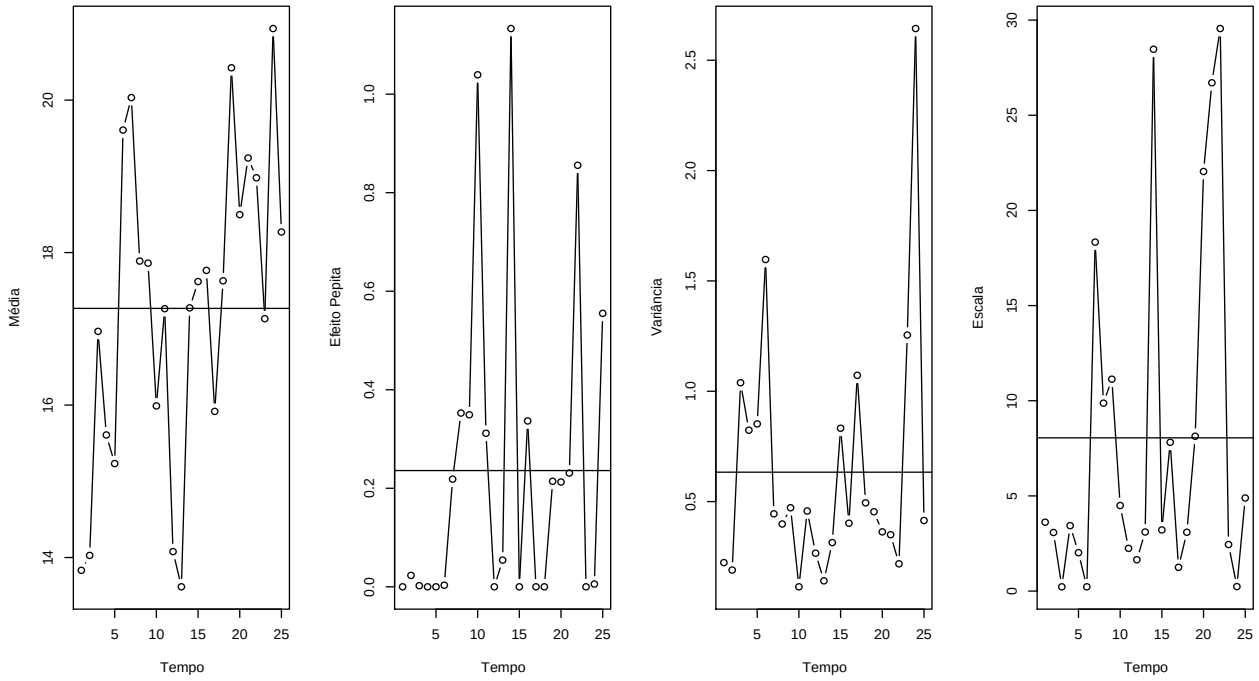

Figura 15 - Estimativas dos parâmetros média, efeito pepita, variância e escala, para cada um dos 25 tempos

4.2.1 Construindo modelos de covariância espaço-temporais

Para a análise espaço-temporal conjunta, foram comparados os mesmos modelos de covariância espaço-temporal utilizados em (4.1.2).

Para os dados referentes à armazenagem de água em um solo com citros, tem-se a seguinte função de covariância espaço-temporal separável ajustada:

$$\boldsymbol{C}(\mathbf{h}, u) = \frac{44.12}{(0.76|u|^{1.65}+1)^{0.08}} exp\{-0.001||\mathbf{h}||^{0.85}\} + \frac{0.25}{(0.76|u|^{1.65}+1)^{0.08}}, \tag{17}$$

em que os parâmetros foram estimados pelo método da máxima verossimilhança.

Já para a função de covariância espaço-temporal não separável, com os parâmetros também estimados pelo método de máxima verossimilhança, tem-se que o modelo ajustado é dado por:

$$\boldsymbol{C}(\mathbf{h},u)=\frac{1}{(0.76|u|^{1.65}+1)^{0.12}}\Big(\frac{44.12}{(0.76|u|^{3.3}+1)^{0.04}}exp\Big\{-0.001\Big[\frac{||\mathbf{h}||}{(0.76|u|^{3.3}+1)^{0.02}}\Big]^{1.65}\Big\}\Big)+\frac{0.25}{(0.76|u|^{1.65}+1)^{0.08}} \quad (18)$$

O valor estimado para o parâmetro β é muito próximo de zero, o que indica que o modelos separável se ajustaria bem aos dados, em outras palavras, não há evidência clara de interação entre o espaço e o tempo. Ressaltando aqui novamente que a estimação deste parâmetro pode ser numericamente instáveis e em uma análise detalhada destes dados, métodos numéricos devem ser cuidadosamente calibrados, bem como é útil se considerar a obtenção da verossimilhança perfilhada. Conforme discutido na análise de dados anterior.

Para verificar o efeito da suposição de separabilidade espaço-temporal foi considerado também o erro médio de predição, 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 17, ilustra estas predições e por este gráfico percebe que o modelo com $\beta = 0$ é o que mais se aproxima do valor real e ainda percebe-se que o aumento do valor de β, causa um diminuição na média e na variância. A Tabela 4 apresenta os erros quadráticos médios, a média e variância da predição, e através dela é possível verificar que o menor valor do erro quadrático médio ocorre quando $\beta = 0$, ou seja, o modelo separável é o mais indicado para o conjunto de dados considerado.

Na Figura 18, são ilustrados os mapas de krigagem simples com parâmetros $\beta = 0, 0.25, 0.95$. As predições foram realizados para o tempo 26, que não foi considerado na estimação. Estas imagens ilustram uma pequena diferença no comportamento das "manchas", mas uma diferença um pouco maior nas escalas. Através das legendas percebe a clara diminuição do valor médio considerado.

Tabela 4 - Erro quadrático médio de predição para funções de covariância espaço-temporais com diferentes valores do parâmetro β

	EQM	Média	Variância
Modelo separável $\beta = 0$	5.561	17.114	0.753
Modelo não separável $\beta = 0.25$	12.674	15.913	0.494
Modelo não separável $\beta = 0.5$	26.060	14.371	0.309
Modelo não separável $\beta = 0.75$	45.661	12.720	0.191
Modelo não separável $\beta = 0.95$	64.436	11.451	0.132

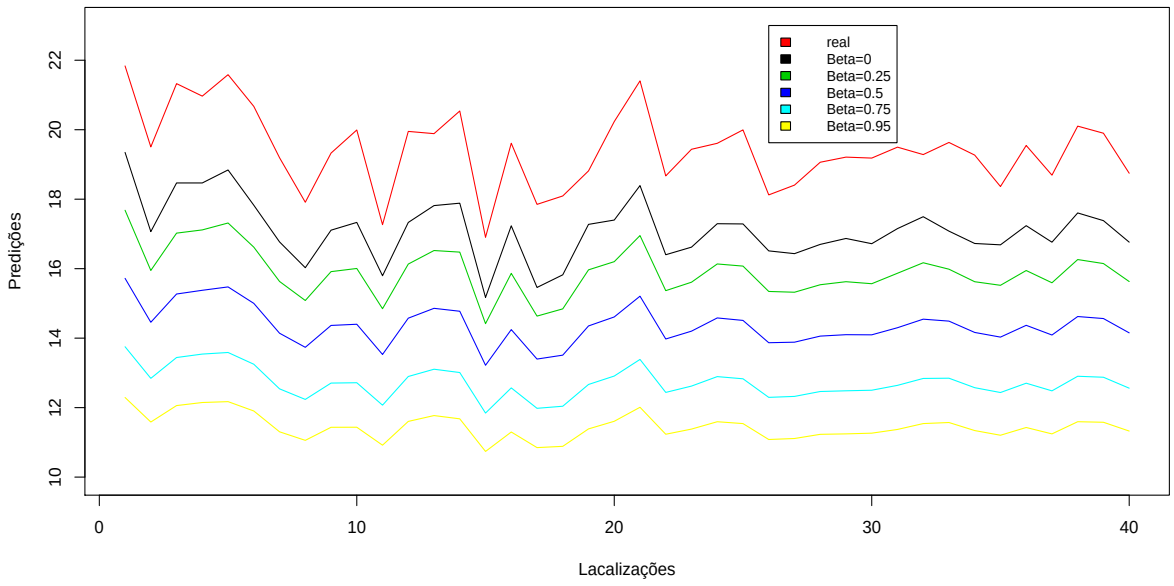

Figura 16 - Distribuição das predições para o tempo 26 e do verdadeiro valor amostrado

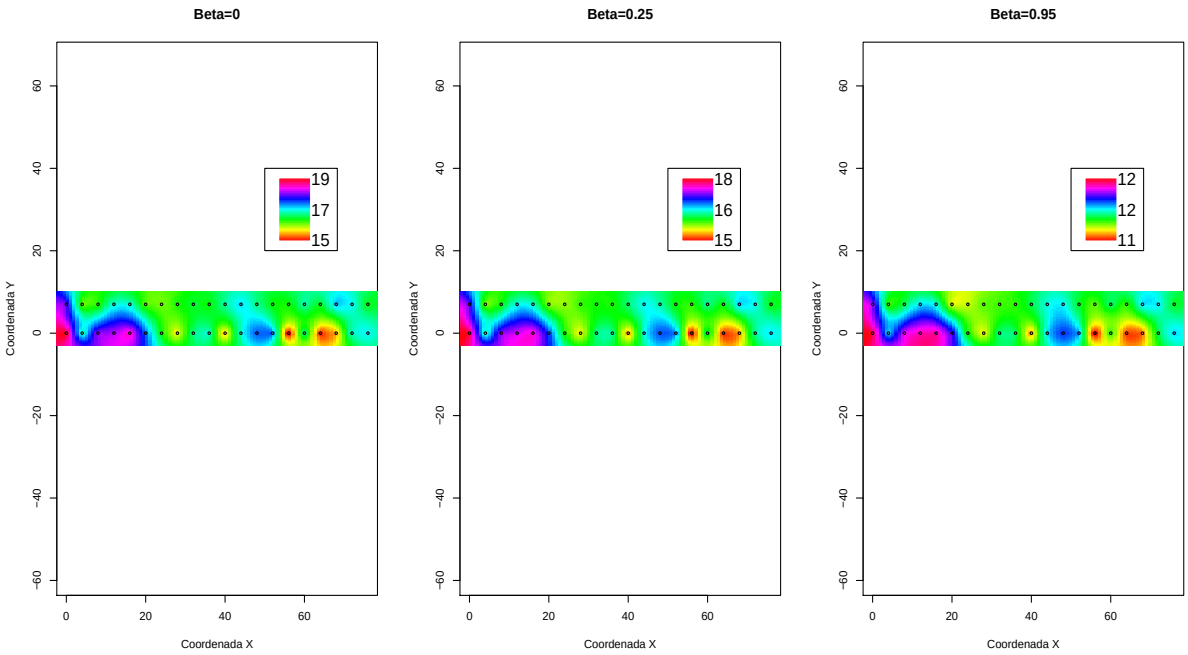

Figura 17 - Krigagem para modelo com $\beta = 0$, $\beta = 0.25$, $\beta = 0.95$, respectivamente

5 CONCLUSÕES

Processos ambientais e geofísicos como concentração de poluentes, campos de precipitação e superfície do vento são caracterizados pela variabilidade contínua no espaço e no tempo. Tais campos espaço-temporais são geralmente tratados como aleatórios, uma vez que as leis que governam seu comportamento não são completamente compreendidas e sua complexidade não segue uma descrição determinística precisa. Modelos estocásticos, por sua vez, são baseados tipicamente em um pequeno número de parâmetros que podem ser modelados e inferidos a partir de observações parciais dos processos.

O objetivo da modelagem de tais processos é informar sobre padrões do comportamento espaço-temporal. Parte-se do princípio de que não existe um modelo que se ajuste perfeitamente aos dados e o que ocorre, é que alguns modelos são mais capazes, que outros, de descrever certas situações. Através da modelagem é possível resumir de forma simples um conjunto de dados complexo, testar hipóteses cientificamente relevantes e fazer predições dos processos em posições não observadas do espaço e/ou tempo.

Fenômenos espaciais e temporais são caracterizados por suas unicidade e não reproducibilidade. Assumindo-se então que é uma amostra parcial simples da realização de um particular mecanismo estocástico. Conseqüentemente, inferências só podem ser feitas a partir da adoção de simplificações e suposições. A princípio a estratégia de análise espaço-temporal apresentada aqui poderiam ser utilizadas simplesmente como uma extensão da análise puramente espacial, pois simplesmente acrescenta uma dimensão referente ao tempo. Entretanto, a diferença física entre espaço e tempo e a ordenação natural dos tempos induzem características particulares e dificuldades adicionais.

O tempo e o espaço são duas noções completamente diferentes, o que dificulta a especificação de funções de covariância espaço-temporais válidas. Além disso estas precisam satisfazer certas condições para serem consideradas válidas. Uma especificação muito comum para a função de covariância espaço-temporal é assegurada por propriedades que dão suporte à família de funções de covariância separáveis. Estas podem ser separadas através da soma e/ou produto de funções de covariância válidas. A grande desvantagem deste método é que

ele não considera na modelagem a interação entre o espaço e o tempo.

Uma alternativa às funções de covariância separáveis são as funções de covariância espaço-temporais não separáveis e neste caso a interação entre o espaço e o tempo é considerada na modelagem. Na literatura existem algumas possibilidades de funções não separável, no presente trabalho foram consideradas a especificação de Cressie e Huang e a família de Gneting.

A especificação de Cressie e Huang se baseia na análise da integral de Fourier o que requer operações utilizando representação Espectral da função de covariância e a inversão de volta do domínio da frequência com o das observações só é obtida em casos particulares. As funções de covariância espaço-temporais propostas por Gneting consideram a combinação de funções completamente monótonas e estritamente crescentes que asseguram funções de covariância válidas.

Nas aplicações deste trabalho foram consideradas funções de covariância espaço-temporais pertencente à família de Gneiting, estacionárias e completamente simétricas, que são uma escolha conveniente, visto que esta especificação não requer qualquer tipo de integrabilidade e possui implementação disponível no pacote RandomFields.

Para os dois conjuntos de dados foi ajustado inicialmente um modelo com estrutura de covariância separável, composta pelo produto entre uma função *cauchy generalizada*, para o componente puramente temporal, e uma função *stable*, para o componente espacial. Ainda para a covariância separável foi adicionado um efeito pepita composto pelo produto entre uma constante e uma função *cauchy generalizada*. Depois de estimados todos os valores para os parâmetros do modelo separável, foi ajustado um modelo de covariância não separável mas com a mesma estrutura de anisotropia estimada para o modelo separável. O modelo não separável considerado é composto pelo produto entre uma função de covariância *cauchy generalizada*, que depende apenas do tempo, e uma função não separável proposta por Gneiting, que depende do tempo e do espaço conjuntamente. Para o efeito pepita foi considerada a mesma estrutura estimada para o modelo separável. É importante ressaltar que o modelo separável proposto é uma caso particular do modelo não separável, quando o parâmetro $\beta = 0$.

Neste trabalho investigou-se as propriedades de simetria e separabilidade através de simulações, e utilizou-se conjuntos de dados reais que levantaram dificuldades práticas na condução das análises. A primeira dificuldade foi a depuração das linhas de comandos do pacote "*RandomFields*", que ainda se encontra em desenvolvimento e poucas são as publicações que se utilização deste software, um outro grande problema relacionado à parte computacional é a demanda por computadores com grande capacidade de memória e processamento, já que o número de observações para análise espaço-temporal cresce muito rapidamente com o aumento de obsevações na coordenada do tempo ou no espaço.

Para os dois conjuntos de dados foram considerados os mesmos pontos em todos os tempos, sendo isto uma obrigatoriedade da implementação disponível no pacote, já que não existe a possibilidade de se especificar coordenadas diferentes para cada tempo. Por conta destas limitações, algumas observações tiveram de ser descartadas da estimação e predição.

Quanto a questão da simetria e separabilidade fica claro, através das simulações, a flexibilidade para se obter diferentes especificações utilizando matriz de anisotropia. Modificações na matriz possibilita simular campos aleatórios espaço-temporais com estrutura de covariância diferentes, ou mesmo estimar os parâmetros supondo determinadas estruturas de covariância. Para estruturas de covariância separáveis é necessário consider duas matrizes de anisotropia diferentes, uma para o componente espacial e outra para o temporal. Já a simetria é considerada, nos valores fora da diagonal principal da matriz de anisotropia. Para as aplicações foram considerada apenas funções de covariância espaço-temporal simétricas, em que todos os valores fora da diagonal principal são iguais a zero. Esta foi uma decisão pragmática considerando-se a redução no número de parâmetros a serem estimados combinada à ausência de informação contextual que justificasse a adoção de modelos assimétricos.

Foram encontradas dificuldades na estimação do parâmetro β da função de covariância proposta por Gneiting, e portanto neste trabalho foram considerados outras formas de assegurar a validade dos valores estimados para este parâmetro. Função de verossimilhança condicional e perfilhada foram consideradas para verificar o comportamento da variabilidade do parâmetro, já que esta informação não é conhecida. Em particular recomenda-se a obtenção da função de verossimilhança perfilhada para uma melhor compreensão da informação contida

na amostra sobre este parâmetro. A curvatura desta função informa sobre a variabilidade do estimador para os dados disponíveis. Em ambas análises foram encontradas estimativa do parâmetro β próximas de zero, apontando para uma fraca interação entre o espaço e o tempo nos dois conjuntos de dados considerados.

Além da exploração da função de verossimilhança, para a garantia da validade do valor estimado para o parâmetro β, foi considerada também a predição de um tempo futuro e comparado com o verdadeiro valor observado para este tempo. O interesse neste caso foi verificar através do erro quadrático médio de predição a validade do ajuste da função de covariância. Na Tabela 2 que é referente aos dados de estoque de peixe, é possível verificar uma pequena diferença no erro quadrático médio para modelos com diferentes valores de β, também é possível verificar pouca diferença entre a média e a variância da predição. Na Tabela 3 que é referente aos dados de armazenagem de água no solo existe uma grande diferença no erro quadrático médio, na média e na variância de predição. Através destas duas tabelas pode-se concluir que a estrutura de covariância separável é a que mais se aproxima dos verdadeiros valores observados, o que pode-se perceber também é a diminuição da média e da variância de predição à medida que o valor do parâmetro β se aproxima de 1.

No desenvolvimento deste trabalho foram feitas ainda várias simulações e análises que serviram para a melhor compreensão do pacote. Em uma destas simulações foi estimado um campo espaço-temporal com um determinado valor do parâmetro β, em seguida foi realizada a estimação dos parâmetros, e verificou-se que a estimativa de verossimilhança não se aproximou de forma satisfatória do valor estimado. Estes resultados são preliminares, e considera-se que estudos desta natureza devem ser efetuados de forma mais detalhada e portanto optou-se por não reportar aqui os resultados deste estudo de simulação. Uma condição que poderia ser considerada é a realização de várias estimações e então seria calculada a média das estimativas. Espera-se que a média dos parâmetros estimados se aproxime do valor dos parâmetros simulados. Com um número suficientemente grande de valores estimados pode-se ter uma melhor idéia da variabilidade de cada parâmetro e com isto a determinação de intervalos de confiança e cobertura poderiam ser obtidos, assim como poderia se avaliar o impacto em predições.

Estas são apenas algumas entre as muitas possibilidades em se aprofundar os estudos nesta área. Neste trabalho foi considerada apenas um resumo das novas metodologias de análise espaço-temporal e com o avanço computacional e o apelo prático desta metodologia fica cada vez mais fácil de se descrever melhor as leis da natureza. Modelos que utilizam a especificação da função de covariância, embora promissores e com fácil interpretação intuitiva, apresentam dificuldades e restrições para uso em dados reais. Estudos complementares são ainda necessário para auxiliar na compreensão das características de identificabilidade e aplicabilidade em situações reais, bem como as exigências e restrições de implementação.

REFERÊNCIAS

ANGULO, J.M.; BUESO, M.C.; ALONSO, F.J. A study on sampling design for optimal prediction of space-time stochastic processes. **Stochastic enviomental research and risk assessment**, Heidelberg, v.14, p.412-427, 2000.

BANERJEE, S.; CARLIN, B.P.; GELFAND, A.E. **Hierarchical modeling and analysis for spatial data**. Boca Raton: Chapman & Hall. 2004. 472 p.

BANERJEE, S. On geodetic distance computational in space modeling, **Biometrics**, Washington, v.61, p.617-625, 2005.

BOCHNER, S. **Lectures on fourier integrals**. Princeton, N.J.: Princeton University Press, 1959. 219 p.

Box, G.E.; Cox, D.R. An analysis of transformations (with discussion), **Journal of Royal Statistical Society**, Oxford, v.26, p.211-256, 1964.

BROWN, P.E.; DIGGLE, P.J.; LORD, M.E.; YOUNG, P.C. Space-time calibration of radar rainfall data. **Applied Statistics**, New York, v.50, p.221-241, 2001.

CHRISTENSEN, O. F.; DIGGLE, P.J.; RIBEIRO Jr., Analysing positive-valued spatial data: the transformed gaussian model, **Geostatistics for environmental applications**, Boston, v.11, p.287-298, 2001.

CRESSIE, N. Statistics for spatial data. New York: Chichester: John Wiley, 1993. 900 p.

CRESSIE, N. An approach to statistical spatial-temporal modeling of meteorological fields: comment. **Journal of American Statistical Association**, Alexandria, v.89, p.379-382, 1994.

CRESSIE, N.; WIKLE, C.K. Space-Time Kalman filter. **Encyclopedia of Environmetrics**, Columbia, v.4, p.2045-2049, 2002.

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

DE CESARE, L.; MYERS, D.E.; POSA, D. Estimating and modeling space-time correlation structures. **Statistics and Probability**, Wisconsin, v.51, p.9-14, 2001.

DE CESARE, L.; MYERS, D.E.; POSA, D. Fortran program for space-time modeling. **Computers and Geosciences**, New York, v.28, p.2005-2512, 2002.

DE IACO, S.; MYERS, D.E.; POSA, D. Space-time analysis using a general product-sum model. **Statistics** & **Probability Letters**, Wisconsin, v.52, p.21-28, 2001.

DE IACO, S.; MYERS, D.E.; POSA, D. The linear coregionalization model and the product-sum space-time variogram. **Mathematical Geology**, Montreal, v.35, p.25-38, 2003.

DE IACO, S.; MYERS, D.E.; POSA, T. Nonseparable space-time covariance models: some parametric families. **Mathematical Geology**, Montreal, v.34, p.23-42, 2002.

DE OLIVEIRA, V.; KEDEM, B.; SHORT, D.A. Bayesian prediction of transformed gaussian random fields, **Journal of the American Statistical Association**, Alexandria, v.92, p.1422-1433, 1997.

DIGGLE, P.J.; RIBEIRO Jr., P.J.; CHRISTENSEN, O.F. **Spatial statistics and computational methods**. New York: Springer, 2003. 202 p.

DIGGLE, P.J.; RIBEIRO Jr., P.J. **Model-Based geoestatistics**, New York: Springer, 2006. 230 p.

ELMATZOGLOU, I. **Spatio-temporal geoestatistical models,with an aplication in fish stock**, 2006. 53 p. Submited for the degree of (Master in statistics) - Lancaster University, Lancaster, 2006.

FERNANDES, M.V.M. **Modelos para processos espaço-temporais inflacionados de zeros**, 2006. 128 p. Dissertção (Mestrado em Estatística) - Intituto de Matemática da Universidade Federal do Rio de Janeiro, 2006.

FUENTES, M; SMITH, R.L. **A new class of stationary spatial models**. North Caroline: Department of Statistics, North Caroline State University, 2001. Technical (Report, 2534).

FUENTES, M. A formal test of nonstationarity of spatial stochastic processes. **Journal of Multivariate Analysis**, California, v.96, p.30-54, 2005.

GNEITING, T. Nonseperable, Stationary Covariance Functions for Space-Time Data. **Journal of the American Statistical Association**, Alexandria, v.97, p.590-600, 2002a.

GNEITING, T. Compactly supported correlation functions. **Journal of the Multivariete Analysis**, California, v.89, p.493-508, 2002b.

GNEITING, T.; SCHLATHER, M. **Stochastic models wich separate fractral dimension and Hurst effect**, 2001. (Technical Report Series, NRCSE-TRS, 69).

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**.Washington: Department of Statistics University of Washington, 2006. 21 p. (Technical Report, 475).

GNEITING, T.; SEVCÍKOVÁ, H.; PERCIVAL, D.B.; SCHLATHER, M. **Fast and exact simulation of large gaussian lattice systems in R^2: exploring the limits**. Washigton: Department of Statistics University of Washington, 2005. 25 p. (Technical Report, 477).

HANDCOCK, M.S. An approach to statistical spatial-temporal modeling of meteorological fields: Rejoinder. **Journal of American Statistical Association**, Alexandria, v.89, p.388-390, 1994.

HANDCOCK, M.S.; WALLIS, J.R. An approach to statistical spatial-temporal modeling of meteorological fields. **Journal of american statistical association**, Alexandria, v.89, p.368-378, 1994.

HIGDON, D. **Quantitative methods for current environmental issues**, Chichester: Wiley, 2002. 185 p.

ICES, Report of International Botton Trawl Survey Working Group, Portugal, 2002. 136 p. (Technical Report Series ICES CM 2002/D:03).

JARDIM, E.; RIBEIRO JUNIOR, P.J. **Geoestatistical assessement of sampling designs for Portuguese botton trawl surveys** (submetido), 2006a.

JARDIM, E.; RIBEIRO JUNIOR, P.J. **Sampling designs for botton trawl surveys: the portuguese autumn survey fields experience** (submetido), 2006b.

JOURNEL, A.G.; HUIJBREGTS, C.J. **Mining geostatistics**. London: Academic Press, 1978. 600 p.

KEPPENNE, C.L.; DETTINGER, M.D.; GHIL, M. An approach to statistical spatial-temporal modeling of meteorological fields: comment. **Journal of the American Statistical Association**, Alexandria, v.89, p.384-387, 1994.

KYRIAKIDIS, P.C.; JOURNEL, A.G. Geostatistical space-time models: a review. **Mathematical Geology**, Montreal, v.31, p.651-684, 1999.

LE, D.N.; ZIDEK, J.V. **Statistical analysis of environmental space-time processes**. New York: Springer, 2006. 327 p.

LOPHAVEN, S. **Design and analysis of enviromental monitoring programs**, 2004. 190 p. Thesis (Ph.D Degree in Engeneering) - Technical University of Denmark, Kongens Lyngby.

MATÉRN, B. **Spatial variation**. Verlag, Berlin: Spinger, 1986. 365 p.

MORETI, D. **Avaliação espaço-temporal de processos do balanço de água em um solo com citros.** 2006. 138 p. Tese (Doutorado) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2006.

SAMPSON, P.D.; GUTTORP, P. Nonparametric estimation of nonstationary spatial covariance structure.**Journal of American Statistical Association**, Alexandria, v.87, p.108-119, 1992.

SCHABENBERGER, O.; GOTWAY, C.A. **Statistical methods for spatial data analysis**, Boca Raton: Chapman & Hall / CRC, 2005. 488 p.

SCHLATHER, M. **An introduction to positive definite functions and a unconditional simulation of random fields**. Lancaster: Lancaster University, 1999. 80 p. (Technical Report ST-99-10).

SCHLATHER, M. Simulation and analysis of ramdom fields. **R News**, Austria, v.1, p.18-20, 2001.

SCHMIDT, A.M.; O'HAGAN, A. Bayesian inference for nonstationary spatial covariance structure via spatial deformations, **Journal of Royal Statistical Society**, Oxford, v.65, p.743-758, 2003.

SCHMIDT, A.M.; SANSÓ, B. Modelagem bayesiana da estrutura de covariância de processos espaciais e espaço-temporais. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 14, 2006. Caxambú, **Minicurso...** São Paulo: Associação Brasileira de Estatística, 2006. 151 p.

STEIN, M.L. **Space-time covariance functions**. Chicago: University of Chicago, 2004. 85 p. (Technical Report, 4).

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

STROUD, J.R.; MÜLLER, P.; SANSÓ, B. Dynamic models for spatiotemporal data. **Journal of Royal Statistical Society**, Oxford, v.63, p.673-689, 2001.

WALLER, L.A.; GOTWAY, C.A. **Applied spatial statistics for public health data**, New York: Wiley, 2004. 520 p.

WIKLE, C.K.; BERLINER, L.M.; CRESSIE, N. Hierarchical Bayesian space-time models. **Environmental and Ecological Statistics**, Yale, v.5, p.117-154, 1998.

R: A Language and Environment for Statistical Computing, R Development Core Team, R Foundation for Statistical Computing, Vienna, Austria, 2006, ISBN 3-900051-07-0, http://www.R-project.org.

geoR : a package for geostatistical analysis. RIBEIRO Jr, P. J. and DIGGLE, P. J., **R-NEWS**, CRAN.R-project.org/doc/Rnews/. v.1, n.2, p.14–18, 2001.

RandomFields: Simulation and Analysis of Random Fields, SCHLATHER, M., **R package version 1.3.29.** www2.hsu-hh.de/schlath/index.html. 2006.

APÊNDICE

Códigos do R

Simulações através do RandomFields de campos aleatórios no $R^1 \times R^1$.

require(RandomFields)

x <− (1:100)/10; T <− c(1,100,1)/10

A1 <− as.matrix(rbind(c(8, 0), c(0, 0)))

A2 <− as.matrix(rbind(c(0, 0), c(0, 4)))

Modelo de covariância separável (soma)

z1 <− GaussRF(x = x, T = T, grid = TRUE, model = list(list (model = ”exp”, var = 1, aniso = A1), ”+”, list(model = ”exp”, var = 1, aniso = A2)))

image(z1, col = gray(seq(1, 0.1, l = 30)), xlab = ”Espaço”, ylab = ”Tempo”, main = ”Soma”)

Modelos de covariância separável (produto)

z2 <− GaussRF(x = x, T = T , grid = TRUE , model = list(list (model = ”exp”, var = 1, aniso = A1), ”*”, list(model = ”exp”, var = 1, aniso = A2)))

image(z2, col = gray(seq(1, 0.1, l = 30)), xlab = ”Espaço”, ylab = ”Tempo”, main = ”Produto”)

Modelo de covariância não separável e completamente simétrico

A3 <− as.matrix(rbind(c(8, 0), c(0, 4)))

z3 <− GaussRF(x = x, T = T, grid = TRUE, model = list(list (model = ”exp”, var = 1, aniso = A3)))

image(z3, col = gray(seq(1, 0.1, l = 30)), xlab = ”Espaço”, ylab = ”Tempo”, main = ”Simétrico”)

Modelo de covariância não separável e não completamente simétrico

A4 <− as.matrix(rbind(c(8, -2), c(-10, 4)))

z4 <− GaussRF(x = x, T = T , grid = TRUE, model = list(list (model = ”exp”, var = 1, aniso = A4)))

image(z4,col = gray(seq(1, 0.1, l = 30)),xlab = ”Espaço”, ylab = ”Tempo”, main = ”Assimétrico”)

Modelo de covariância não separável Gneiting

x < − y < − (0:4) ; T < − c(1,3,1)*1

k < − c(1.2 , phi=1 , 1.5 , 0.95 , psi=1 , dim=2)

mm < − list(list(m="nsst", v=1 , k=k , a = diag(c(2,2,2))))

z < − GaussRF(x=x , y=y , T=T, grid=TRUE , model=mm)

par(mfrow=c(1,3)); rl < − function() if (interactive())

readline("Press return") for (i in 1:dim(z)[3]) image(z[,,i],col = gray(seq(1, 0.1, l = 30))); rl();

Simulações através do RandomFields de campos aleatórios no $R^2 \times R^1$.

require(RandomFields)

x < − y < − c(1,1000,1)/10; T < − c(1,3,1)

Modelo de covariância separável e simétrico

A1=as.matrix(rbind(c(2, 0, 0), c(0, 2, 0), c(0, 0, 0)))

A2=as.matrix(rbind(c(0, 0, 0), c(0, 0, 0), c(0, 0, 10)))

z1 < − GaussRF(x = x, y = y, T = T, grid = TRUE, gridtriple = TRUE, model = list(list(model = "exp", var = 1, aniso = A1),"+", list(model = "exp", var = 1, aniso = A2)))

h = [x, y, t]

C(h,u) = exp (- $||hA1||$) + exp (- $||hA2||$)

C(h,u) = exp (- [2x 2y 0]) + exp (- [0 0 10t])

Modelo de covariância separável e assimétrico NÃO existe pois:

A1 = as.matrix(rbind(c(2, 0, a), c(0, 2, b), c(c, d, 0)))

A2 = as.matrix(rbind(c(0, 0, 0), c(0, 0, 0), c(0, 0, 10)))

C(h,u) = exp (- $||hA1||$) + exp (- $||hA2||$)

C(h,u) = exp (- [2x+ct 2y+dt ax+by]) + exp (- [0 0 10t])

Deixa de existir a separação entre o espaço e o tempo, ou seja o modelo não é mais separável.

Modelo de covariância não separável e simétrico.

N1 = as.matrix(rbind(c(2, 0, 5), c(0, 2, 5), c(5, 5, 0)))

N2 = as.matrix(rbind(c(0, 0, 5), c(0, 0, 5), c(5, 5, 10)))

z2 < − GaussRF(x = x, y = y, T = T, grid = TRUE, gridtriple = TRUE, model = list(list(model = ”exp”, var = 1, aniso = N1),”+”, list(model = ”exp”, var = 1, aniso = N2)), meth = ”direct matrix decomposition”)

par(mfrow=c(1,3));rl < − function() if (interactive())

readline(”Press return”) for (i in 1:dim(z2)[3]) image(z2[,,i],col = gray(seq(1, 0.1, l = 30)),asp=1); rl();

C(h,u) = exp (- $||hN1||$) + exp (- $||hN2||$)

C(h,u) = exp (- [2x+5t 2y+5t 5x+5y]) + exp (- [5t 5t 5x+5y+10t])

outra forma poderia ser

U1=as.matrix(rbind(c(2, 0, 5), c(0, 2, 5), c(5, 5, 10)))

z3 < − GaussRF(x = x, y = y, T = T, grid = TRUE, gridtriple = TRUE, model = list(list(model = ”exp”, var = 1, aniso = U1)), meth = ”direct matrix decomposition”)

par(mfrow=c(1,3));rl < − function() if (interactive())

readline(”Press return”) for (i in 1:dim(z3)[3]) image(z3[,,i],col = gray(seq(1, 0.1, l = 30)),asp=1); rl();

C(h,u) = exp (- $||hU1||$)

C(h,u) = exp (- [2x+5t 2y+5t 5x+5y+10t])

Modelo não separável assimétrico.

NA1 = as.matrix(rbind(c(2, 0, 5), c(0, 2, 3), c(1, 5, 0)))

NA2 = as.matrix(rbind(c(0, 0, 2), c(0, 0, 4), c(1, 3, 10)))

z4 < − GaussRF(x = x, y = y, T = T, grid = TRUE, gridtriple = TRUE, model =

list(list(model = ”exp”, var = 1, aniso = NA1),”+”, list(model = ”exp”, var = 1, aniso = NA2)), meth = ”direct matrix decomposition”)

par(mfrow=c(1,3)); rl < − function() if (interactive())

readline(”Press return”) for (i in 1:dim(z4)[3]) image(z4[„i],col = gray(seq(1, 0.1, l = 30)),asp=1); rl();

C(h,u) = exp (- $||hNA1||$) + exp (- $||hNA2||$)

C(h,u) = exp (- [2x+t 2y+5t 5x+3y]) + exp (- [5t 3t 2x+4y+10t])

outra forma

NU1=as.matrix(rbind(c(2, 0, 3), c(0, 2, 5), c(1, 4, 10)))

z3 < − GaussRF(x = x, y = y, T = T, grid = TRUE, gridtriple = TRUE, model = list(list(model = ”exp”, var = 1, aniso = NU1)) , meth = ”direct matrix decomposition”)

par(mfrow=c(1,3)); rl < − function() if (interactive())

readline(”Press return”) for (i in 1:dim(z3)[3]) image(z3[„i],col = gray(seq(1, 0.1, l = 30)),asp=1); rl();

C(h,u) = exp (- $||hNU1||$)

C(h,u) = exp (- [2x+t 2y+4t 3x+5y+10t])

Estimação dos parâmetros dos modelos de covariância.

Modelos separável

f < − function(param){

param[7] < − param[3]

param[25] < − param[3]

param[29] < − param[3]

param[33] < − param[23]

param[35] < − param[13]

param[36] < − param[14]

```
param[45] < − param[23]
return(param) }

        SEP.M < − list(
        list( model = ”stable”, var = NA, k = list(NA), aniso = list( NA, 0, 0, 0, NaN,0,0,0,0) ) ,”*”,
        list( model = ”gencauchy”, var = 1 ,k = list(NA, NA), aniso = list( 0, 0, 0, 0, 0, 0, 0, 0, NA) ) ,”+”,
        list( model = ”nugget”, var = NA , aniso = list(NaN, 0, 0, 0, NaN, 0, 0, 0, NaN) ),”*”,
        list( model = ”gencauchy”, var = 1 ,k = list(NaN, NaN) , aniso = list( 0, 0, 0, 0, 0, 0, 0, 0, NaN ) ) )

        fitS = fitvario( x = cbind(Hx, Hy), T=TT, data = data, model = SEP.M, cross.me = NULL, table=T, transform=f)

        Modelo não separável
f2 < − function(param){
param[4] < − param[18]
param[12] < − param[8]
param[16] < − param[28]
param[30] < − param[8]
param[34] < − param[8]
param[38] < − param[28]
param[40] < − param[18]
param[41] < − param[19]
param[50] < − param[28]
return(param) }

        k = c( a= 1 , phi=1 , c = NaN , d= NA, psi=1 , dim=2 )
```

NSST = list(

list(model = ”nsst”, var = NA, k = k , aniso = list(NA, 0, 0, 0, NaN, 0, 0, 0, NaN)) ,”*”,

list(model = ”gencauchy”, var = 1, k = c(NA, NA), aniso = list(0, 0, 0, 0, 0, 0, 0, 0, NA)) ,”+”,

list(model = ”nugget”, var = NA, aniso = list(NaN, 0, 0, 0, NaN, 0, 0, 0, NaN)),”*”, list(model = ”gencauchy”, var = 1, k = c(NaN, NaN), aniso = list(0, 0, 0, 0, 0, 0, 0, 0, NaN)))