



Universidade de Brasília  
Instituto de Ciências Exatas  
Departamento de Estatística

**KRIGAGEM E COKRIGAGEM UNIVERSAL:  
INCORPORANDO A HETEROGENEIDADE  
ESPACIAL**

Roberto Lazarte Kaqui

10/0020241

Brasília

**2014**

Roberto Lazarte Kaqui

10/0020241

**KRIGAGEM E COKRIGAGEM UNIVERSAL:  
INCORPORANDO A HETEROGENEIDADE  
ESPACIAL**

Relatório apresentado à disciplina Estágio Supervisionado II do curso de graduação em Estatística, Departamento de Estatística, Instituto de Exatas, Universidade de Brasília, como parte dos requisitos necessários para o grau de Bacharel em Estatística.

Orientador: Prof. Dr. Alan Ricardo da Silva

Brasília

**2014**

# Agradecimentos

Agradeço primeiramente à Universidade de Brasília (UnB), principal espaço responsável pela minha formação profissional. Mais especificamente, agradeço a toda equipe do Departamento de Estatística, que em vários momentos me auxiliou e sanou minhas dúvidas.

Agradeço à ESTAT Consultoria, por todas as experiências que ela me proporcionou durante o período que fiz parte da empresa. A oportunidade de viver o dia a dia na empresa júnior não representou apenas um crescimento profissional, mas também um grande crescimento pessoal que foi muito importante na formação da pessoa que sou hoje.

Agradeço a todos os meus amigos e colegas, pessoas fantásticas que tive o prazer de conhecer e conviver nesses anos de curso e com as quais compartilhei vários momentos. Não citarei nomes para não ser injusto, mas cada amizade que construí tem grande importância e com certeza carregarei para sempre comigo.

Agradeço a todos os/as professores/as que contribuíram direta e indiretamente na minha formação. Sempre serei grato pelo conhecimento transmitido por cada um deles. Em especial, agradeço ao meu orientador, professor Alan, grande amante do R e da estatística Bayesiana. Desde a época do PIBIC sempre estive solícito para me ajudar e aconselhar. Pela dedicação e apoio mostrados na concretização deste trabalho e por tudo que me ensinou, muito obrigado Professor!

Por último e com muito carinho, agradeço imensamente à minha família. Sem o apoio incondicional deles eu não teria conseguido concluir mais esta etapa. A saudade do convívio diário com todos eles foi muito grande nesses últimos quatro anos e meio, amenizada em parte pela presença da minha querida irmã aqui em Brasília. Entretanto, todo esse tempo longe mostrou que o amor que nos une e que sempre está presente em nossa família é imenso e supera qualquer distância que por ventura esteja nos separando. Amo vocês!

# Resumo

A geoestatística e suas técnicas atualmente exercem papel importante no entendimento de fenômenos espaciais, sendo fundamentais em estudos que envolvem processos que ocorrem no espaço. As técnicas de krigagem e cokrigagem destacam-se entre os principais métodos de estimação espacial, de modo que o desenvolvimento de algoritmos capazes de realizar esse tipo de análise de superfície torna-se cada vez mais necessário.

Este trabalho tem como objetivo realizar a implementação e validação de um algoritmo utilizando a linguagem IML do SAS 9.3 que permita a aplicação de técnicas de krigagem e cokrigagem, com foco na aplicação do caso universal como meio de incorporar a heterogeneidade dos dados no modelo de estimação e também como alternativa para obtenção de estimativas com melhor precisão quando comparado com os modelos de krigagem e cokrigagem ordinária.

Os resultados obtidos validaram o algoritmo proposto e mostraram que o modelo de krigagem universal gerou em um dos casos analisados uma melhora na qualidade da estimação em relação à krigagem ordinária. Já para as técnicas de cokrigagem, o modelo universal não apresentou melhora na qualidade da estimação. Também foi constatado que o efeito da adição de uma covariável gerou uma melhora na estimação principalmente quando comparamos os modelos de krigagem e cokrigagem ordinária. Para o caso universal essa melhora não foi tão significativa.

**Palavras-chaves:** krigagem, cokrigagem, variograma, universal.

# Lista de Tabelas

4.1	Algoritmo para ajuste automático do <i>lag</i> e <i>lagd</i> na construção do semivariograma . . . . .	31
5.1	Estatísticas Descritivas do log de zinco . . . . .	37
5.2	Estatísticas Descritivas das Estimativas - Krigagem Ordinária . . . . .	39
5.3	Estatísticas Descritivas das Diferenças entre Estimativas e Erro Padrão	39
5.4	Estatísticas Descritivas das Estimativas - Krigagem Universal . . . . .	42
5.5	Estatísticas Descritivas das Diferenças entre Estimativas e Erro Padrão	42
5.6	Estatísticas Descritivas do log do material orgânico . . . . .	43
5.7	Estatísticas Descritivas do log de cobre . . . . .	50
5.8	Matriz de Correlações . . . . .	57

# Lista de Figuras

2.1	Esquema de amostragem em duas dimensões . . . . .	6
2.2	Semivariograma . . . . .	6
2.3	Covariograma e Variograma . . . . .	11
4.1	Rio Meuse . . . . .	27
5.1	Estimativas de Krigagem Ordinária para <b>ltzn</b> . . . . .	38
5.2	Erro Padrão das Estimativas de Krigagem Ordinária . . . . .	38
5.3	Estimativas de Krigagem Universal para <b>ltzn</b> . . . . .	40
5.4	Erro Padrão das Estimativas de Krigagem Universal . . . . .	41
5.5	Modelos teóricos simulados para ajuste dos semivariogramas . . . . .	44
5.6	Modelos teóricos ajustados para os semivariogramas . . . . .	45
5.7	Níveis de concentração de material orgânico no solo estimados . . . . .	46
5.8	Erro Padrão das concentrações de material orgânico estimadas . . . . .	47
5.9	Diferença entre estimativas e entre erros dos métodos universal e or- dinária para <b>ltom</b> . . . . .	49
5.10	Modelos teóricos simulados para ajuste dos semivariogramas . . . . .	51
5.11	Modelos teóricos ajustados para os semivariogramas . . . . .	52
5.12	Níveis de concentração de cobre no solo estimados . . . . .	53

5.13	Erro Padrão das concentrações de cobre estimadas . . . . .	54
5.14	Diferença entre estimativas e entre erros dos métodos universal e ordinária para <b>ltpb</b> . . . . .	55
5.15	Semivariogramas e Pseudo-Variograma Cruzado dos dados originais ajustados a partir do modelo Seno (Covariável <b>ltzn</b> ) . . . . .	58
5.16	Semivariogramas e Pseudo-Variograma Cruzado dos resíduos ajustados a partir do modelo Seno (Covariável <b>ltzn</b> ) . . . . .	59
5.17	Níveis de concentração de cobre no solo estimados por cokrigagem - (Covariável <b>ltzn</b> ) . . . . .	60
5.18	Erro Padrão das concentrações de cobre estimadas por cokrigagem - (Covariável <b>ltzn</b> ) . . . . .	61
5.19	Diferença entre estimativas e entre erros dos métodos universal e ordinária para <b>ltpb</b> - Covariável <b>ltzn</b> . . . . .	62
5.20	Diferença entre os erros de krigagem e cokrigagem - (Covariável <b>ltzn</b> )	64

# Sumário

<b>Resumo</b>	<b>iii</b>
<b>1 INTRODUÇÃO</b>	<b>1</b>
1.1 Objetivos . . . . .	3
<b>2 TÉCNICAS DE KRIGAGEM</b>	<b>4</b>
2.1 Introdução . . . . .	4
2.2 Variograma . . . . .	5
2.2.1 Variograma para caso Universal . . . . .	8
2.3 Covariograma . . . . .	9
2.4 Krigagem Simples . . . . .	11
2.5 Krigagem Ordinária . . . . .	14
2.6 Krigagem Universal . . . . .	15
<b>3 TÉCNICAS DE COKRIGAGEM</b>	<b>19</b>
3.1 Introdução . . . . .	19
3.2 Variograma Cruzado, Covariograma Cruzado e Pseudo-Variograma Cruzado . . . . .	20
3.3 Cokrigagem Ordinária . . . . .	21

3.4	Cokrigagem Universal . . . . .	22
<b>4</b>	<b>MATERIAL E MÉTODOS</b>	<b>26</b>
4.1	Introdução . . . . .	26
4.2	Rio Meuse . . . . .	26
4.3	Algoritmo IML . . . . .	27
4.3.1	Macro Variogram . . . . .	28
4.3.2	Macro Crossvariogram . . . . .	32
4.3.3	Macro Cokriging . . . . .	34
<b>5</b>	<b>ANÁLISE DOS RESULTADOS</b>	<b>36</b>
5.1	INTRODUÇÃO . . . . .	36
5.2	BALIZAMENTO DO ALGORITMO . . . . .	37
5.2.1	Krigagem Ordinária . . . . .	37
5.2.2	Krigagem Universal . . . . .	40
5.3	ANÁLISE UTILIZANDO OS MÉTODOS DE KRIGAGEM OR- DINÁRIA E KRIGAGEM UNIVERSAL . . . . .	42
5.3.1	Estimação do nível de concentração de Material Orgânico - <b>ltom</b> . . . . .	43
5.3.2	Estimação do nível de concentração de Cobre - <b>ltpb</b> . . . . .	50
5.4	ANÁLISE UTILIZANDO OS MÉTODOS DE COKRIGAGEM OR- DINÁRIA E COKRIGAGEM UNIVERSAL . . . . .	56
5.4.1	Estimação do nível de concentração de Cobre - <b>ltpb</b> utilizando o nível de concentração de zinco - <b>ltzn</b> como covariável . . . . .	57

5.5	COMPARAÇÃO ENTRE OS MÉTODOS ORDINÁRIO E UNIVERSAL PARA KRIGAGEM E COKRIGAGEM . . . . .	63
<b>6</b>	<b>CONCLUSÕES</b>	<b>66</b>
	<b>Referências</b>	<b>69</b>

# Capítulo 1

## INTRODUÇÃO

Hoje em dia cada vez mais se tem interesse em explicar ou descrever fenômenos espaciais, assim como avaliar a interação desses processos que ocorrem no espaço. Vários desses fenômenos podem ser observados e medidos em infinitos lugares, mas na prática, eles são medidos em poucos pontos, principalmente por questões econômicas. Para conhecer o valor dessa variável de interesse em uma localidade específica é necessário estimar essa informação a partir das medições pontuais que se tem acesso (Webster and Oliver, 2007). Nesse sentido, a geoestatística e suas técnicas são muito importantes, pois nos permitem estudar os fenômenos ao longo do espaço geográfico.

A krigagem é uma técnica geoestatística que permite estimar a superfície de uma variável contínua a partir de uma amostra dessa variável, logo essas estimativas podem ser entendidas como combinações lineares dos dados observados (Cressie, 1991). Para melhorar a estimação da variável de interesse, pode-se fazer uso de uma ou mais variáveis densamente amostradas que apresentem correlação com essa variável em estudo. Esse processo é possível através da técnica de cokrigagem.

Segundo Webster and Oliver (2007), a cokrigagem é uma extensão multivariada

da krigagem, que permite adicionar covariáveis na modelagem da variável de interesse. Em alguns casos essas covariáveis são amostradas nos mesmos pontos em que a variável principal e em outros casos em pontos diferentes (Cressie, 1991).

Entre os principais tipos de krigagem e cokrigagem, destacam-se três: Simples, Ordinária e Universal. Ao contrário dos casos Simples e Ordinária, onde a média da variável em estudo é constante, o caso Universal apresenta uma situação de não estacionariedade, já que a média da variável em estudo não é constante e assume uma certa tendência. Desse modo, os modelos de krigagem e cokrigagem universal são os mais apropriados para se trabalhar com dados que apresentam certa heterogeneidade, e a implementação de um algoritmo que permita a aplicação dessas técnicas torna-se muito pertinente visto que grande parte dos dados reais apresentam essa característica não-estacionária.

Dessa forma, este trabalho visa implementar a krigagem e cokrigagem universal na linguagem SAS/IML, gerando uma ferramenta que possibilite a análise de dados que apresentam perfil heterogêneo, incorporando esse fator de não-estacionariedade na estimação da variável em estudo.

## 1.1 Objetivos

O objetivo geral do trabalho é fazer a implementação e validação de um algoritmo utilizando a linguagem IML do SAS 9.3 que permita a aplicação de técnicas de krigagem e cokrigagem, com foco no caso universal.

Os objetivos específicos são:

- comparar os resultados obtidos pelo algoritmo proposto com os resultados das funções do pacote *gstat* do R (Rossiter, 2007) em aplicações do método de krigagem;
- utilizando o algoritmo proposto, fazer a aplicação das técnicas de krigagem e cokrigagem tanto para o método ordinário como para o método universal e comparar o efeito da utilização do modelo com componente de tendência na estimação;
- comparar os resultados obtidos pelos modelos de krigagem e cokrigagem e avaliar o efeito das covariáveis adicionadas na estimação.

# Capítulo 2

## TÉCNICAS DE KRIGAGEM

### 2.1 Introdução

A teoria da Krigagem tem sua origem ligada aos trabalhos de Daniel G. Krige, engenheiro de minas sul-africano que foi um pioneiro no campo da geoestatística. Seus estudos foram fundamentais para que a partir de 1963 na França, o matemático e geólogo Georges Matheron, em parceria com seus colegas da Escola de Minas de Paris, pudesse desenvolver e formalizar a técnica de estimação espacial que ficaria conhecida como Krigagem, termo escolhido em homenagem ao seu precursor (Stein and Corsten, 1991).

Segundo Cressie (1991), a krigagem consiste em fazer inferências em valores não observados de uma variável aleatória a partir de valores observados em localizações espaciais conhecidas. Portanto, essa técnica faz o melhor uso do conhecimento existente através da análise de sua variabilidade no espaço (Webster and Oliver, 2007).

O termo Krigagem abrange um conjunto de métodos, dentre quais pode-se definir quatro tipos principais: a krigagem simples, situação onde a média dos valores é constante e conhecida; a krigagem ordinária, onde a média dos valores é constante

e desconhecida; e a krigagem universal, onde a média dos valores é desconhecida e não constante. O quarto tipo conhecido pode ser considerado uma extensão da krigagem, a cokrigagem, método que adiciona ao modelo de estimação uma covariável altamente correlacionada à variável principal com o objetivo de melhorar a qualidade das estimativas (Cressie, 1991). Neste trabalho trataremos especificamente dos casos de krigagem ordinária e universal.

## 2.2 Variograma

A análise variográfica é uma etapa muito importante no estudo de dados com características espaciais, pois é com ela que se torna possível realizar a modelagem da estrutura de covariância. Nesse sentido, o variograma é uma importante ferramenta que permite avaliar quantitativamente a variabilidade de um determinado atributo regionalizado no espaço (Cressie, 1991).

O cálculo do variograma depende de um esquema de amostragem em duas dimensões. A representação dessa amostragem está na Figura 2.1, onde  $Z(u_1)$  e  $Z(u_1 + h)$  são dois pontos amostrados separados por um vetor de distância  $h$ .

A partir da Figura 2.1, pode-se avaliar o nível de dependência entre dois pontos amostrados através do cálculo do variograma  $2\gamma(h)$ :

$$2\gamma(h) = E[Z(u) - Z(u + h)]^2 = Var[Z(u) - Z(u + h)] \quad (2.1)$$

Como estimador do variograma, tem-se o semivariograma:

$$2\hat{\gamma}(h) = \frac{1}{N(h)} \sum_{N(h)} [Z(u_i) - Z(u_i + h)]^2 \quad (2.2)$$

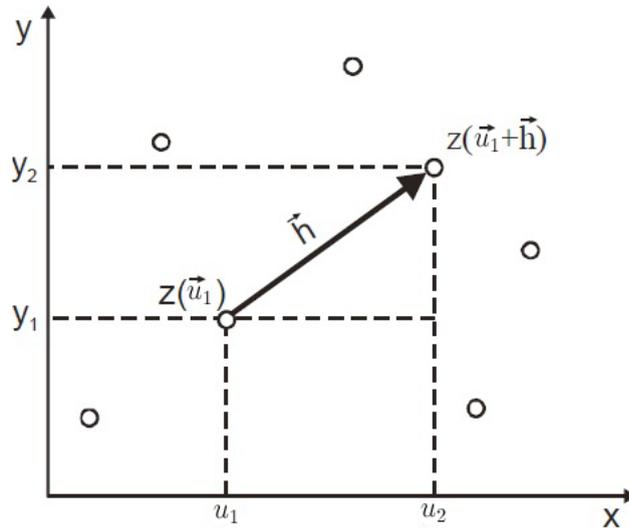


Figura 2.1: Esquema de amostragem em duas dimensões  
 Fonte: Barreto (2011) com modificações

onde  $N(h)$  é o número de pares de pontos amostrados. Usualmente utiliza-se o semivariograma  $\hat{\gamma}(h)$ :

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{N(h)} [Z(u_i) - Z(u_i + h)]^2 \quad (2.3)$$

A representação do semivariograma pode ser vista na Figura 2.2, caracterizada pelos seguintes parâmetros:

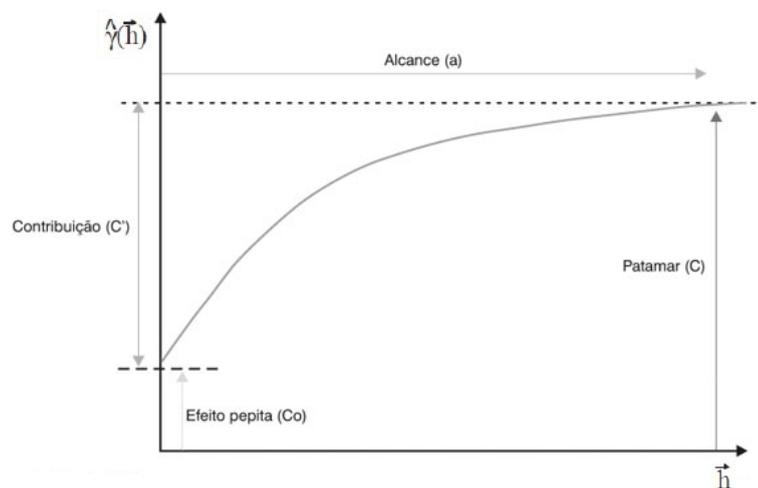


Figura 2.2: Semivariograma

- Patamar (C) : é o valor do semivariograma correspondente ao seu alcance.

A partir deste ponto considerasse que não há mais dependência espacial entre os pontos amostrados, já que a variabilidade da diferença entre os pares observados ( $Var[Z(u) - Z(u + h)]$ ) torna-se constante.

- Alcance (a) : distância em que as amostras apresentam correlação espacial.
- Efeito Pepita ( $C_0$ ): em um semivariograma ideal  $\gamma(0) = 0$ . Entretanto, a medida que  $h$  tende a 0,  $\gamma(h)$  tende a um valor positivo conhecido como Efeito Pepita, o que revela a descontinuidade do semivariograma para distâncias menores do que a distância  $h$  entre os pontos da amostra.
- Contribuição ( $C'$ ): é a diferença entre o patamar ( $C$ ) e o Efeito Pepita ( $C_0$ ).

A representação gráfica do semivariograma deixa claro a interpretação básica dessa função. Quanto mais próximos estiverem dois pontos, maior será a semelhança entre eles, portanto menor a semivariância. Logo, quanto mais distantes estiverem dois pontos amostrados, menor será a semelhança e, conseqüentemente, maior será a variabilidade entre eles.

Para a definição do semivariograma, deve ser determinado um *lag* de distância para que haja a construção das classes de distância. O valor para o *lag* pode ser obtido através de histogramas que permitem avaliar se a distribuição dos pontos em cada classe está adequada (Cressie, 1991).

Outra análise que deve ser levada em conta é o efeito que a estrutura espacial sofre quando as medições ocorrem em diferentes direções. Se o fenômeno em estudo for isotrópico, a determinação do semivariograma depende apenas da distância  $h$

entre as observações. Na situação em que além da distância, a direção entre as amostras também é levada em consideração no estudo, tem-se o caso de anisotropia dos dados.

Uma vez definido o semivariograma experimental, o seguinte passo na representação da variabilidade do fenômeno é o ajuste de um modelo teórico que melhor caracterize a estrutura de variância apresentada pelos dados (Cressie, 1991).

Entre os modelos teóricos utilizados, destacam-se três básicos: esférico, gaussiano e o seno.

Modelo Esférico:

$$Sph(|h|) = \begin{cases} 0 & , |h| = 0 \\ 1, 5\left(\frac{|h|}{a}\right) - 0, 5\left(\frac{|h|}{a}\right)^3 & , 0 < |h| \leq a \\ 1 & , |h| > a \end{cases} \quad (2.4)$$

Modelo Gaussiano:

$$Gau(|h|) = \begin{cases} 0 & , |h| = 0 \\ 1 - exp\left(\frac{|h|}{a}\right)^2 & , |h| \neq 0 \end{cases} \quad (2.5)$$

Modelo Seno:

$$Sin(|h|) = \begin{cases} 0 & , |h| = 0 \\ 1 - exp\left(\frac{sen(\pi|h|/a)}{\pi|h|/a}\right)^2 & , |h| > 0 \end{cases} \quad (2.6)$$

### 2.2.1 Variograma para caso Universal

Em estudos com aplicação dos métodos de krigagem e cokrigagem universal, o variograma a ser utilizado nos cálculos da estimação deve ser referente aos resíduos dos dados (Webster and Oliver, 2007), já que nesses casos o variograma deve refletir a

estrutura de variabilidade dos dados sem o efeito da tendência, que será incorporado posteriormente na estimação.

Segundo Kitadinis (1993), a maneira mais simples e a mais eficiente computacionalmente para retirar o efeito da tendência e obter os resíduos dos dados é aplicar o método dos Mínimos Quadrados ordinários. Usualmente pode-se representar uma variável regionalizada no espaço  $Z(\mathbf{u})$  com efeito de tendência segundo:

$$Z(\mathbf{u}) = \beta_0 + \beta_1 x + \beta_2 y + \varepsilon(\mathbf{u}) \quad (2.7)$$

onde  $x$  e  $y$  são as coordenadas dos pontos, os  $\beta$ 's são os coeficientes de tendência desconhecidos e  $\varepsilon(\mathbf{u})$  o resíduo dos dados.

Trabalhando com a notação matricial, os resíduos dos dados podem ser obtidos através de:

$$\mathbf{z}_r = \mathbf{z} - \mathbf{X}\hat{\boldsymbol{\beta}} = \left( \mathbf{I} - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \right) \mathbf{z} \quad (2.8)$$

onde

$$\mathbf{z} = \begin{bmatrix} z(u_1) \\ z(u_2) \\ \dots \\ z(u_n) \end{bmatrix}; \mathbf{X} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \dots & \dots & \dots \\ 1 & x_n & y_n \end{bmatrix}$$

Toda a formulação utilizada para caracterizar a variável regionalizada no espaço  $Z(\mathbf{u})$  com a presença de tendência será explicada de maneira detalhada mais a frente no tópico referente a krigagem universal.

## 2.3 Covariograma

Dentro do estudo de fenômenos espaciais, variáveis regionalizadas tendem a estar relacionadas. O covariograma é um importante método que permite avaliar a estru-

tura de dependência espacial, uma vez que determina o nível de relação entre pares de observações de um processo estacionário (Cressie, 1991). Seja  $Z(\cdot)$  um processo estacionário qualquer, então a covariância entre dois pontos  $Z(s_1)$  e  $Z(s_2)$  é definida como:

$$Cov[Z(s_1), Z(s_2)] = C(s_1 - s_2) \quad (2.9)$$

A função  $C(\cdot)$  é chamada de covariograma ou função de covariância estacionária (Cressie, 1991).

Pode-se relacionar o covariograma ao variograma. Se  $\|s_1 - s_2\| = h$  então por (2.9):

$$Cov[Z(s_1), Z(s_2)] = C(h) \quad (2.10)$$

Note que  $Var(Z(s_1)) = Cov[Z(s_1), Z(s_1)]$ , então:

$$Var(Z(s_1)) = Var(Z(s_2)) = C(0) \quad (2.11)$$

Partindo da relação:

$$Var(Z(s_1) - Z(s_2)) = Var(Z(s_1)) + Var(Z(s_2)) - 2Cov[Z(s_1), Z(s_2)] \quad (2.12)$$

Temos para  $\|s_1 - s_2\| = h$  de (2.11) e da definição do variograma:

$$2\lambda(h) = 2(C(0) - C(h)) \quad (2.13)$$

Percebe-se que quando  $h \rightarrow \infty$ , então  $2\lambda(h) \rightarrow 2C(0)$ , de maneira que  $C(0)$  é o patamar do variograma.

Pode-se definir o estimador do covariograma por:

$$\hat{C}(h) = \frac{1}{|N(h)|} \sum_{N(h)} (Z(s_i) - \bar{Z})(Z(s_j) - \bar{Z}) \quad (2.14)$$

onde:

$$\bar{Z} = \sum_{i=1}^n \frac{Z(s_i)}{n} \quad (2.15)$$

é o estimador da média e  $N(h)$  é o número de pares observados.

A Figura 2.3 faz uma comparação gráfica entre o variograma e o covariograma. Ao contrário do variograma, o covariograma tende a aumentar à medida que a distância entre dois pontos for menor, já que quanto mais próximos estiverem duas medições, maior a correlação entre elas.

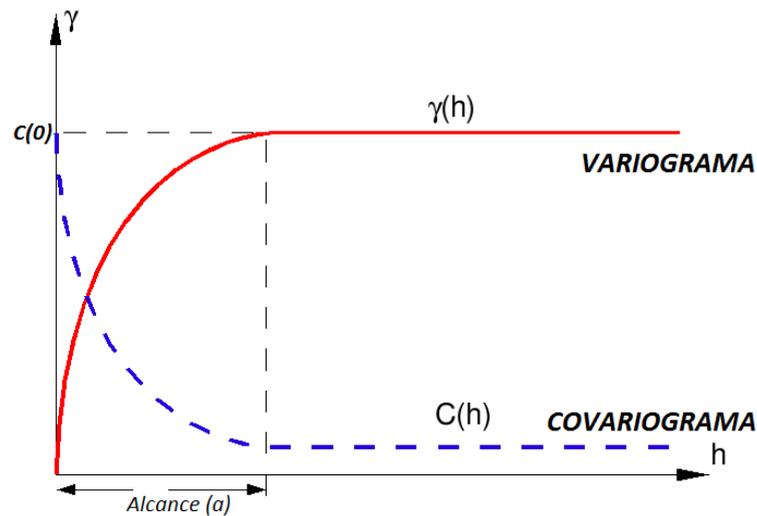


Figura 2.3: Covariograma e Variograma

## 2.4 Krigagem Simples

Analisando uma área que possui uma determinada característica de interesse  $Z$ , em  $n$  pontos amostrados com suas respectivas coordenadas representadas pelo vetor  $\mathbf{u}$ , obtém-se o conjunto de valores:

$$S = \{Z(\mathbf{u}_i) : i = 1, 2, \dots, n\}$$

onde  $\mathbf{u}_i$  caracteriza uma localidade representada pelas coordenadas  $(x_i, y_i)$ . Para

estimar o valor dessa característica de interesse  $Z$  em um determinado ponto  $u_0$ ,  $Z(u_0)$ , utiliza-se a fórmula geral dada por:

$$Z^*(u_0) = \lambda_0 + \sum_{i=1}^n \lambda_i Z(u_i) \quad (2.16)$$

onde  $Z^*(u_0)$  é o valor estimado para a característica de interesse,  $\lambda_0$  é uma constante que depende do ponto  $u_0$  e  $\lambda_i$  é o peso a ser atribuído ao valor  $Z(u_i)$ .

No estudo de krigagem simples, considera-se que a média  $m(u)$  dos valores amostrados é constante e conhecida. Se o objetivo seja estimar um valor da característica  $Z$  no ponto  $u_0$ , utiliza-se o estimador dado por (2.16). Neste caso, um estimador não-viesado é obtido quando os valores de  $\lambda_0$  e  $\lambda_i$  minimizam o erro quadrático médio (EQM) dado por:

$$E(Z^*(u_0) - Z(u_0))^2 = Var(Z^*(u_0) - Z(u_0)) + [E(Z^*(u_0) - Z(u_0))]^2 \quad (2.17)$$

Na expressão (2.17), apenas o viés  $E(Z^*(u_0) - Z(u_0))$  depende de  $\lambda_0$ , portanto minimizar o EQM corresponde a anular o viés:

$$E(Z^*(u_0) - Z(u_0)) = 0 \quad (2.18)$$

Assim, obtém-se:

$$\begin{aligned} E(Z^*(u_0)) &= Z(u_0) \\ E\left(\lambda_0 + \sum_{i=1}^n \lambda_i Z(u_i)\right) &= E(Z_0) \\ \lambda_0 + \sum_{i=1}^n \lambda_i m(u_i) &= m(u_0) \\ \lambda_0 &= m(u_0) - \sum_{i=1}^n \lambda_i m(u_i) \end{aligned} \quad (2.19)$$

Portanto, a partir de (2.19), o estimador  $Z^*(u_0)$  torna-se:

$$Z^*(u_0) = m(u_0) + \sum_{i=1}^n \lambda_i [Z(u_i) - m(u_i)] \quad (2.20)$$

Obtendo o EQM minimizado, tem-se:

$$E (Z^*(u_0) - Z(u_0))^2 = Var (Z^*(u_0) - Z(u_0)),$$

que expresso em termos de covariância  $\sigma(x, y)$  de  $Z(u)$ , torna-se:

$$E (Z^*(u_0) - Z(u_0))^2 = \sum_i \sum_j \lambda_i \lambda_j \sigma_{ij} - 2 \sum_j \lambda_j \sigma_{i0} + \sigma_{00} \quad (2.21)$$

Desse modo, para obter a variância mínima é preciso minimizar a expressão (2.21). Para realizar essa minimização, deriva-se parcialmente a expressão e se iguala o resultado a zero.

$$\frac{\partial}{\partial \lambda_i} E (Z_0^* - Z_0)^2 = 2 \sum_j \lambda_j \sigma_{ij} - 2 \sigma_{i0} = 0$$

Com isso, os pesos  $\lambda_i$  serão as soluções do sistema de equações, que também é conhecido como sistema de krigagens simples:

$$\sum_i \lambda_j \sigma_{ij} = \sigma_{i0}, i = 1, 2, \dots, n \quad (2.22)$$

Para obter a variância da krigagem simples, substitui-se (2.22) em (2.21):

$$E (Z_0^* - Z_0)^2 = \sigma_{00} - \sum_i \lambda_i \sigma_{i0}$$

e usando a igualdade  $Var (Z_0^* - Z_0) = E (Z_0^* - Z_0)^2$ , conclui-se:

$$\sigma_{ks}^2 = \sigma_{00} - \sum_i \lambda_i \sigma_{i0} \quad (2.23)$$

## 2.5 Krigagem Ordinária

A krigagem ordinária pode ser considerada uma extensão natural da krigagem simples (Bailey and Gatrell, 1995). Esse método assume que a média dos valores dos pontos é desconhecida, porém constante, ou seja, trata-se de um processo estacionário.

Pela suposição do método ordinário, para que a média seja constante deve-se fazer  $\lambda_0 = 0$  e  $\sum_{i=1}^n \lambda_i = 1$ , conseqüentemente o estimador será:

$$Z^*(u_0) = \sum_{i=1}^n \lambda_i Z(u_i) \quad (2.24)$$

Para obter o melhor estimador deve-se minimizar a variância do erro ( $Var(Z_0^* - Z_0)$ ) sujeito a condição  $\sum_{i=1}^n \lambda_i = 1$ . Partindo do variograma definido em (2.1), deve-se minimizar:

$$E\left(\sum_{i=1}^n \lambda_i Z(u_i) - Z(u_0)\right)^2 - 2\alpha\left(\sum_{i=1}^n \lambda_i - 1\right) \quad (2.25)$$

onde  $\alpha$  é o multiplicador de Lagrange que garante  $\sum_{i=1}^n \lambda_i = 1$ . Essa condição implica:

$$\left(\sum_{i=1}^n \lambda_i Z(u_i) - Z(u_0)\right)^2 = -\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j [Z(u_i) - Z(u_j)]^2 + 2\sum_{i=1}^n \lambda_i [Z(u_0) - Z(u_i)]^2 \quad (2.26)$$

Dessa forma, substituindo (2.26) em (2.25), obtém-se:

$$-\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(u_i - u_j) + 2\sum_{i=1}^n \lambda_i \gamma(u_0 - u_i) - 2\alpha\left(\sum_{i=1}^n \lambda_i - 1\right) \quad (2.27)$$

Derivando em relação a  $\lambda_1, \dots, \lambda_n$  e igualando a 0, obtém-se a equação:

$$-\sum_{j=1}^n \lambda_j \gamma(u_i - u_j) + \gamma(u_0 - u_i) - \alpha = 0 \quad (2.28)$$

Portanto os pesos  $\lambda_i$  podem ser calculados através do sistema de Equações (2.29), conhecido como sistema de krigagem ordinária:

$$\begin{cases} -\sum_{i=1}^n \lambda_i \gamma(u_i - u_j) + \gamma(u_0 - u_i) - \alpha = 0 \\ \sum_{i=1}^n \lambda_i = 1 \end{cases} \quad (2.29)$$

onde  $i = 1, 2, \dots, n$ ;  $\gamma(u_i - u_j)$  e  $\gamma(u_0 - u_i)$  são respectivamente a semivariância entre os pontos  $u_i$  e  $u_j$  e entre os pontos  $u_0$  e  $u_i$  e  $\alpha$  é o multiplicador de Lagrange. A variância da krigagem ordinária é dada por:

$$\sigma_{ko}^2 = \lambda(0) - \sum_{i=1}^n \lambda_i \gamma(u_0 - u_i) - \alpha \quad (2.30)$$

## 2.6 Krigagem Universal

A krigagem universal é utilizada quando a média dos valores amostrados é desconhecida e não é constante, portanto pode ser entendida como uma extensão da krigagem ordinária para casos onde o processo estudado apresenta característica não-estacionária. Assim como no caso da krigagem ordinária, as estimativas produzidas na krigagem universal são resultado de uma combinação linear ponderada dos dados amostrados. A tendência apresentada pelos dados é calculada simultaneamente e estimada implicitamente (Bailey and Gatrell, 1995).

O modelo básico de krigagem universal é dado por:

$$Z(u) = m(u) + \varepsilon(u) \quad (2.31)$$

onde  $Z(u)$  é o valor da variável em estudo,  $m(u)$  é a tendência e  $\varepsilon(u)$  é o resíduo com média zero e variograma  $\gamma(u)$ . Nessas circunstâncias,  $E[Z(u) - Z(u+h)]^2$  não é igual a  $E[\varepsilon(u) - \varepsilon(u+h)]^2$ , portanto as semivariâncias calculada pela equação (2.2)

serão estimativas viesadas de  $\gamma(u)$ . Para estimar  $\gamma(u)$  sem viés é necessário separar os resíduos  $\varepsilon(u)$  da tendência  $m(u)$ .

Segundo Chiles and Delfiner (1999), do ponto de vista matemático a tendência é bem definida como  $E[Z(u)] = m(u)$ . Entretanto, a tendência  $m(u)$  não é observável exceto no caso em que há repetições, permitindo assim o cálculo do valor esperado a partir de várias observações do processo espacial.

Quando o fenômeno é único e não é possível a realização de repetições, a tendência pode ser modelada através de uma função matemática que representa pequenas variações locais sobre uma clara tendência geral. Desse modo, restaurando a estacionariedade dos dados, é possível calcular o variograma dos resíduos, que difere pouco do variograma verdadeiro.

A tendência  $m(u)$  geralmente pode ser expressa como uma forma simples e funcional:

$$m(u) = \sum_{k=0}^K \beta_k f_k(u) \quad (2.32)$$

onde  $\beta_k$  são coeficientes desconhecidos e  $f_k(u)$  são funções conhecidas de  $u$ . Combinando as Equações (2.31) e (2.32), pode-se representar o modelo de krigagem universal como:

$$Z(u) = m(u) + \varepsilon(u) = \sum_{k=0}^K \beta_k f_k(u) + \varepsilon(u) \quad (2.33)$$

Usualmente uma tendência espacial pode ser modelada como um polinômio de pequena ordem em função das coordenadas geográficas  $(x, y)$ . Desse modo, o caso mais simples de tendência linear pode ser representado pela equação:

$$Z(\mathbf{u}) = \beta_0 + \beta_1 x + \beta_2 y + \varepsilon(\mathbf{u}) \quad (2.34)$$

considerando  $f_0(u) = 1$  e  $\mathbf{u}$  o vetor de coordenadas geográficas.

O estimador para a krigagem universal proposto por Matheron é dado por:

$$Z^*(u) = \sum_{i=1}^n \lambda_i f_k(u_i) \quad (2.35)$$

O valor esperado é igual a:

$$E[Z^*(u)] = \sum_{k=0}^K \sum_{i=0}^n \beta_k \lambda_i f_k(u_i) \quad (2.36)$$

de modo que o estimador será não-viesado caso:

$$\sum_{i=0}^n \lambda_i f_k(u_i) = f_k(u_0) \quad (2.37)$$

para todo  $k = 0, 1, \dots, K$ .

Assim como na krigagem ordinária, Matheron propôs para o cálculo dos pesos  $\lambda_i$  o sistema de krigagem universal dado por:

$$\begin{cases} \sum_{i=1}^n \lambda_i \gamma(u_i, u_j) + \alpha_0 + \sum_{k=0}^K \alpha_k f_k(u_j) = \gamma(u_0, u_j) \\ \sum_{i=1}^n \lambda_i = 1 \\ \sum_{i=1}^n \lambda_i f_k(u_i) = f_k(u_0) \end{cases} \quad (2.38)$$

onde  $\gamma(u_i, u_j)$  e  $\gamma(u_0, u_j)$  são respectivamente a semivariância dos resíduos entre os pontos  $u_i$  e  $u_j$  e a semivariância entre os pontos  $u_0$  e  $u_j$  e  $\alpha_0$  e  $\alpha_k$  são os multiplicadores Lagrange.

O sistema de krigagem universal, assim como o sistema de krigagem ordinária, é um conjunto de equações lineares que pode ser representada na forma matricial, o que facilita a sua implementação computacional.

Seja  $\Gamma_u$  a matriz de semivariância da variável  $u$  de interesse, amostrada em  $n$  locais respectivamente. Essa matriz possui dimensão  $n \times n$ :

$$\Gamma_{\mathbf{u}} = \begin{bmatrix} \gamma_u(x_1, x_1) & \gamma_u(x_1, x_2) & \cdots & \gamma_u(x_1, x_n) \\ \gamma_u(x_2, x_1) & \gamma_u(x_2, x_2) & \cdots & \gamma_u(x_2, x_n) \\ \vdots & \vdots & \cdots & \vdots \\ \gamma_u(x_n, x_1) & \gamma_u(x_n, x_2) & \cdots & \gamma_u(x_n, x_n) \end{bmatrix}$$

Denota-se por  $b_u$  o vetor de semivariância entre os pontos amostrados e o ponto  $u_0$  que deseja-se estimar:

$$\mathbf{b}_u = \begin{bmatrix} \gamma_u(x_1, x_u) \\ \gamma_u(x_2, x_u) \\ \vdots \\ \gamma_u(x_n, x_u) \end{bmatrix}$$

Portanto, a equação matricial possui a forma:

$$\begin{bmatrix} & & & & 1 & x_1 & y_1 \\ & & & & 1 & x_2 & y_2 \\ & & \Gamma_u & & \vdots & \vdots & \vdots \\ & & & & 1 & x_n & y_n \\ 1 & 1 & \cdots & 1 & 0 & 0 & 0 \\ x_1 & x_2 & \cdots & x_n & 0 & 0 & 0 \\ y_1 & y_2 & \cdots & y_n & 0 & 0 & 0 \end{bmatrix} X \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_u \\ 1 \\ x_u \\ y_u \end{bmatrix}$$

Denotando essa equação matricial como  $\mathbf{A}\boldsymbol{\lambda} = \mathbf{b}$ , temos que o vetor de pesos e dos multiplicadores de Lagrange  $\boldsymbol{\lambda}$  pode ser obtido através:

$$\boldsymbol{\lambda} = \mathbf{A}^{-1}\mathbf{b} \quad (2.39)$$

Os pesos serão utilizados na Equação (2.35) e a variância da krigagem universal será dada por:

$$\sigma_{UK}^2 = \mathbf{b}^T \boldsymbol{\lambda} \quad (2.40)$$

# Capítulo 3

## TÉCNICAS DE COKRIGAGEM

### 3.1 Introdução

A cokrigagem é uma extensão multivariada da krigagem, onde são adicionadas covariáveis no modelo de estimação espacial. Além da verificação prévia da correlação entre as variáveis, outro aspecto importante na utilização da técnica é o método de amostragem das covariáveis e da variável principal, o que determina a análise de variância utilizada na modelagem do estimador. Quando estes são medidos no mesmo ponto, utiliza-se o variograma cruzado, já no caso de serem amostrados em pontos diferentes, utiliza-se o pseudo-variograma cruzado (Webster and Oliver, 2007).

Partindo do argumento de que as covariáveis estão correlacionadas espacialmente com a variável principal, pode-se estimar o valor dessa correlação utilizando todas as observações de pontos onde essas variáveis foram amostradas. Através da correlação é possível melhorar a estimação da variável principal em localidades em que ela não foi observada utilizando os valores dessas covariáveis. Portanto, a cokrigagem tem como principal objetivo melhorar a estimação da variável principal sub-amostrada

através da informação de outras variáveis mais densamente amostradas (Bailey and Gatrell, 1995).

Segundo Memarsadeghi et al. (2005), existem vários tipos de cokrigagem, entre as quais podemos destacar a cokrigagem simples, a cokrigagem ordinária e a cokrigagem universal. Neste trabalho trataremos da cokrigagem ordinária e da cokrigagem universal, que representa uma extensão da krigagem universal.

### 3.2 Variograma Cruzado, Covariograma Cruzado e Pseudo-Variograma Cruzado

A cokrigagem faz uso de mais de uma variável na sua modelagem. Nessa situação, torna-se necessário desenvolver um método formal para estimar e modelar a correlação entre as variáveis. Essa análise é possível através da extensão da ideia do variograma e do covariograma referente a uma variável, para o variograma cruzado e o covariograma cruzado entre duas variáveis (Bailey and Gatrell, 1995).

Sejam  $Z_1(u_i)$  com  $i = 1, 2, \dots, n_1$  o valor da variável principal  $Z_1$  em  $n_1$  pontos  $u_i$  e  $Z_2(u_j)$  com  $j = 1, 2, \dots, n_2$  o valor da covariável  $Z_2$  em  $n_2$  pontos  $u_j$ . O variograma cruzado é definido por:

$$2\gamma_{Z_1 Z_2}(h) = E((Z_1(u+h) - Z_1(u))(Z_2(u+h) - Z_2(u))) \quad (3.1)$$

onde  $h$  é um vetor distância definido.

De mesma forma, o covariograma cruzado é definido por:

$$C_{Z_1 Z_2}(h) = E((Z_1(u+h) - \mu_{Z_1})(Z_2(u) - \mu_{Z_2})) \quad (3.2)$$

O variograma cruzado e o covariograma cruzado aplicam-se em situações onde

a variável principal e a covariável são medidas no mesmo ponto. Entretanto, esse tipo de situação não ocorre usualmente, o que dificulta sua utilização em análises práticas (Webster and Oliver, 2007). Para atender essa necessidade, formulou-se o pseudo-variograma cruzado (*pseudo-cross-variogram*):

$$\gamma(\mathbf{h})_{Z_1, Z_2}^C = \frac{1}{2} E \{ [Z_1(\mathbf{u}) - Z_2(\mathbf{u} + \mathbf{h})]^2 \} \quad (3.3)$$

que é utilizado quando a variável principal e a covariável são medidas em pontos diferentes.

Para um processo estacionário de segundo ordem  $2\gamma_{Z_1 Z_2}^p(h)$  está relacionado à função de covariância da forma:

$$2\gamma_{Z_1 Z_2}^p(h) = \frac{1}{2} [C_{Z_1 Z_1}(0) + C_{Z_2 Z_2}(0)] - C_{Z_1 Z_2}(h) \quad (3.4)$$

### 3.3 Cokrigagem Ordinária

Supondo  $V$  variáveis,  $l = 1, 2, \dots, V$  e  $Z_1$  como sendo a variável principal. O estimador da cokrigagem ordinária para  $Z_1$  em um ponto  $u_0$  será:

$$Z_1^*(u_0) = \sum_{l=1}^V \sum_{i=1}^{n_l} \lambda_{il} Z_l(x_i) \quad (3.5)$$

onde  $i = 1, 2, \dots, l$  são os pontos onde a variável  $l$  foi amostrada. A cokrigagem ordinária tem as seguintes suposições para os pesos  $\lambda_{il}$ :

$$\sum_{i=1}^{n_l} \lambda_{il} = \begin{cases} 1 & \text{se } l = Z_1 \\ 0 & \text{se } l \neq Z_1 \end{cases} \quad (3.6)$$

Sujeito às condições de (3.6), obtém-se o estimador da variância de  $Z_1(u_0)$  minimizando:

$$\sum_{i=1}^{n_l} \lambda_{il} \gamma_{lv}(x_i, x_j) + \alpha_v = \bar{\gamma}_{z_1 v}(x_j, u_0) \quad (3.7)$$

para todo  $v = 1, 2, \dots, V$  e  $j = 1, 2, \dots, n_v$ .

Defini-se  $\gamma_{lv}(x_i, x_j)$  como o semivariograma cruzado entre as variáveis  $l$  e  $v$  nos locais  $i$  e  $j$  separados pelo vetor  $x_i - x_j$ ,  $\bar{\gamma}_{z_1 v}(x_j, u_0)$  como o semivariograma cruzado médio entre o local  $j$  e o local  $u_0$  e  $\alpha_v$  como o multiplicador de Lagrange para a  $v$ -ésima variável.

Resolvendo o sistema, obtém-se o estimador para a variância de  $Z_1$ :

$$\sigma_{z_1}^2 = \sum_{l=1}^V \sum_{i=1}^{n_l} \lambda_{il} \gamma_{lv}(x_i, x_j) + \alpha_v \quad (3.8)$$

### 3.4 Cokrigagem Universal

A cokrigagem universal naturalmente é uma extensão multivariada de krigagem universal. Neste caso, tanto a variável principal como as covariáveis apresentam características não-estacionárias, portanto cada variável terá uma componente de tendência em sua fórmula.

Segundo Chiles and Delfiner (1999), para aplicação da técnica de cokrigagem universal é necessário especificar como as médias da variável principal e das covariáveis se relacionam. Três casos específicos podem ser considerados: caso onde as tendências das variáveis são algebricamente independentes, caso onde as tendências apresentam dependência linear e por fim um caso misto. Existe um quarto caso possível, que ocorre quando há uma relação através de funções não-lineares entre as tendências. Essa caso não pode ser tratado por cokrigagem universal e trans-

formações e reparametrizações são necessárias.

Neste trabalho será tratado o caso em que as tendências das variáveis estudadas apresentam independência algébrica, portanto assume-se que tanto a variável principal como as covariáveis possuem tendência própria e os coeficientes que compõe essas tendências não possuem relação. Desse modo, cada variável possui a forma:

$$Z_l(u) = \sum_{k=0}^K \beta_{lk} f_{lk}(u) + \varepsilon_l(u) \quad (3.9)$$

Por uma questão de facilidade, a partir deste momento trabalharemos com uma notação matricial. Definindo:

$$\mathbf{Z}_l = (Z_l(x_1), Z_l(x_2), \dots, Z_l(x_{n_l}))'$$

o vetor coluna formado pelas observações  $Z_l(x_i)$  da variável  $l$ ;

$$\boldsymbol{\lambda}_l = (\lambda_{1l}, \lambda_{2l}, \dots, \lambda_{n_l l})'$$

o vetor coluna formado pelos pesos  $\lambda_{il}$  da variável  $l$ ;

$$\mathbf{F}_l = (f_{lk}(x))$$

a matriz das funções  $f_{lk}$  composta por  $k$  colunas e  $n_l$  linhas;

$$\boldsymbol{\beta}_l = (\beta_{lk})$$

o vetor coluna composto pelos coeficientes  $\beta_{lk}$  da tendência referente a variável  $l$ ;

$$\mathbf{f}_{10} = (f_{1k}(u_0))$$

a matriz das funções  $f_{1k}$  referentes a variável principal aplicadas no ponto  $u_0$ .

Supondo novamente  $V$  variáveis,  $l = 1, 2, \dots, V$  e  $Z_1$  como sendo a variável principal, o interesse é obter uma estimativa para  $Z_1$  em um ponto  $u_0$ :

$$Z_1^*(u_0) = \sum_{l=1}^V \sum_{i=1}^{n_l} \lambda_{il} Z_l(x_i) = \sum_{l=1}^V \boldsymbol{\lambda}'_l \mathbf{Z}_l \quad (3.10)$$

Para avaliar o estimador, temos:

$$E(Z_1^*(u_0) - Z_1(u_0)) = \sum_{l=1}^V \boldsymbol{\lambda}'_l \mathbf{F}_l \boldsymbol{\beta}_l - \mathbf{f}'_{10} \boldsymbol{\beta}_1 \quad (3.11)$$

de modo que o estimador será não-viesado caso atenda as condições:

$$\boldsymbol{\lambda}'_l \mathbf{F}_l = \delta_l \mathbf{f}'_{10} \quad \text{onde} \quad \delta_l = \begin{cases} 1 & ; \quad l = 1 \\ 0 & ; \quad l \neq 1 \end{cases}$$

Minimizando a variância seguindo as restrições impostas, o sistema de cokrigagem universal obtido para o cálculo do vetor  $\boldsymbol{\lambda}_l$  de pesos e do vetor  $\boldsymbol{\alpha}_l$  de multiplicadores Lagrange é igual a:

$$\begin{cases} \sum_{j=1}^V \Gamma_{lj} \boldsymbol{\lambda}_j + \mathbf{F}_l \boldsymbol{\alpha}_l = \mathbf{b}_l & l = 1, 2, \dots, V \\ \mathbf{F}'_l \boldsymbol{\lambda}_l = \delta_l \mathbf{f}_{10} & l = 1, 2, \dots, V \end{cases} \quad (3.12)$$

Essas equações podem ser representadas na forma matricial. Para facilitar a questão de notação, será tratado um exemplo com apenas uma covariável. Casos com mais de uma covariável são obtidos por simples extensão.

Seja  $\Gamma_{uv}$  a matriz de semivariância cruzada das variáveis  $u$  e  $v$ , onde  $u$  é a variável principal, amostradas em  $n_u$  e  $n_v$  locais respectivamente. Essa matriz possui dimensão  $n_u \times n_v$ :

$$\Gamma_{\mathbf{uv}} = \begin{bmatrix} \gamma_{uv}(x_1, x_1) & \gamma_{uv}(x_1, x_2) & \cdots & \gamma_{uv}(x_1, x_{n_v}) \\ \gamma_{uv}(x_2, x_1) & \gamma_{uv}(x_2, x_2) & \cdots & \gamma_{uv}(x_2, x_{n_v}) \\ \vdots & \vdots & \cdots & \vdots \\ \gamma_{uv}(x_{n_u}, x_1) & \gamma_{uv}(x_{n_u}, x_2) & \cdots & \gamma_{uv}(x_{n_u}, x_{n_v}) \end{bmatrix}$$

Denota-se por  $b_{uu}$  e  $b_{uv}$  os vetores de semivariância e semivariância cruzada entre os pontos amostrados e o ponto  $u_0$  que deseja-se estimar:

$$\mathbf{b}_{uu} = \begin{bmatrix} \gamma_{uu}(x_1, u_0) \\ \gamma_{uu}(x_2, u_0) \\ \vdots \\ \gamma_{uu}(x_{n_u}, u_0) \end{bmatrix}, \mathbf{b}_{uv} = \begin{bmatrix} \gamma_{uv}(x_1, u_0) \\ \gamma_{uv}(x_2, u_0) \\ \vdots \\ \gamma_{uv}(x_{n_v}, u_0) \end{bmatrix}$$

Portanto, a equação matricial possui a forma:

$$\begin{bmatrix} \Gamma_{uu} & \Gamma_{uv} & \mathbf{F}_u & \mathbf{0} \\ \Gamma_{vu} & \Gamma_{vv} & \mathbf{0} & \mathbf{F}_v \\ \mathbf{F}'_u & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}'_v & \mathbf{0} & \mathbf{0} \end{bmatrix} X \begin{bmatrix} \boldsymbol{\lambda}_u \\ \boldsymbol{\lambda}_v \\ \boldsymbol{\alpha}_u \\ \boldsymbol{\alpha}_v \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{uu} \\ \mathbf{b}_{uv} \\ \mathbf{f}_{10} \\ \mathbf{0} \end{bmatrix}$$

Denotando a matriz aumentada de  $\Gamma$ 's por  $\mathbf{A}$ , o vetor de pesos e do multiplicadores de Lagrange por  $\mathbf{X}$  e a matriz resultante do produto por  $\mathbf{B}$ , pode-se escrever a solução da equação como:

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B} \quad (3.13)$$

e a variância como:

$$\sigma_{UC}^2 = \mathbf{b}_0 - \mathbf{X}'\mathbf{B} \quad (3.14)$$

# Capítulo 4

## MATERIAL E MÉTODOS

### 4.1 Introdução

Neste capítulo serão apresentados os materiais e métodos a serem utilizados no trabalho. O banco de dados que será utilizado contém informações sobre a concentração no solo de quatro metais pesados medidos ao longo das margens do rio Meuse. Os dados estão disponibilizados no pacote *gstat* do software R (Rossiter, 2007). Para este trabalho, serão selecionados os dados referentes à concentração de cobre, zinco e de material orgânico no solo.

Com esses dados, será utilizado o algoritmo desenvolvido na linguagem SAS/IML para realizar a estimação do nível de concentração dessas variáveis de interesse através das técnicas de krigagem e cokrigagem explicadas anteriormente.

### 4.2 Rio Meuse

O rio Meuse é um rio de 950 Km de comprimento que tem sua nascente na França, no planalto de Langres. O Meuse banha importantes cidades francesas, atravessa a Bélgica e a Holanda em direção do mar do Norte, banhando parte da Alemanha.

O conjunto de dados disponibilizados no pacote *gstat* do software R apresenta



Figura 4.1: Rio Meuse

155 observações com informações sobre os níveis de concentrações de metais pesados (ppm) na superfície do solo, juntamente com uma série de variáveis do solo e da paisagem. Estes dados foram coletados em uma planície de inundação do rio Meuse, perto da aldeia de Stein, na Alemanha. As concentrações de metais pesados são amostradas em massa, em uma área de aproximadamente  $15m \times 15m$ .

Entre os metais pesados disponibilizados, serão analisados os dados referentes aos níveis de concentração de cobre e zinco, além dos níveis de concentração de material orgânico. Por questão de praticidade e facilidade no ajuste dos modelos teóricos e interpretação dos resultados, será trabalhado com a transformação logarítmica dessas variáveis.

### 4.3 Algoritmo IML

O algoritmo desenvolvido para este trabalho foi construído utilizando-se o procedimento IML (*Interactive Matrix Language*), pois os códigos implementados em IML são facilmente adaptáveis para outras linguagens de programação.

O algoritmo é composto por diversas macros, dentre as quais se destacam três principais que envolvem os procedimentos essenciais na aplicação dos métodos de krigagem e cokrigagem, que são a construção da estrutura variográfica dos dados e a estimação da variável de estudo para toda a região de interesse.

### 4.3.1 Macro Variogram

A macro *variogram* realiza a construção do semivariograma experimental assim como o ajuste do modelo teórico para essa estrutura de variância espacial. Dois componentes são importantes no ajuste da curva do semivariograma, que são a definição do número de classes de distância (*lag*) e a amplitude dessas classes de distância (*lag distance*).

A macro permite o ajuste da curva do semivariograma definindo arbitrariamente essas duas componentes ou caso não sejam especificadas, o algoritmo realiza o ajuste automático adequando os valores de *lag* e *lag distance* conjuntamente através de um processo iterativo. A idéia desse processo iterativo é utilizar uma porcentagem de afrouxamento, que pode variar entre 0% a 100% (na prática entre 1% e 60%), em que a curva pode descer ou subir. Além dessa porcentagem de afrouxamento, deve-se especificar dois valores de incremento, o primeiro que especifica a porcentagem sobre a menor distância observada entre pontos que será utilizada na definição do *lag distance* inicial e o segundo que será o incremento sobre *lag distance* para cada iteração do processo.

Os parâmetros necessários para execução da macro são:

```
%macro variogram(tab=,var=,long=,lat=,lag=,lagd=,model=,sill=,range=,  
nugget=,tabout=,maxint=,pct=,method=,incri=0.1,incr=1,alpha=0.05)
```

onde: **tab** é o banco de dados que contém as observações da variável em estudo coletadas na amostra, bem como suas coordenadas geográficas; **var** é o nome da variável em estudo contida no banco de dados; **long** é a variável correspondente à coordenada de longitude; **lat** é a variável correspondente à coordenada de latitude; **lag** especifica o número máximo de classes de distâncias na construção do semivariograma; **lagd** especifica a amplitude de cada classe de distâncias, por exemplo, se **lagd**=2 os valores correspondentes em cada classe serão múltiplos de 2; **model** especifica qual modelo será utilizado no ajuste do semivariograma. Os modelos implementados são: o modelo esférico (spherical), o modelo gaussiano (gaussian), o modelo exponencial (exponential), o modelo cubico (cubic), o modelo pentaesférico (pentaspherical) e o modelo seno (sine); **sill**, **range** e **nugget** especificam os valores iniciais para Patamar, Alcance e Efeito Pepita utilizados na minimização e ajuste do modelo teórico do semivariograma. Caso não sejam especificados valores para esses parâmetros, são utilizados os valores iniciais definidos em Jian et al. (1996); **maxint** especifica o número máximo de iterações no ajuste automático do semivariograma; **pct** especifica a porcentagem de afrouxamento sobre qual a curva do semivariograma pode subir ou descer; **incri** especifica a porcentagem sobre a menor distância observada entre pontos que será utilizada na definição do *lag distance* inicial do processo iterativo. Caso não seja especificado, o valor *default* é 0,1; **incr** especifica o incremento sobre o *lag distance* para cada iteração do ajuste do semivariograma. Caso não seja especificado o valor *default* é 1; **method** especifica se o semivariograma deve ser referente aos dados originais ou sobre os resíduos, neste caso a entrada no

parâmetro deve ser “universal”; **alpha** especifica o nível de significância a ser considerado no teste sobre os betas em caso de ajuste dos resíduos; **tabout** corresponde ao nome que se quer dá a tabela de saída com as medidas obtidas;

Para o caso do ajuste do semivariograma ser referente aos resíduos, aplica-se a estimação dos mesmos através do método de mínimos quadrados ordinários.

```
%if %upcase(&method)=UNIVERSAL %then %do;  
_x_=j(nrow(z),1,1)||COORD;  
beta=inv(_x_'*_x_)*_x_'*z;  
z=z-_x_*beta;  
%end;
```

A condicional representada pelo parâmetro **method** indicará se essa transformação dos dados originais para o resíduo deve ser feita. Independente desse aspecto, poderia-se manter esse processo de estimação por mínimos quadrado como regra geral, apenas indicando no caso de ajuste do semivariograma dos dados originais que a matriz **\_x\_** seja formada apenas por uma coluna de 1's, neste caso o modelo teria apenas um  $\beta_0$ , que seria a própria média e conseqüentemente se manteria os dados originais.

Posteriormente, com as distâncias entre os pontos calculadas, é realizada a construção da matriz semivariograma, que conterá os limites inferior e superior de cada classe de distâncias, o número de pares contidos em cada classe, a soma dos quadrados das distâncias entre os pares de pontos na classe, o ponto médio de cada classe e o valor calculado para a semivariância.

No caso do ajuste automático do semivariograma, o processo iterativo pode ser representado pelo esquema apresenta na Tabela 4.1. Com o semivariograma experimental calculado, é realizado a minimização dos parâmetros de Alcance, Patamar

e Efeito Pepita para o ajuste do modelo teórico que melhor caracterize a estrutura de variância obtida. Para mais detalhes desse processo e informações sobre a macro *variogram* consultar Barreto (2011) e Galvão (2012).

Tabela 4.1: Algoritmo para ajuste automático do *lag* e *lagd* na construção do semi-variograma

---



---

```

Inicializar I e  $\Delta$  iguais a 1
Enquanto (  $\Delta > 0$  e  $I > M$  )
{
  Para k variando de 1 até nlag
  {
    Calcule  $\gamma(h_k)$ 
    Se  $k > 1$  faça
    {
       $\Delta = \gamma(h_k) - \gamma(h_{k-1})$ 
       $pctl = \gamma(h_{k-1}) \times pct$ 
      Se  $k > 3$  faça
      {
         $\Delta = \gamma(h_k) - \gamma(h_{k-3})$ 
         $pctl = \gamma(h_{k-3}) \times pct$ 
      }
      Se  $\Delta < -pctl$  e  $f_{k-1} < 2$  faça
      {
         $Lagd = Lagd + dmin \times incri$ 
         $I = I + incr$ 
         $\Delta = 99999$ 
         $k = lag$ 
      }
    }
  }
  Se  $I = M$  faça
  {
     $\Delta = 99999$ 
     $lag = lag - 1$ 
     $lagd = dmin$ 
     $I = 1$ 
  }
  Se  $\Delta < 99999$  faça {
     $\Delta = -1$  }
}

```

---



---

### 4.3.2 Macro Crossvariogram

A macro *crossvariogram* realiza a construção dos semivariogramas necessários para a aplicação do método de cokrigagem. Além dos semivariogramas experimentais para a variável principal e as covariáveis, a macro calcula os variogramas cruzados. O algoritmo gera automaticamente o variograma cruzado, para a situação onde as observações da variável principal e das covariáveis são coletadas nos mesmos pontos, e o pseudo-variograma cruzado, para a situação onde as observações da variável principal e das covariáveis são observadas em pontos diferentes.

A macro trabalha com o ajuste de um Modelo Linear de Corregionalização (*Linear Model of Coregionalization*), onde todos os modelos (semivariogramas e variogramas cruzados) possuem o mesmo modelo teórico e Alcance, mas com diferentes valores de Patamar e Efeito Pepita, portanto é necessário especificar apenas o modelo teórico de ajuste ideal para a variável principal. Foi adicionado um parâmetro que permite um relaxamento nos pressupostos de igualdade para o modelo teórico e Alcance, realizando o ajuste de modelos com valores diferentes para o Alcance.

Os parâmetros necessários para execução da macro são:

```
%macro crossvariogram(taby=,vary=,longy=,laty=,tabx=,varx=,longx=,latx=,  
lag=,lagd=,model=,tabout=,maxint=,pct=,condrange=,method=,incri=0.1,  
incr=1,alpha=0.05);
```

onde: **taby** é o banco de dados que contém as observações da variável principal do estudo, bem como suas coordenadas geográficas; **vary** é a nome da variável principal em estudo contida no banco de dados; **longy** é a variável correspondente à coordenada de longitude da variável principal; **laty** é a variável correspondente à coordenada de latitude; **tabx**, **varx**, **longx** e **latx** correspondem às mesmas variáveis

citadas anteriormente para as covariáveis; **lag** especifica o número máximo de classes de distâncias na construção do variograma cruzado; **lagd** especifica o tamanho de cada classe de distâncias, por exemplo, se  $\text{lagd}=2$  os valores correspondentes em cada classe serão múltiplos de 2; **model** especifica qual modelo será utilizado no ajuste do variograma cruzado. Os modelos implementados são: o modelo esférico (spherical), o modelo gaussiano (gaussian), o modelo exponencial (exponencial), o modelo cubico (cubic), o modelo pentaesférico (pentaspherical) e o modelo seno (sine); **tabout** corresponde ao nome que se quer dá a tabela de saída com as medidas obtidas; **maxint** especifica o número máximo de iterações no ajuste automático dos semivariogramas da variável principal e das covariáveis; **pct** especifica a porcentagem de afrouxamento sobre qual a curva do semivariograma pode subir ou descer; **condrange** especifica se os semivariogramas das covariáveis e os variogramas cruzados devem ter ou não o mesmo Alcance do que o semivariograma da variável principal. Para utilizar o mesmo Alcance, a entrada no parâmetro deve ser “equal”; **incri** especifica a porcentagem sobre a menor distância observada entre pontos que será utilizada na definição do *lag distance* inicial do processo iterativo no ajuste dos semivariogramas. Caso não seja especificado, o valor *default* é 0,1; **incr** especifica o incremento sobre o *lag distance* para cada iteração do ajuste dos semivariogramas. Caso não seja especificado o valor *default* é 1; **method** especifica se os modelos de semivariograma e variogramas cruzados devem ser refente aos dados originais ou sobre os resíduos, neste caso a entrada no parâmetro deve ser “universal”; **alpha** especifica o nível de significância a ser considerado no teste sobre os betas em caso

de ajuste dos resíduos;

Para mais detalhes e informações sobre a macro *crossvariogram* consultar Galvão (2012).

### 4.3.3 Macro Cokriging

A macro *cokriging* realiza a estimação da variável de estudo para a região de interesse. O algoritmo permite a estimação através dos métodos de krigagem e cokrigagem, quando forem especificadas covariáveis, tanto para o caso ordinário como para o caso universal.

A construção da matriz expandida dos variogramas para o cálculo das estimativas segue uma regra geral, que é a construção da matriz para o caso universal. No caso de estimação por krigagem ou cokrigagem ordinária, as linhas e colunas referentes às coordenadas  $x$  e  $y$  serão preenchidas com zeros. Como existe a possibilidade de construção de uma matriz com colunas e linhas de zero, a estimação é realizada trabalhando com a inversa generalizada da matriz expandida dos variogramas.

Os parâmetros necessários para execução da macro são:

```
%macro cokriging(tabby=,vary=,longy=,laty=,tabx=,varx=,longx=,latx=,model=,  
sill=,range=,nugget=,map=,by=,tabout=,method=,alpha=0.05);
```

onde: **taby** é o banco de dados que contém as observações da variável principal do estudo, bem como suas coordenadas geográficas; **vary** é a nome da variável principal em estudo contida no banco de dados; **longy** é a variável correspondente à coordenada de longitude da variável principal; **laty** é a variável correspondente à coordenada de latitude; **tabx**, **varx**, **longx** e **latx** correspondem às mesmas variáveis citadas anteriormente para as covariáveis; **model** especifica qual modelo será utili-

zado no ajuste do variograma cruzado. Os modelos implementados são: o modelo esférico (spherical), o modelo gaussiano (gaussian), o modelo exponencial (exponential), o modelo cubico (cubic), o modelo pentaesférico (pentaspherical) e o modelo seno (sine); **sill** especifica o parâmetro patamar encontrado na análise variográfica; **range** especifica o parâmetro de alcance encontrado na análise variográfica; **nugget** especifica o parâmetro efeito pepita encontrado na análise variográfica; **map** especifica o banco de dados que contém o shape do mapa que se deseja montar; **by** é usado na especificação da grade com a localização espacial para os pontos estimados. O **by** especifica a que taxa a coordenada muda a partir do menor valor da coordenada até chegar ao maior; **tabout** especifica o nome da tabela de dados de saída.; **method** especifica qual método de estimação deve ser aplicado, ordinário ou universal. No caso universal, a entrada no parâmetro deve ser “universal”; **alpha** especifica o nível de significância a ser considerado no teste sobre os betas em caso de krigagem ou cokrigagem universal ;

Para mais detalhes e informações sobre a macro *cokriging* consultar Galvão (2012).

# Capítulo 5

## ANÁLISE DOS RESULTADOS

### 5.1 INTRODUÇÃO

Neste capítulo serão apresentados os resultados obtidos pelas análises de krigagem e cokrigagem realizadas através do algoritmo implementado. Primeiramente foi realizado uma comparação dos resultados obtidos para análises de krigagem ordinária e krigagem universal utilizando o algoritmo implementado em SAS/IML e utilizando as funções contidas no pacote *gstat* do software R, com o intuito de balizar o algoritmo proposto com uma função já implementada em outro pacote estatístico.

Em seguida, são apresentadas as estimativas obtidas para as variáveis de estudo através dos métodos de krigagem ordinária e krigagem universal, comparando os resultados dessas duas técnicas para avaliar a influência sobre a qualidade da estimativa ao considerarmos o modelo com a componente de tendência. Serão esclarecidos os critérios de escolha dos modelos de semivariograma utilizados na estimação.

Serão mostrados também os resultados obtidos para as análises de cokrigagem ordinária e cokrigagem universal, considerando os níveis de concentração de cobre como variável principal e as outras concentrações como possíveis covariáveis. Os modelos de semivariograma e semivariograma cruzado utilizados na estimação serão

apresentados.

Para concluir, será realizada uma comparação entre as técnicas de krigagem e cokrigagem para os métodos ordinário e universal com o intuito de avaliar o efeito da inclusão da covariável na qualidade da estimativa.

## 5.2 BALIZAMENTO DO ALGORITMO

Para avaliar o algoritmo desenvolvido em SAS/IML, utilizou-se um exemplo de aplicação da krigagem ordinária e krigagem universal para estimação da variável zinco, presente no manual de ajuda do pacote *gstat* do software R. A Tabela 5.1 apresenta as medidas descritivas para o logaritmo da variável zinco (**ltzn**).

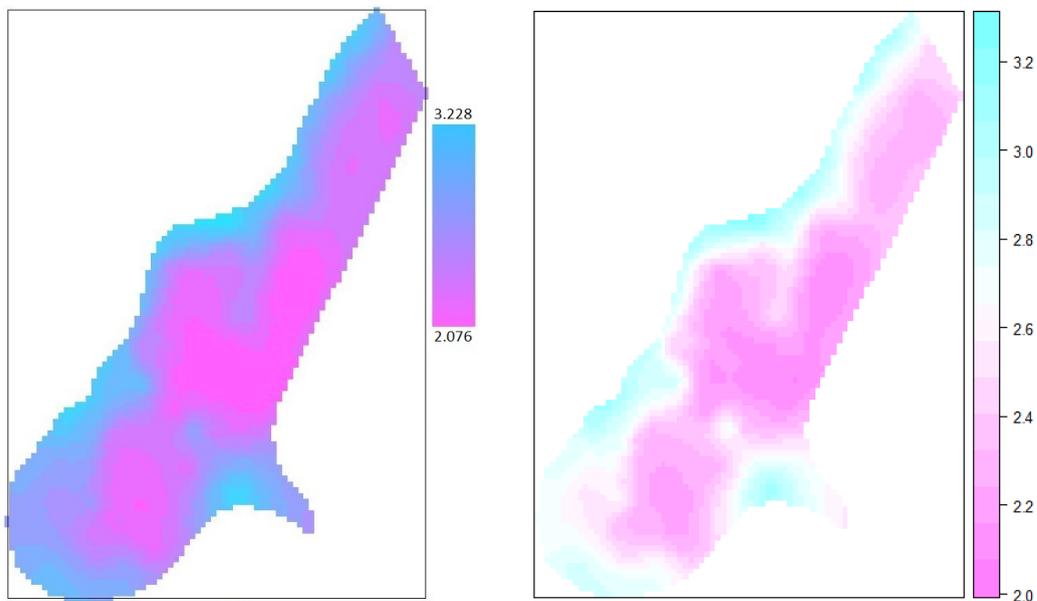
Tabela 5.1: Estatísticas Descritivas do log de zinco

Estatística	<b>ltzn</b>
N	155
Média	2,55
Desvio Padrão	0,31
Mínimo	2,05
Máximo	3,26

Os modelos teóricos utilizados assim como os parâmetros estimados para o ajuste do modelo de semivariograma foram os mesmos indicados no manual de ajuda.

### 5.2.1 Krigagem Ordinária

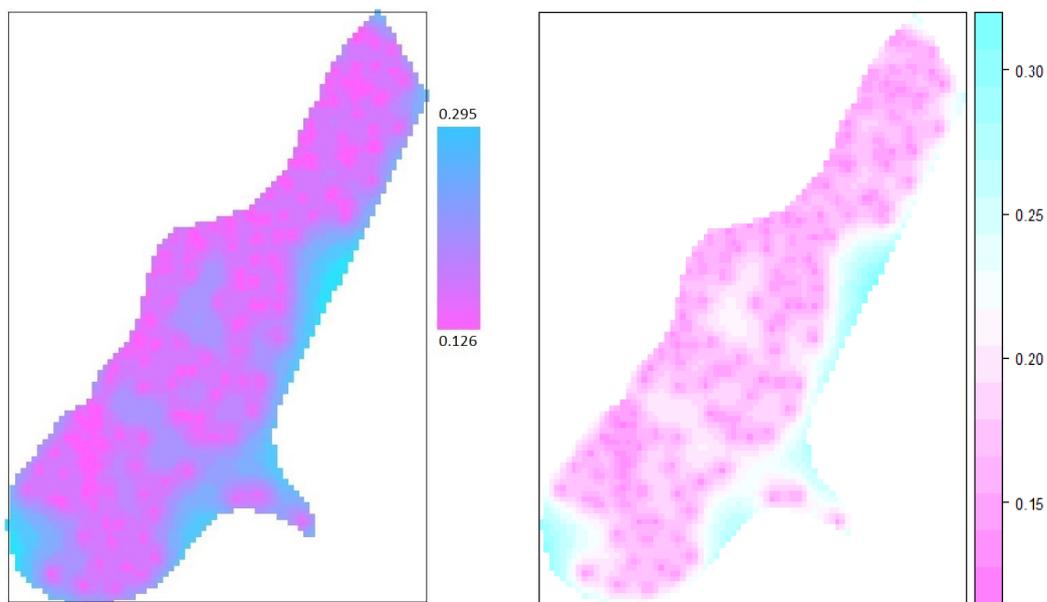
Para aplicação da krigagem ordinária, foram estimados 3.103 pontos. O modelo teórico utilizado para ajustar o semivariograma foi o modelo Esférico. A Figura 5.1 apresenta as estimativas obtidas pelo algoritmo em SAS/IML e pela função *krige* do pacote *gstat* do R. A Figura 5.2 apresenta o erro padrão associado a cada estimativa.



(a) Valores estimados para **ltzn** - Algoritmo SAS/IML

(b) Valores estimados para **ltzn** - R

Figura 5.1: Estimativas de Krigagem Ordinária para **ltzn**



(a) Erro Padrão - Algoritmo SAS/IML

(b) Erro Padrão - R

Figura 5.2: Erro Padrão das Estimativas de Krigagem Ordinária

Analisando os mapas de estimativas de **ltzn** apresentados na Figura 5.1, nota-se que o comportamento dos valores estimados foi bem semelhante nos dois casos,

com o mesmo padrão de crescimento e decrescimento para **ltzn** ao longo da região analisada. Os mapas referentes ao erro padrão das estimativas também mostraram comportamento semelhante em ambos os casos. Nos dois casos é possível perceber regiões específicas onde o erro padrão é menor, que são justamente as localidades onde foram coletadas os níveis de concentração de zinco.

A Tabela 5.2 apresenta as medidas descritivas das estimativas e erro padrão associado obtidas em cada caso. Os resultados das estatísticas descritivas mostram que as estimativas de **ltzn** obtidas tanto pelo algoritmo proposto como pelas funções contidas no pacote *gstat* do R foram bem semelhantes, com diferenças apenas na ordem da terceira casa decimal. De mesmo modo, o erro padrão das estimativas também teve resultados bem próximos nos dois casos considerados.

Tabela 5.2: Estatísticas Descritivas das Estimativas - Krigagem Ordinária

Estatística	Estimativa		Erro Padrão	
	Algoritmo SAS/IML	Pacote R	Algoritmo SAS/IML	Pacote R
Mínimo	2,076	2,074	0,125	0,126
1° Quartil	2,275	2,274	0,157	0,160
Mediana	2,421	2,420	0,170	0,174
Média	2,478	2,478	0,179	0,184
3° Quartil	2,680	2,680	0,193	0,199
Máximo	3,227	3,231	0,295	0,307

Finalmente, foram calculadas as diferenças das estimativas e dos erros padrão obtidos em cada caso para cada um dos 3.103 pontos estimados. A Tabela 5.3 apresenta os resultados obtidos para essas diferenças.

Tabela 5.3: Estatísticas Descritivas das Diferenças entre Estimativas e Erro Padrão

Estatística	Diferença entre Estimativas	Diferença entre Erros Padrão
Média	0,0001	-0,003
Desvio Padrão	0,001	0,004

Pelos resultados apresentados pela Tabela 5.3, fica evidente que os resultados da krigagem ordinária para estimação de **ltzn** obtidos através do algoritmo proposto em SAS/IML e pela função *krige* do pacote *gstat* foram praticamente iguais.

### 5.2.2 Krigagem Universal

Para aplicação da krigagem universal, foram estimados 3.103 pontos. O modelo teórico utilizado para ajustar o semivariograma, neste caso referente aos resíduos, foi o modelo Exponencial. A Figura 5.3 apresenta as estimativas obtidas pelo algoritmo em SAS/IML e pela função *krige* do pacote *gstat*. A Figura 5.4 apresenta o erro padrão associado a cada estimativa.

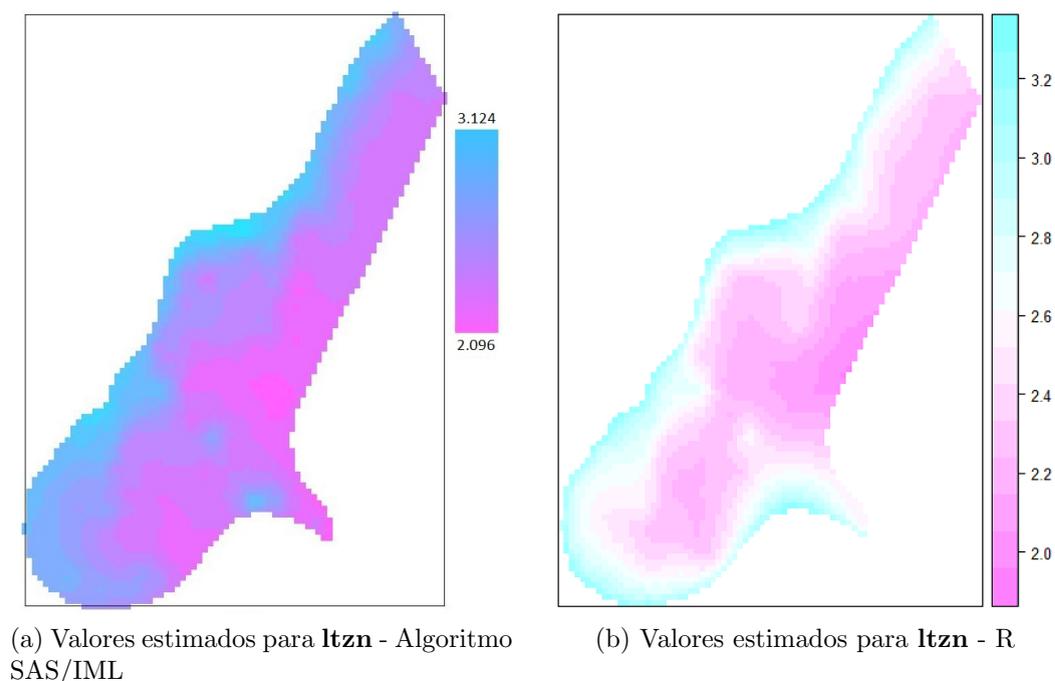


Figura 5.3: Estimativas de Krigagem Universal para **ltzn**

Assim como no caso da krigagem ordinário, os mapas de estimativas de **ltzn** apresentados na Figura 5.3 foram semelhantes, com o comportamento dos valores estimados bem próximos nos dois casos e o mesmo padrão de crescimento e decres-

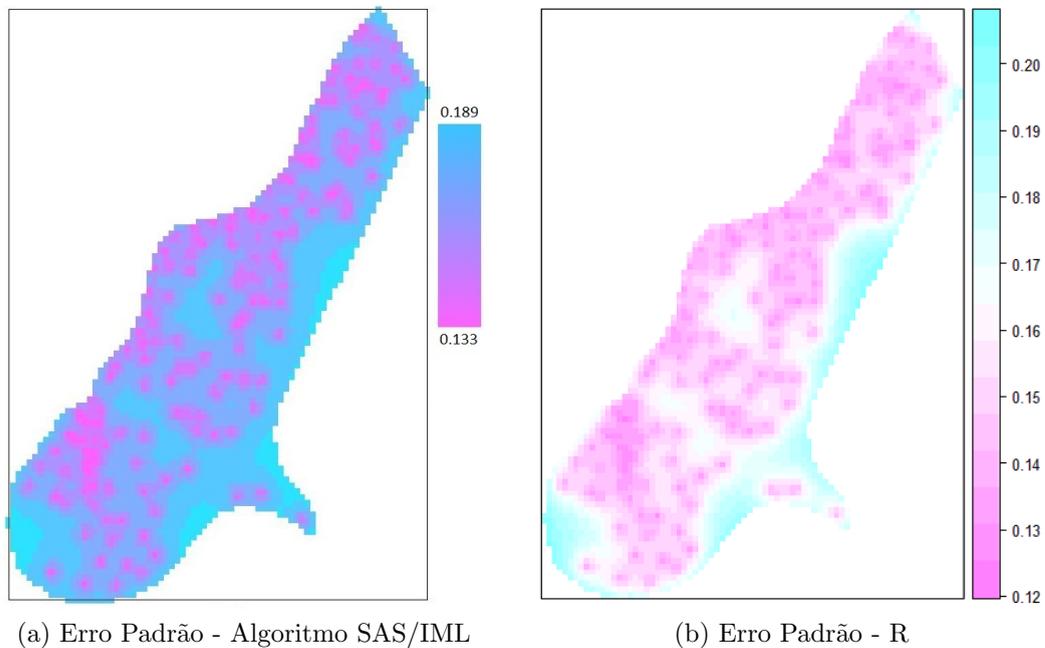


Figura 5.4: Erro Padrão das Estimativas de Krigagem Universal

cimento para **ltzn** ao longo da região analisada. Os mapas referentes ao erro padrão das estimativas também tiveram comportamento bem semelhante em ambos os casos. Nos dois casos é possível perceber regiões específicas onde o erro padrão tem queda acentuada, que são justamente as localidades onde foram coletadas os níveis de concentração de zinco.

A Tabela 5.4 apresenta as medidas descritivas das estimativas e erro padrão associado obtidas em cada caso. Os resultados das estatísticas descritivas mostram que as estimativas de **ltzn** obtidas tanto pelo algoritmo proposto como pelas funções do R em média são iguais, mas existe uma variação relativamente pequena nos valores extremos estimados. De mesmo modo, o erro padrão das estimativas também teve uma variação relativamente pequena em relação ao valores extremos obtidos e as médias também tiveram uma diferença relativamente pequena.

Tabela 5.4: Estatísticas Descritivas das Estimativas - Krigagem Universal

Estatística	Estimativa	Estimativa	Erro Padrão	Erro Padrão
	Algoritmo SAS/IML	Pacote R	Algoritmo SAS/IML	Pacote R
Mínimo	2,096	1,955	0,132	0,125
1° Quartil	2,319	2,263	0,159	0,143
Mediana	2,423	2,417	0,167	0,150
Média	2,476	2,476	0,167	0,154
3° Quartil	2,617	2,667	0,176	0,163
Máximo	3,124	3,269	0,189	0,202

Para melhor comparação, foram calculadas as diferenças das estimativas e dos erros padrão obtidos em cada caso para cada um dos 3.103 pontos estimados. A Tabela 5.5 apresenta os resultados obtidos para essas diferenças.

Tabela 5.5: Estatísticas Descritivas das Diferenças entre Estimativas e Erro Padrão

Estatística	Diferença entre Estimativas	Diferença entre Erros Padrão
Média	-0,0004	0,01
Desvio Padrão	0,15	0,007

Pelos resultados apresentados pela Tabela 5.5, fica evidente que em média os resultados da krigagem universal para estimação de **ltzn** obtidos através do algoritmo proposto em SAS/IML e pela função *krige* do R são bem próximos, com uma diferença relativamente bem pequena. Existe uma variação da ordem de 0,1 para as diferenças dessas estimativas, mas aparentemente trata-se de uma oscilação não tão significativa.

### 5.3 ANÁLISE UTILIZANDO OS MÉTODOS DE KRIGAGEM ORDINÁRIA E KRIGAGEM UNIVERSAL

A aplicação dos métodos de krigagem visam a obtenção das melhores estimativas para a variável de interesse. O sucesso na aplicação desse método está diretamente

ligado no melhor ajuste possível de um modelo teórico para a estrutura de variância espacial dos dados. Nesse sentido, é muito importante avaliar dentre todos os modelos teóricos disponíveis aquele que melhor ajuste o semivariograma dos dados.

Para definir o modelo teórico que melhor caracteriza a estrutura de variância espacial mostrada pelo semivariograma experimental, foram simulados diversos ajustes considerando diferentes modelos teóricos, porcentagens de afrouxamento e valores de incremento para *lag distance*. Para comparar a qualidade de ajuste utilizou-se o Critério de Informação de Akaike (AIC). Nesta análise estimou-se os valores para os níveis de concentração no solo de material orgânico e cobre.

### 5.3.1 Estimação do nível de concentração de Material Orgânico - *ltom*

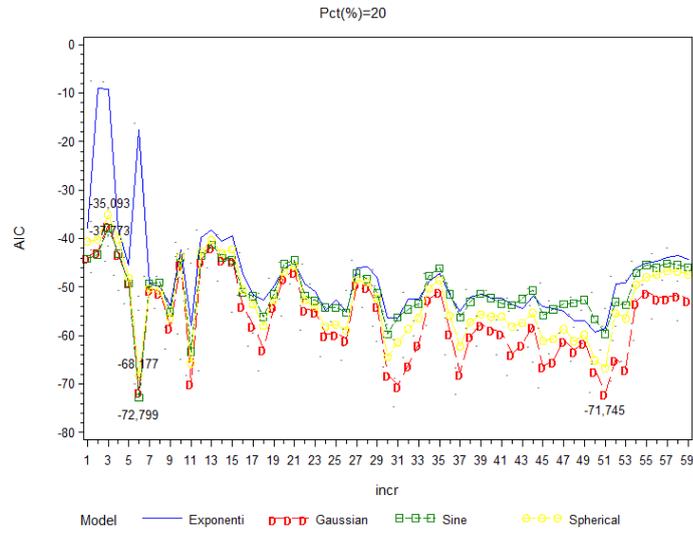
Inicialmente foi realizado uma análise descritiva para os valores amostrados de concentração de material orgânico no solo, apresentada na Tabela 5.6. Neste estudo de caso trabalhou-se com a transformação logarítmica da variável original (***ltom***).

Tabela 5.6: Estatísticas Descritivas do log do material orgânico

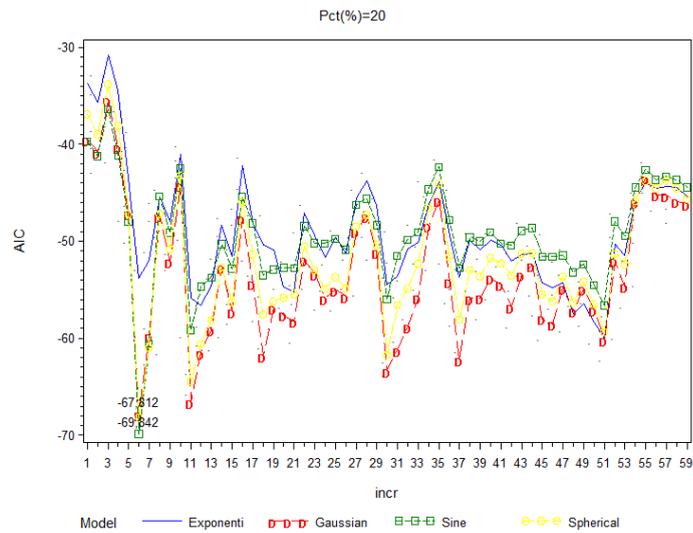
Estatística	<b><i>ltom</i></b>
N	155
Média	0,82
Desvio Padrão	0,21
Mínimo	0,14
Máximo	1,23

A Figura 5.5 apresenta a relação de modelos teóricos testados no ajuste do semivariograma, indicando a variação do AIC do modelo para diversos valores de incremento na distância dos lags. Tanto para o ajuste do semivariograma dos dados originais como para o ajuste do semivariograma dos resíduos, a porcentagem de

afrouxamente que apresentou os melhores resultados foi igual a 20%.



(a) Modelos simulados para ajuste do semivariograma dos dados originais

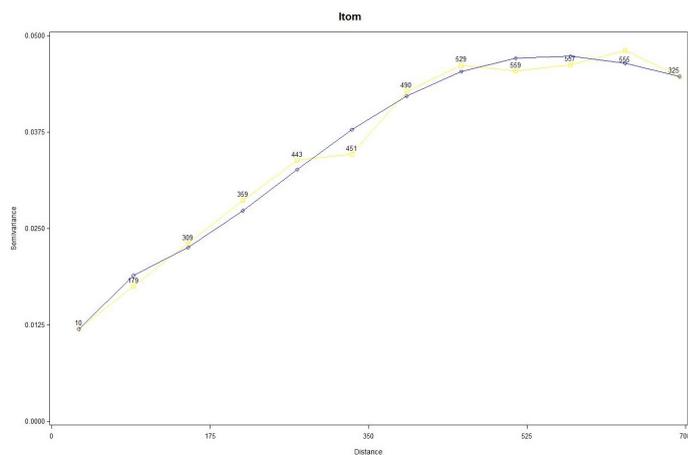


(b) Modelos simulados para ajuste do semivariograma dos resíduos

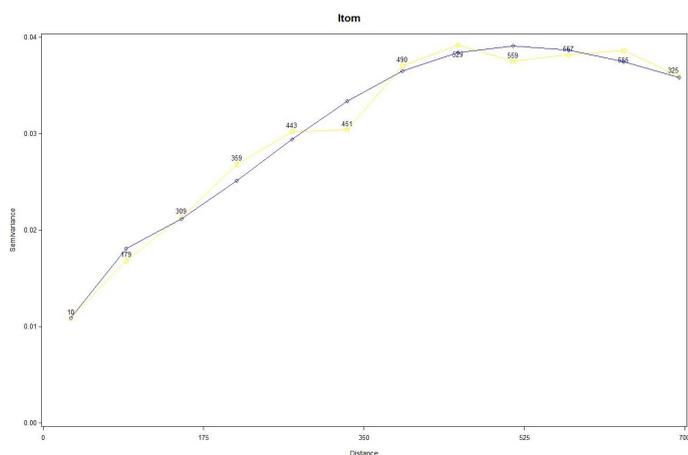
Figura 5.5: Modelos teóricos simulados para ajuste dos semivariogramas

Os gráficos apresentados na Figura 5.5 indicam que o modelo teórico que apresentou menor AIC tanto para o semivariograma dos dados originais como para o semivariograma dos resíduos foi o modelo Seno, iguais a  $-72,79$  e  $-69,84$  respectivamente. Em ambos os casos o incremento que apresentou melhores resultados foi igual a 6. Com a definição do modelo teórico e dos valores ótimos para porcentagem

de afrouxamento e nível de incremento da distância dos lags, ajustou-se os melhores modelos para os semivariogramas experimentais, apresentados na Figura 5.6.



(a) Modelos Seno ajustado para o semivariograma dos dados originais



(b) Modelo Seno ajustado para o semivariograma dos resíduos

Figura 5.6: Modelos teóricos ajustados para os semivariogramas

Os modelos ajustados tiveram resultados bastante satisfatórios, uma vez que conseguem acompanhar bem o comportamento dos semivariogramas experimentais. Com a definição da estrutura formal que caracteriza a variância espacial dos dados, partiu-se para estimação dos níveis de concentração de material orgânico no solo através dos métodos de krigagem ordinária e universal.

A Figura 5.7 apresenta os valores estimados para os níveis de concentração de ma-

terial orgânico. Os pontos em preto são referentes aos locais amostrados. Os mapas de estimativas para **Itom** mostram que ao longo da região onde foram amostradas as concentrações de material orgânico, o comportamento dos valores estimados tanto pela krigagem ordinária como a krigagem universal foram de maneira geral semelhantes. Já nas regiões em que não há pontos amostrados é mais evidente a diferença entre os valores estimados pelos dois métodos, principalmente nas regiões extremas do mapa. A amplitude dos valores estimados foi bem semelhante nos dois casos, com a krigagem universal estimando um limite superior um pouco acima do valor máximo observado na amostra coletada.

## Estimativas de Krigagem Ordinária - Itom      Estimativas de Krigagem Universal - Itom

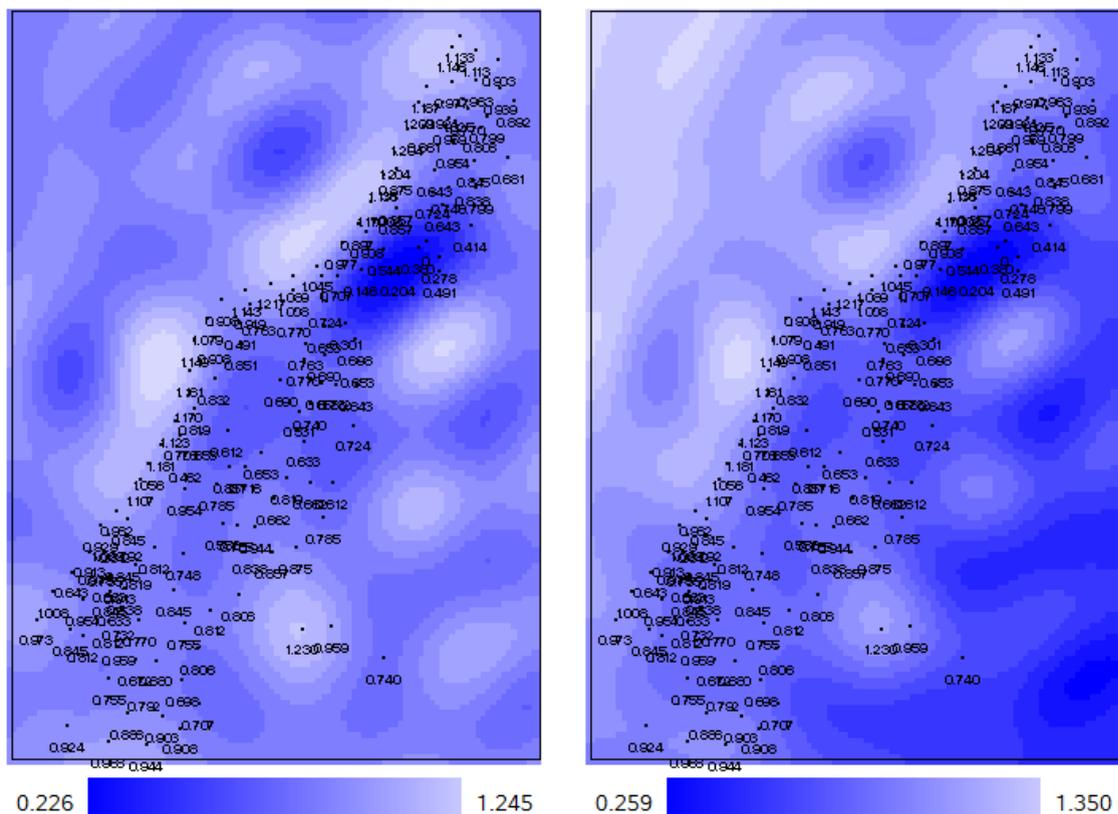


Figura 5.7: Níveis de concentração de material orgânico no solo estimados

A Figura 5.8 apresenta o erro padrão das estimativas obtidas em cada método. Assim como os mapas de estimativas, os mapas do erro padrão também tiveram comportamento semelhante ao longo da região com valores amostrados para os dois métodos. Os menores erros foram observados justamente nessas localidades, com a amplitude de valores um pouco maior para o caso universal, que consegue ampliar a área de erros mínimos.

### Erro Padrão das Estimativas de Krigagem Ordinária

### Erro Padrão das Estimativas de Krigagem Universal

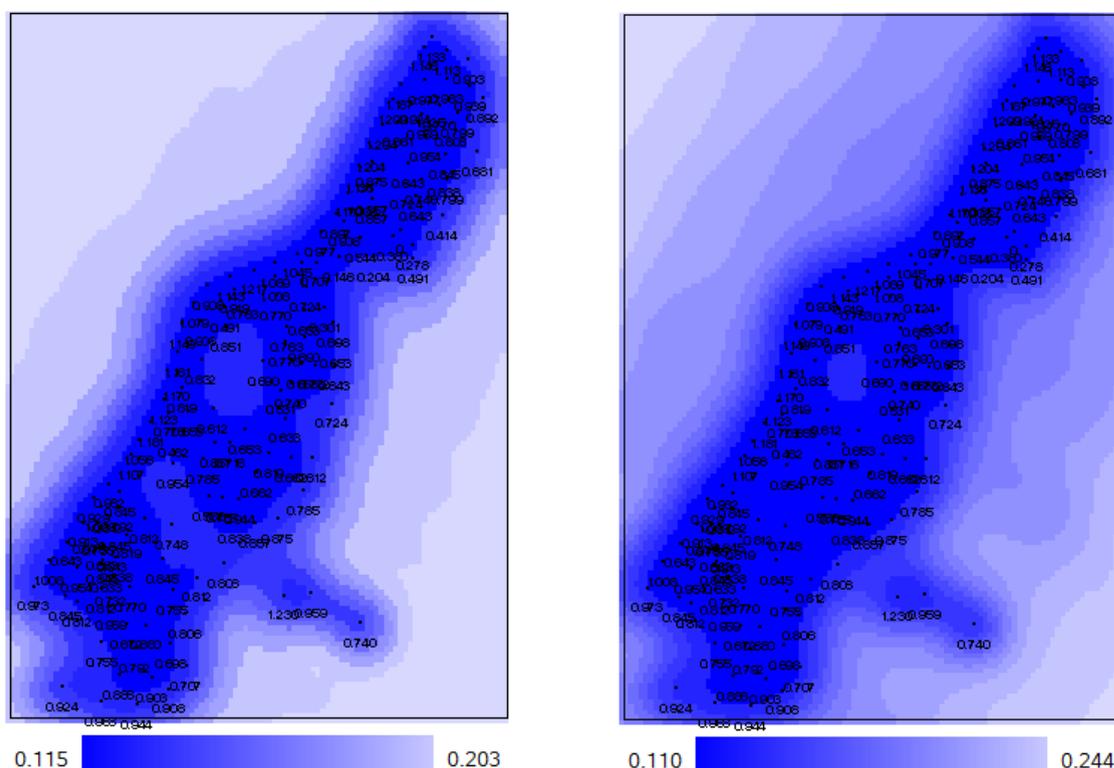


Figura 5.8: Erro Padrão das concentrações de material orgânico estimadas

Nota-se que os erros para krigagem ordinária tendem a estabilizar nas regiões mais extremas do mapa enquanto que os erros da krigagem universal apresentam faixas de crescimento gradativo ao longo do mapa. Essa característica do caso universal pode ser devido a propriedade do modelo, que considera a presença de uma

tendência nos dados.

### **Comparação entre krigagem universal e krigagem ordinária**

Para avaliar o efeito da inclusão da componente de tendência na estimação, foram calculadas as diferenças entre as estimativas e os erros obtidos pelos métodos de krigagem universal e ordinário. As diferenças são definidas como:

$$\text{Diferença Estimativas} = \text{Estimativa Krig.Universal} - \text{Estimativa Krig.Ordinária}$$

$$\text{Diferença Erros} = \text{Erro Krig.Universal} - \text{Erro Krig.Ordinária}$$

Os mapas das diferenças apresentados na Figura 5.9 mostram que para as estimativas, as diferenças seguiram um comportamento mais homogêneo. Na região mais escura do mapa, onde há maior concentração de diferenças negativas, a krigagem ordinário apresentou estimativas maiores do que a krigagem universal. Cenário oposto pode-se observar no outro extremo do mapa, na região mais clara, onde há maior concentração de diferenças positivas, neste caso as estimativas da krigagem universal foram maiores. A região central do mapa, onde se concentram os valores amostrados, nota-se que as diferenças são próximas de zero.

Avaliando as diferenças entre os erros, nota-se que os métodos universal e ordinário obtiveram resultados parecidos, já que as diferenças estiveram de maneira geral bem próximas de zero. Esse resultado mostra que em questão de qualidade de estimativa, as krigagens ordinária e universal tiveram resultados bem similares.



### 5.3.2 Estimação do nível de concentração de Cobre - **ltpb**

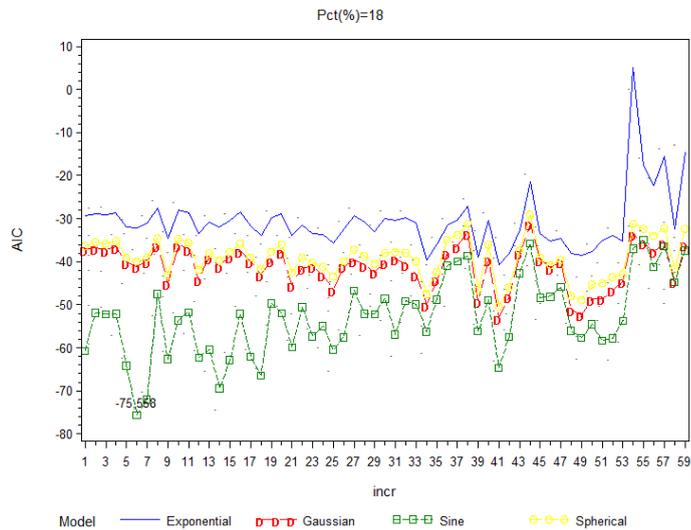
Da amostra original de 155 observações foi selecionada uma amostra de 52 observações para simular uma situação de subamostragem. A Tabela 5.7 apresenta a análise descritiva para os valores amostrados de concentração de cobre no solo nessa subamostra. Neste estudo de caso trabalhou-se com a transformação logarítmica da variável original (**ltpb**).

Tabela 5.7: Estatísticas Descritivas do log de cobre

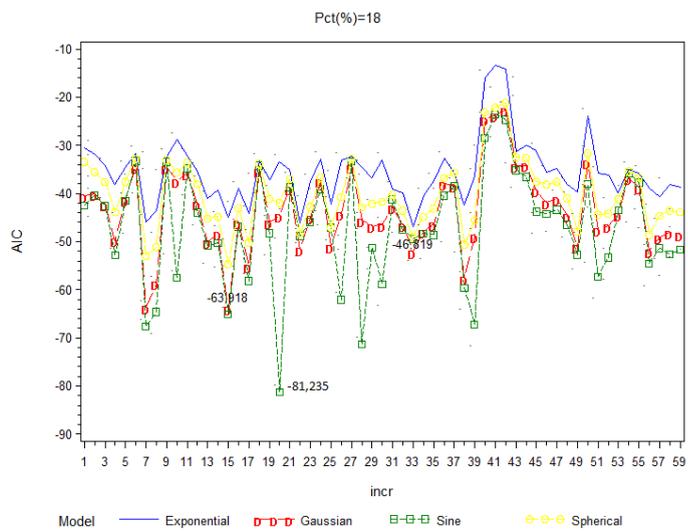
Estatística	<b>ltpb</b>
N	52
Média	2,11
Desvio Padrão	0,30
Mínimo	1,68
Máximo	2,68

A Figura 5.10 apresenta a relação de modelos teóricos testados no ajuste do semivariograma, indicando a variação do AIC do modelo para diversos valores de incremento na distância dos lags. Tanto para o ajuste do semivariograma dos dados originais como para o ajuste do semivariograma dos resíduos a porcentagem de afrouxamento que apresentou os melhores resultados foi igual a 18%.

Os resultados das simulações apresentados na Figura 5.10 indicam que o modelo teórico que apresentou menor AIC para o semivariograma dos dados originais e para o semivariograma dos resíduos foi o modelo Seno, iguais a  $-75,55$  e  $-81,23$  respectivamente. Para o caso do ajuste do semivariograma dos dados originais o incremento que apresentou melhor resultado foi igual a 6, já para o semivariograma dos resíduos o incremento que gerou melhor resultado foi igual a 20. Com a definição do modelo teórico e dos valores ótimos para porcentagem de afrouxamento e nível de



(a) Modelos simulados para ajuste do semivariograma dos dados originais

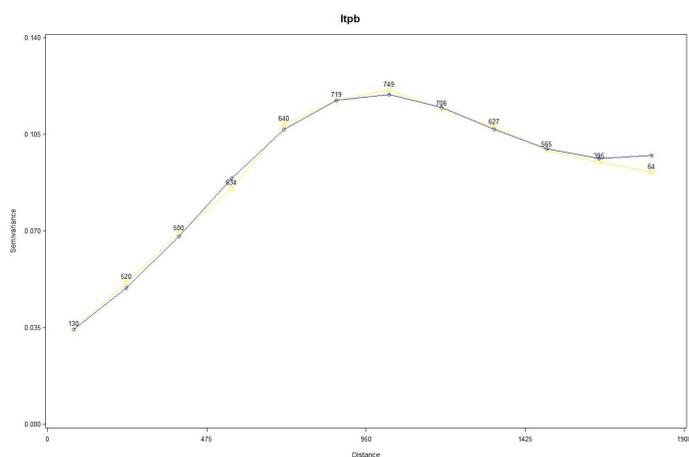


(b) Modelos simulados para ajuste do semivariograma dos resíduos

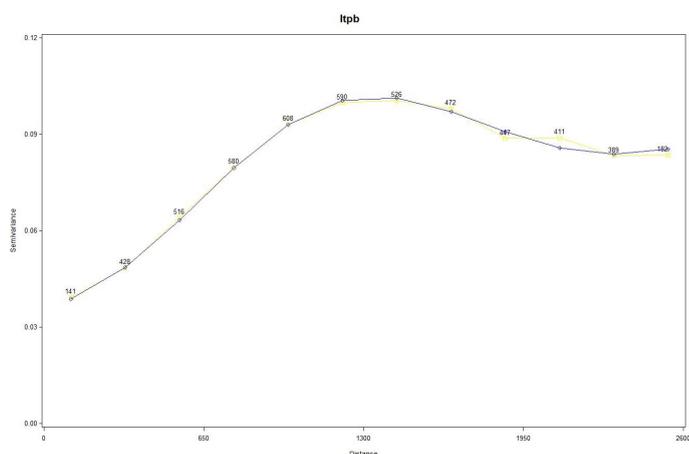
Figura 5.10: Modelos teóricos simulados para ajuste dos semivariogramas

incremento, ajustou-se os melhores modelos para os semivariogramas experimentais, apresentados na Figura 5.11.

Os modelos ajustados tiveram resultados bastante satisfatórios, uma vez que conseguem acompanhar bem o comportamento dos semivariogramas experimentais. Com a definição da estrutura formal que caracteriza a variância espacial dos dados, partiu-se para estimação dos níveis de concentração de cobre no solo através dos



(a) Modelos Seno ajustado para o semivariograma dos dados originais



(b) Modelo Seno ajustado para o semivariograma dos resíduos

Figura 5.11: Modelos teóricos ajustados para os semivariogramas

métodos de krigagem ordinária e universal.

A Figura 5.12 apresenta os valores estimados para os níveis de concentração de cobre. Os pontos em preto são referentes aos locais amostrados. Os mapas de estimativas para **ltpb** mostram que os métodos ordinário e universal geraram resultados distintos. No mapa da krigagem ordinária as estimativas variam entre 1,64 e 2,55 e é possível visualizar regiões de valores bem definidas, com a formação de áreas circulares ao longo do mapa. Já no mapa da krigagem universal a amplitude das estimativas foi superior, variando entre 1,2 e 3,18. Ao contrário do método ordinário,

o comportamento das estimativas apresentaram uma característica ressonante, com a formação de faixas de valores bem definidas que seguem uma tendência de crescimento ao longo da região analisada. Esse resultado pode ser atribuído a propriedade básica do modelo universal, que considera um fator de tendência nos dados.

## Estimativas de Krigagem Ordinária - *ltpb*      Estimativas de Krigagem Universal - *ltpb*

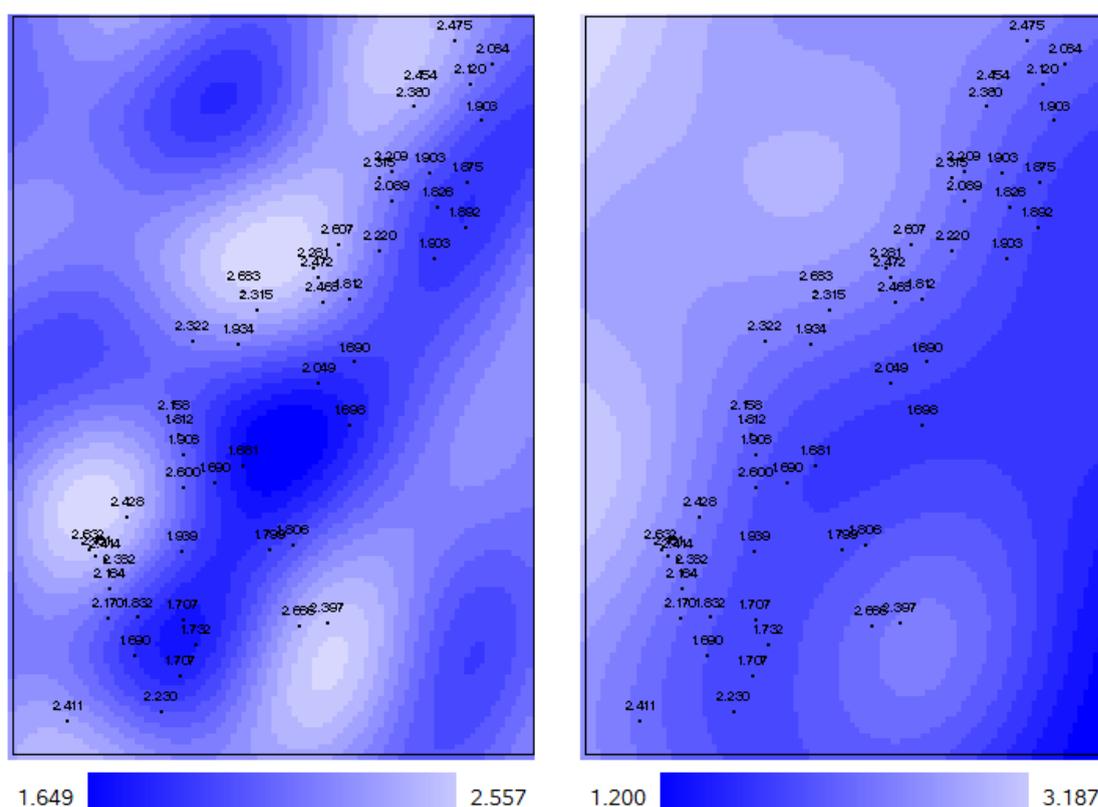


Figura 5.12: Níveis de concentração de cobre no solo estimados

A Figura 5.13 apresenta o erro padrão das estimativas obtidas em cada método. Os mapas do erro padrão também apresentaram comportamentos distintos para os dois métodos. Os erros obtidos pela krigagem ordinária variaram entre 0,19 e 0,32, sendo que os menores erros foram obtidos nas regiões que cercam os pontos observados de *ltpb*. Já para o método de krigagem universal, os menores erros também

foram observados ao longo da região em que foram amostrados observações de concentração de cobre, mas com uma abrangência de área maior, conseguindo esticar um pouco essa regiões de valores mínimos. Os valores de erro no caso universal variaram entre 0,20 e 0,54, ou seja, houve um aumento na amplitude de valores ao considerarmos o modelo com componente de tendência. Entretanto, esse aumento não é predominante no mapa e observa-se apenas nas regiões mais afastadas.

### Erro Padrão das Estimativas de Krigagem Ordinária

### Erro Padrão das Estimativas de Krigagem Universal

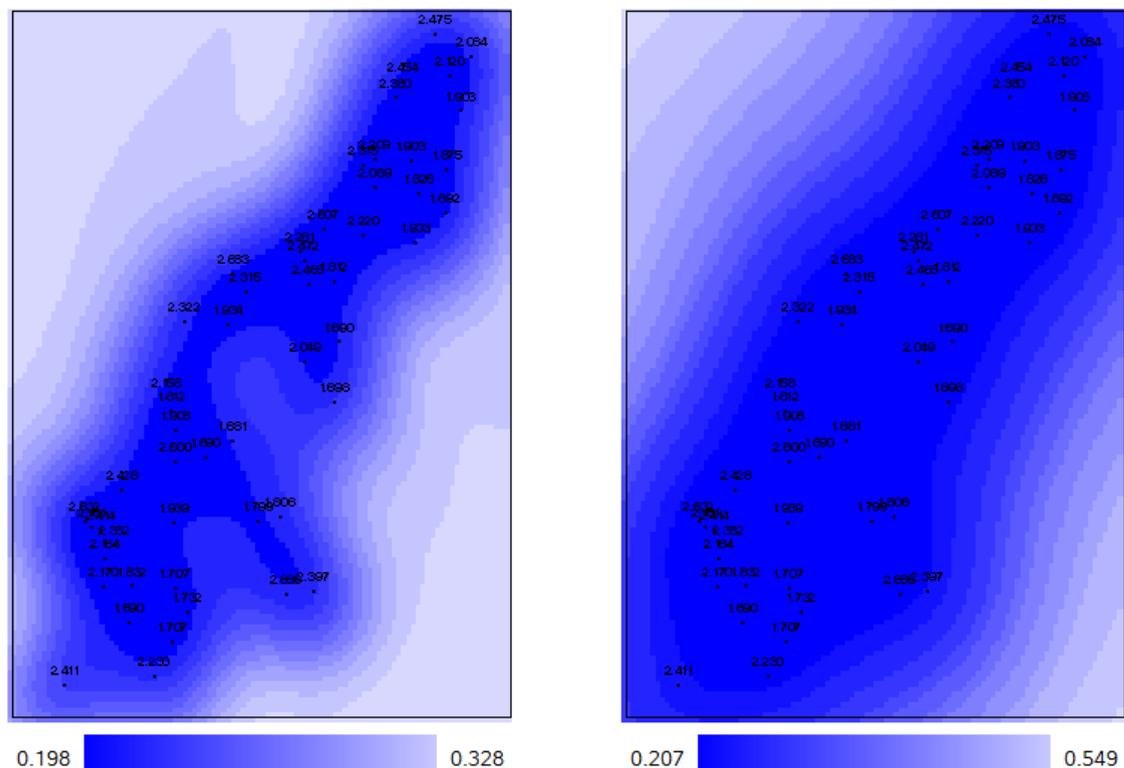


Figura 5.13: Erro Padrão das concentrações de cobre estimadas

Nota-se que os erros para krigagem ordinária tendem a estabilizar nas regiões mais extremas do mapa enquanto que os erros da krigagem universal apresentam a característica ressonante que já havia sido mostrada no mapa de estimativas, com faixas de crescimento gradativas ao longo do mapa. Novamente pode ser enten-

dido essa característica do caso universal como consequência da característica não estacionário do seu modelo.

### Comparação entre krigagem universal e krigagem ordinária

Para avaliar o efeito da inclusão da componente de tendência na estimação, foram calculadas as diferenças entre as estimativas e os erros obtidos pelos métodos de krigagem universal e ordinária. As diferenças são definidas como:

$$\text{Diferença Estimativas} = \text{Estimativa Krig.Universal} - \text{Estimativa Krig.Ordinária}$$

$$\text{Diferença Erros} = \text{Erro Krig.Universal} - \text{Erro Krig.Ordinária}$$

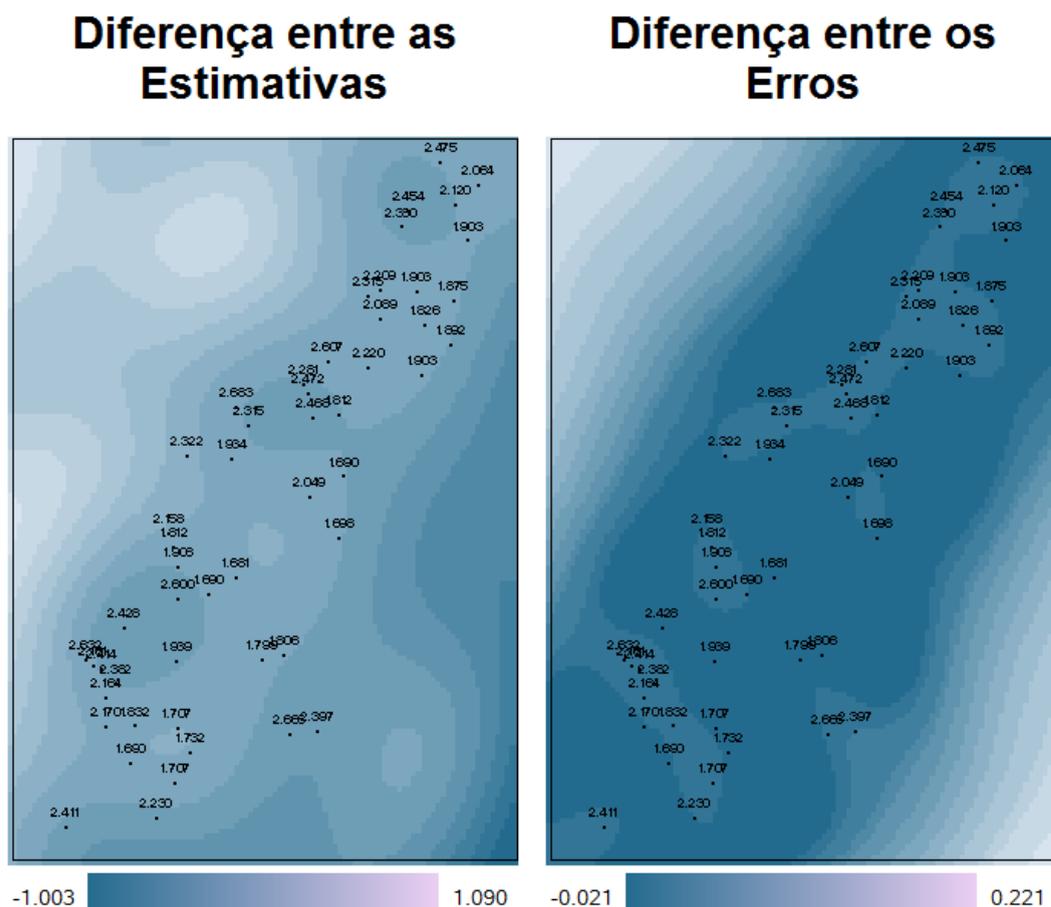


Figura 5.14: Diferença entre estimativas e entre erros dos métodos universal e ordinária para **ltpb**

Os mapas das diferenças apresentados na Figura 5.14 mostram que para as estimativas, as diferenças observadas entre os dois métodos foram significativas, praticamente variando entre  $-1$  e  $1$ . Fica claro que as maiores diferenças se encontram nas regiões extremas do mapa, onde prevalecem áreas mais escuras, locais em que as estimativas de krigagem ordinária foram maiores, e áreas mais claras, onde as estimativas de krigagem universal foram maiores. Na região central, onde estão os valores amostrados de **ltpb**, fica mais evidente que essa diferença diminui e se aproxima de zero, o que indica nesses casos que os dois métodos obtiveram estimativas mais próximas.

Avaliando as diferenças entre os erros, nota-se que os valores variaram entre  $-0,021$  e  $0,221$ , com a predominância de uma área central bem escura no mapa, o que indica a maioria de diferenças negativas. Nessa região mais escura, as estimativas de krigagem universal obtiveram melhor qualidade, já que apresentaram menores erros. Fica claro que nas regiões ao redor das observações de **ltpb**, a coloração fica mais clara, o que indica que as diferenças tendem a zero ou a um valor positivo. Pode-se entender que nessas localizações o ganho de qualidade por parte da krigagem universal diminui e os dois métodos geram resultados equivalentes.

## 5.4 ANÁLISE UTILIZANDO OS MÉTODOS DE COKRIGAGEM ORDINÁRIA E COKRIGAGEM UNIVERSAL

A aplicação dos métodos de cokrigagem visam a obtenção das melhores estimativas para a variável de interesse com a inclusão de uma ou mais covariáveis na

estimação. Assim como no caso da krigagem, é muito importante realizar um ajuste satisfatório das estruturas de variograma e variograma cruzado das variáveis, assim como selecionar covariáveis que possuam alta correlação com a variável de interesse para justificar sua inclusão na estimação.

Nesta análise, definiu-se a concentração de cobre (**ltpb**) como variável de interesse para estimação e será utilizada como covariável os níveis de concentração de zinco (**ltzn**).

#### 5.4.1 Estimação do nível de concentração de Cobre - ltpb utilizando o nível de concentração de zinco - ltzn como covariável

Inicialmente foi calculada a matriz de correlação entre as variáveis **ltom** e **ltzn** com a variável de interesse **ltpb**, com o objetivo de definir a covariável a ser utilizada na cokrigagem.

Tabela 5.8: Matriz de Correlações

	Coeficientes de correlação p-valor sob $H_0 : Rho = 0$		
	<b>ltpb</b>	<b>ltom</b>	<b>ltzn</b>
<b>ltpb</b>	1,00	0,52980 < 0,0001	0,96716 < 0,0001
<b>ltom</b>	0,52980 < 0,0001	1,00	0,58669 < 0,0001
<b>ltzn</b>	0,96716 < 0,0001	0,58669 < 0,0001	1,00

A Tabela 5.8 mostra que a variável Zinco é a que apresenta a maior correlação com a variável de interesse **ltpb**. A concentração de Material Orgânica também possui correlação significativa com **ltpb**, mas em menor intensidade. Como os valores de **ltzn** e **ltom** foram amostrados nos mesmo pontos, definiu-se apenas a variável **ltzn**

como covariável a ser utilizada na cokrigagem. Para a análise foi selecionada uma amostra de 103 observações de **ltzn** observadas em localidades diferentes das 52 observações consideradas para a variável de interesse **ltpb**.

Para o ajuste dos modelos teóricos para os variogramas e os variogramas cruzados, utilizou-se os parâmetros ótimos para **ltpb** obtidos na simulação mostrada na Figura 5.10. No caso da cokrigagem, o modelo teórico utilizado no ajuste do semivariograma da variável principal também é utilizada no ajuste do semivariograma da covariável e do variograma cruzado, neste caso, o modelo Seno. Como as observações da covariável foram amostradas em locais diferentes das observações de **ltpb**, calculou-se o pseudo-variograma cruzado.

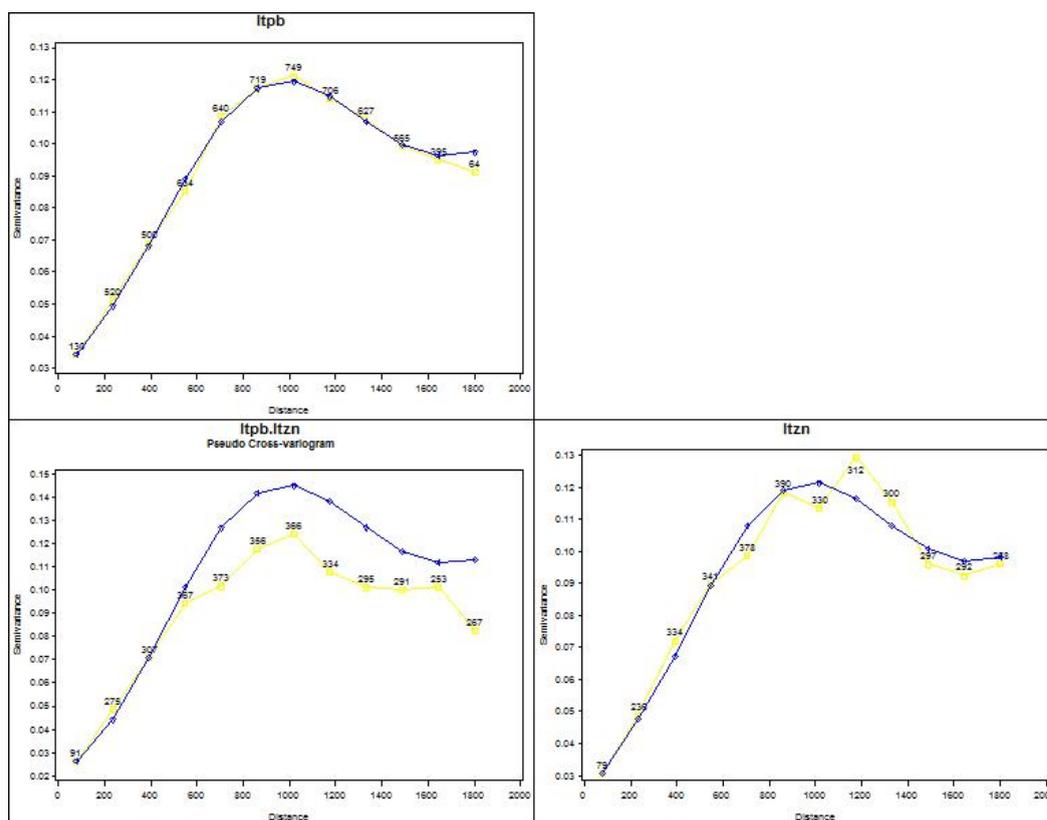


Figura 5.15: Semivariogramas e Pseudo-Variograma Cruzado dos dados originais ajustados a partir do modelo Seno (Covariável **ltzn**)

Os modelos ajustados para os semivariogramas apresentados nas Figuras 5.15 e 5.16 tiveram resultados bastante satisfatórios. Já os modelos ajustados para os pseudo-variogramas cruzado tanto para os dados originais como para os resíduos apresentaram patamares superiores ao patamar observado nas estruturas experimentais de variância cruzada, representada pela linha amarela. Com a definição dos modelos teóricos, realizou-se a estimação de **ltpb**.

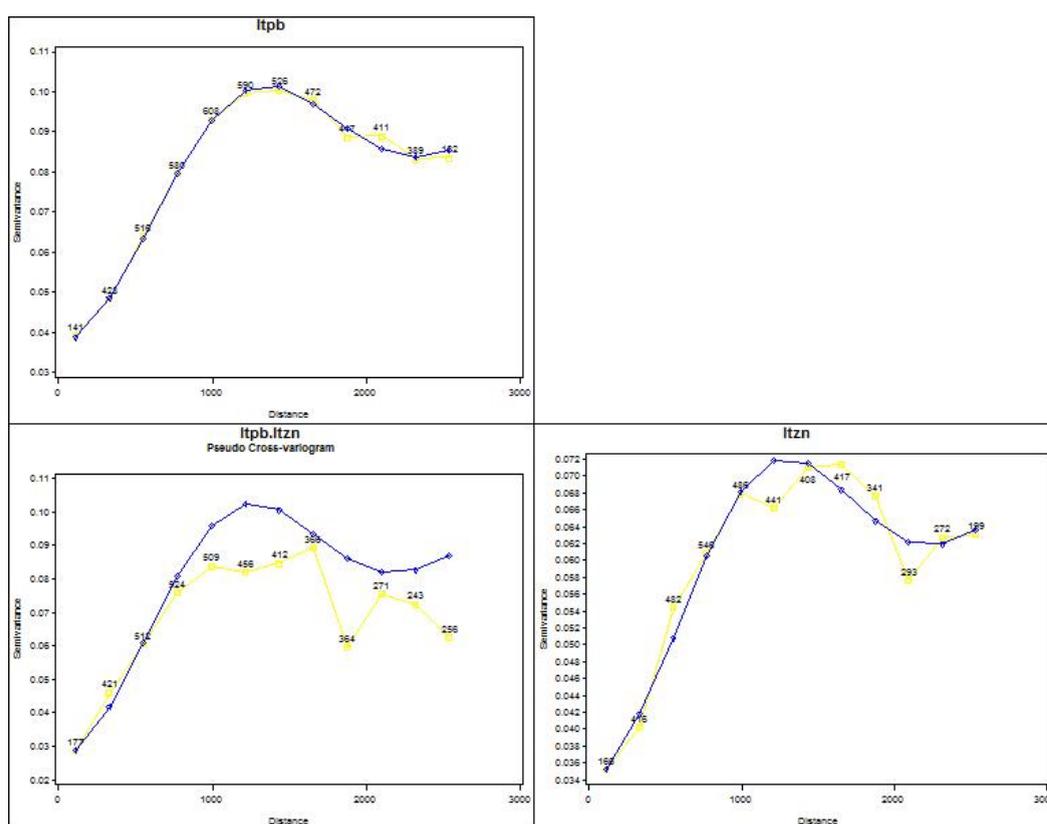


Figura 5.16: Semivariogramas e Pseudo-Variograma Cruzado dos resíduos ajustados a partir do modelo Seno (Covariável **ltn**)

A Figura 5.17 apresenta as estimativas de cokrigagem para os métodos ordinário e universal. Os pontos em preto representam os valores amostrados de **ltpb**, enquanto que os pontos em vermelho representam as observações da covariável **ltn**.

Assim como no caso da krigagem, os mapas de estimativas dos dois métodos

## Estimativas de CoKrigagem Ordinária      Estimativas de CoKrigagem Universal

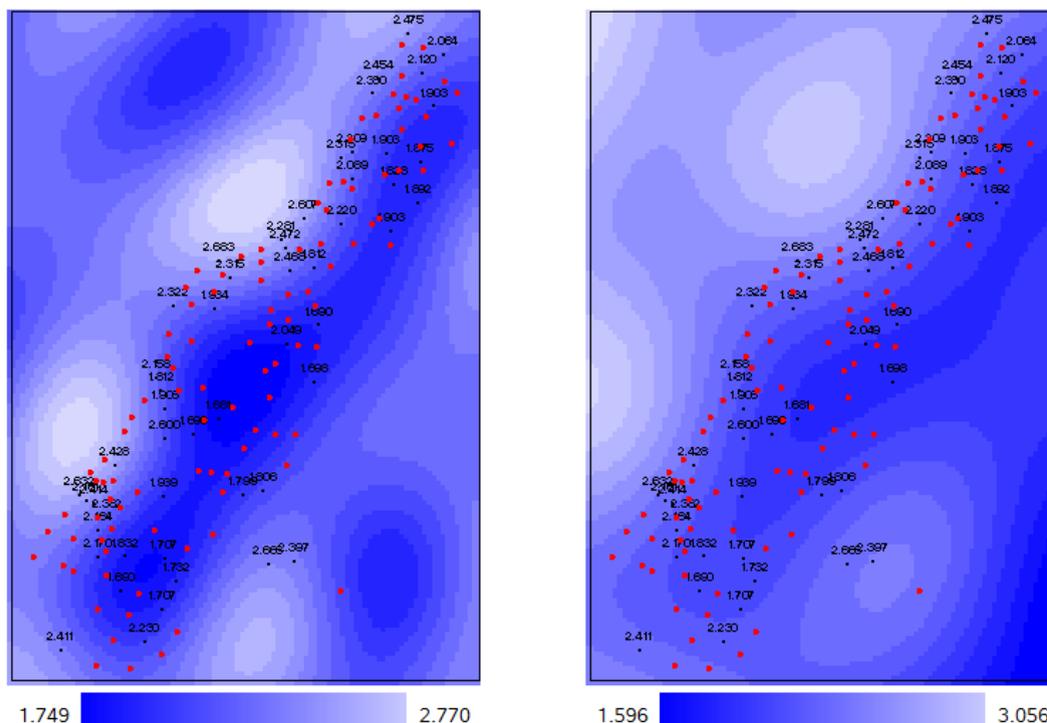


Figura 5.17: Níveis de concentração de cobre no solo estimados por cokrigagem - (Covariável **Itzn**)

tiveram comportamentos distintos. A cokrigagem ordinária obteve estimativas variando entre 1,74 e 2,77. Nota-se a formação de áreas de valores bem definidas, com a presença de regiões circulares no mapa, padrão que havia sido observada nas estimativas de krigagem ordinária. A cokrigagem universal obteve estimativas variando entre 1,59 e 3,05, portanto a amplitude de valores foi maior do que o observado no método ordinário. O comportamento das estimativas apresentaram uma característica ressonante, com a formação de faixas de valores bem definidas que seguem uma tendência de crescimento ao longo da região analisada, bem parecida com as estimativas da krigagem universal.

A Figura 5.18 apresenta o erro padrão das estimativas obtidas em cada método.

Os mapas do erro padrão também apresentaram comportamentos distintos para os dois métodos. Os erros obtidos pela cokrigagem ordinária variaram entre 0,20 e 0,33, com os menores valores sendo observados nas localizações onde foram amostradas as concentrações de **ltpb** e da covariável **ltzn**. Já para a cokrigagem universal, os menores erros também foram observados ao longo da região em que foram amostrados observações da variável de interesse e da covariável, mas com uma abrangência maior para essa área de valores mínimos. Os valores de erro no caso universal variaram entre 0,22 e 0,46, ou seja, houve uma aumento nos erros associados às estimativas.

### Erro Padrão das Estimativas de CoKrigagem Ordinária      Erro Padrão das Estimativas de CoKrigagem Universal

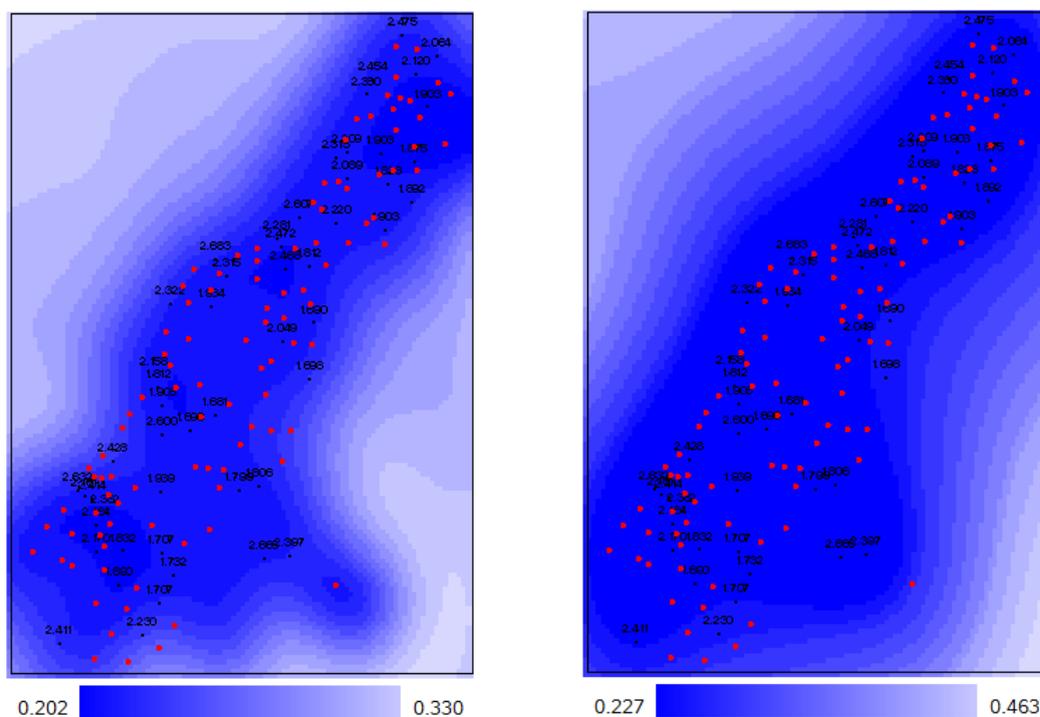


Figura 5.18: Erro Padrão das concentrações de cobre estimadas por cokrigagem - (Covariável **ltzn**)

### Comparação entre cokrigagem universal e cokrigagem ordinária

Para avaliar o efeito da inclusão da componente de tendência na estimação,

foram calculadas as diferenças entre as estimativas e os erros obtidos pelos métodos de cokrigagem universal e ordinária. As diferenças são definidas como:

$$\text{Diferença Estimativas} = \text{Estimativa Cokrig. Universal} - \text{Estimativa Cokrig. Ordinária}$$

$$\text{Diferença Erros} = \text{Erro Cokrig. Universal} - \text{Erro Cokrig. Ordinária}$$

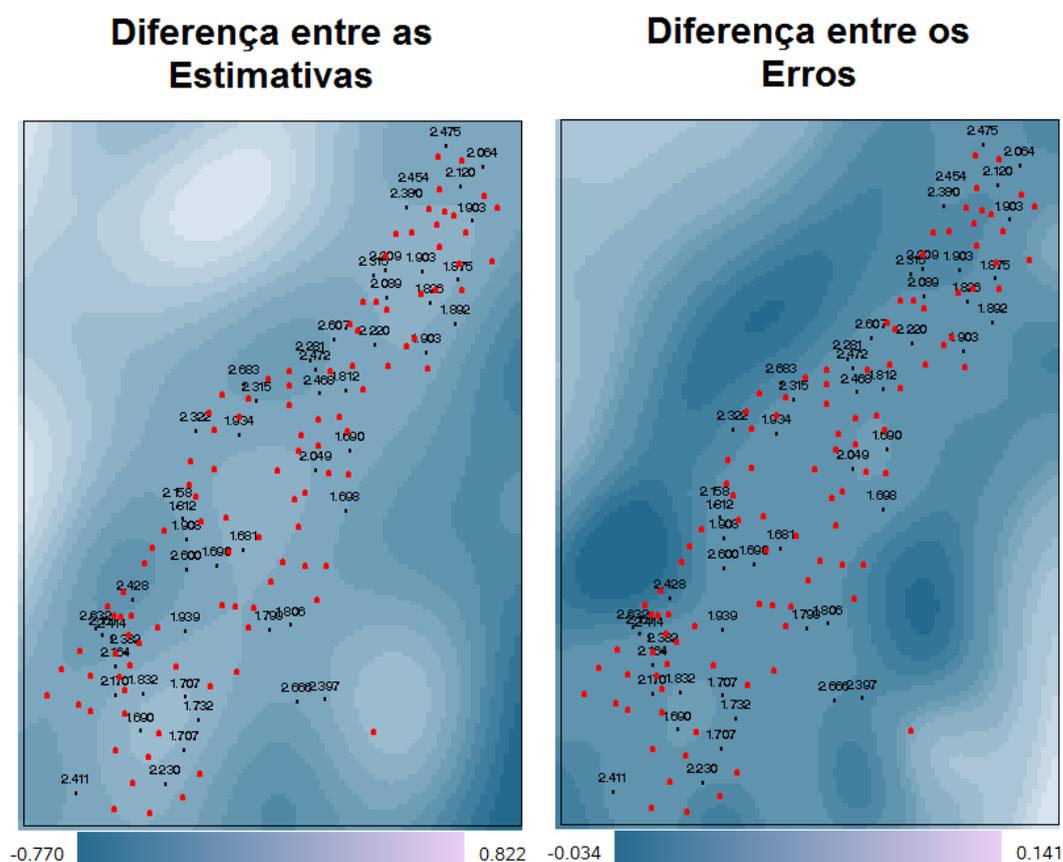


Figura 5.19: Diferença entre estimativas e entre erros dos métodos universal e ordinária para **ltpb** - Covariável **ltzn**

Os mapas das diferenças apresentados na Figura 5.19 mostram que para as estimativas, as diferenças observadas entre os dois métodos foram significativas, praticamente variando entre  $-0,7$  e  $0,8$ . Não é possível identificar um padrão de comportamento que se destaque para essas diferenças. De certo modo, podemos identificar que as diferenças mais extremas, tanto positivas como negativas, estão nas regiões

mais afastadas do mapa. Na região central, onde estão os valores observados de **ltpb** e **ltzn**, nota-se uma coloração mais intermediária, o que mostra indícios que nessa área as estimativas para dois métodos tendem a ser mais próximas.

Avaliando as diferenças entre os erros, nota-se que os valores variaram entre  $-0,034$  e  $0,221$ , com a predominância de uma área central mais clara no mapa, o que indica a maioria de diferenças positivas nessa região. As regiões mais escuras, que indicam diferenças negativas, são bem limitadas. Fica claro que nas regiões ao redor das observações de **ltpb** e **ltzn**, onde há predominância de diferenças positivas, as estimativas obtidas por cokrigagem ordinária possuem menores erros e portanto melhor qualidade do que as estimativas obtidas pela cokrigagem universal. Aparentemente a melhora na qualidade das estimativas com o uso do modelo universal mostrada na análise de krigagem não foi observada neste caso pois os modelos de cokrigagem já possuem a contribuição de uma covariável para melhorar a estimação em locais em que não se observa a variável principal, desse modo a contribuição gerada pelo modelo com tendência não é tão sentida.

## 5.5 COMPARAÇÃO ENTRE OS MÉTODOS ORDINÁRIO E UNIVERSAL PARA KRIGAGEM E COKRIGAGEM

Foi realizada uma comparação entre a qualidade das estimativas obtidas para **ltpb**, para isso avaliou-se as diferenças entre os erros observados na aplicação das técnicas de estimação.

Nos mapas contidos na Figura 5.20 foram calculadas as diferenças dos erros

entre as técnicas de krigagem e cokrigagem, tanto para o método ordinário como para o método universal. Na situação onde os valores da diferença são negativos, a técnica de cokrigagem apresentou melhor qualidade, já para diferenças positivas, é a krigagem que apresenta melhores resultados. Novamente a variável principal está na cor preta e a covariável está na cor vermelha.

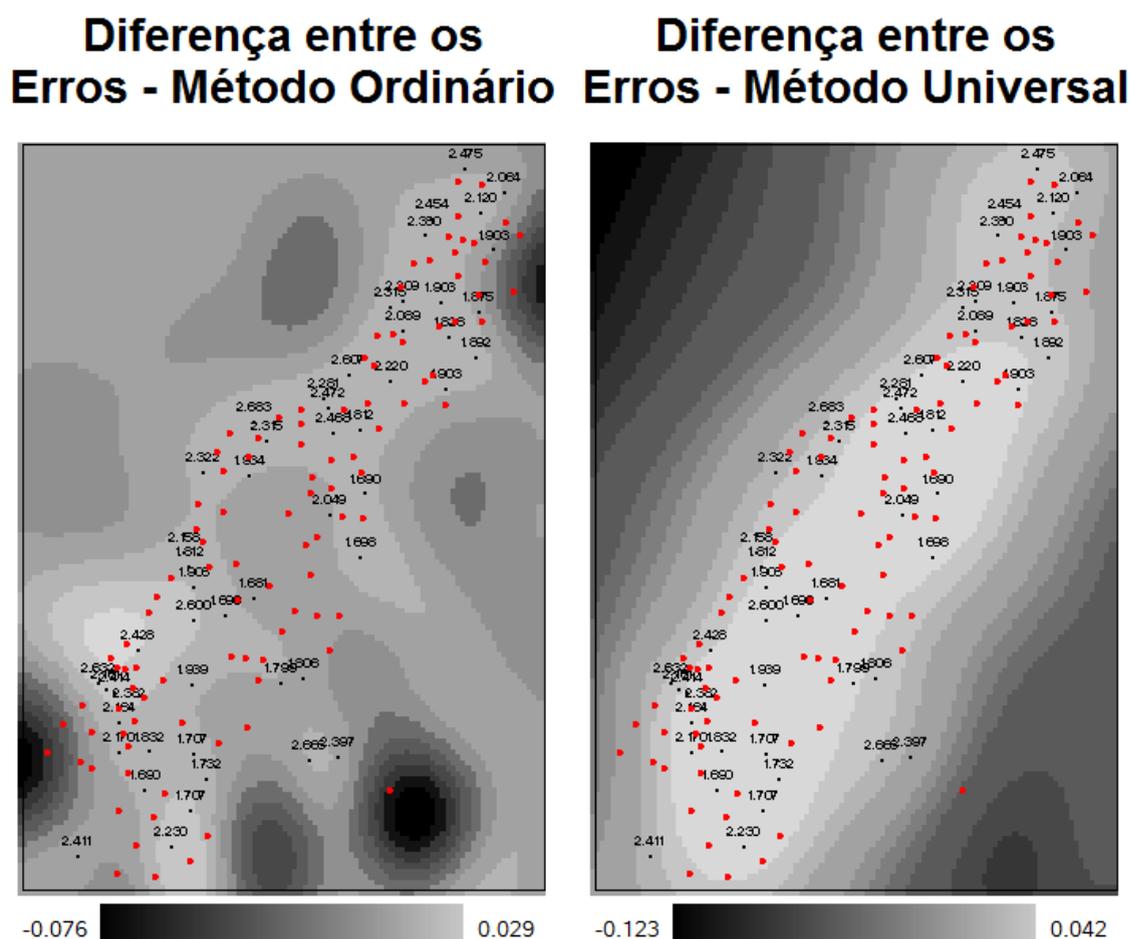


Figura 5.20: Diferença entre os erros de krigagem e cokrigagem - (Covariável **ltzn**)

Para o método ordinário, as diferenças variaram entre  $-0,07$  e  $0,02$ . Mesmo com a variação dessas diferenças estando bem próxima de zero, as maiores diferenças foram negativas, o que mostra uma indicação geral de que as estimativas da cokrigagem apresentaram melhor qualidade. Nas regiões em que há observações da variável

principal, nota-se que o o ganho de qualidade pela técnica de cokrigagem diminui, já que essas regiões estão mais claras. Para as regiões onde há menor concentração de observações da variável principal e maior presença de observações da covariável, o mapa tende a escurecer, o que nesses casos indica maior ganho de qualidade na estimativa por parte da técnica de cokrigagem. Esse comportamento era esperado, já que a ideia principal da cokrigagem é melhorar a qualidade da estimação justamente nas regiões onde se tem pouca informação da variável de interesse utilizando a informação disponível de outra covariável.

Para o método universal, as diferenças variaram entre  $-0.123$  e  $0,042$ . Fica claro que de maneira geral as estimativas de cokrigagem tiveram melhor qualidade já que as maiores diferenças foram negativas. Essa melhor qualidade obtida pela técnica de cokrigagem é mais acentuada nas regiões mais extremas do mapa, onde não há observações da variável principal e da covariável. Na região central, onde estão as observações da variável principal, novamente se observa que o ganho de qualidade por parte da cokrigagem diminui. Ao contrário do método ordinário, em que as regiões mais claras ficaram bem delimitadas ao redor das observações de **ltpb**, a krigagem universal conseguiu ampliar um pouco mais a área em que o ganho da cokrigagem não é tão grande.

# Capítulo 6

## CONCLUSÕES

A implementação de um algoritmo para aplicação das técnicas de krigagem e co-krigagem, tanto para o método ordinário como para o método universal, representa um grande avanço em relação aos procedimentos existentes no *software* SAS 9.3, que possui apenas procedimento para estimação por krigagem ordinária. Os códigos desenvolvidos em SAS/IML neste trabalho são de grande valia pois ampliam as opções de técnicas para estimação espacial, são de fácil conversão para outras linguagens e permitem a possibilidade de inclusão de outros modelos teóricos.

Os resultados obtidos no balizamento do algoritmo validam o código proposto, já que as estimativas obtidas para as krigagens ordinária e universal pelo algoritmo em SAS/IML foram iguais ou bem próximos às estimativas obtidas pelas funções de estimação espacial já implementadas no R.

Os resultados obtidos na aplicação das técnicas de krigagem mostraram que o método universal gera estimativas com amplitude de valores iguais ou maiores do que o método ordinário, o que mostra uma vantagem por parte da krigagem universal ao ampliar o intervalo de valores estimados. Tanto as estimativas como os erros no caso universal mostraram uma característica ressonante, com formação de

faixas de valores ao longo da região estimada. Esse comportamento pode-se entender como consequência da suposição da componente de tendência que existe no modelo universal. Foi possível observar que a krigagem universal, quando comparada com a krigagem ordinária, gera uma melhora na qualidade das estimativas principalmente nas regiões onde estão amostradas as observações da variável de estudo, já que o modelo com tendência consegue ampliar as áreas de erro mínimo ao redor desses pontos.

As análises utilizando as técnicas de cokrigagem ordinária e cokrigagem universal mostraram cenário semelhante ao observado nas análises por krigagem. As estimativas obtidas pelo método universal tiveram amplitude de valores maior do que as estimativas observadas no método ordinário. O comportamento ressonante com formação de faixas de valores novamente foi característico para o modelo universal, o que reafirma a influência da componente de tendência sobre o comportamento das estimativas e erros gerados. Ao contrário do que foi constatado nas aplicações da krigagem, a cokrigagem universal não gerou uma melhora na qualidade das estimativas. Aparentemente, como a cokrigagem ordinária já incorpora a informação de uma covariável para melhorar a qualidade da estimação, a utilização do modelo universal com a componente de tendência não proporciona melhora substancial.

Foi possível observar na comparação das técnicas de krigagem e cokrigagem que de maneira geral a adição da covariável no modelo de estimação gerou uma pequena melhora na qualidade das estimativas obtidas. Para o modelo ordinário, fica claro que nas regiões onde há pouca informação da variável principal, a presença da co-

variável melhora a precisão da estimação. Já para o caso universal, nota-se que o ganho por parte da cokrigagem nessas regiões próximas das observações da variável principal diminui uma vez que a krigagem universal já consegue ampliar as áreas de erro mínimo sem a necessidade da informação de uma covariável.

# Referências Bibliográficas

- Bailey, T. C. & Gatrell, A. C. (1995). *Interactive Spatial Data Analysis*. Prentice Hall.
- Barreto, L. T. (2011). Estimaco da temperatura da regio amaznica via interpoladores geoestatísticos. Technical report, Departamento de Estatística, Universidade de Braslia.
- Camargo, E. C. G., Felgueiras, C. A., & Monteiro, A. M. V. (2001). A importncia da modelagem da anisotropia na distribuio espacial de variveis ambientais utilizando procedimentos geoestatísticos. Technical report, INPE.
- Chiles, J. P. & Delfiner, P. (1999). *Geostatistics Modeling Spatial Uncertainty*. John Wiley & Sons.
- Cressie, N. & Wikle, C. K. (1998). The variance-based cross-variogram: You can add apples and oranges. *Mathematical Geology*, 30(7):789–799.
- Cressie, N. A. C. (1991). *Statistic for Spatial Data*. Wiley.
- de Barros Silva, A. (2003). *Sistemas de Informaces Geo-referenciadas Conceitos e Fundamentos*. Unicamp.
- Galvo, J. A. M. (2012). Estimaco da temperatura mdia na regio amaznica por meio de tcnicas de cokrigagem. Technical report, Departamento de Estatística, Universidade de Braslia.
- Goovaerts, P. (1998). Ordinary cokriging revisited. *Mathematical Geology*, 30(1):21–42.
- Jian, X., Olea, R. A., & Yu, Y.-S. (1996). Semivariogram modeling by weighted least squares. *Computers and Geosciences*, 22:381–391.
- Kitadinis, P. K. (1993). Generalized covariance functions in estimation. *Mathematical Geology*, 25:525–540.

- Memarsadeghi, N., Moigne, J. L., Mount, D. M., & Morisette, J. (2005). A new approach to image fusion based on cokriging. In: *7th International Conference on Information Fusion*.
- Myers, D. E. (1982). Matrix formulation of co-kriging. *Mathematical Geology*, 14(3):249–257.
- R Core Team (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. ISBN 3-900051-07-0.
- Rossiter, D. (2007). Technical note: Co-kriging with the gstat package of the r environment for statistical computing. Technical report, International Institute for Geo-information Science & Earth Observation.
- SAS (2011). *SAS on line doc, Versão 9.3*. Cary, NC: SAS Institute Inc.
- Silva, A. B. (2003). *Sistema de Informações Geo-referenciadas Conceitos e Fundamentos*. Unicamp.
- Stein, A. & Corsten, L. C. A. (1991). Universal kriging and cokriging as a regression procedure. *Biometrics*, 47(2):575–587.
- Webster, R. & Oliver, M. A. (2007). *Geostatistics for Environmental Scientists*, (2nd edition ed.). John Wiley & Sons, Ltd.