



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

**ESTIMAÇÃO DA TEMPERATURA MÉDIA
NA REGIÃO AMAZÔNICA POR MEIO DE
TÉCNICAS DE COKRIGAGEM**

João Antônio Martins Galvão

09/97366

Brasília

2012

João Antônio Martins Galvão

09/97366

ESTIMAÇÃO DA TEMPERATURA MÉDIA NA REGIÃO AMAZÔNICA POR MEIO DE TÉCNICAS DE COKRIGAGEM

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

2012

Agradecimentos

Ao SAS *Institute* Brasil por possibilitar a utilização desse software por meio de parceria acadêmica com o Departamento de Estatística da UnB. Ao Instituto Nacional de Meteorologia - INMET pelo banco de dados sobre temperatura na Amazônia. Ao Instituto Nacional de Pesquisas Espaciais pelo banco de dados sobre a precipitação média e área desmatada na Amazônia.

Agradeço aos meus familiares pelo apoio, atenção e paciência destinados a mim em todos os momentos cruciais deste período. Aos meus colegas de curso cujo a companhia e amizade me ajudaram em muitos momentos de decisão. Aos professores por todo conhecimento transmitido e em especial ao meu professor orientador, Alan Ricardo da Silva, pela ajuda e tempo dedicados a este trabalho.

João Antônio Martins Galvão

Resumo

Ocupando quase metade do território brasileiro e com uma grande biodiversidade a Amazônia é motivo de preocupação mundial, sendo o controle da temperatura na região de vital importância.

Este trabalho busca estimar a temperatura na região da Amazônia no ano de 2010. Mais do que isso, busca na implementação da técnica de Cokrigagem, na linguagem SAS/IML, apresentar uma alternativa mais abrangente nas estimações de variáveis amostradas espacialmente comparativamente à técnica de Krigagem.

Pode-se observar que a técnica de Cokrigagem não é sempre melhor que a Krigagem, porém vale ressaltar que nos locais onde a Cokrigagem foi melhor, os erros padrões da temperatura estimada foram muito menores do que quando a Krigagem foi melhor.

Sumário

RESUMO	iii
1 INTRODUÇÃO	1
1.1 OBJETIVOS	3
2 TÉCNICAS DE KRIGAGEM	4
2.1 INTRODUÇÃO	4
2.2 VARIOGRAMA	5
2.3 COVARIOGRAMA	7
2.4 KRIGAGEM ORDINÁRIA	8
2.5 COKRIGAGEM	10
2.5.1 Variograma Cruzado, Covariograma Cruzado e Pseudo- Variograma Cruzado	11
2.5.2 Cokrigagem Ordinária	12
2.6 ANISOTROPIA	14
2.6.1 Modelagem Anisotrópica	16
3 MATERIAL E MÉTODOS	19
3.1 INTRODUÇÃO	19

3.2	ESTAÇÕES METEOROLÓGICAS	19
3.3	SATÉLITES	21
3.3.1	BDQ	21
3.3.2	DETER	23
3.4	SISTEMAS DE COORDENADAS	25
3.4.1	Sistema Lat/Long	25
3.5	ALGORITMO IML	28
3.5.1	Macro Crossvariogram	28
3.5.2	Macro Cokriging	34
4	ANÁLISE DOS RESULTADOS	41
4.1	INTRODUÇÃO	41
4.2	ANÁLISE DESCRITIVA	41
4.3	ANÁLISE UTILIZANDO O MÉTODO DE COKRIGAGEM OR- DINÁRIA	46
4.4	COMPARAÇÃO ENTRE AS COVARIÁVEIS	58
4.5	RESULTADOS UTILIZANDO O MÉTODO DA KRIGAGEM OR- DINÁRIA	60
4.6	COMPARAÇÃO ENTRE KRIGAGEM E COKRIGAGEM	63
5	CONCLUSÕES	68
	Referências	70

Capítulo 1

INTRODUÇÃO

Considerada a maior floresta tropical do mundo, a Amazônia sul-americana ocupa quase metade do território brasileiro (aproximadamente 6,5 milhões de quilômetros quadrados). Além de sua grande extensão, a Amazônia brasileira também é dona de uma enorme biodiversidade, possui 1,5 milhões de espécies vegetais catalogadas, três mil espécies de peixes, 950 tipos de pássaros, além de insetos, répteis, anfíbios e mamíferos (IBAMA, 2010).

A grande importância da floresta, além da enorme diversidade vegetal, animal, e mineral está na sua capacidade de absorção de gás carbônico (CO_2) fundamental no equilíbrio climático do planeta. Apesar de sua grande exuberância, a formação geológica e relevo diferenciado resultaram em diferentes formações de classes de solo que, sob as características do clima equatorial quente superúmido e úmido, proporcionaram baixa fertilidade natural do solo. A Amazônia também apresenta um delicado equilíbrio ecológico muito sensível a ações do homem (IBAMA, 2010).

A ocupação da Amazônia teve início no ano de 1540 com os colonizadores europeus, porém até o começo da década de 1970, a floresta permanecia sem grandes alterações em sua cobertura vegetal. Após a inauguração da rodovia Transamazônica

e a política de integração, a ocupação foi intensificada e, conseqüentemente, o desmatamento também (Fearnside, 2005).

A utilização do solo na Amazônia é caracterizado pelas atividades de extrativismo vegetal e animal, pela pecuária e pela produção de grãos. Tanto a pecuária quanto a produção de grãos estão altamente relacionados com os incêndios florestais, utilizados como forma barata e eficiente de limpeza de solo, o que causa aumento nas emissões de gases do efeito estufa (Fearnside, 2005).

A emissão e acumulação de gases poluentes na atmosfera é conhecida como efeito estufa. Apesar de muitas controvérsias, algumas entidades ecológicas consideram as queimadas da Amazônia como um dos principais causadores do aquecimento global, fenômeno que, quando ocorre em excesso, provoca um superaquecimento do planeta resultando, entre outros fatores, em desequilíbrio climático (Souza, 2011)

Segundo Barreto (2011), a temperatura média na região amazônica aumentou cerca de 1°C ao longo das décadas de 1970 a 2010, porém não foi possível detectar uma relação entre este aumento de temperatura e a incidência de desmatamento na região, por exemplo. Assim, este trabalho busca analisar essa questão através de uma técnica mais abrangente, a cokrigagem, que utiliza covariáveis nas suas estimações.

1.1 OBJETIVOS

O objetivo geral do trabalho é estimar a temperatura média mensal por meio da técnica de cokrigagem utilizando como covariáveis a precipitação média mensal e a área de desmatamento na região amazônica.

Os objetivos específicos são:

- utilizando o *software* SAS 9.2, desenvolver um algoritmo em IML para aplicação do modelo de cokrigagem;
- comparar os modelos de krigagem com o modelo de cokrigagem .

Capítulo 2

TÉCNICAS DE KRIGAGEM

2.1 INTRODUÇÃO

A estatística espacial está sendo desenvolvida desde o século passado. A partir de 1963, na França, George Matheon, em parceria com outros pesquisadores desenvolveu uma interessante técnica de estimação conhecida como krigagem, termo escolhido em homenagem ao sulafricano, Daniel G. Krige, pioneiro no uso do modelo (Stein and Corsten, 1991).

Segundo Cressie (1991), a palavra krigagem é sinônimo de predição ótima, ou seja, a krigagem consiste em fazer inferências sobre valores não observados de uma variável aleatória a partir de valores observados em localizações espaciais desconhecidas.

A krigagem divide-se em quatro tipos principais: a krigagem simples em que a média dos valores observados é conhecida e constante; a krigagem ordinária em que a média é desconhecida porém constante; e a krigagem universal em que a média é desconhecida e não constante. Existe ainda um quarto tipo, uma extensão da krigagem, a cokrigagem, em que uma covariável altamente correlacionada à variável

em estudo é adicionada, melhorando a qualidade das estimativas. Neste trabalho serão explicados os métodos de krigagem ordinária e, mais adiante, a técnica de cokrigagem. Estudaremos também o caso de anisotropia em que a estrutura de variância não depende somente da distância entre os pontos mas, também da direção, que deve ser levada em consideração.

Porém, para entendermos o modelo de krigagem ordinária, é necessário entender como funcionam duas importantes ferramentas geoestatísticas: o variograma e o covariograma.

2.2 VARIOGRAMA

O variograma é uma importante ferramenta geoestatística, com ele é possível modelar a estrutura de covariância espacial dos dados em estudo. Utilizaremos Z como a variável em estudo, \mathbf{u} como a coordenada no espaço e portanto $Z(\mathbf{u})$ será a variável Z observada no ponto \mathbf{u} . O cálculo do variograma ($2\gamma(h)$) entre dois pontos $Z(\mathbf{u})$ e $Z(\mathbf{u} + \mathbf{h})$ separados por um vetor de distância \mathbf{h} é definido como:

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

O estimador do variograma será:

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

onde $N(h)$ é o número de pares de valores medidos. Na prática, utilizamos o estimador do semivariograma, definido por:

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

A figura a seguir ilustra um semivariograma experimental e seus parâmetros:

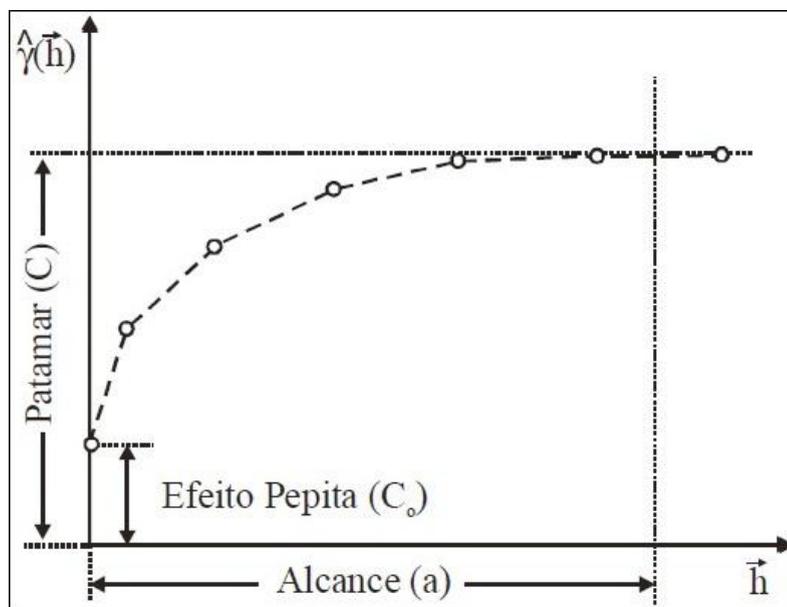


Figura 2.1: Semivariograma

- Alcance (a): distância em que as amostras apresentam correlação espacial.
- Patamar (C): valor do semivariograma correspondente a seu alcance. A partir deste ponto considera-se que não existe dependência espacial, pois a variância da diferença entre pares de amostras ($Var[Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})]$) torna-se constante.
- Efeito Pepita (C_0): em um semivariograma ideal, $\gamma(0) = 0$ porém, na prática, à medida que \mathbf{h} tende para zero, $\gamma(h)$ se aproxima de um valor maior que zero chamado Efeito Pepita. Esse parâmetro revela a descontinuidade do semivariograma para distâncias menores que a menor distância \mathbf{h} .

Para a construção do semivariograma, deve ser determinado um *lag* de distância para a construção de classes de distâncias. Esse *lag* pode ser obtido através de histogramas para avaliar se o número de pontos em cada classe é satisfatório.

Existem vários modelos teóricos para ajustar a tendência de variabilidade apre-

sentada pelo semivariograma, dentre eles, os principais são os modelos gaussiano, esférico e seno. Para mais informações sobre tais modelos, consultar Barreto (2011).

2.3 COVARIOGRAMA

O covariograma e sua forma normalizada são os métodos mais intuitivos de sintetizar a estrutura de dependência espacial. Se $Z(\cdot)$ é um processo estacionário, então a covariância entre dois pontos $Z(s_1)$ e $Z(s_2)$ é definida como:

$$Cov[Z(s_1), Z(s_2)] = C(s_1 - s_2) \quad (2.4)$$

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

Podemos relacionar o covariograma ao variograma. Se $\|s_1 - s_2\| = h$, então por (2.4):

$$Cov[Z(s_1), Z(s_2)] = C(h) \quad (2.5)$$

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

$$Var(Z(s_1)) = Var(Z(s_2)) = C(0) \quad (2.6)$$

A partir da relação:

$$Var(Z(s_1) - Z(s_2)) = Var(Z(s_1)) + Var(Z(s_2)) - 2Cov[Z(s_1), Z(s_2)] \quad (2.7)$$

Temos, para $\|s_1 - s_2\| = h$, de (2.6) e da definição de variograma:

$$2\gamma(h) = 2(C(0) - C(h)) \quad (2.8)$$

Note que quando $h \rightarrow \infty$, então $2\gamma(h) \rightarrow 2C(0)$, dessa forma $C(0)$ é o patamar do variograma.

O estimador do covariograma é dado por:

$$\hat{C}(h) = \frac{1}{|N(h)|} \sum_{N(h)} (Z(s_i) - \bar{Z})(Z(s_j) - \bar{Z}) \quad (2.9)$$

onde:

$$\bar{Z} = \sum_{i=1}^n \frac{Z(s_i)}{n} \quad (2.10)$$

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

O gráfico do covariograma fica o contrário do variograma, enquanto a variância aumenta à medida que a distância aumenta, a covariância diminui.

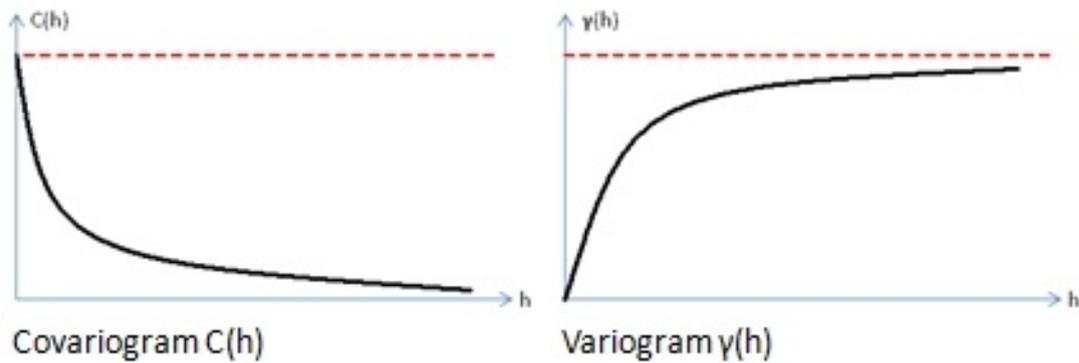


Figura 2.2: Covariograma e Variograma

2.4 KRIGAGEM ORDINÁRIA

A krigagem ordinária é a extensão natural da krigagem simples em que a média dos valores dos pontos é desconhecida, porém constante, ou seja, é um processo estacionário. O estimador geral para o valor da característica Z no ponto u_o é dado por:

$$Z^*(\mathbf{u}_0) = \lambda_0 + \sum_{i=1}^n \lambda_i Z(\mathbf{u}_i) \quad (2.11)$$

Na krigagem ordinária para que a média seja igual, devemos fazer $\lambda_0 = 0$ e $\sum_{i=1}^n \lambda_i = 1$, logo o estimador será:

$$Z^*(\mathbf{u}_0) = \sum_{i=1}^n \lambda_i Z(\mathbf{u}_i) \quad (2.12)$$

com $\sum_{i=1}^n \lambda_i = 1$.

Devemos agora minimizar a variância do erro ($Var(Z_0^* - Z_0)$) sujeito a $\sum_{i=1}^n \lambda_i =$

1. Utilizando o variograma definido em (2.1) temos que minimizar:

$$E \left(\sum_{i=1}^n \lambda_i Z(u_i) - Z(u_0) \right)^2 - 2\alpha \left(\sum_{i=1}^n \lambda_i - 1 \right) \quad (2.13)$$

onde α é o multiplicador de Lagrange que garante $\sum_{i=1}^n \lambda_i = 1$. Desta mesma condição temos:

$$\left(\sum_{i=1}^n \lambda_i Z(u_i) - Z(u_0) \right)^2 = - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j [Z(u_i) - Z(u_j)]^2 + 2 \sum_{i=1}^n \lambda_i [Z(u_0) - Z(u_i)]^2. \quad (2.14)$$

Substituindo (2.14) em (2.13) temos:

$$- \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(u_i - u_j) + 2 \sum_{i=1}^n \lambda_i \gamma(u_0 - u_i) - 2\alpha \left(\sum_{i=1}^n \lambda_i - 1 \right) \quad (2.15)$$

derivando em relação a $\lambda_1, \dots, \lambda_n$ e igualando a zero, obtém-se:

$$- \sum_{j=1}^n \lambda_j \gamma(u_i - u_j) + \gamma(u_0 - u_i) - \alpha = 0. \quad (2.16)$$

Obtemos os pesos λ_i através do sistema de equações, conhecido como sistema de krigagem ordinária:

$$\begin{cases} - \sum_{j=1}^n \lambda_j \gamma(u_i - u_j) + \gamma(u_0 - u_i) - \alpha = 0 \\ \sum_{i=1}^n \lambda_i = 1 \end{cases} \quad (2.17)$$

onde $i = 1, \dots, n$, $\gamma(u_i - u_j)$ e $\gamma(u_0 - u_i)$ são, respectivamente, a semivariância entre os pontos u_i e u_j e entre os pontos u_i e u_0 e α é o multiplicador de Lagrange.

A variância da krigagem ordinária é dada por:

$$\sigma_{ko}^2 = \lambda(0) - \sum_{i=1}^n \lambda_i \gamma(u_0 - u_i) - \alpha. \quad (2.18)$$

2.5 COKRIGAGEM

A cokrigagem é uma extensão mais abrangente da técnica de krigagem, em que são utilizadas covariáveis nas estimações. Em alguns casos, essas covariáveis são amostradas nos mesmos pontos em que a variável principal e, em outros casos, em pontos diferentes.

O principal objetivo da cokrigagem está em, através das informações das covariáveis, melhorar a estimativa da variável principal em pontos que ela não foi amostrada. O argumento utilizado é que, se as covariáveis estão correlacionadas com a variável principal, então podemos utilizar observações de pontos onde todas as variáveis foram amostradas para estimar essa correlação. Através da correlação, valores das covariáveis em pontos onde a variável principal não foi observada nos ajudam a melhorar as estimativas (Bailey and Gatrell, 1995).

Segundo Memarsadeghi et al. (2005), existem vários tipos de cokrigagem, sendo os três tipos mais usuais: a cokrigagem simples, a cokrigagem ordinária e a cokrigagem padronizada (*standardized*). A diferença está em como são impostas as restrições. Neste trabalho trataremos da cokrigagem ordinária e, para facilitar a notação utilizaremos apenas uma covariável.

2.5.1 Variograma Cruzado, Covariograma Cruzado e Pseudo-Variograma Cruzado

Para modelar a cokrigagem, devemos primeiro desenvolver um método formal para estimar e modelar a correlação entre as variáveis. Faremos isto através da extensão da idéia de variograma e covariograma com uma variável para o variograma cruzado e o covariograma cruzado entre duas variáveis (Bailey and Gatrell, 1995). Existe ainda o pseudo-variograma cruzado que é utilizado quando as variáveis não são medidas nos mesmos pontos.

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

$$C_{Z_1 Z_2}(h) = E[(Z_1(u+h) - \mu_{Z_1})(Z_2(u) - \mu_{Z_2})] \quad (2.19)$$

em que h é um vetor de distância qualquer, μ_{Z_1} e μ_{Z_2} são as médias das variáveis Z_1 e Z_2 , respectivamente.

Da mesma forma, o variograma cruzado será definido como:

$$2\gamma_{Z_1 Z_2}(h) = E[(Z_1(u+h) - Z_1(u))(Z_2(u+h) - Z_2(u))] \quad (2.20)$$

Segundo Webster and Oliver (2007), é evidente que o covariograma cruzado e o variograma cruzado só podem ser obtidos quando a variável principal e a covariável forem amostradas no mesmo ponto, entretanto, em muitos problemas práticos, isso não acontece, como em nosso problema, em que os dados foram obtidos de fontes diferentes, formas de medições diferentes e, conseqüentemente, em pontos diferentes.

Assim, para solucionar este problema, foi definido o pseudo-variograma cruzado (*pseudo-cross-variogram*) como:

$$2\gamma_{Z_1Z_2}^p(h) = Var((Z_1(u) - Z_2(u+h))) \quad (2.21)$$

Para um processo estacionário de segunda ordem, $2\gamma_{Z_1Z_2}^p(h)$ está relacionado à função de covariância da forma:

$$2\gamma_{Z_1Z_2}^p(h) = \frac{1}{2}[C_{Z_1Z_1}(0) + C_{Z_2Z_2}(0)] - C_{Z_1Z_2}(h) \quad (2.22)$$

2.5.2 Cokrigagem Ordinária

Suponha agora l variáveis, $l = 1, 2, \dots, V$ e denotando uma delas por Z_1 , que será a variável principal. O estimador da cokrigagem ordinária para Z_1 no ponto u_0 será:

$$Z_1^*(u_0) = \sum_{l=1}^V \sum_{i=1}^{n_l} \lambda_{il} Z_l(x_i) \quad (2.23)$$

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

$$\sum_{i=1}^{n_l} \lambda_{il} = \begin{cases} 1 & \text{se } l = Z_1 \\ 0 & \text{se } l \neq Z_1 \end{cases} \quad (2.24)$$

Sujeito à essas condições, o estimador da variância de $Z_1(u_0)$ será obtido minimizando:

$$\sum_{i=1}^{n_l} \lambda_{il} \gamma_{lv}(x_i, x_j) + \alpha_v = \bar{\gamma}_{Z_1v}(x_j, u_0) \quad (2.25)$$

para todo $v = 1, 2, \dots, V$ e $j = 1, 2, \dots, n_v$, em que $\gamma_{lv}(x_i, x_j)$ é o semivariograma cruzado entre as variáveis l e v nos locais i e j separados pelo vetor $x_i - x_j$, $\bar{\gamma}_{Z_1v}(x_j, u_0)$ é o semivariograma cruzado médio entre o local j e o local u_0 e α_v é o multiplicador de Lagrange para a v -ésima variável.

Resolvendo o sistema, obtemos os estimador para a variância de Z_1 :

$$\sigma_{Z_1}^2 = \sum_{i=1}^{n_i} \lambda_{il} \gamma_{lv}(x_i, x_j) + \alpha_v \quad (2.26)$$

Essas equações podem ser representadas matricialmente, para facilitar a notação, utilizaremos apenas uma covariável. O caso com mais de uma covariável é obtido por simples extensão. Seja Γ_{uv} a matriz de semivariância das variáveis u e v , observadas em n_u e n_v locais respectivamente. Essa matriz possui dimensão $n_u \times n_v$:

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

Denotamos também por b_{uu} e b_{uv} os vetores de semivariância e semivariância

cruzada:

$$\mathbf{b}_{\mathbf{uu}} = \begin{bmatrix} \gamma_{uu}(x_1, x_0) \\ \gamma_{uu}(x_2, x_0) \\ \vdots \\ \gamma_{uu}(x_{n_u}, x_0) \end{bmatrix}, \quad \mathbf{b}_{\mathbf{uv}} = \begin{bmatrix} \gamma_{uv}(x_1, x_0) \\ \gamma_{uv}(x_2, x_0) \\ \vdots \\ \gamma_{uv}(x_{n_v}, x_0) \end{bmatrix}$$

A equação matricial fica na forma:

$$\begin{bmatrix} \Gamma_{\mathbf{uu}} & \Gamma_{\mathbf{uv}} & 1 & 0 \\ & & 1 & 0 \\ & & \vdots & \vdots \\ \Gamma_{\mathbf{vu}} & \Gamma_{\mathbf{vv}} & 1 & 0 \\ & & 0 & 1 \\ & & 0 & 1 \\ & & \vdots & \vdots \\ & & 0 & 1 \\ 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_{1u} \\ \lambda_{2u} \\ \vdots \\ \lambda_{n_u u} \\ \lambda_{1v} \\ \lambda_{2v} \\ \vdots \\ \lambda_{n_v v} \\ \alpha_u \\ \alpha_v \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{\mathbf{uu}} \\ \mathbf{b}_{\mathbf{uv}} \\ 1 \\ 0 \end{bmatrix}$$

Denotando a matriz aumentada de Γ 's por \mathbf{G} , o vetor dos pesos e multiplicadores de lagrange por λ e o vetor do lado direito da igualdade por \mathbf{b} , podemos escrever a solução da equação como:

$$\lambda = \mathbf{G}^{-1} \mathbf{b} \quad (2.27)$$

e a variância será:

$$\hat{\sigma}_u^2(x_0) = \mathbf{b}^T \boldsymbol{\lambda} \quad (2.28)$$

2.6 ANISOTROPIA

A anisotropia está muito presente na natureza, ocorre quando a variabilidade da distribuição espacial de certas características não depende somente da distância, mas também da direção entre elas. É o caso oposto a isotropia, na qual a variação só depende da distância.

Uma forma de identificar a anisotropia é verificar diferenças nos semivariogramas para diferentes direções. Existem três tipos de anisotropia: a anisotropia zonal em que os Patamares (C) dos semivariogramas são os mesmos, porém os Alcances (a) diferentes; a anisotropia geométrica, em que os Alcances (a) são iguais e os Patamares (C) são diferentes; e a anisotropia combinada, a mais comum, que é uma combinação das anisotropias zonal e geométrica, onde os Patamares (C) e Alcances (a) são diferentes (Camargo et al., 2001).

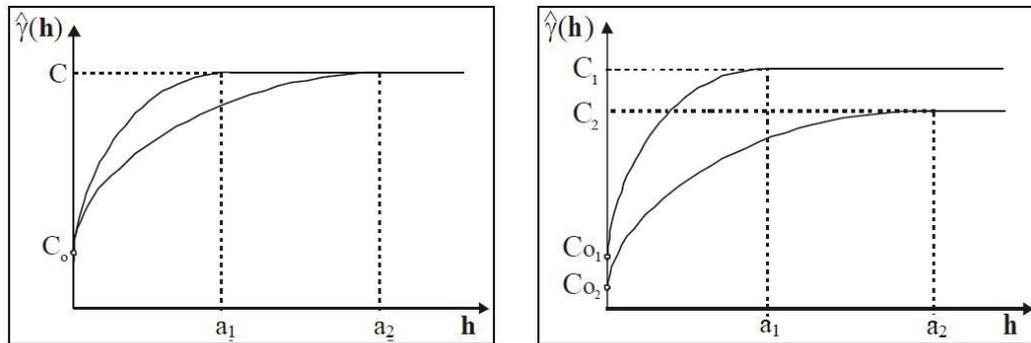


Figura 2.3: Anisotropia Zonal e Anisotropia Combinada (Camargo et al., 2001)

Além de verificar a diferença entre semivariogramas para várias direções, é

Podemos normalizar h em relação ao Alcance (a), assim:

$$\left| \frac{h}{a} \right| = \sqrt{\left(\frac{h_x}{a_x} \right)^2 + \left(\frac{h_y}{a_y} \right)^2} \quad (2.30)$$

Substituindo agora o valor de h no modelo teórico escolhido para cada direção, assim obtemos novo semivariograma com a anisotropia corrigida:

$$\gamma'(h) = \gamma(h_x, h_y) \quad (2.31)$$

Para a anisotropia zonal, utilizamos como patamar o maior valor de patamar dentre os semivariogramas:

$$\gamma'(h) = \omega \gamma'(h') \quad (2.32)$$

onde ω é o patamar e $h' = \frac{h}{a}$.

Para a anisotropia combinada, corrigi-se a anisotropia zonal e geométrica:

$$\gamma'(h) = \omega_1 \gamma'_1(h') + \omega_2 \gamma'_2(h') \quad (2.33)$$

onde ω_1 é o patamar do semivariograma com maior alcance, ω_2 é o patamar do semivariograma com menor alcance e $h' = \sqrt{\left(\frac{h_x}{a_x} \right)^2 + \left(\frac{h_y}{a_y} \right)^2}$

Podemos então, por exemplo, definir a anisotropia (Ω) para o modelo esférico como:

$$\Omega(\theta) = \{A^2 \cos^2(\theta - \varphi) + B^2 \sin^2(\theta - \varphi)\}^{\frac{1}{2}} \quad (2.34)$$

em que A é o maior alcance, B é o menor alcance, φ é a direção do maior alcance e ϑ é a direção do *lag*.

2.6.1 Modelagem Anisotrópica

Quando tem-se a anisotropia geométrica, após identificar os eixos de maior e menor continuidade e para iniciar a modelagem da anisotropia, o primeiro passo é

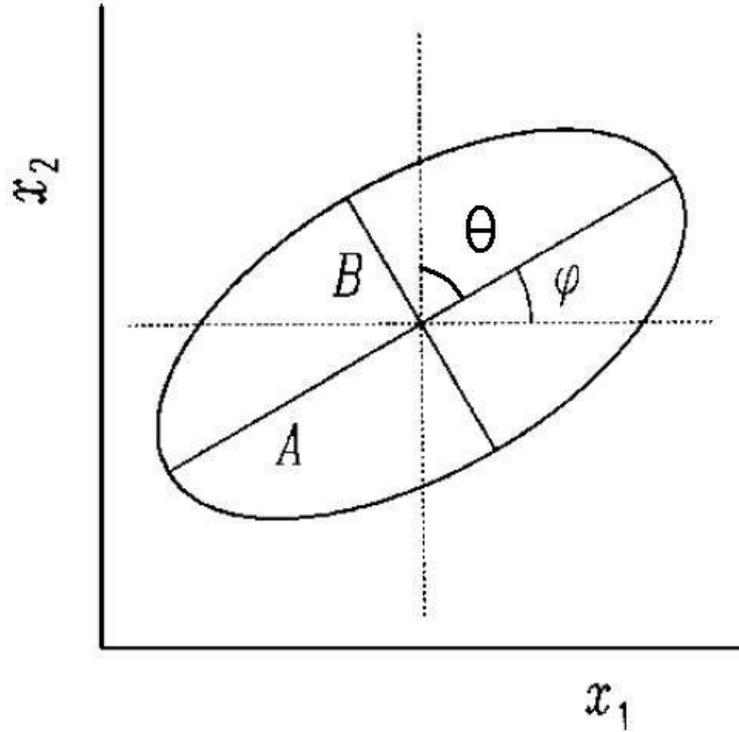


Figura 2.5: Representação geométrica da anisotropia na qual a elipse descreve o alcance do modelo esférico

decompor o módulo do vetor distância \mathbf{h} :

$$|\mathbf{h}| = \sqrt{(h_x)^2 + (h_y)^2} \quad (2.35)$$

onde h_x é a componente ao longo do eixo de menor variabilidade x e h_y é a componente ao longo do eixo de maior variabilidade y .

Normalizando (2.35) em relação ao alcance (a), tem-se:

$$\left| \frac{\mathbf{h}}{\mathbf{a}} \right| = \sqrt{\left(\frac{h_x}{a_x} \right)^2 + \left(\frac{h_y}{a_y} \right)^2}. \quad (2.36)$$

Após o cálculo acima, substituiu-se o valor de \mathbf{h} no modelo teórico escolhido anteriormente para cada direção. Tem-se assim o novo semivariograma experimental com a anisotropia corrigida, que pode ser definido por:

$$\gamma'(h) = \gamma(h_x, h_y) \quad (2.37)$$

No caso da anisotropia zonal, usa-se a seguinte forma de modelagem, que considera como patamar o maior valor de patamar apresentado entre os semivariogramas construídos :

$$\gamma'(h) = \omega\gamma(h') \quad (2.38)$$

onde ω é o patamar e $h' = \frac{h}{a}$, em que a é o alcance apresentado.

Já na anisotropia combinada é proposto o seguinte modelo, que consiste em corrigir a anisotropia geométrica e depois a anisotropia zonal:

$$\gamma'(h) = \omega_1\gamma_1(h') + \omega_2\gamma_2(h') \quad (2.39)$$

onde ω_1 representa o patamar do semivariograma com maior alcance, ω_2 o patamar do semivariograma de menor alcance e

$$h' = \sqrt{\left(\frac{h_x}{a_x}\right)^2 + \left(\frac{h_y}{a_y}\right)^2}$$

sendo h_x e h_y as componentes ao longo do eixo de menor e maior variabilidade, respectivamente, e a_x e a_y os alcances na direção x e y , respectivamente.

Assim, conclui-se de (2.39) que, ao encontrar o semivariograma ideal, determina-se os pesos ideais para se ter uma correta estimação do atributo em questão.

Capítulo 3

MATERIAL E MÉTODOS

3.1 INTRODUÇÃO

Neste capítulo são apresentados os materiais e métodos a serem usados no trabalho. O material utilizado será um banco de dados com a temperatura média em diferentes áreas da região Amazônica no ano 2010, sendo esses dados obtidos através de 66 estações meteorológicas instaladas na região. Para isso, utiliza-se o sistema de coordenadas lat/long. Utilizaremos também dados de precipitação e desmatamento obtidos através de imagens de satélites. Depois de obtidos os dados, será utilizado um algoritmo implementado em linguagem IML do *software* SAS 9.2 para realizar a estimação da temperatura nas regiões não amostradas pelas estações através dos métodos explicados anteriormente.

3.2 ESTAÇÕES METEOROLÓGICAS

Para se fazer estudos meteorológicos são necessários certos instrumentos específicos tais como termômetros, barômetros, pluviômetro, entre outros. Uma estação meteorológica é constituída pela reunião desses instrumentos em um mesmo local, e um conjunto de estações em uma determinada região é definido como rede

de estações (INMET, 2011).

Existem três tipos de estações meteorológicas: a estação de Observação de Altitude ou de Radiossonda, a Estação Meteorológica de Observação de Superfície Automática e a Estação Meteorológica de Observação de Superfície Convencional. A radiossonda é um conjunto de instrumentos e sensores para medir a temperatura do ar, umidade relativa e pressão atmosférica, enquanto é elevada na atmosfera até alturas típicas da ordem de 30 Km, por um balão inflado com gás hélio. O deslocamento da sonda é registrado por uma antena GPS que permite a medida da direção e velocidade do vento. Os dados observados, minuto a minuto, são enviados via rádio para a estação receptora no solo que os processa, gera uma mensagem codificada e a envia para o Centro Coletor onde ocorrerá a distribuição global.

A estação meteorológica convencional é composta de vários sensores isolados que registram continuamente os parâmetros meteorológicos (pressão atmosférica, temperatura e umidade relativa do ar, precipitação, radiação solar, direção e velocidade do vento, etc), que são lidos e anotados por um observador a cada intervalo e este os envia a um centro coletor por um meio de comunicação qualquer. A distribuição espacial das estações meteorológicas convencionais estão na Figura 3.1, onde a região destacada em verde representa a região da Amazônia Legal. Para mais informações consultar Barreto (2011).



Figura 3.1: Rede de estações convencionais.

3.3 SATÉLITES

Utilizaremos também dados gerados a partir de satélites. Esses dados são divulgados pelo Instituto Nacional de Pesquisas Espaciais (INPE). Serão consultados dois bancos de dados: o Banco de Dados de Queimadas (BDQ) e o Sistema de Detecção de Desmatamento em Tempo Real (DETER). Agora discutiremos a metodologia empregada na coleta de cada um deles.

3.3.1 BDQ

A Divisão de Processamento de Imagens do INPE criou um banco de dados geográficos que contem várias informações relativas a queimadas, tais como: focos de calor dos mais variados satélites, imagens de vários satélites e com vários níveis de

resolução, base cartográfica, dados meteorológicos, mapa de desmatamento, mapa das unidades de conservação.

São utilizados todos os satélites que possuem sensores óticos operando na faixa termal-média de 4um e que o INPE consegue receber. São processadas operacionalmente as imagens AVHRR dos satélites polares NOAA-15, NOAA-16, NOAA-17, NOAA-18 e NOAA-19, as imagens MODIS dos satélites polares NASA TERRA e AQUA, as imagens dos satélites geoestacionários GOES-12 e MSG-2. Cada satélite de órbita polar produz pelo menos um conjunto de imagens por dia, e os geoestacionários geram algumas imagens por hora, sendo que no total o INPE processa mais de 100 imagens por dia especificamente para detectar focos de queima da vegetação.

Para os satélites de órbita polar (NOAAs a 800 km de distância, e TERRA e AQUA a 730 km), trabalhos de validação de campo indicam que uma frente de fogo com cerca de 30 m de extensão por 1 m de largura, ou maior, será detectada. Para os geoestacionários, a 25 mil km de distância, a frente precisa ter o dobro de tamanho para ser localizada. Entretanto, como o elemento de resolução espacial (píxel) do satélite tem 1 km x 1 km ou mais, uma queimada de algumas dezenas de m^2 será identificada como tendo pelo menos 1 km^2 . Nas imagens dos satélites geoestacionários, onde o píxel tem 4km x 4km, esta pequena queimada passará a ser indicada por uma área de 16 km^2 ou mais. Assim, um foco de queima, que aqui é a mesma coisa que um píxel de queima, pode indicar tanto uma pequena queimada assim como várias pequenas queimadas ou uma muito grande no seu interior. Ou seja, este sistema do INPE detecta a existência de fogo na vegetação

sem ter condições de avaliar o tamanho da área que está queimando ou o tipo de vegetação afetada. Em casos com muitos píxeis de queima juntos, e com a presença de uma nuvem de fumaça grande, pode-se inferir que a queimada terá a dimensão dos píxeis de queima detectados (INPE, 2011).

Neste banco que utilizaremos, são medidas algumas variáveis como: número de dias sem chuva, risco de queimada e precipitação média mensal. Este último será utilizado em nosso estudo. A Figura 3.2 mostra um exemplo de imagem de um satélite e os focos encontrados.

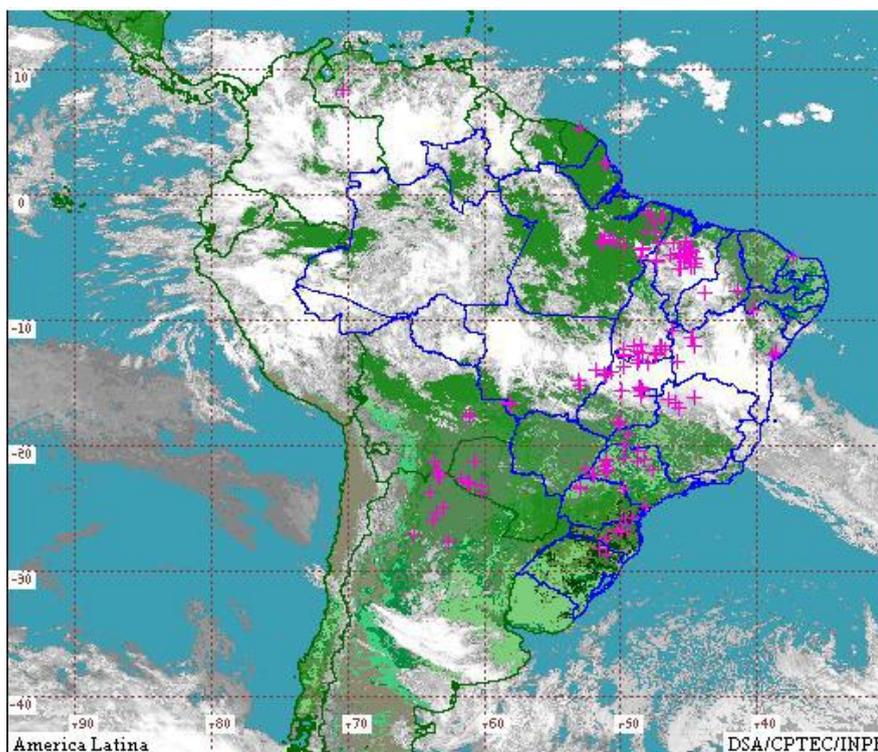


Figura 3.2: Exemplo de focos detectados pelos satélites (INPE, 2011).

3.3.2 DETER

O sistema de Detecção de Desmatamento em Tempo Real (DETER) é um sistema de apoio a fiscalização e controle de desmatamento na Amazônia, como o DETER o

INPE divulga mensalmente um mapa de alertas com áreas maiores que 25 ha. Esses mapas indicam áreas totalmente desmatadas e as áreas em processo de degradação desde 2004.

Este sistema utiliza imagens dos sensores MODIS, a bordo do satélite TERRA da NASA, e imagens do WFI, a bordo do satélite brasileiro CBERS-2B do INPE. Esses sensores cobrem a Amazônia com alta frequência, de dois e cinco dias, respectivamente, porém com resolução muito limitada, entre 250 e 260 metros. Com baixa resolução é possível a detecção de áreas desmatadas superiores a 25 ha, porém a limitação de resolução é compensada pela alta frequência de observação.

A identificação de desmatamento é feita a partir da foto-interpretação da imagem MODIS, considerando apenas partes da imagem onde supostamente ainda existe cobertura vegetal.

Como não é possível detectar os desmatamentos na presença de nuvens, também é feito um mapa de nuvens que contém as informações de cobertura de nuvens de todas as imagens utilizadas, sendo avaliadas quinzenalmente.

A identificação do padrão de alteração da cobertura florestal é feita com base em três principais elementos para a foto-interpretação: tonalidade, textura e contexto (INPE, 2008).

A Figura 3.3 mostra um exemplo de detecção de focos de desmatamento, onde máscara é a área anteriormente desmatada.

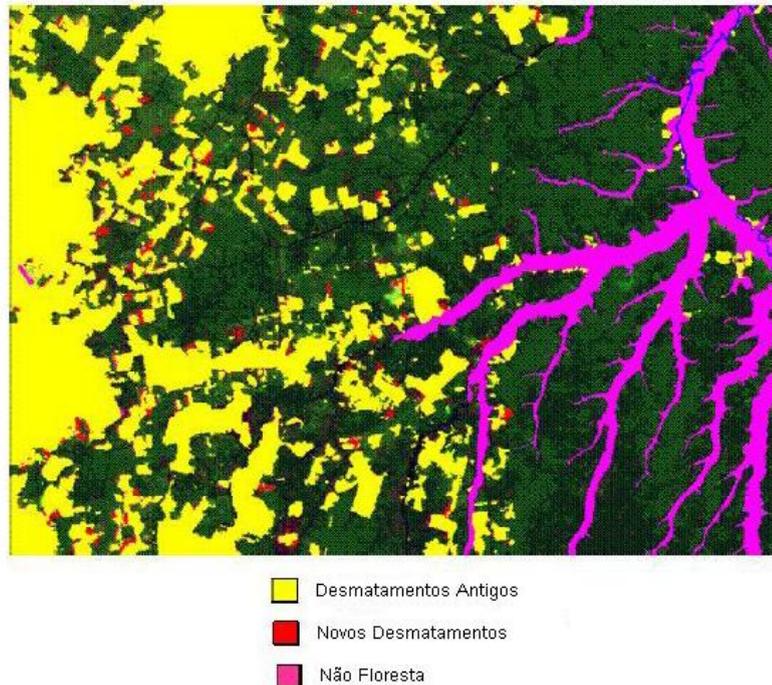


Figura 3.3: Exemplo de foto de satélite do sistema DETER (INPE, 2011).

3.4 SISTEMAS DE COORDENADAS

Os sistemas de coordenadas são usados para que se possa definir a posição de pontos na superfície da Terra. Esses pontos são representados por valores lineares e/ou angulares. Os principais tipos de sistemas de coordenadas são x, y, z , Sistema Lat/Long e Sistema Universo Transverso de Mercator - UTM, sendo esses dois últimos os mais utilizados. Neste trabalho apresentaremos o Sistema Lat/Long.

3.4.1 Sistema Lat/Long

O sistema de coordenadas Lat/Long é considerado um sistema de coordenadas esféricas que se baseia nos conceitos de Latitude e Longitude. A latitude é um método que descreve as posições de pontos situados a norte ou a sul do equador. O

equador é uma linha perpendicular ao eixo terrestre que divide o planeta em dois hemisférios: norte e sul. A latitude varia entre 0° (Equador) e 90° (pólos). As linhas que possuem a mesma latitude são chamadas de paralelos. A longitude constitui-se um método que descreve as posições de pontos situados a leste e a oeste do meridiano de Greenwich, situado na Inglaterra. Os meridianos são semicírculos que possuem extremidades nos pólos e são perpendiculares aos paralelos, sendo o meridiano de Greenwich definido como o meridiano central. A longitude varia entre 0° (meridiano de Greenwich) e 180° a leste ou a oeste .

A latitude e a longitude podem ser usadas em um sistema de coordenadas onde a origem se situa no ponto de 0° de latitude e 0° de longitude, podendo ser definido assim quatro quadrantes: NE (latitude e longitude positivas), NW (latitude positiva e longitude negativa), SE (latitude negativa e longitude positiva) e SW (latitude e longitude negativas). O conjunto de meridianos e paralelos formam uma grade de linhas imaginárias ao redor do planeta, constituindo as coordenadas geográficas . Cada grau de latitude e longitude pode ser dividido em 60 minutos e cada minuto em 60 segundos, com isso é possível ter uma localização precisa de um ponto qualquer da superfície terrestre (de Barros Silva, 2003).

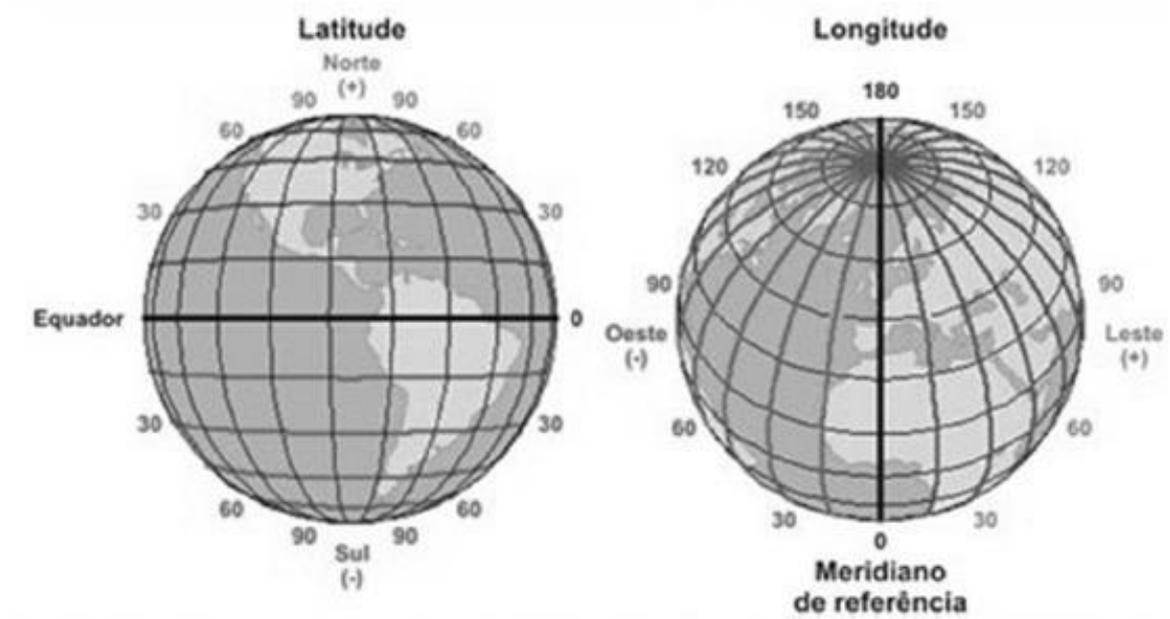


Figura 3.4: Linha do Equador e Meridiano de Greenwich.

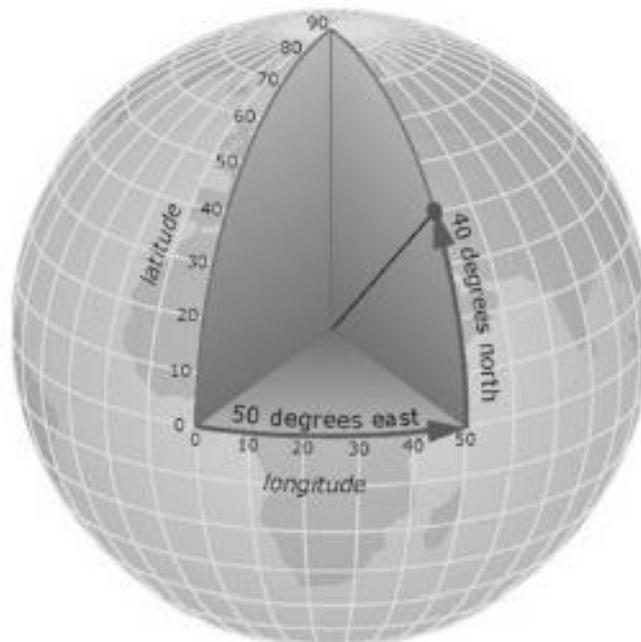


Figura 3.5: Ponto na superfície terrestre representado pela latitude e longitude.

3.5 ALGORITMO IML

O algoritmo a seguir foi construído utilizando-se o procedimento IML (*Interactive Matrix Language*), do *software SAS 9.2*, pois os códigos implementados no IML são facilmente transformados para outras linguagens de programação. O algoritmo é dividido em duas macros: uma para construir a estrutura variográfica dos dados, levando em consideração a covariável (variograma cruzado), e outra para realizar a estimação para toda a superfície aplicando a cokrigagem. Não existem procedimentos semelhantes já implementados no *software SAS 9.2*.

3.5.1 Macro Crossvariogram

Os parâmetros necessários para execução da macro são:

```
%macro crossvariogram(taby=,vary=,longy=,laty=,tabx=,varx=,longx=,latx=,lag=,lagd=,model=,tabout=);
```

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; **varx**, **longx** e **latx** correspondem às mesmas variáveis para a covariável; **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 **lagd**=2 os valores correspondentes em cada classe serão múltiplos de 2; **model** especifica qual modelo será utilizado no ajuste da variância cruzada. 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.

Primeiramente é executada a macro *variogram* para o cálculo do variograma de cada variável separadamente, esta macro está apresentada em Barreto (2011). Em seguida, as variáveis são padronizadas para eliminar uma possível variação oriunda de diferença de escala:

```
proc greplay igout=work.gseg nofs;
delete s_&vary s_&varx s_y_x st_y_x;
run;quit;
%variogram(tab=&taby,var=&vary,long=&longy,lat=&laty,lag=&lag,lagd=&lagd,
model=&model,tabout=variogram_&vary);
%variogram(tab=&tabx,var=&varx,long=&longx,lat=&latx,lag=&lag,lagd=&lagd,
model=&model,tabout=variogram_&varx);
z=(z-z[:])/sqrt((z-z[:])*'(z-z[:])/(nrow(z)-1));
x=(x-x[:])/sqrt((x-x[:])*'(x-x[:])/(nrow(x)-1));
```

A seguir é criada a matriz **dist**, que contém as distâncias calculadas entres os pontos dois a dois e os ângulos existentes entre eles:

```
d=j(1,4,0);
do i=1 to n;
do j=1 to m;
d[1]=i;
d[2]=j;
d[3]=sqrt((COORDy[i,1]-COORDx[j,1])**2+(COORDy[i,2]-COORDx[j,2])**2);
if d[3]=0 then d[4]=0;
else if (COORDy[i,1]-COORDx[j,1])=0 then d[4]=90;
else if (COORDy[i,2]-COORDx[j,2])=0 then d[4]=0;
else d[4]=atan((COORDy[i,2]-COORDx[j,2])/(COORDy[i,1]-COORDx[j,1]))/arcos(-1)*180;
if d[4]<0 then d[4]=180-(-d[4]);
else if d[4]<90 then d[4]=90-d[4];
else d[4]=275-d[4];
if i=1 & j=1 then dist=d;else dist=dist//d;
end;
end;
```

A seguir tem-se a distância máxima obtida entre dois pontos amostrados e as especificações de *default* para o parâmetro **lagd** e no caso deste não ter sido especificado.

```
max=max(dist[,3]);
print max;

lag=&lag;
%if &lagd= %then %do;
lagd=max/lag;
%end;
%else %do;
lagd=&lagd;
%end;
```

Agora, tem-se a construção da matriz variograma cruzado, que conterà os limites inferior e superior de cada classe de distâncias, o números 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:

```
classes=j(lag,4,0);
do j=1 to lag;
do i=1 to nrow(dist);
if j=1 then do;
if dist[i,3]<lagd then do;
classes[j,1]=0;
classes[j,2]=lagd;
classes[j,3]=classes[j,3]+1;
classes[j,4]=classes[j,4]+(z[dist[i,1]]-x[dist[i,2]])**2;
end;
end;
else do;
classes[j,1]=classes[j-1,2];
classes[j,2]=classes[j,1]+lagd;
if (dist[i,3]>=classes[j,1] & dist[i,3]<classes[j,2]) then do;
classes[j,3]=classes[j,3]+1;
classes[j,4]=classes[j,4]+(z[dist[i,1]]-x[dist[i,2]])**2;
end;
end;
end;
end;
variogram=(classes[,4]/(2*classes[,3]));
```

```

columns={"l" "u" "freq" "zi_zj" "distance" "semivariogram"};
variogram=classes[,1:4]||((classes[,1]+classes[,2])/2)||variogram;
create &tabout from variogram[colname=columns];
append from variogram;
close &tabout;

```

Na próxima etapa, obtemos os valores iniciais para os parâmetros do variograma cruzado a ser ajustado e é feita uma correção afim de obter a melhor modelagem. Esse ajuste consiste em verificar se o primeiro valor de semivariância é maior que o segundo, se isso ocorrer, ignoramos este primeiro valor e o fazemos igual ao segundo. Esta medida permite um melhor ajuste da curva a ser modelada. Estes valores são aplicados ao modelo escolhido previamente.

```

variogram2=variogram;
if variogram[1,6]>variogram[2,6] then variogram[1,6]=variogram[2,6];
nug=max(0,variogram[1,6]-(variogram[1,5]/
(variogram[2,5]-variogram[1,5]))*(variogram[2,6]-variogram[1,6]));
c0=(variogram[lag-2,6]+variogram[lag-1,6]+variogram[lag,6])/3-nug;
a0=0.5*variogram[lag,5];
adjvariogram=j(lag,1,0);
do ll=1 to nrow(variogram);
%if %upcase(&model)=GAUSSIAN %then %do;
adjvariogram[ll]=nug+(c0-nug)*(1-exp(-variogram[ll,5]**2/(a0**2)));
%end;
%if %upcase(&model)=EXPONENTIAL %then %do;
adjvariogram[ll]=nug+(c0-nug)*(1-exp(-variogram[ll,5]/a0));
%end;
%if %upcase(&model)=SINE %then %do;
adjvariogram[ll]=nug+(c0-nug)*(1-sin(arcos(-1)*variogram[ll,5]/a0)/
(arcos(-1)*variogram[ll,5]/a0));
%end;
%if %upcase(&model)=PENTASPHERICAL %then %do;
if variogram[ll,5]<=a0 then adjvariogram[ll]=nug+(c0-nug)*(1.875*variogram[ll,5]/
a0 -1.25*(variogram[ll,5]/a0)**3 +0.375*(variogram[ll,5]/a0)**5);
else adjvariogram[ll]=c0;
%end;
%if %upcase(&model)=CUBIC %then %do;
if variogram[ll,5]<=a0 then adjvariogram[ll]=nug+(c0-nug)*(7*(variogram[ll,5]/
a0)**2 -8.75*(variogram[ll,5]/a0)**3 +3.5*(variogram[ll,5]/a0)**5 - 0.75*
(variogram[ll,5]/a0)**7);
else adjvariogram[ll]=c0;
%end;

```

```

%if %upcase(&model)=SPHERICAL %then %do;
if variogram[11,5]<=a0 then adjvariogram[11]=nug+(c0-nug)*(1.5*variogram[11,5]/
a0 -0.5*(variogram[11,5]/a0)**3);
else adjvariogram[11]=c0;
%end;
end;

```

Abaixo está o algoritmo de minimização dos pesos para estimação da semivariância

ajustada:

```

start minvariog(lambda) global(variogram,adjvariogram);
r=0;
c0=lambda[1];a0=lambda[2];nug=lambda[3];
do ll=1 to nrow(variogram);
%if %upcase(&model)=GAUSSIAN %then %do;
adjvariogram[11]=nug+(c0-nug)*(1-exp(-variogram[11,5]**2/(a0**2)));
%end;
%if %upcase(&model)=EXPONENTIAL %then %do;
adjvariogram[11]=nug+(c0-nug)*(1-exp(-variogram[11,5]/a0));
%end;
%if %upcase(&model)=SINE %then %do;
adjvariogram[11]=nug+(c0-nug)*(1-sin(arcos(-1)*variogram[11,5]/
a0)/(arcos(-1)*variogram[11,5]/a0));
%end;
%if %upcase(&model)=PENTASPHERICAL %then %do;
if variogram[11,5]<=a0 then adjvariogram[11]=nug+(c0-nug)*(1.875*variogram[11,5]/
a0 -1.25*(variogram[11,5]/a0)**3 +0.375*(variogram[11,5]/a0)**5);
else adjvariogram[11]=c0;
%end;
%if %upcase(&model)=CUBIC %then %do;
if variogram[11,5]<=a0 then adjvariogram[11]=nug+(c0-nug)*(7*(variogram[11,5]/
a0)**2 -8.75*(variogram[11,5]/a0)**3 +3.5*(variogram[11,5]/a0)**5 - 0.75*
(variogram[11,5]/a0)**7);
else adjvariogram[11]=c0;
%end;
%if %upcase(&model)=SPHERICAL %then %do;
if variogram[11,5]<=a0 then adjvariogram[11]=nug+(c0-nug)*(1.5*variogram[11,5]/
a0 -0.5*(variogram[11,5]/a0)**3);
else adjvariogram[11]=c0;
%end;
r=r+variogram[11,3]*(variogram[11,6]/adjvariogram[11]-1)##2;
end;
r=r/2;
return(r);
finish minvariog;

```

Agora são acrescentados à matriz do variograma cruzado, os valores das semi-variâncias ajustadas pelo modelo escolhido previamente e o Critério de Informação de Akaike (AIC) que será utilizado na obtenção do melhor modelo. Também é construída uma matriz com os parâmetros do variograma cruzado ajustado. Essa matriz será utilizada posteriormente para as estimações no método de cokrigagem.

```
lambda=c0||a0||nug;
optn={0};
call nlpqn(rc,xr,"minvariog",lambda,optn);
lambda=xr;
WSSE=((1/c0)*(variogram[,6]-adjvariogram))*((1/c0)*(variogram[,6]-adjvariogram));
AIC=nrow(variogram)*log(WSSE/nrow(variogram))+2*ncol(lambda);
print "Variables: &vary..&varx";
print "Model: &model",, lambda[colname={"sill" "range" "nugget"}];
print "AIC: " AIC;
columns=columns||"adjvariogram";
variogram=variogram2||adjvariogram;
create &tabout from variogram[colname=columns];
append from variogram;
create _parameters_&vary._&varx._ from
lambda[colname={"_sill_" "_range_" "_nugget_"}];
append from lambda;
quit;
data _par_;set _parameters_&vary._(in=a) _parameters_&varx._(in=b)
_parameters_&vary._&varx._(in=c);
if c then _var_="&vary._&varx._";if b then _var_="&varx";if a then _var_="&vary";
run;
```

Por fim são criados os gráficos dos variogramas, separadamente, e o gráfico do variograma cruzado.

```
goptions reset=all;
title h=1.5 "&vary..&varx";
proc gplot data=&tabout;
plot semivariogram*distance adjvariogram*distance /overlay
vaxis=axis2 haxis=axis1 frame name="s_y_x";
symbol1 pointlabel=( '#freq' ) i=join c=yellow v=square;
symbol2 pointlabel=none i=join c=blue v=diamond;
axis1 minor=none label=(c=black 'Distance');
axis2 minor=none label=(angle=90 rotate=0 c=black 'Semivariance');
run;
quit;
```

```

proc greplay igout=work.gseg gout=work.gseg nofs tc=sashelp.templt
template=l2r2;
treplay 1:s_&vary 2:s_y_x 4:s_&varx name="st_y_x";
run;
quit;
%end;
%mend crossvariogram;

```

3.5.2 Macro Cokriging

Os parâmetros necessários para execução da macro são:

```

%macro cokriging(taby=,vary=,longy=,laty=,tabx=,varx=,longx=,latx=,model=,
sill=,range=,nugget=,map=,by=,tabout=);

```

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; **varx**, **longx** e **latx** correspondem às mesmas variáveis para a covariável; **model** especifica qual modelo será utilizado no ajuste da variância cruzada. 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*); **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; e **tabout**

especifica o nome da tabela de dados de saída.

Primeiramente verificamos se existe uma matriz de dados para a covariável. Se ela não existir, então as estimações serão feitas pelo método de krigagem. Logo depois são verificados os parâmetros do variograma cruzado. As variáveis **sill**, **range** e **nugget** podem ser colocadas como números ou como tabelas, sendo que pode-se utilizar a tabela de parâmetros gerada na macro *crossvariogram*.

```
%if &tabx ne %then %do;
C=j((n+1)+(m+1),(n+1)+(m+1),0);
%end;
%else %do;
C=j(n+1,n+1,0);
%end;

c0={&sill};
if type(c0)='C' then do;
use &sill ;
read all var{&sill_} into c0;
end;
close &sill;

a0={&range};
if type(a0)='C' then do;
use &range ;
read all var{&range_} into a0;
end;
close &range;

nug={&nugget};
if type(nug)='C' then do;
use &nugget ;
read all var{&nugget_} into nug;
end;
close &nugget;
```

A seguir, tem-se a construção da matriz de variância espacial a partir do uso do modelo teórico ajustado à variabilidade dos dados na análise variográfica.

```
do v=1 to nvar+1;
  if v=1 then do;
```

```

    C1=j(n,n,0);
    do i=1 to n;
    do j=i+1 to n;
    dist=sqrt((COORDy[i,1]-COORDy[j,1])**2+(COORDy[i,2]-COORDy[j,2])**2);
    %if %upcase(&model)=GAUSSIAN %then %do;
    C1[i,j]=nug[v]+(c0[v]-nug[v])*(1-exp(-dist**2/(a0[v]**2)));
    %end;
    %if %upcase(&model)=EXPONENTIAL %then %do;
    C1[i,j]=nug[v]+(c0[v]-nug[v])*(1-exp(-dist/a0[v]));
    %end;
    %if %upcase(&model)=SINE %then %do;
    C1[i,j]=nug[v]+(c0[v]-nug[v])*(1-sin(arcos(-1)*dist/a0[v])/
    (arcos(-1)*dist/a0[v]));
    %end;
    %if %upcase(&model)=PENTASPHERICAL %then %do;
    if dist<=a0[v] then C1[i,j]=nug[v]+(c0[v]-nug[v])*
    (1.875*dist/a0[v] -1.25*(dist/a0[v])**3 +0.375*(dist/a0[v])**5);
    else C1[i,j]=c0[v];
    %end;
    %if %upcase(&model)=CUBIC %then %do;
    if dist<=a0[v] then C1[i,j]=nug[v]+(c0[v]-nug[v])*
    (7*(dist/a0[v])**2 -8.75*(dist/a0[v])**3 +3.5*(dist/a0[v])**5
    - 0.75*(dist/a0[v])**7);
    else C1[i,j]=c0[v];
    %end;
    %if %upcase(&model)=SPHERICAL %then %do;
    if dist<=a0[v] then C1[i,j]=(nug[v]+(c0[v]-nug[v])*
    (1.5*dist/a0[v]-0.5*(dist/a0[v])**3));else C1[i,j]=c0[v];
    %end;
    end;
    C1[i,i]=nug[v];
    end;
    C[1:n,1:n]=C1;
    C[n+m+v,1:n]=j(1,n,1);C[1:n,n+m+v]=j(n,1,1);
    end;
    else if v=nvar+1 then do;
    C1=j(n,m,0);
    do i=1 to n;
    do j=1 to m;
    dist=sqrt((COORDy[i,1]-COORDx[j,1])**2+(COORDy[i,2]-COORDx[j,2])**2);
    %if %upcase(&model)=GAUSSIAN %then %do;
    C1[i,j]=nug[v]+(c0[v]-nug[v])*(1-exp(-dist**2/(a0[v]**2)));
    %end;
    %if %upcase(&model)=EXPONENTIAL %then %do;
    C1[i,j]=nug[v]+(c0[v]-nug[v])*(1-exp(-dist/a0[v]));
    %end;
    %if %upcase(&model)=SINE %then %do;
    C1[i,j]=nug[v]+(c0[v]-nug[v])*(1-sin(arcos(-1)*dist/a0[v])/
    (arcos(-1)*dist/a0[v]));

```

```

%end;
%if %uppercase(&model)=PENTASPHERICAL %then %do;
if dist<=a0[v] then C1[i,j]=nug[v]+(c0[v]-nug[v])*
(1.875*dist/a0[v] -1.25*(dist/a0[v])**3 +0.375*(dist/a0[v])**5);
else C1[i,j]=c0[v];
%end;
%if %uppercase(&model)=CUBIC %then %do;
if dist<=a0[v] then C1[i,j]=nug[v]+(c0[v]-nug[v])*
(7*(dist/a0[v])**2 -8.75*(dist/a0[v])**3 +3.5*(dist/a0[v])**5
- 0.75*(dist/a0[v])**7);
else C1[i,j]=c0[v];
%end;
%if %uppercase(&model)=SPHERICAL %then %do;
if dist<=a0[v] then C1[i,j]=(nug[v]+(c0[v]-nug[v])*
(1.5*dist/a0[v]-0.5*(dist/a0[v])**3));else C1[i,j]=c0[v];
%end;
end;
C1[i,i]=nug[v];
    end;
    C[1:n,n+1:n+m]=C1;C[n+1:n+m,1:n]=C1';
    end;
    else do;
    C1=j(m,m,0);
    do i=1 to m;
do j=i+1 to m;
dist=sqrt((COORDx[i,1]-COORDx[j,1])**2+(COORDx[i,2]-COORDx[j,2])**2);
%if %uppercase(&model)=GAUSSIAN %then %do;
C1[i,j]=nug[v]+(c0[v]-nug[v])*(1-exp(-dist**2/(a0[v]**2)));
%end;
%if %uppercase(&model)=EXPONENTIAL %then %do;
C1[i,j]=nug[v]+(c0[v]-nug[v])*(1-exp(-dist/a0[v]));
%end;
%if %uppercase(&model)=SINE %then %do;
C1[i,j]=nug[v]+(c0[v]-nug[v])*(1-sin(arcos(-1)*dist/a0[v])/
(arcos(-1)*dist/a0[v]));
%end;
%if %uppercase(&model)=PENTASPHERICAL %then %do;
if dist<=a0[v] then C1[i,j]=nug[v]+(c0[v]-nug[v])*
(1.875*dist/a0[v] -1.25*(dist/a0[v])**3 +0.375*(dist/a0[v])**5);
else C1[i,j]=c0[v];
%end;
%if %uppercase(&model)=CUBIC %then %do;
if dist<=a0[v] then C1[i,j]=nug[v]+(c0[v]-nug[v])*
(7*(dist/a0[v])**2 -8.75*(dist/a0[v])**3 +3.5*(dist/a0[v])**5
- 0.75*(dist/a0[v])**7);
else C1[i,j]=c0[v];
%end;
%if %uppercase(&model)=SPHERICAL %then %do;
if dist<=a0[v] then C1[i,j]=(nug[v]+(c0[v]-nug[v])*

```

```

(1.5*dist/a0[v]-0.5*(dist/a0[v])**3));else C1[i,j]=c0[v];
%end;
end;
C1[i,i]=nug[v];
end;
C[n+1:n+m,n+1:n+m]=C1;
C[n+m+v,n+1:n+m]=j(1,m,1);C[n+1:n+m,n+m+v]=j(m,1,1);
end;
end;

```

A seguir, tem-se parte da construção do mapa com os valores estimados. Aqui são atribuídas as coordenadas geográficas de cada ponto que será estimado. As coordenadas são inseridas nas colunas 1 e 2 da matriz **POINTS** criada.

```

use &map;
read all var{x y} into p;
maxx=max(p[,1]);minx=min(p[,1]);
maxy=max(p[,2]);miny=min(p[,2]);
close &map;
free p;
POINTS1=j(1,4,0);
_v_=0;
do xx=minx to maxx by &by;
do yy=miny to maxy by &by;
POINTS1[1]=xx;
POINTS1[2]=yy;

```

Por fim, tem-se a construção da matriz de variâncias dos pontos estimados baseada no modelo teórico escolhido na análise variográfica. Essa matriz é utilizada, juntamente com a matriz anterior, para atribuir valores aos pesos de cada ponto da amostra, fazer o cálculo das estimativas da variável principal em estudo e o seu erro padrão.

```

_v_=_v_+1;
%if &tabx ne %then %do;
c_u=j((n+1)+(m+1),1,1);
%end;
%else %do;
c_u=j((n+1),1,1);
%end;

```

```

do v=1 to nvar+1;
  if v=1 then do;
    c_u1=j(n,1,1);
    do j=1 to n;
      dist=sqrt((POINTS1[1]-COORDy[j,1])**2+(POINTS1[2]-COORDy[j,2])**2);
      %if %upcase(&model)=GAUSSIAN %then %do;
c_u1[j]=nug[v]+(c0[v]-nug[v])*(1-exp(-dist**2/(a0[v]**2)));
      %end;
      %if %upcase(&model)=EXPONENTIAL %then %do;
c_u1[j]=nug[v]+(c0[v]-nug[v])*(1-exp(-dist/a0[v]));
      %end;
      %if %upcase(&model)=SINE %then %do;
c_u1[j]=nug[v]+(c0[v]-nug[v])*(1-sin(arcos(-1)*dist/a0[v])/
(arcos(-1)*dist/a0[v]));
      %end;
      %if %upcase(&model)=PENTASPHERICAL %then %do;
if dist<=a0[v] then c_u1[j]=nug[v]+(c0[v]-nug[v])*
(1.875*dist/a0[v] -1.25*(dist/a0[v])**3 +0.375*(dist/a0[v])**5);
else c_u1[j]=c0[v];
      %end;
      %if %upcase(&model)=CUBIC %then %do;
if dist<=a0[v] then c_u1[j]=nug[v]+(c0[v]-nug[v])*
(7*(dist/a0[v])**2 -8.75*(dist/a0[v])**3 +3.5*(dist/a0[v])**5
- 0.75*(dist/a0[v])**7);
else c_u1[j]=c0[v];
      %end;
      %if %upcase(&model)=SPHERICAL %then %do;
if dist<=a0[v] then c_u1[j]=(nug[v]+(c0[v]-nug[v])*
(1.5*dist/a0[v]-0.5*(dist/a0[v])**3));else c_u1[j]=c0[v];
      %end;
    end;
    c_u[1:n]=c_u1;
  end;
  else if v=nvar+1 then do;
    c_u1=j(m,1,1);
    do j=1 to m;
      dist=sqrt((POINTS1[1]-COORDx[j,1])**2+(POINTS1[2]-COORDx[j,2])**2);
      %if %upcase(&model)=GAUSSIAN %then %do;
c_u1[j]=nug[v]+(c0[v]-nug[v])*(1-exp(-dist**2/(a0[v]**2)));
      %end;
      %if %upcase(&model)=EXPONENTIAL %then %do;
c_u1[j]=nug[v]+(c0[v]-nug[v])*(1-exp(-dist/a0[v]));
      %end;
      %if %upcase(&model)=SINE %then %do;
c_u1[j]=nug[v]+(c0[v]-nug[v])*(1-sin(arcos(-1)*dist/a0[v])/
(arcos(-1)*dist/a0[v]));
      %end;
      %if %upcase(&model)=PENTASPHERICAL %then %do;
if dist<=a0[v] then c_u1[j]=nug[v]+(c0[v]-nug[v])*

```

```

(1.875*dist/a0[v] -1.25*(dist/a0[v])**3 +0.375*(dist/a0[v])**5);
else c_u1[j]=c0[v];
%end;
%if %upcase(&model)=CUBIC %then %do;
if dist<=a0[v] then c_u1[j]=nug[v]+(c0[v]-nug[v])*
(7*(dist/a0[v])**2 -8.75*(dist/a0[v])**3 +3.5*(dist/a0[v])**5
- 0.75*(dist/a0[v])**7);
else c_u1[j]=c0[v];
%end;
%if %upcase(&model)=SPHERICAL %then %do;
if dist<=a0[v] then c_u1[j]=(nug[v]+(c0[v]-nug[v])*
(1.5*dist/a0[v]-0.5*(dist/a0[v])**3));else c_u1[j]=c0[v];
%end;
end;
c_u[(v-nvar)*n+1:n+m]=c_u1;
end;
end;
%if &tabx ne %then %do;
c_u[n+m+2:n+m+nvar]=0;
%end;
lambda=inv(C)*c_u;*print lambda;
sigma2=(c_u[1:n]/c_u[n+m+1])'*(lambda[1:n]/lambda[n+m+1]);
POINTS1[3]=lambda[1:n]'*(z);
POINTS1[4]=sqrt(abs(sigma2));
if _v_=1 then POINTS=POINTS1;else POINTS=POINTS//POINTS1;
end;
end;
end;

```

Capítulo 4

ANÁLISE DOS RESULTADOS

4.1 INTRODUÇÃO

Neste capítulo serão apresentados os resultados obtidos das análises. Primeiramente foi realizada uma análise descritiva dos dados verificando a quantidade de pontos e sua distribuição no espaço, a fim de evitar problemas computacionais. Depois são mostradas as estimativas obtidas a partir do método de Cokrigagem, esclarecendo os critérios de escolhas dos modelos. Por fim será feita uma comparação com as estimativas encontradas utilizando Krigagem ordinária. Todos os resultados foram gerados utilizando o *software* SAS 9.2.

4.2 ANÁLISE DESCRITIVA

Verificando o banco de dados de cada variável, observamos que a variável temperatura está distribuída de uma forma satisfatória pela região e a quantidade de pontos para esta uma nos permite uma boa estimação, o que não acontece com a variável área desmatada.

Agora, observando a variável precipitação média mensal, verificamos que a grande quantidade de satélites fornece uma grande quantidade de pontos que, além

de dificultar a realização das estimativas computacionalmente, são muito próximos, cobrindo uma grande superfície fazendo com que a estimação nesta área não seja necessária. Na Figura 4.1, temos os pontos observados por todos os satélites no mês de dezembro. Podemos notar uma grande concentração de pontos entre o Maranhão e o Pará.

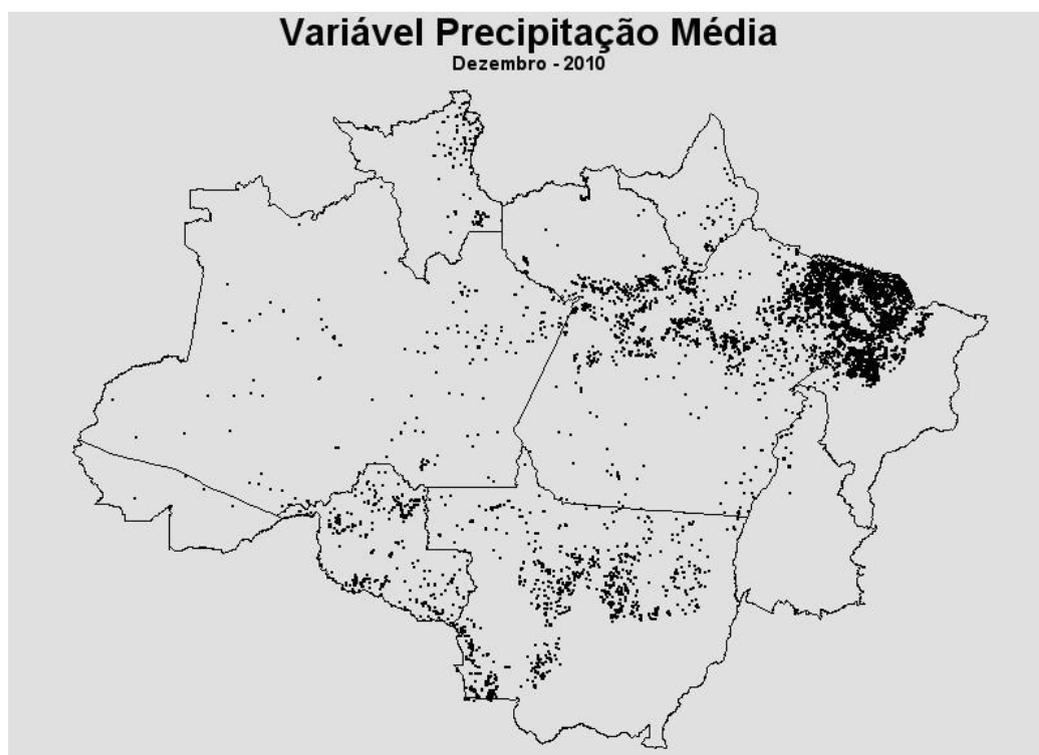


Figura 4.1: Pontos de todos os satélites onde a variável precipitação média é observada no mês de Dezembro.

Na Figura 4.2, podemos observar como os pontos estão próximos. Na Tabela 4.1 temos um exemplo de como as frequências de todos os satélites estão distribuídas no mês de Janeiro. Para este mês, teríamos um total de 2907 pontos, os quais, para o cálculo do variograma das distâncias, deveríamos calcular as distâncias de todas as combinações de pontos par a par. Ou seja, deveríamos calcular C_2^{2907} que é mais que 4 milhões de distâncias, o que é inviável computacionalmente.

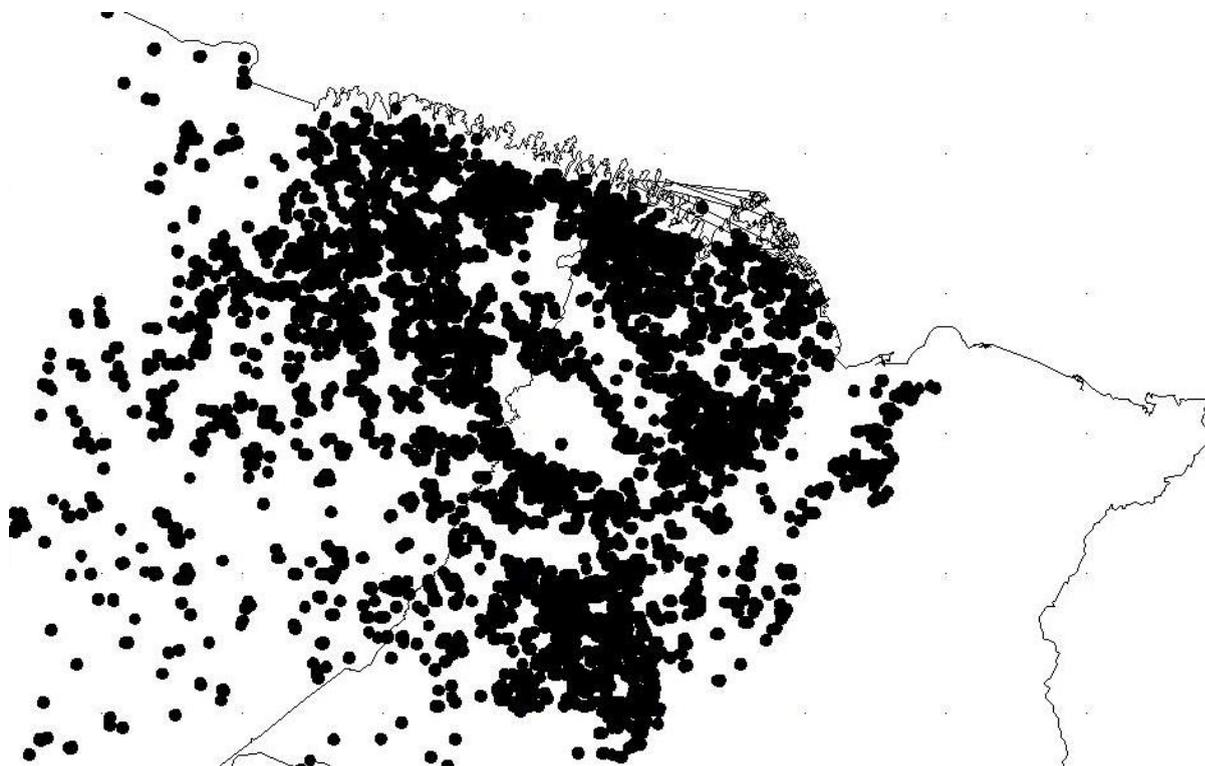


Figura 4.2: Ampliação da figura anterior na região entre o Maranhão e o Pará.

Assim, para podermos utilizar esse banco de dados nas estimações, foi necessário escolher algum satélite para tentar reduzir o número de pontos de observação.

Segundo o INPE (2008), o satélite de referência temporal, ou seja, o satélite cujos dados diários de focos detectados são usados para compor a série temporal ao longo dos anos, permitindo assim a análise de tendências em períodos de interesse, é o satélite NOAA-15, por isso utilizaremos os pontos deste em nossas estimativas.

Apesar de escolhido um satélite, ainda verificamos que existiam muitos pontos, pontos em que as coordenadas de latitude e longitude se diferenciavam somente na quarta ou quinta casa decimal. Por este motivo agrupamos os pontos próximos fazendo uma aproximação das coordenadas para uma casa decimal e calculamos a média da precipitação em cada grupo.

Tabela 4.1: Frequência do número de pontos de todos os satélites para o mês de Janeiro de 2010

Satélite	Frequência	Frequência Acumulada
AQUA-T	167	167
AQUA-M-M	36	203
AQUA-M-T	1199	1402
GOES-12	310	1712
METEOSAT-02	4	1716
NOAA-15	327	2043
NOAA-15D	14	2057
NOAA-16	103	2160
NOAA-16N	5	2165
NOAA-17	20	2185
TERRA-M	11	2196
TERRA-T	28	2224
TERRA-M-M	194	2418
TERRA-M-T	489	2907

Dessa forma conseguimos diminuir o número de pontos para os meses de Janeiro e Dezembro sem prejudicar a qualidade dos dados. Porém a quantidade de pontos para o mês de Julho ainda é elevada (Tabela 4.2) e assim utilizaremos em nosso estudo somente os meses de Janeiro e Dezembro.

Tabela 4.2: Frequência do número de pontos agrupados do satélite NOAA-15 nos meses 2010

Mês	Frequência	Frequência Acumulada
1	209	209
2	105	314
3	235	549
4	113	662
5	203	865
6	299	1164
7	886	2050
8	1802	3852
9	3148	7000
10	2222	9222
11	758	9980
12	198	10178

Podemos agora verificar a distribuição dos pontos de observação de cada variável

na região Amazônica para os meses de interesse:

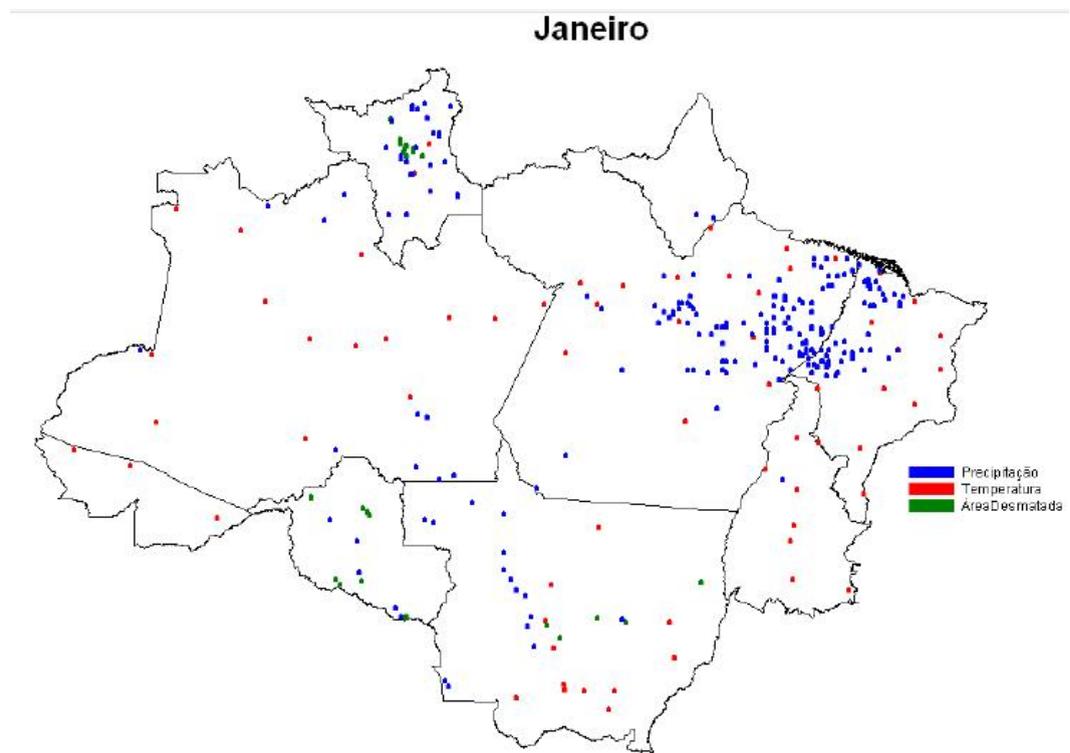


Figura 4.3: Distribuição dos pontos observados de cada variável no mês de Janeiro

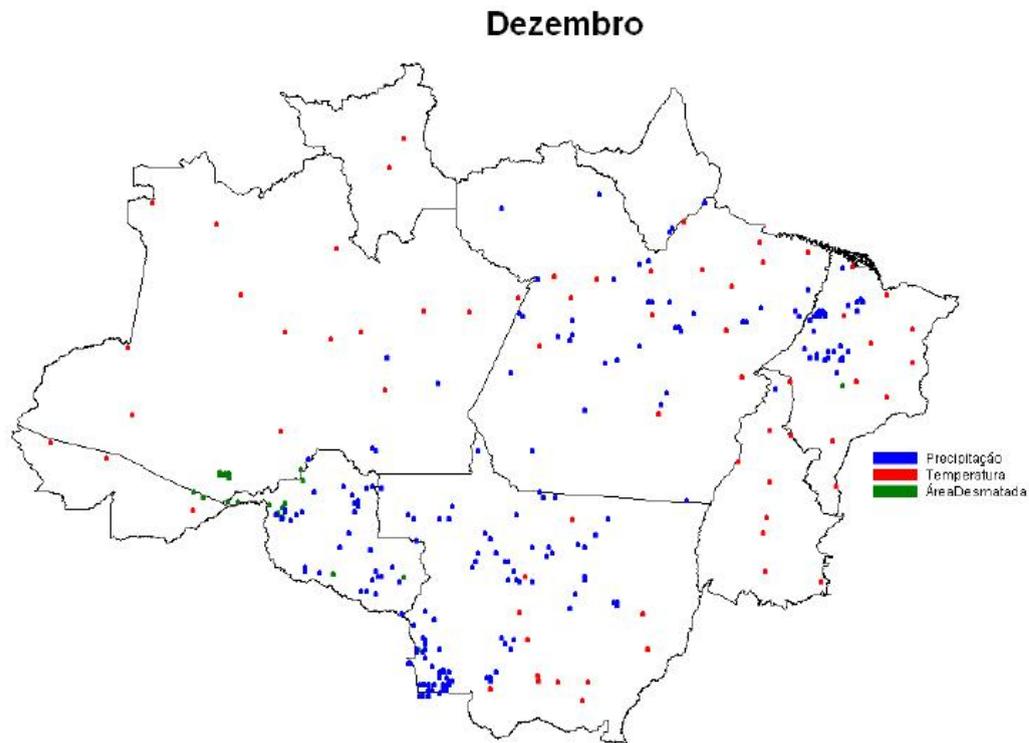


Figura 4.4: Distribuição dos pontos observados de cada variável no mês de Dezembro

4.3 ANÁLISE UTILIZANDO O MÉTODO DE COKRIGAGEM ORDINÁRIA

Para a obtenção da melhor estimativa, devemos modelar a variância espacial cruzada entre as variáveis em estudo. O modelo encontrado deve ajustar bem a covariância e também as variâncias separadamente, porém deve-se buscar o melhor ajuste para a variável principal, que neste estudo é a temperatura. Para isso um estudo das distribuições de distâncias é muito importante na melhor definição do número de classes e do tamanho dessas classes para o cálculo das semivariâncias. Esses dois valores devem ser definidos de forma que um número suficiente de pontos esteja em cada classe. Neste estudo de caso utilizaremos as covariáveis Precipitação média mensal e Área de Desmatamento na cokrigagem, fazendo as comparações com

o método de krigagem nos meses de janeiro e dezembro de 2010.

Comportamento da temperatura média utilizando como covariável a Precipitação Média Mensal

Observando a distribuição das distâncias dos pontos para a temperatura no mês de Janeiro (Figura 4.5), verificamos que um número de classes suficientes para o cálculo da semivariância é 8 e o tamanho dessas classes em torno de 2.

Distribuição dos Pares de Distâncias

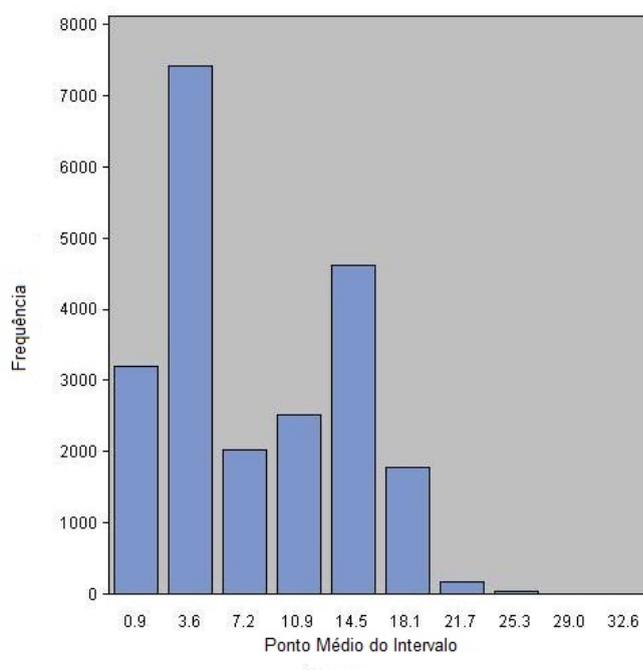


Figura 4.5: Distribuição dos pares de distâncias da Precipitação para o mês de Janeiro de 2010

Agora devemos selecionar o modelo que melhor ajusta as variâncias, para isso será utilizado o Critério de Informação de Akaike (AIC). Na Tabela 4.3 temos esses valores para cada modelo.

Observando os valores AIC podemos perceber que os modelos que melhor ajustam os três variogramas são o modelo seno e o modelo gaussiano. Agora devemos escolher

Tabela 4.3: AIC de cada modelo para os variogramas de cada variável e o variograma cruzado no mês de janeiro (Covariável Precipitação)

Modelo	Temperatura	Precipitação Média	Temperatura x Precipitação
Seno	-30.07437	-14.33925	-24.8805
Gaussiano	-29.86502	-5.107684	-25.23472
Esférico	-25.71046	-5.107405	4.384774
Pentaesférico	-25.63466	-5.106628	-4.384774
Exponencial	18.230263	15.86941	31.212908
Cúbico	-29.4361	-5.10768	-4.384774

o modelo que apresenta menor desvio padrão nas estimativas. Foi possível perceber que a amplitude do desvio padrão das estimativas no modelo seno é menor e portanto este modelo será o escolhido. Na Figura 4.6 podemos verificar o ajuste do variograma cruzado.

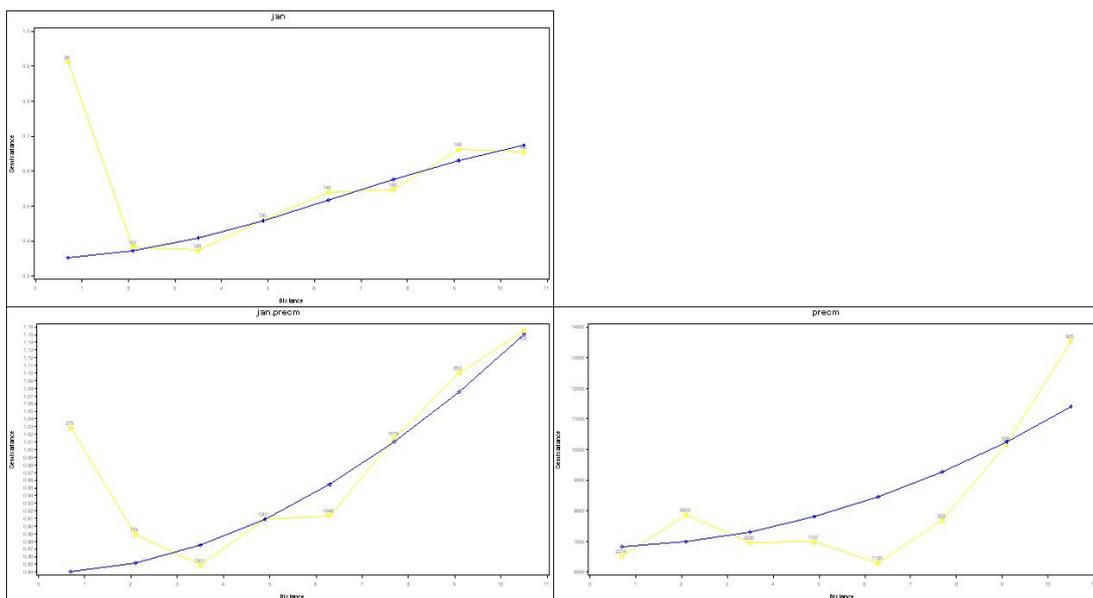


Figura 4.6: Variograma cruzado ajustado a partir do modelo seno (Covariável Precipitação)

Na Figura 4.7 temos as estimativas para a temperatura no mês de Janeiro de 2010 e seus erros padrões, os pontos em azul são da covariável e os pontos em preto são da variável principal.

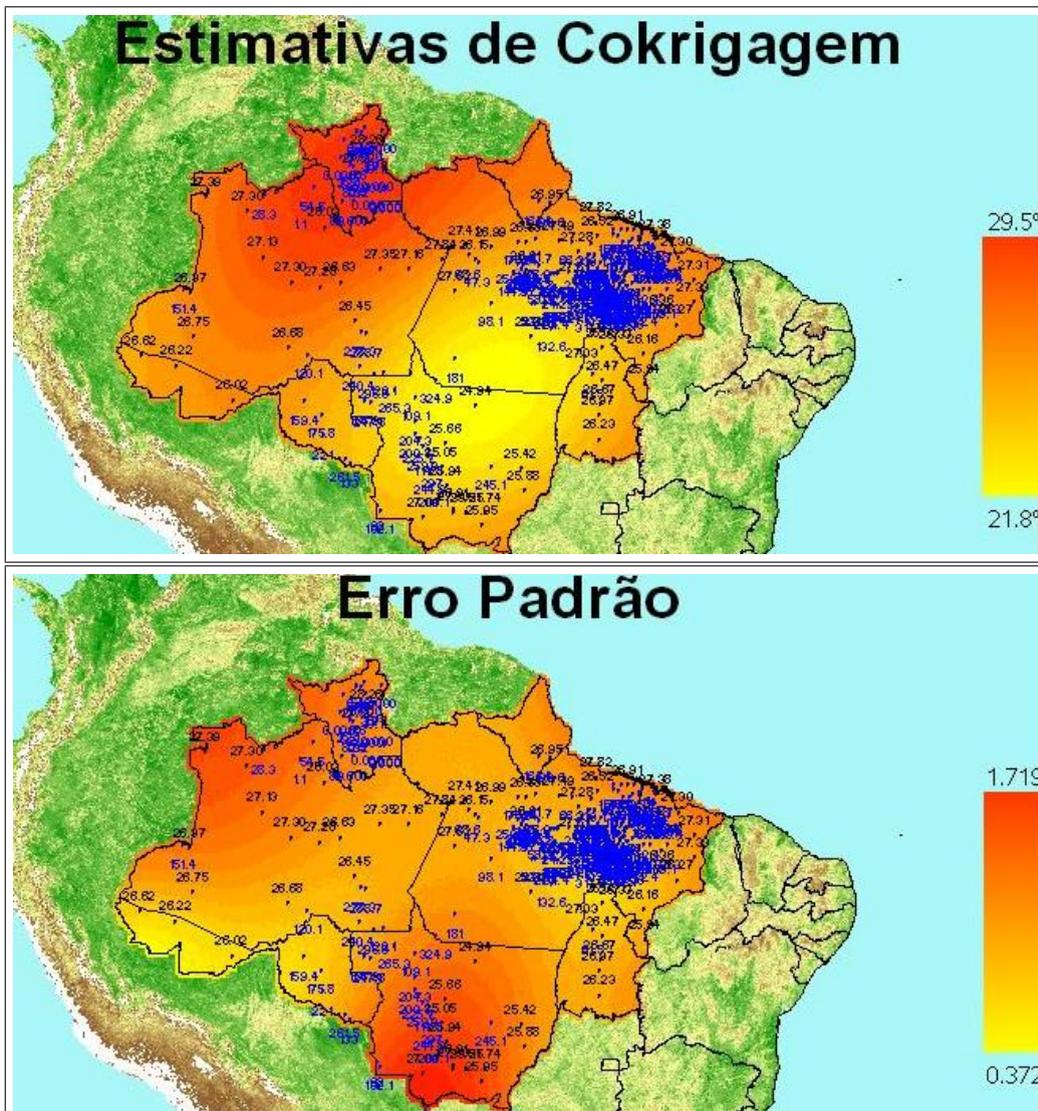


Figura 4.7: Temperatura estimada na região amazônica no mês de janeiro no ano de 2010 e seus erros padrões (Covariável Precipitação).

Observando a distribuição das distâncias dos pontos para a temperatura no mês de Dezembro (Figura 4.8), verificamos que um número de classes suficientes para o cálculo da semivariância é 10, e com tamanho dessas classes em torno de 2.

Agora devemos selecionar o modelo que melhor ajusta as variâncias, para isso será utilizado o Critério de Informação de Akaike (AIC). Na Tabela 4.4 temos esses valores para cada modelo.

Observando os valores AIC, podemos perceber que os modelos que melhor ajustam

Distribuição dos Pares de Distâncias

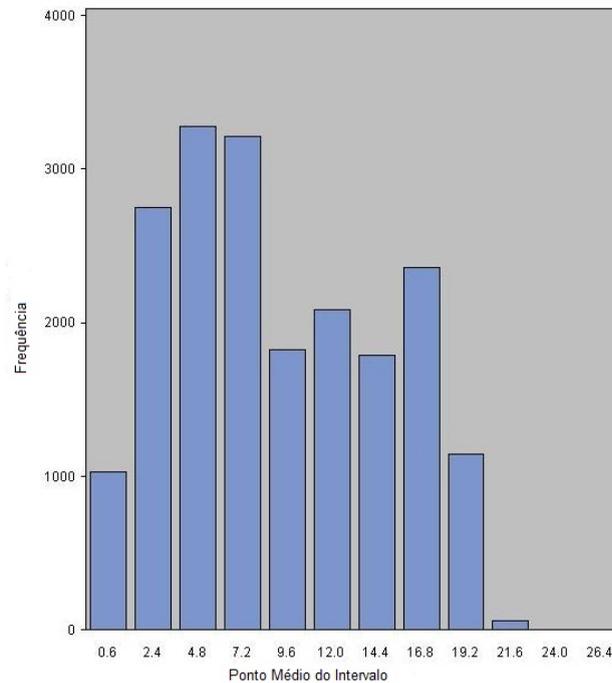


Figura 4.8: Distribuição dos pares de distâncias da variável Precipitação para o mês de Dezembro de 2010

Tabela 4.4: AIC de cada modelo para os variogramas de cada variável e o variograma cruzado no mês de Dezembro (Covariável Precipitação)

Modelo	Temperatura	Precipitação Média	Temperatura x Precipitação
Seno	-22.17775	-22.58054	-15.19399
Gaussiano	-19.87028	-37.69577	-12.0402
Esférico	-19.98542	-37.98729	-11.18124
Pentaesférico	-18.74764	-37.77881	243.77088
Exponencial	14.626337	4.0779226	28.197938
Cúbico	-20.19569	-37.84094	238.8346

tam os três variogramas são o modelo seno, o modelo gaussiano e o modelo esférico.

Agora devemos escolher o modelo que apresenta menor desvio padrão nas estimativas. Foi possível perceber que a amplitude do desvio padrão das estimativas no modelo esférico é menor e portanto este modelo será o escolhido. Na Figura 4.9 podemos verificar o ajuste do variograma cruzado.

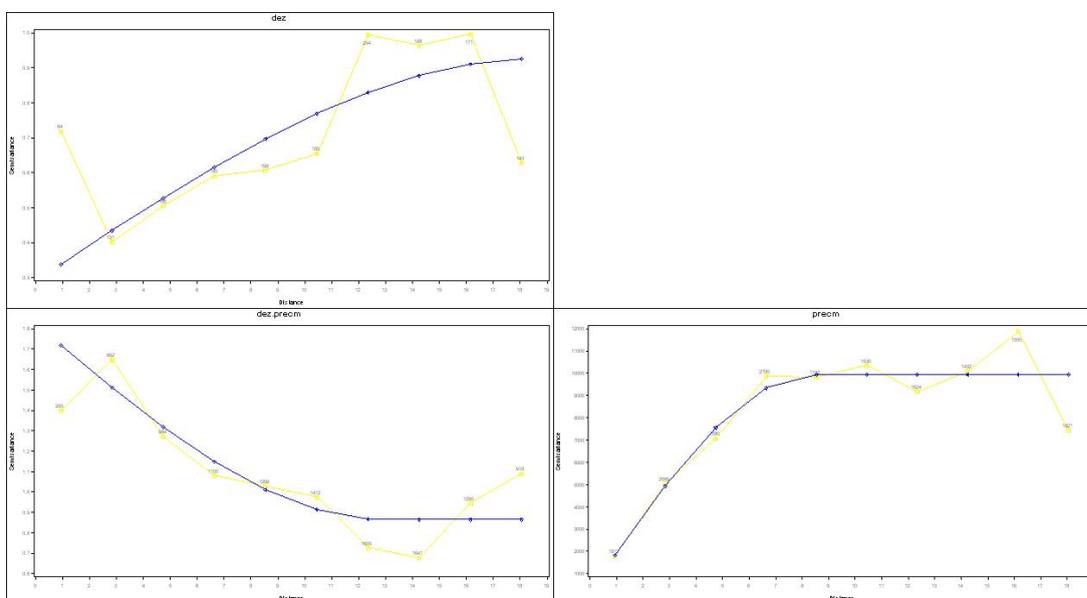


Figura 4.9: Variograma cruzado ajustado a partir do modelo esférico (Covariável Precipitação)

Na Figura 4.10 temos as estimativas para a temperatura no mês de Dezembro de 2010 e seus erros padrões.

Comportamento da temperatura média utilizando como covariável a Área de Desmatamento

Observando a distribuição das distâncias dos pontos para a área no mês de Janeiro (Figura 4.11), verificamos que um número de classes suficientes para o cálculo da semivariância assim como para a variável precipitação é 8 e com o tamanho dessas classes em torno de 2.

Selecionamos agora o modelo que melhor ajusta as variâncias, para isso será utilizado o Critério de Informação de Akaike (AIC). Na Tabela 4.5 são apresentados esses valores para cada modelo.

Observando os valores AIC, percebemos que o modelo que melhor ajusta as três variâncias é, assim como para a covariável precipitação, o modelo seno. A seguir

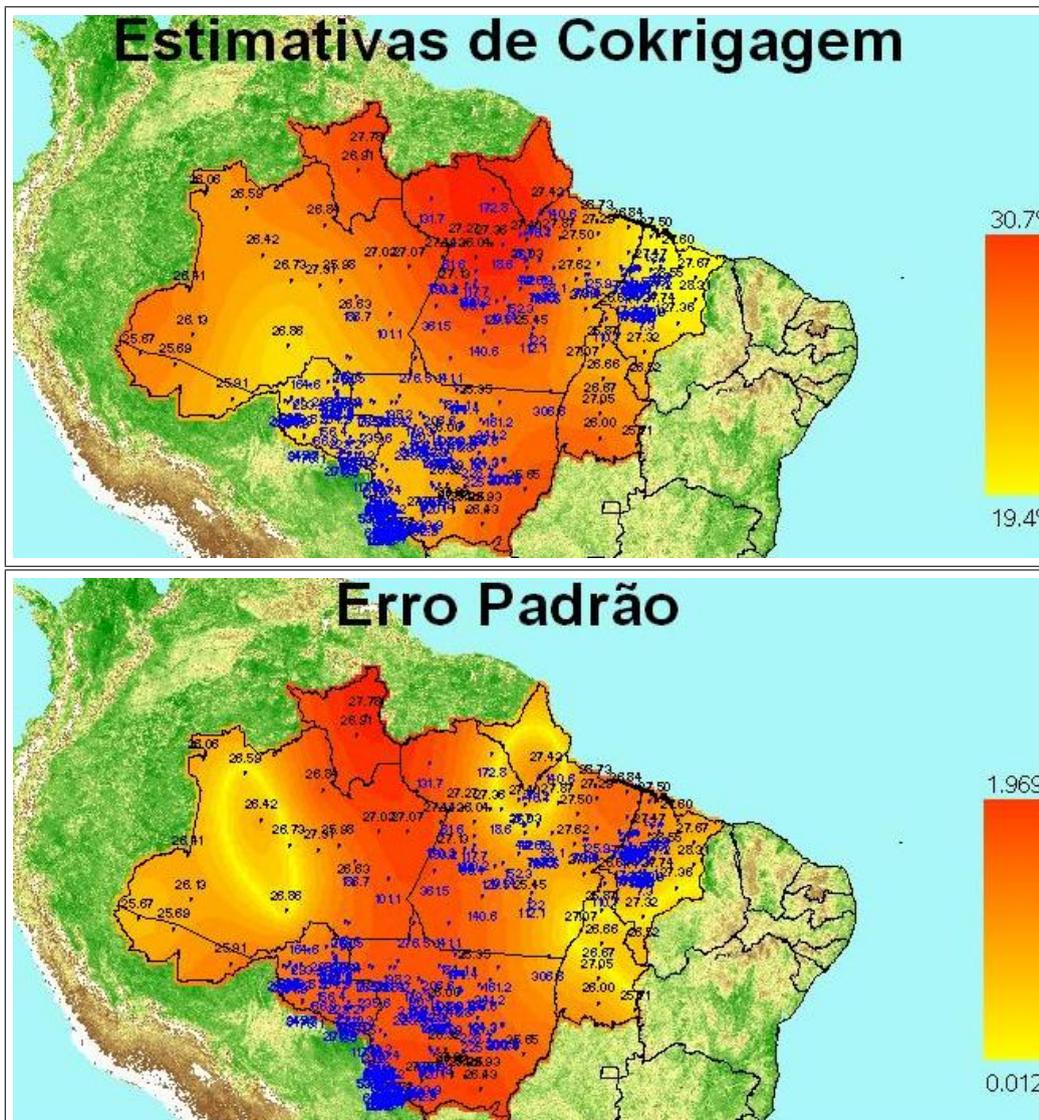


Figura 4.10: Temperatura estimada na região amazônica no mês de dezembro no ano de 2010 e seus erros padrões (Covariável Precipitação).

(Figura 4.12) temos o respectivo variograma cruzado ajustado.

Na Figura 4.13 temos as estimativas para a temperatura no mês de Janeiro de 2010 e seus erros padrões.

Observando a distribuição das distâncias dos pontos para a área no mês de Dezembro (Figura 4.14), verificamos que existem classes que não possuem pontos. Dessa forma o estudo da variância será feito nas primeiras 6 classes e o tamanho de

Distribuição dos Pares de Distâncias

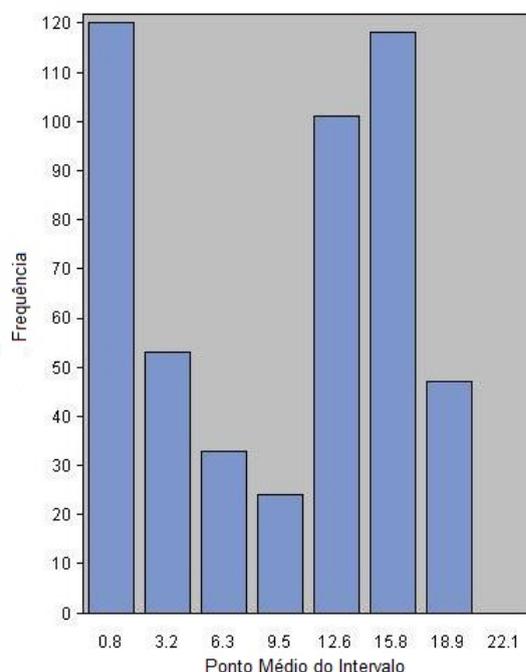


Figura 4.11: Distribuição dos pares de distâncias da variável Área de desmatamento para o mês de Janeiro de 2010

cada classe ficará em torno de 1.

Selecionamos agora o modelo que melhor ajusta as variâncias, para isso será utilizado o Critério de Informação de Akaike (AIC). Na Tabela 4.6 estão esses valores para cada modelo.

Observando os valores AIC, percebemos que o modelo que melhor ajusta as três

Tabela 4.5: AIC de cada modelo para os variogramas de cada variável e o variograma cruzado no mês de Janeiro (Covariável Área de desmatamento)

Modelo	Temperatura	Área de Desmatamento	Temperatura x Área
Seno	-30.07437	-11.94229	12.052311
Gaussiano	-29.86502	-8.107064	21.497483
Esférico	-25.71046	-7.010447	169.37514
Pentaesférico	-25.63466	-5.09448	209.6697
Exponencial	18.230263	5.9313132	28.326118
Cúbico	-29.4361	-8.107064	21.4974

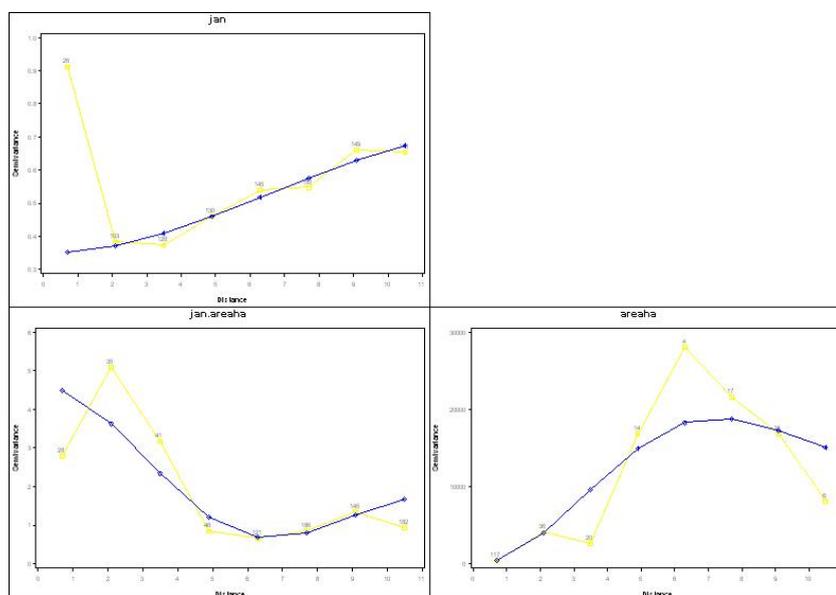


Figura 4.12: Variograma cruzado ajustado a partir do modelo seno (Covariável Área de desmatamento)

Tabela 4.6: AIC de cada modelo para os variogramas de cada variável e o variograma cruzado mês de Dezembro (Covariável Área de desmatamento)

Modelo	Temperatura	Área de Desmatamento	Temperatura x Área
Seno	15.196293	2.8576044	-6.450642
Gaussiano	31.393866	5.8271822	-7.161542
Esférico	25.477877	5.8287456	-6.874834
Pentaesférico	11.245449	5.8296984	-6.613972
Exponencial	15.932699	24.29327	5.8023047
Cúbico	11.245451	5.8273202	-7.221495

variâncias é o modelo seno, porém fica claro que esses valores são muito maiores que os valores dos ajustes anteriores, o que mostra um ajuste pior, que é consequência da distribuição dos pares de distâncias. A seguir (Figura 4.15), temos o variograma cruzado ajustado.

Na Figura 4.16 temos as estimativas para a temperatura no mês de Dezembro de 2010 e seus erros padrões.

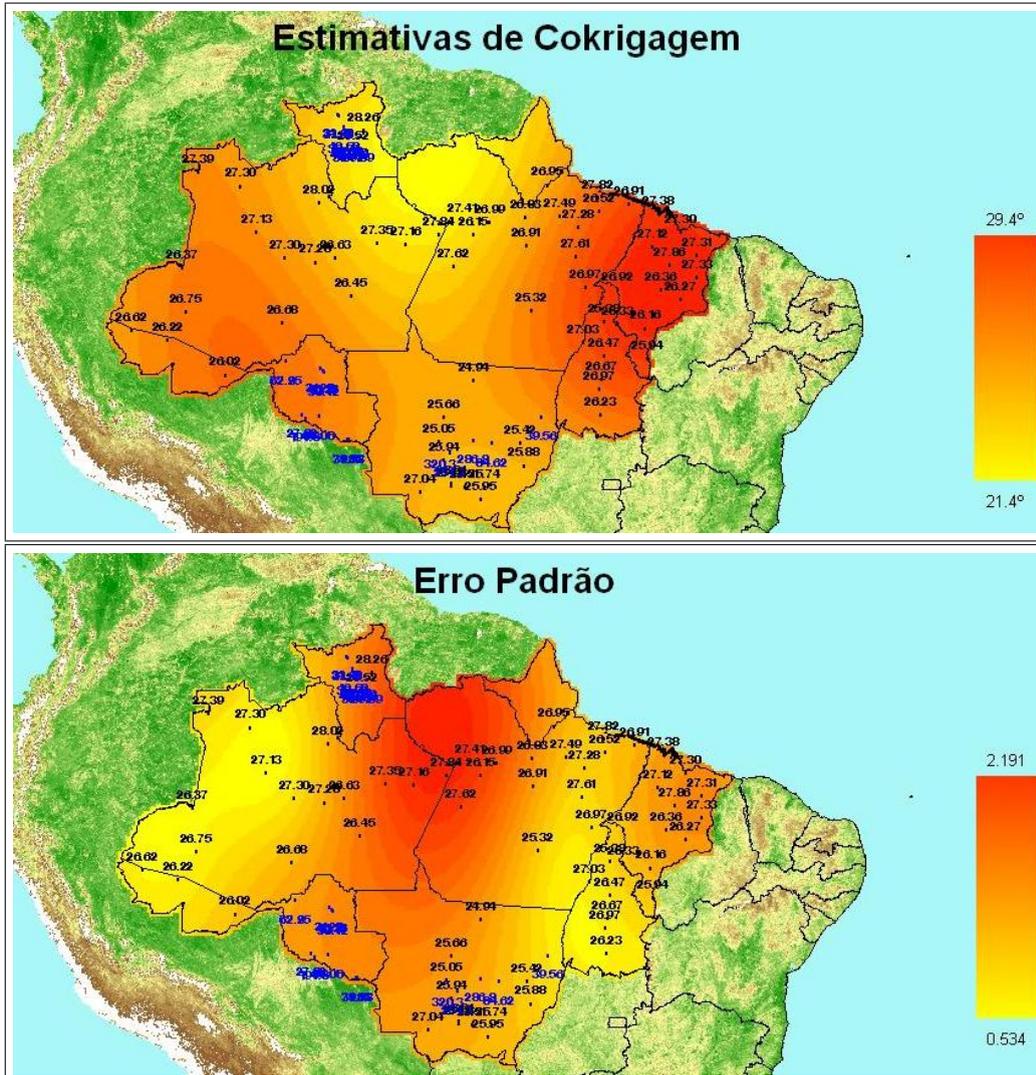


Figura 4.13: Temperatura estimada na região amazônica no mês de janeiro no ano de 2010 e seus erros padrões (Covariável Área de desmatamento).

Distribuição dos Pares de Distâncias

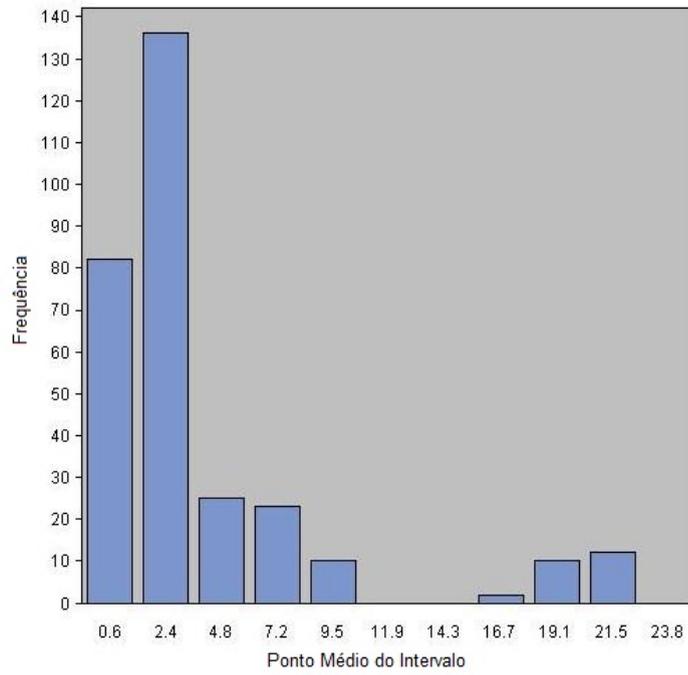


Figura 4.14: Distribuição dos pares de distâncias da variável Área de Desmatamento para o mês de Dezembro de 2010

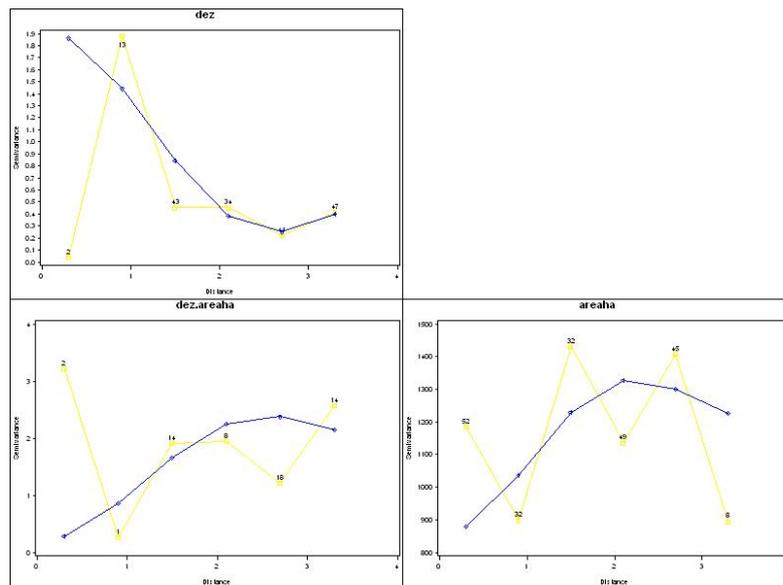


Figura 4.15: Variograma cruzado ajustado a partir do modelo seno (Covariável Área de desmatamento)

4.4 COMPARAÇÃO ENTRE AS CO-VARIÁVEIS

Agora podemos realizar uma comparação entre as covariáveis a fim de descobrir aquela que se associa mais a temperatura e consequentemente aquela que traz mais benefícios nas estimativas. Vamos mostrar como exemplo o mês de Janeiro. Na Figura 4.17 podemos comparar as estimativas.

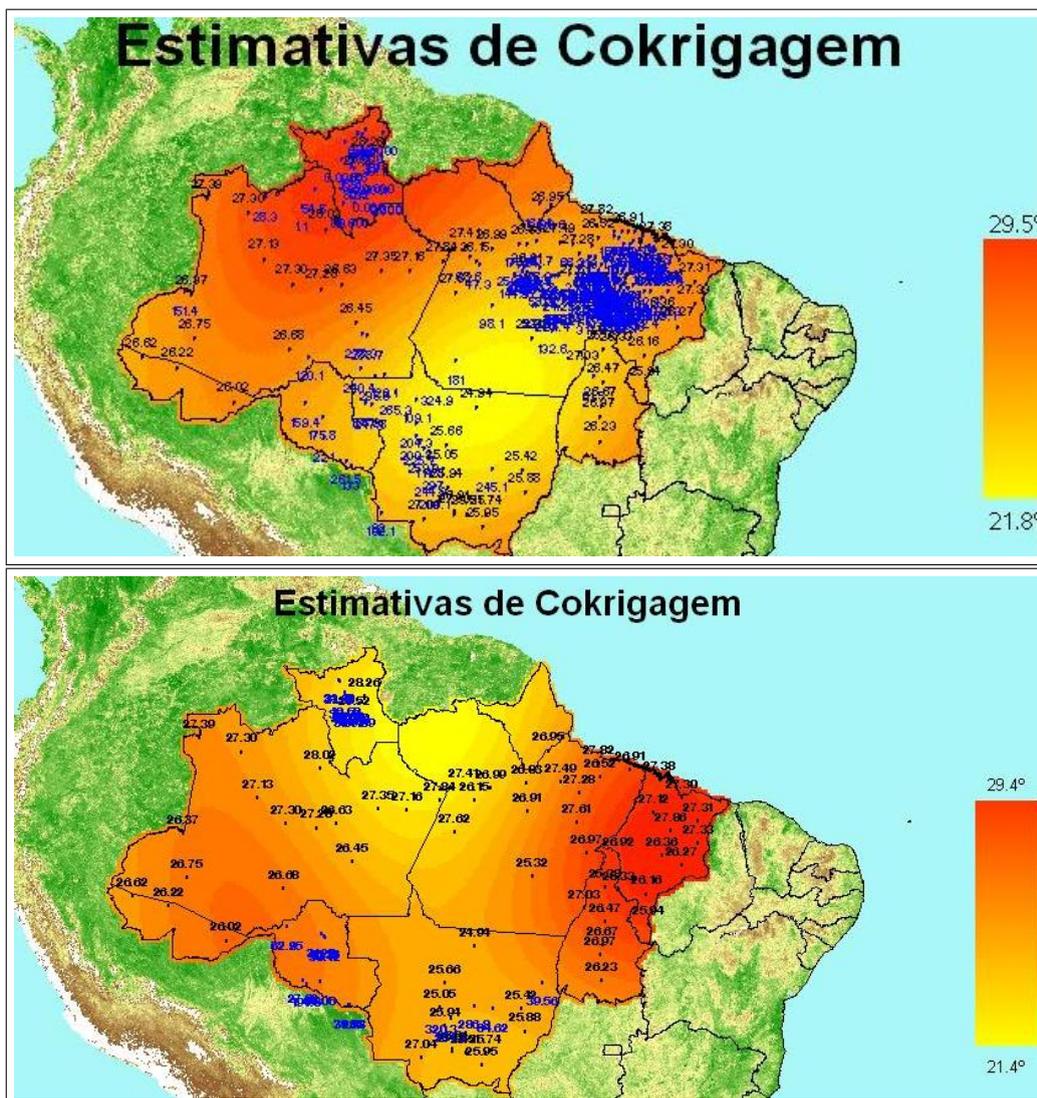


Figura 4.17: Temperatura estimada na região amazônica no mês de janeiro no ano de 2010 para as covariáveis Precipitação e Área Desmatada respectivamente.

Agora podemos comparar os desvios padrões Figura 4.18:

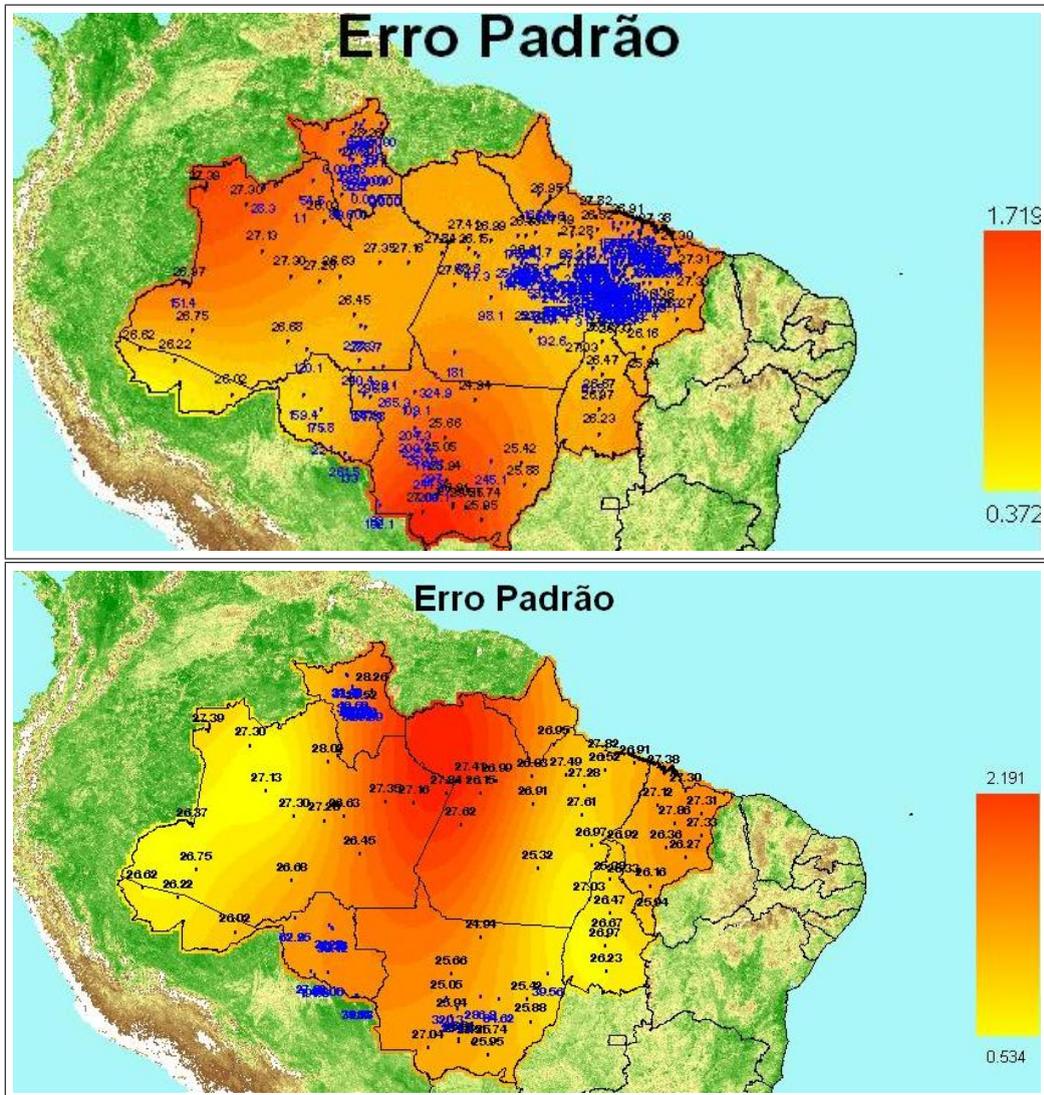


Figura 4.18: Erros Padrões da temperatura estimada na região amazônica no mês de janeiro no ano de 2010 para as covariáveis Precipitação e Área Desmatada respectivamente.

Podemos perceber que o erro padrão da cokrigagem utilizando a covariável Precipitação Média mensal tem menor amplitude e magnitude, portanto esta parece ser a variável mais associada a temperatura. Assim podemos avaliar como essa correlação se comporta.

A partir da Figura 4.19 onde temos as estimativas da temperatura com a co-

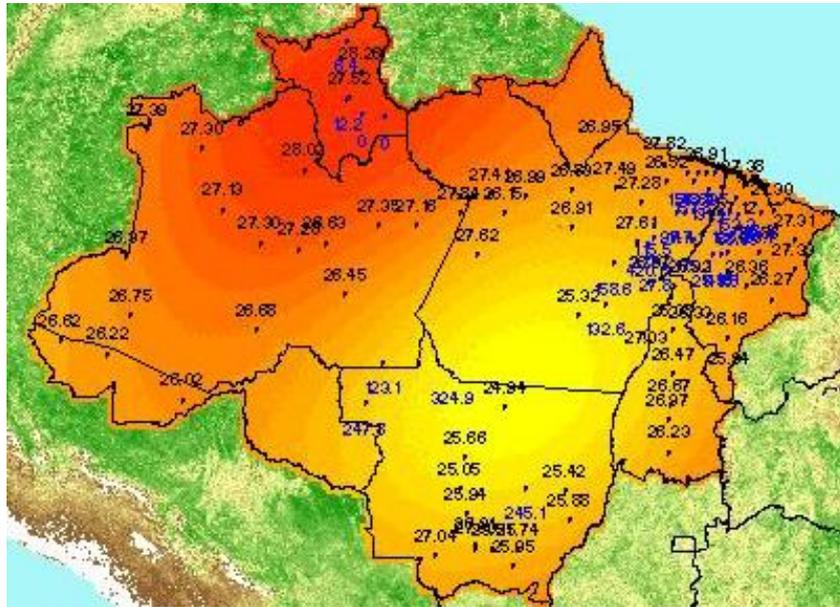


Figura 4.19: Temperatura estimada na região amazônica no mês de janeiro no ano de 2010 com menos pontos da covariável (Covariável Precipitação)

variável Precipitação Média mensal com apenas 30 pontos, podemos perceber que, na área mais clara, onde a estimativa da temperatura é menor, a precipitação é maior, e para onde a área é mais escura, a precipitação parece ser menor.

4.5 RESULTADOS UTILIZANDO O MÉTODO DA KRIGAGEM ORDINÁRIA

Agora utilizaremos os resultados das estimações por krigagem para posteriormente realizar as comparações com a cokrigagem. Para isso, a krigagem será feita como em Barreto (2011), a partir do modelo gaussiano.

Comportamento da temperatura média no mês de janeiro utilizando o modelo gaussiano

Na Figura 4.20 tem-se os resultados provenientes da estimacão por krigagem ordinária para o mês de janeiro de 2010.

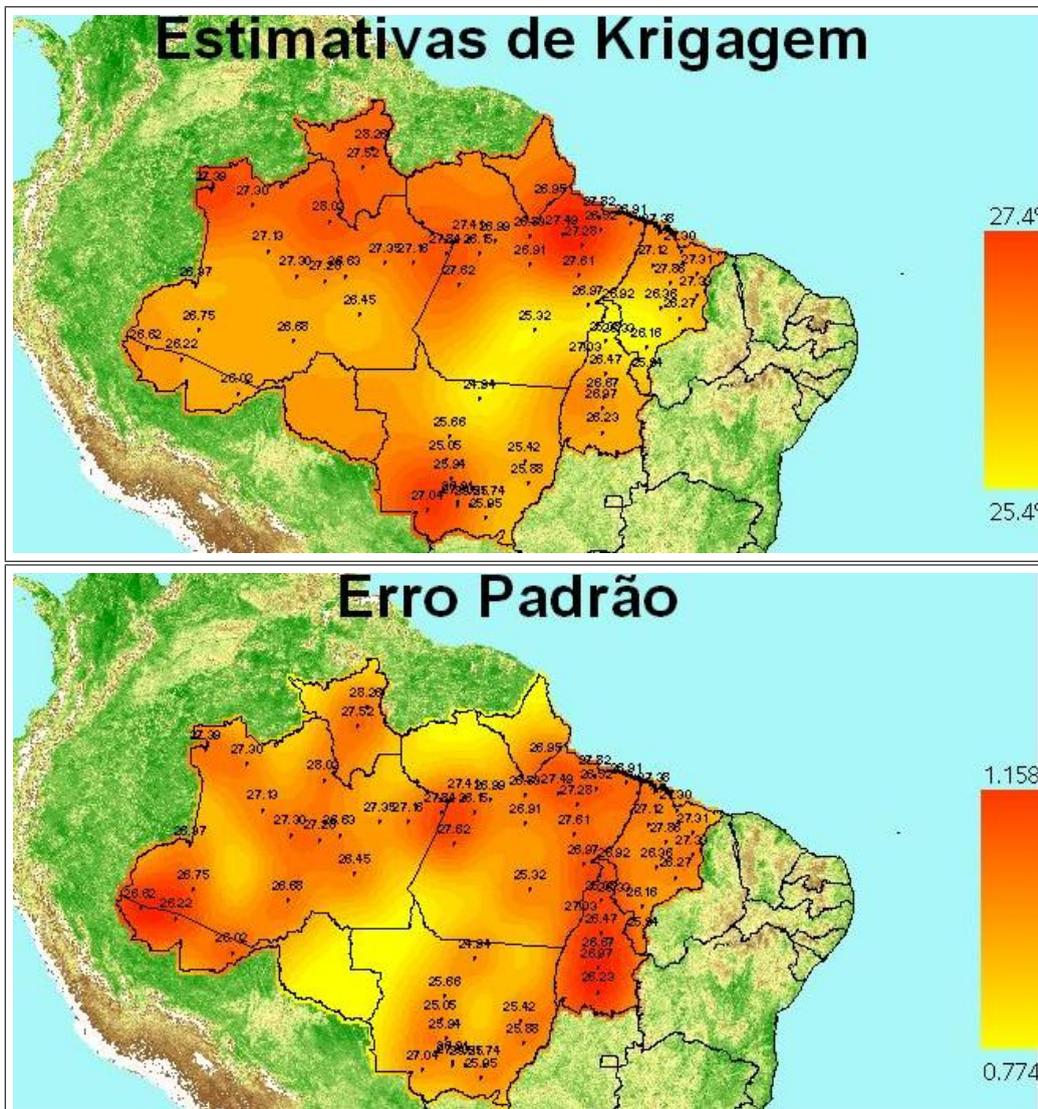


Figura 4.20: Temperatura estimada na região amazônica no mês de janeiro no ano de 2010 e seus erros padrões (Krigagem) .

Comportamento da temperatura média no mês de dezembro utilizando o modelo gaussiano

Na Figura 4.21 tem-se os resultados provenientes da estimação por krigagem ordinária para o mês de dezembro no ano de 2010.

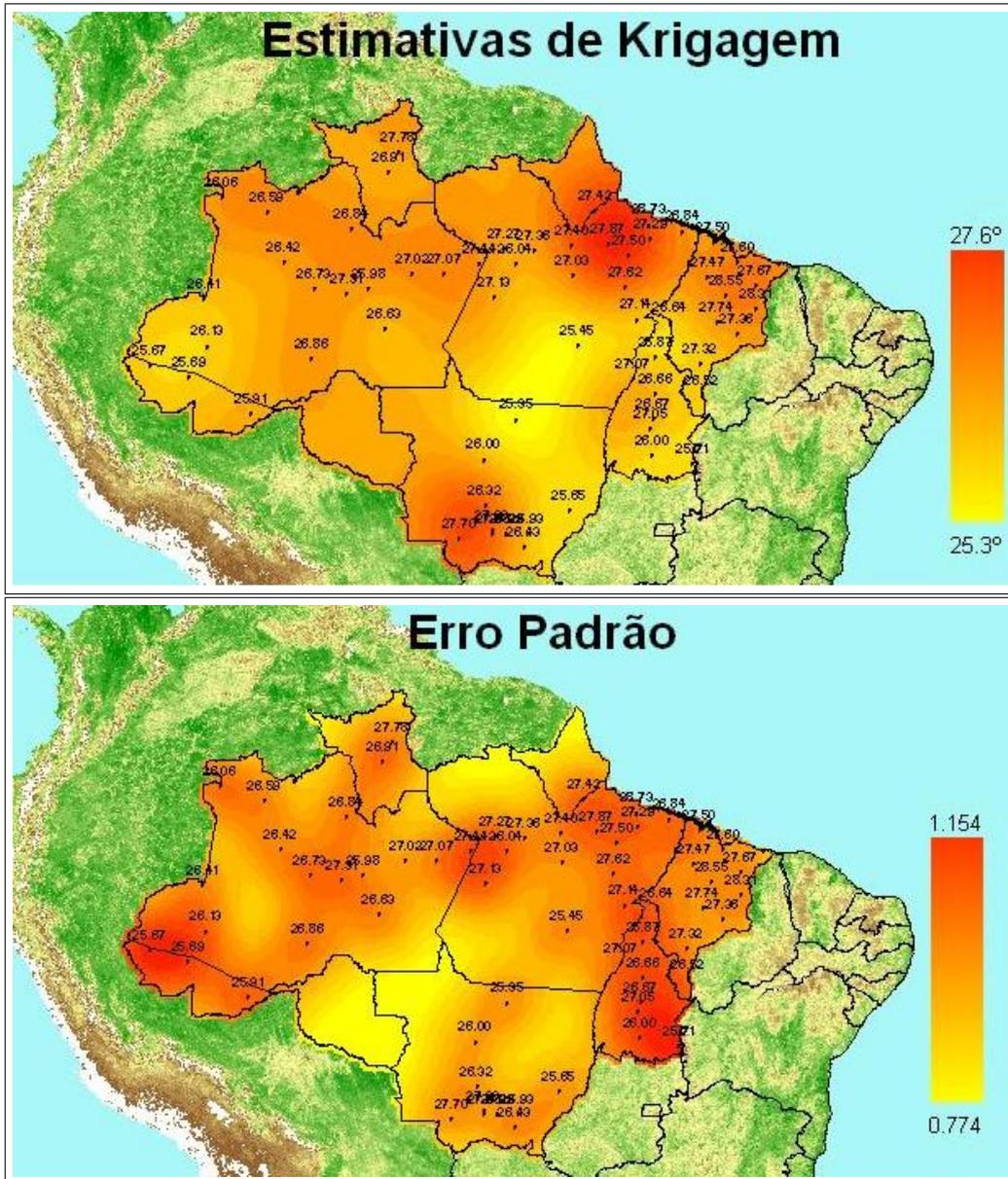


Figura 4.21: Temperatura estimada na região amazônica no mês de janeiro no ano de 2010 e seus erros padrões (Krigagem).

4.6 COMPARAÇÃO ENTRE KRIGAGEM E COKRIGAGEM

Agora faremos a comparação entre a qualidade das estimativas, para isso é necessário avaliar a diferença entre os erros.

Nos mapas abaixo temos a diferença dos erros entre a técnica de Krigagem e a técnica de Cokrigagem. Onde os valores são positivos, observamos que as estimativas foram melhores utilizando a Cokrigagem e, quando os valores são negativos, a Krigagem foi melhor. Neste mapas a variável principal está na cor azul e a covariável na cor vermelha.

de janeiro (Figura 4.22), uma área muito grande apresenta a cokrigagem como técnica com menor desvio padrão, ou seja, quando a diferença entre o desvio da krigagem e o desvio da cokrigagem é maior que zero. A magnitude da diferença é muito parecida.

Para o mês de dezembro (Figura 4.23), observamos que a magnitude das diferenças também são semelhantes e, da mesma forma, a área onde a cokrigagem é melhor, é maior.

Covariável Área de Desmatamento

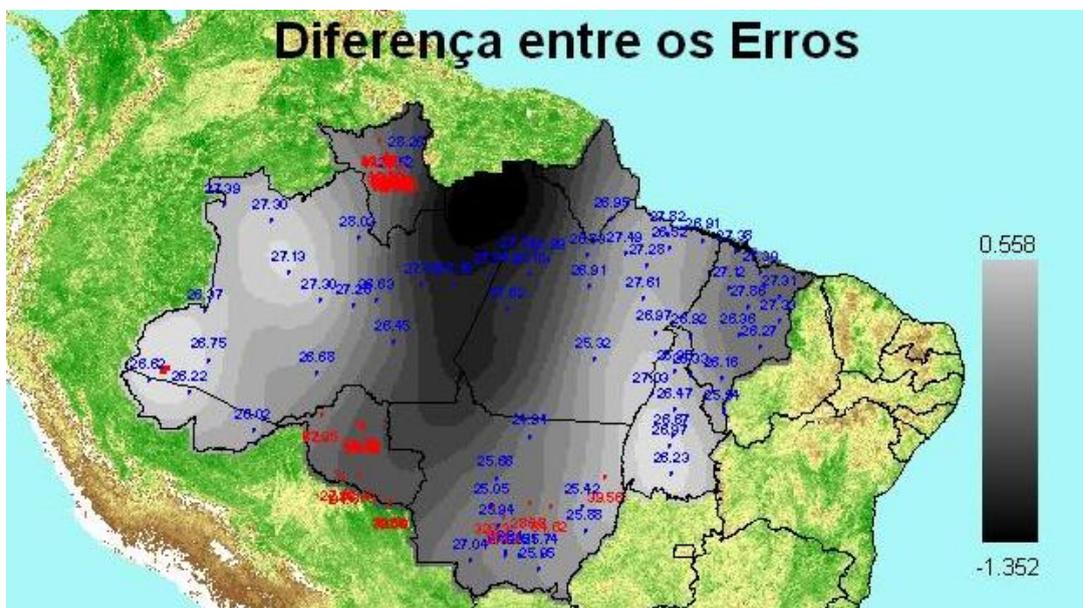


Figura 4.24: Diferença entre erros para o mês de janeiro de 2010 (Covariável Área Desmatada)



Figura 4.25: Diferença entre erros para o mês de dezembro de 2010 (Covariável Área Desmatada)

Para a variável área de desmatamento no mês de janeiro (Figura 4.24), verifi-

camos que a área em que a cokrigagem é melhor, é maior, porém a magnitude do desvio padrão é maior quando a krigagem é melhor, ou seja, a diferença onde a cokrigagem é melhor (0,558) é menor que o módulo da diferença quando a krigagem é melhor (1,352).

Para o mês de dezembro (Figura 4.25), notamos que a cokrigagem foi pior nos dois sentidos, tanto na magnitude da diferença dos desvios padrões quanto no tamanho da área em que ela é melhor.

Essa piora nas estimativas pode estar associada a um ajuste ruim da variância da variável área de desmatamento e ao baixo número de pontos amostrados.

Capítulo 5

CONCLUSÕES

A implementação do algoritmo para as estimações por cokrigagem não existe em nenhum procedimento do *software* SAS 9.2, logo os códigos desenvolvidos em SAS/IML para a obtenção do variograma cruzado e da cokrigagem são de grande valia para a aplicação deste método. Outras vantagens são a fácil conversão em outras linguagens e a possibilidade de inclusão de outros modelos teóricos.

A partir destes dois estudos pudemos verificar que a cokrigagem nem sempre será melhor, porém apresenta ganhos significativos em algumas regiões de estimações. Também foi possível perceber que um ajuste ruim para a variância da covariável pode resultar em estimativas com desvios padrões maiores que os da krigagem.

Foi possível também, observar que a covariável Precipitação Mensal está mais relacionada à temperatura e, portanto, é a covariável que mais ajuda na melhoria das estimativas. Além disso, é possível perceber que a correlação parece seguir um sentido negativo, ou seja, onde a temperatura é maior a precipitação é menor e onde esta é menor, aquela é maior. Não foi possível encontrar um sentido da relação entre temperatura e a variável Área de Desmatamento devido à falta de pontos para o auxílio nas estimativas e à dificuldade de ajustamento.

Uma sugestão para melhores estimações seria tentar combinar as duas técnicas utilizando a krigagem e o ganho da cokrigagem, mas sempre partindo de um bom ajuste variográfico da 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 Brasília.
- Camargo, E. C. G., Felgueiras, C. A., & Monteiro, A. M. V. (2001). A importncia da modelagem da anisotropia na distribuio espacial de variáveis ambientais utilizando procedimentos geoestatísticos. Technical report, INPE.
- 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.
- Fearnside, P. M. (2005). Desmatamento na amaznia brasileira: histria, índices e consequências. *Megadiversidade*, 1(1):113–123.
- IBAMA (2010). Ecossistemas brasileiros: Amaznia. Technical report. URL <http://www.ibama.gov.br/ecossistemas/amazonia.htm>. Acesso em 16 out. 2010.
- INMET (2011). Rede de estaes. Technical report, Instituto nacional de meteorologia. Acesso em 12 de outubro fde 2011.
- INPE (2008). Monitoramento da cobertura florestal da amaznia por satélites. Technical report, Instituto Nacional de Pesquisas Espaciais. Acesso em 12 de outubro de 2011.
- INPE (2011). Diviso de satélites e sistemas ambientais. Technical report, Instituto Nacional de Pesquisas Espaciais. Acesso em 12 de outubro de 2011.
- Memarsadeghi, N., Moigne, J. L., Mount, D. M., & Morissette, J. (2005). A new approach to image fusion based on cokriging. In: *7th International Conference on Information Fusion*.

- Souza, N. J. (2011). Influência das queimadas da amazônia sobre o efeito estufa. Technical report, Laboratório de Proteção Florestal - Universidade Federal do Paraná. URL <http://www.floresta.ufpr.br/~lpf/efeitoestufa.html>. Acesso em 13 jul. 2011.
- 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.