

ÉRICA DOS SANTOS RIOS

## **O EFEITO DE BORDA NA GEOESTATÍSTICA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

VIÇOSA  
MINAS GERAIS – BRASIL  
2018

Ficha catalográfica preparada pela Biblioteca Central da Universidade Federal de  
Viçosa - Campus Viçosa

T

R586e  
2018 Rios, Érica dos Santos, 1993-  
O efeito de borda na Geoestatística / Érica dos Santos Rios. -  
Viçosa, MG, 2018.  
vii, 43f. : il. (algumas color.) ; 29 cm.

Inclui anexo.

Orientador: Gérson Rodrigues dos Santos.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f.31.

1. Geologia - Métodos estatísticos. 2. Análise espacial  
(Estatística). I. Universidade Federal de Viçosa. Departamento de  
Estatística. Programa de Pós-graduação em Estatística Aplicada e  
Biometria. II. Título.

CDD 22 ed. 551

ÉRICA DOS SANTOS RIOS


**O EFEITO DE BORDA NA GEOESTATÍSTICA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

APROVADA: 19 de fevereiro de 2018.

  
Adriana Maria Rocha T. Santos

  
Lidiane Maria Ferraz Rosa

  
Gérson Rodrigues dos Santos  
(Orientador)

*“Dedico este trabalho à minha família.”*

## AGRADECIMENTOS

A Deus, cujo sustento e proteção me mantém seguindo adiante. Mesmo em Viçosa, longe da família e amigos, Ele esteve sempre comigo.

A minha família por sempre acreditar no meu sucesso acadêmico e profissional.

Ao meu namorado Marcello Lira por fazer dos meus problemas, os nossos problemas. O seu amor por mim me ajuda a entender como Cristo me vê e o quanto Ele me ama.

A meus amigos João Paulo, Luiza Amélia e Jéssica Tamara que oraram por mim nos momentos que eu mais precisei.

À Rosielle Farias (Rosita) que com seu jeitinho especial de ser alegrou os meus dias em Viçosa.

À Universidade Federal de Viçosa, principalmente aos professores do Programa de Pós-Graduação em Estatística Aplicada e Biometria, por todos os ensinamentos que me ajudaram a aprimorar meus conhecimentos nessa ciência que eu tanto admiro chamada Estatística.

Aos secretários da Pós-Graduação em Estatística Aplicada e Biometria Carla Zinato e Júnior Pires, pelo atendimento rápido e eficiente sempre que precisei.

Ao Departamento de Engenharia Civil da UFV, por ter cedido o conjunto de dados utilizado nesse trabalho.

Aos professores Gérson Rodrigues dos Santos e Afonso de Paula dos Santos cujos conselhos e orientação foram imprescindíveis na realização desse trabalho.

À Lidiane Maria Ferraz Rosa, Adriana Maria Rocha Trancoso e Nilcilene das Graças Medeiros por terem aceitado o convite para compor a banca.

A CAPES pelo apoio financeiro, sem o qual meu mestrado seria apenas um sonho.

*“Portanto, quer comais, quer bebais ou façais outra coisa  
qualquer, fazei tudo para a glória de Deus.”  
I Coríntios 10:31*

## SUMÁRIO

<b>RESUMO</b> .....	<b>vi</b>
<b>ABSTRACT</b> .....	<b>vii</b>
<b>1. INTRODUÇÃO</b> .....	<b>1</b>
<b>2. REFERENCIAL TEÓRICO</b> .....	<b>3</b>
<b>2.1. Amostragem</b> .....	<b>3</b>
<b>2.2. Semivariograma</b> .....	<b>4</b>
2.2.1 Modelo exponencial de semivariograma .....	7
2.2.1 Modelo esférico de semivariograma.....	8
2.2.3 Modelo gaussiano de semivariograma .....	8
<b>2.3. Krigagem</b> .....	<b>8</b>
2.3.1 Krigagem simples .....	9
2.3.1 Variância de krigagem.....	10
<b>3. MATERIAL E MÉTODOS</b> .....	<b>11</b>
<b>4. RESULTADOS E DISCUSSÃO</b> .....	<b>13</b>
<b>4.1. Resultados da seleção regular</b> .....	<b>13</b>
4.1.1 Cálculo da variância de krigagem .....	16
4.1.2 Resultados da variância de krigagem .....	18
<b>4.2. Resultados da seleção irregular</b> .....	<b>21</b>
4.2.1 Resultados da variância de krigagem .....	25
<b>4.3. Comparação entre os resultados da seleção regular e da irregular</b> .....	<b>29</b>
<b>CONCLUSÃO</b> .....	<b>30</b>
<b>REFERÊNCIAS</b> .....	<b>31</b>
<b>ANEXO</b> .....	<b>32</b>

## RESUMO

RIOS, Érica dos Santos, M.Sc., Universidade Federal de Viçosa, fevereiro de 2018. **O efeito de borda na Geoestatística.** Orientador: Gérson Rodrigues dos Santos. Coorientadores: Afonso de Paula dos Santos e João Marcos Louzada.

A Geoestatística é uma ciência que se dedica a predição de valores de um fenômeno regionalizado a partir de pontos amostrados em uma área de interesse. Apesar da existência de métodos ótimos de predição, como a krigagem, ainda há alguns problemas a serem solucionados nessa ciência. Um desses problemas é o efeito de borda que ocorre na predição de pontos localizados na região mais extrema da área de interesse. Ele acontece porque ao se prever nessa região há pouca ou nenhuma vizinhança amostrada. Dessa forma, o erro de predição é maior na borda do que na região central da área de interesse. Uma estratégia bastante empregada para diminuir o efeito de borda é a utilização de pontos amostrados fora da área de interesse na predição. No entanto, pouca informação é encontrada a respeito da eficácia dessa estratégia. Portanto, esse trabalho busca avaliar e mensurar a melhora obtida nas predições ao se utilizar dados externos à área de interesse. De acordo com os resultados obtidos, a utilização de dados externos à área de interesse é eficaz na diminuição do efeito de borda. Essa eficácia se mostrou maior para uma amostragem irregular do que para uma amostragem regular. O ganho na precisão das predições se torna pequeno na utilização de pontos distantes da área de interesse para a interpolação.

## ABSTRACT

RIOS, Érica dos Santos, M.Sc., Universidade Federal de Viçosa, February, 2018. **The border effect in Geostatistics.** Adviser: Gérson Rodrigues dos Santos. Co-advisers: Afonso de Paula dos Santos and João Marcos Louzada.

Geostatistics is a science that is concerned about the prediction of values of a regionalized phenomenon. Despite the existence of good interpolation methods, such as kriging, there are still a lot of problems to be solved in this science. One of these problems is the border effect, which occurs in the prediction of points located near the boundary of the area of interest. It happens because when the value of a point at the boundary is predicted there are a few or none sampling points in the neighborhood. Therefore, the prediction error is bigger at the boundary of the area of interest than in the center of this area. A strategy widely known to reduce the border effect is to use sampling points outside the area of interest; however, only little information is found regarding the efficacy of this strategy. Hence, this study aims to evaluate and measure the improvement in prediction obtained by using data outside the area of interest. According to the results obtained, the use of data outside the area of interest is efficient on reducing the border effect. This efficacy was higher for an irregular sampling than for a regular sampling. The gain in precision of the predictions becomes small when points away from the area of interest are used for interpolation.

## 1. INTRODUÇÃO

A Geoestatística passou a ser amplamente utilizada no início da década de 1980, quase trinta anos após seu surgimento (YAMAMOTO e LANDIM, 2015). Essa ciência surgiu na África do Sul em 1951, quando o engenheiro de minas Daniel G. Krige concluiu que as variâncias dos dados de concentração de ouro só faziam sentido ao se considerar a distância entre os pontos amostrais (VIEIRA et al., 2000).

De acordo com Yamamoto e Landim (2015), a partir do trabalho de Krige e dos trabalhos de H.J. de Wijs, o professor George Matheron apresentou na década de 1960 vários trabalhos que formalizaram a Teoria das Variáveis Regionalizadas. Como o objeto de estudo da Geoestatística são as variáveis regionalizadas, Matheron é considerado o criador dessa ciência.

Segundo Vieira et al. (2000), Matheron afirma que uma variável regionalizada é uma função espacial numérica, cujo valor varia de um local para outro, que apresenta continuidade aparente e sua variação não pode ser representada por uma função matemática simples. Logo, para essas variáveis o local onde cada ponto foi amostrado influencia diretamente no valor da variável.

Uma variável regionalizada possui um aspecto aleatório (imprevisível) e um aspecto estruturado (previsível), que reflete o padrão estrutural do fenômeno regionalizado. Na Geoestatística, a teoria sobre essas variáveis é utilizada para descrever a correlação espacial de um fenômeno regionalizado e prever valores de uma variável regionalizada com base nos pontos amostrais (YAMAMOTO e LANDIM, 2015).

Por ser relativamente nova, quando comparada a outros ramos da estatística, há ainda muitos aspectos da Geoestatística a serem explorados. Um deles é o efeito de borda, problema comumente relatado na predição de valores de uma variável regionalizada com base nos pontos amostrais. O surgimento desse problema é consequência de um dos princípios dessa ciência – a primeira Lei da Geografia. Segundo Tobler (1970), essa lei diz que “Todas as coisas são relacionadas, mas coisas mais próximas são mais relacionadas”.

De acordo com Harzheim e Warnecke (2010), o efeito de borda significa que a acurácia do modelo decresce na borda da área de interesse devido ao fato dos pontos amostrais nesse local não estarem rodeados de outros pontos amostrais na vizinhança. Isso

ocorre porque a estimativa de um ponto não amostrado depende dos valores dos pontos amostrais, principalmente daqueles que estão mais próximos.

De acordo com Walvoort, Brus e Gruijter (2010), esse problema intensifica-se quando a região de estudo tem um formato irregular ou quando a variável regionalizada em estudo não permite que uma amostragem em malha regular seja feita. A fim de minimizar esse efeito de borda, Harzheim e Warnecke (2010) sugerem coletar pontos amostrais fora da área de interesse para a qual se deseja prever, mas perto da mesma.

Lucas, Abreu e Parizzi (2016) utilizando o método de interpolação do inverso do quadrado da distância (IQD), avaliaram visualmente a diferença obtida ao utilizar-se a estratégia sugerida por Harzheim e Warnecke (2010), relatando apenas uma pequena melhora. Entretanto, não há muitas pesquisas iguais que avaliam a eficácia da utilização dessa estratégia.

Assim, objetiva-se com este trabalho mensurar o efeito de borda quanto à interpolação por krigagem, tendo como variante o aumento do número de pontos amostrais (em formato de quatro *buffers*) nas extremidades da área de interesse. Como indicadores de qualidade, adotam-se as medidas estatísticas média e desvio-padrão das variâncias de krigagem. Deseja-se estudar também a influência do tipo de amostragem (regular ou irregular) no efeito de borda.

## **2. REFERENCIAL TEÓRICO**

### **2.1. Amostragem**

Em uma pesquisa estatística, é necessário definir a população de elementos que apresentam o atributo da variável que se deseja estudar. Na maior parte dos casos, essa população é extremamente grande o que torna inviável, por uma questão de tempo e/ou custo, obter o valor da variável para toda a população (BUSSAB e MORETTIN, 2010). Portanto, para estudar o comportamento dessa variável deve-se realizar uma amostragem a fim de se selecionar elementos que representem a população.

Na Geoestatística, uma amostra representativa é um conjunto de valores da variável regionalizada que é capaz de reproduzir a distribuição e a variabilidade espaciais em tamanho, número de pontos amostrais, e em termos de distribuição dos pontos no domínio a ser estudado (YAMAMOTO e LANDIM, 2015).

Yamamoto e Landim (2015) apresentam três tipos de amostragem utilizados na Geoestatística: a amostragem aleatória simples, a amostragem aleatória estratificada e a amostragem sistemática. Na amostragem aleatória simples, as coordenadas geográficas dos pontos que serão amostrados são escolhidas aleatoriamente dentro de todo o domínio da região de estudo (WACKERNAGEL, 2003). Esse tipo de amostragem é também conhecida como amostragem irregular.

Na realização de uma amostragem aleatória estratificada de tamanho  $n$ , a região de estudo é dividida em  $n$  áreas e se seleciona aleatoriamente uma coordenada geográfica para cada ponto amostral dentro de cada área (WACKERNAGEL, 2003). Na amostragem sistemática, as coordenadas geográficas dos pontos que serão amostrados são definidos sobre os nós de uma malha regular (YAMAMOTO e LANDIM, 2015).

Dentre os três tipos de amostragem, o que oferece melhores resultados é a amostragem sistemática, também chamada de regular (YAMAMOTO e LANDIM, 2015). A distribuição espacial dos pontos amostrais é melhor na amostragem sistemática, uma vez que os pontos cobrem a região de estudo de maneira uniforme (WACKERNAGEL, 2003). Portanto, não há agrupamento de pontos amostrais nesse tipo de amostragem, o que pode ocorrer na amostragem aleatória simples e na amostragem aleatória estratificada (YAMAMOTO e LANDIM, 2015). Infelizmente, às vezes o acesso, a vegetação e os acidentes geográficos impedem que a amostragem sistemática seja feita para a variável regionalizada em estudo

(YAMAMOTO e LANDIM, 2015). Nesses casos, a alternativa é fazer a amostragem aleatória simples ou a estratificada.

## 2.2. Semivariograma

Após a amostragem, deve-se realizar uma análise exploratória dos dados, onde são fornecidos gráficos e medidas de resumo, ou seja, estatísticas descritivas que retratam as variáveis e que sejam representativas dos dados. De acordo com Bussab e Morettin (2010) gráficos como o histograma, medidas de posição central como a média e medidas de dispersão como a variância, são capazes de fornecer muito mais informação sobre o comportamento de uma variável do que a tabela original de dados. Em particular, na Geoestatística, um gráfico que faz parte da análise exploratória dos dados é o semivariograma experimental.

Existem dois tipos de semivariograma: o semivariograma teórico e o experimental. O semivariograma experimental capta a estrutura de covariância dos dados observados. Já o semivariograma teórico, que faz parte da etapa de modelagem, estabelece-se por meio da escolha da função matemática que melhor se ajusta ao semivariograma experimental (WEBSTER e OLIVER, 2007).

O semivariograma é uma ferramenta de muito valor para a Geoestatística, pois de acordo com Vieira et al. (2000) ele mede e caracteriza a dependência espacial da variável regionalizada em estudo. Ele é representado pelo símbolo  $\gamma$ .

Para a realização de uma análise geoestatística e o cálculo do semivariograma, considera-se que qualquer ponto  $x$  pertencente a região de estudo da variável regionalizada  $Z$  é uma realização do conjunto de variáveis aleatórias  $Z(x)$  (VIEIRA et al., 2000). Essa família de variáveis aleatórias é denominada função aleatória e denotada por  $Z(x)$  (WACKERNAGEL, 2003).

No cálculo do semivariograma utilizam-se pares de pontos, separados por um vetor de distância  $\mathbf{h}$ , ou seja, as variáveis aleatórias  $Z(x)$  e  $Z(x + \mathbf{h})$ . Ele é definido pela Equação (2.1) (WEBSTER e OLIVER, 2007).

$$\gamma(\mathbf{h}) = \frac{1}{2} E\{[Z(x) - Z(x + \mathbf{h})]^2\} \quad (2.1)$$

De acordo com Webster e Oliver (2007), o estimador do semivariograma é obtido por:

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

onde  $N(\mathbf{h})$  é o número de pares de valores observados,  $Z(x_i)$  e  $Z(x_i + \mathbf{h})$  são os valores da variável regionalizada  $Z$  em pontos amostrais separados por um vetor de distância  $\mathbf{h}$ .

Segundo Cressie (1993), na construção do semivariograma é preciso levar em conta se o fenômeno estudado é unidirecional ou anisotrópico. Um fenômeno é unidirecional quando somente a distância entre os pontos amostrais influencia na correlação entre as observações. Um fenômeno é dito anisotrópico se além da distância, a direção entre os pontos amostrais também influenciarem na correlação entre as observações.

Os quatro parâmetros que o semivariograma possui estão apresentados na Figura 1. O conceito de cada um desses será apresentado de acordo com (WACKERNAGEL, 2003; YAMAMOTO e LANDIM, 2015).

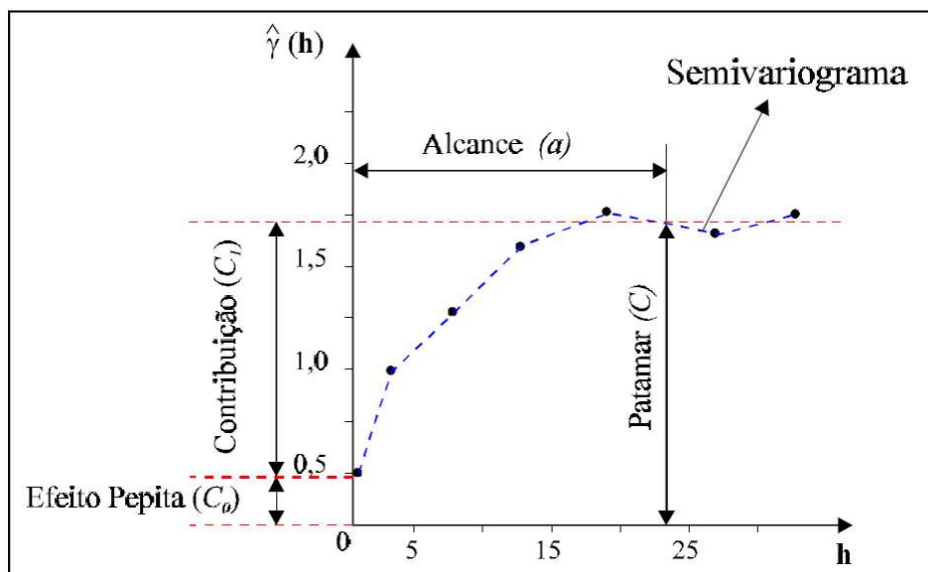


Figura 1: Parâmetros do semivariograma.  
Fonte: Câmara e Medeiros (1998)

Descrevem-se os parâmetros do semivariograma por:

- \* **Efeito pepita ( $C_0$ ):** é o valor do semivariograma para distâncias muito pequenas, sendo, portanto, o limite do semivariograma quando  $\mathbf{h}$  tende a 0. O efeito pepita modela a descontinuidade do semivariograma na origem, que pode ser resultado tanto

da escala de amostragem quanto da variabilidade do fenômeno regionalizado em estudo.

- \* **Alcance ( $a$ ):** é a distância  $h$  máxima para a qual os pares de valores  $Z(x)$  e  $Z(x + h)$  são considerados correlacionados no espaço, de maneira que  $Z(x)$  e  $Z(x + b)$  para  $b > a$  são pares de valores independentes.
- \* **Patamar ( $C$ ):** é o valor do semivariograma para distâncias iguais ou maiores que o alcance, ou seja, quando  $h \geq a$ . À medida que a distância  $h$  aumenta, a variância entre os pares de valores  $Var[Z(x) - Z(x + h)]$  também aumenta. Entretanto, na maioria dos casos, para distâncias elevadas essa variância torna-se aproximadamente constante, atingindo assim uma variância máxima representada pelo patamar.
- \* **Contribuição ( $C_1$ ):** é a diferença entre o patamar ( $C$ ) e o efeito pepita ( $C_0$ ). Portanto, enquanto o efeito pepita é causado pela variância aleatória, a contribuição representa a parte da variabilidade dos dados causada pela dependência espacial.

Após a construção do semivariograma experimental por meio da Equação (2.2), é preciso buscar o modelo, a função matemática, que melhor adequa-se a esse. São diversos os modelos teóricos utilizados para esse ajuste. Os mais utilizados são os modelos exponencial, esférico e gaussiano que estão representados na Figura 2. É importante ressaltar que para essa representação considerou-se um efeito pepita nulo.

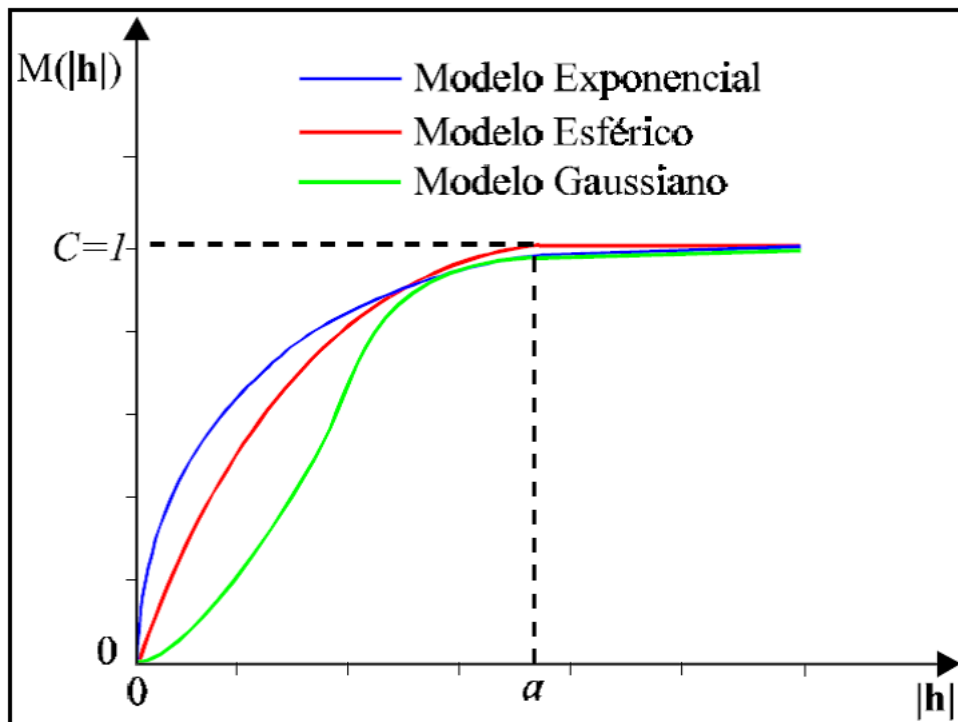


Figura 2: Modelos teóricos do semivariograma  
 Fonte: Camargo (1998) modificado de Isaaks e Srivastava (1989)

De acordo com Cressie (1993) é necessário verificar se o modelo escolhido demonstrou bem a tendência de variabilidade apresentada pelo semivariograma experimental em relação aos valores de  $\mathbf{h}$ , uma vez que um modelo mais adequado pode fornecer predições do valor de pontos não amostrados mais exatas.

### 2.2.1 Modelo exponencial de semivariograma

Pela Figura 2, nota-se que o modelo exponencial apresenta um crescimento muito rápido na origem. Esse modelo aproxima-se de seu patamar assintoticamente e é dado pela Equação (2.3):

$$\gamma(\mathbf{h}) = \begin{cases} 0 & , \text{se } |\mathbf{h}| = 0 \\ C \left[ 1 - \exp\left(-\frac{|\mathbf{h}|}{r}\right) \right] & , \text{se } |\mathbf{h}| > 0 \end{cases} \quad (2.3)$$

onde  $C$  é o patamar,  $\mathbf{h}$  é a distância entre pontos e  $r$  é um parâmetro de distância. O valor do alcance prático é aproximadamente três vezes o valor do parâmetro  $r$  (WEBSTER e OLIVER, 2007).

### 2.2.1 Modelo esférico de semivariograma

Pela Figura 2, percebe-se que o crescimento desse modelo na origem é rápido e linear. O modelo esférico é um dos modelos teóricos de semivariograma mais utilizados na Geoestatística, sendo representado pela Equação (2.4):

$$\gamma(\mathbf{h}) = \begin{cases} 0 & , se |\mathbf{h}| = 0 \\ C \left[ \frac{3}{2} \left( \frac{|\mathbf{h}|}{a} \right) - \frac{1}{2} \left( \frac{|\mathbf{h}|}{a} \right)^3 \right] & , se 0 < |\mathbf{h}| \leq a \\ C & , se |\mathbf{h}| > a \end{cases} \quad (2.4)$$

onde  $C$  é o patamar,  $\mathbf{h}$  é a distância entre pontos e  $a$  é o alcance do modelo (WEBSTER e OLIVER, 2007).

### 2.2.3 Modelo gaussiano de semivariograma

Pela Figura 2, observa-se que o modelo gaussiano apresenta um crescimento lento na origem. Esse modelo aproxima-se do seu patamar assintoticamente e é dado pela Equação (2.5):

$$\gamma(\mathbf{h}) = \begin{cases} 0 & , se |\mathbf{h}| = 0 \\ C \left[ 1 - \exp\left(-\frac{|\mathbf{h}|^2}{r^2}\right) \right] & , se |\mathbf{h}| > 0 \end{cases} \quad (2.5)$$

onde  $C$  é o patamar,  $\mathbf{h}$  é a distância entre pontos e  $r$  é um parâmetro de distância. O valor do alcance é aproximadamente  $\sqrt{3}r$  (WEBSTER e OLIVER, 2007).

## 2.3. Krigagem

Denominada por Matheron como método de predição linear espacial ótima, a krigagem é utilizada para prever o valor da variável regionalizada nos pontos que não foram amostrados, por meio dos valores dos pontos amostrais (CRESSIE, 1993). A krigagem é preferível a outros métodos de interpolação por apresentar predições não tendenciosas e a mínima variância associada ao valor predito (YAMAMOTO e LANDIM, 2015). Esse método pressupõe que a variável em estudo possui estacionariedade de segunda ordem (CRESSIE, 1993).

Segundo Vieira et al. (2000) uma variável regionalizada é estacionária de segunda ordem se o primeiro e o segundo momentos da variável aleatória  $Z(x + \mathbf{h})$  forem os mesmos para qualquer vetor de distância  $\mathbf{h}$ . Logo, uma função aleatória  $Z(x)$  é dita estacionária de segunda ordem quando as Equações (2.6) e (2.7) forem satisfeitas (VIEIRA et al., 2000).

$$E[Z(x)] = \mu \quad (2.6)$$

$$C(\mathbf{h}) = E[Z(x)Z(x + \mathbf{h})] - \mu^2 \quad (2.7)$$

A Equação (2.6) indica que o valor esperado  $E[Z(x)]$  existe e independe da posição do ponto  $x$ . Já a Equação (2.7) aponta que para cada par de variáveis aleatória  $Z(x)$ ,  $Z(x + \mathbf{h})$  a função de covariância  $C(\mathbf{h})$  existe e depende apenas do vetor de distância  $\mathbf{h}$ .

### 2.3.1 Krigagem simples

A krigagem ordinária é a mais usual, mas além dela existem outros tipos de krigagem, como a krigagem simples, a krigagem da média e a krigagem universal (YAMAMOTO e LANDIM, 2015). Entretanto, a krigagem simples deveria ser a mais utilizada, pois de acordo com Santos et al. (2011) ela é um preditor mais preciso do que as outras krigagens.

A krigagem simples assume que a média da variável regionalizada é conhecida (YAMAMOTO e LANDIM, 2015). Desse modo, o valor predito não viesado da variável aleatória  $Z$  no ponto  $x_0$  é dado pela Equação (2.8).

$$\hat{Z}_{KS}(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) + \left\{ 1 - \sum_{i=1}^n \lambda_i \right\} \mu \text{ para } i = 1, 2, \dots, n \quad (2.8)$$

onde  $\lambda_i$  representa o peso de cada ponto amostrado  $x_i$ , ou seja, a influência que cada ponto  $x_i$  tem sobre o ponto  $x_0$ ,  $n$  é o número de pontos amostrais,  $Z(x_i)$  o valor da variável aleatória  $Z$  no ponto  $x_i$  e  $\mu$  é a média dos valores da variável  $Z$  (WEBSTER e OLIVER, 2007).

Para encontrar os respectivos valores dos pesos  $\lambda_i$  é necessário solucionar o sistema de equações lineares (2.9).

$$\sum_{i=1}^n \lambda_i C(x_i, x_j) = C(x_0, x_j) \text{ para } i, j = 1, 2, \dots, n \quad (2.9)$$

onde  $C(x_i, x_j)$  é a função de covariância entre o  $i$ -ésimo e o  $j$ -ésimo ponto amostrado,  $C(x_0, x_j)$  é a função de covariância entre o  $j$ -ésimo ponto amostrado e aquele que se deseja prever (WEBSTER e OLIVER, 2007).

### 2.3.1 Variância de krigagem

Denominada de variância mínima da predição por Vieira et al. (2000) e conhecida como variância do erro por Yamamoto e Landim (2015), a variância de krigagem é uma medida que pode ser usada na comparação de krigagens quanto à incerteza na predição, escolhendo aquela que apresentar menor variância média de krigagem (SANTOS et al., 2011).

De acordo com Wackernagel (2003) essa medida é ainda mais importante quando uma amostragem irregular for realizada, pois fornece uma análise da variação da precisão das predições causada pela disposição irregular dos pontos amostrais. A variância de krigagem para a krigagem simples é dada pela Equação (2.10)

$$\sigma_{KS}^2(x_0) = C(0) - \sum_{i=1}^n \lambda_i C(x_i, x_0) \text{ para } i = 1, 2, \dots, n \quad (2.10)$$

onde  $C(0)$  é a variância da variável regionalizada  $Z$ ,  $\lambda_i$  representa o peso de cada ponto amostrado  $x_i$  e  $C(x_i, x_0)$  é a função de covariância entre o  $i$ -ésimo ponto e aquele que se deseja prever (WEBSTER e OLIVER, 2007).

### 3. MATERIAL E MÉTODOS

A base de dados utilizada neste estudo, consiste em um Modelo Digital de Superfície (MDS) de uma região do município de Viçosa-MG, abrangendo uma área de aproximadamente 5,7 km<sup>2</sup>. Este MDS, cedido pelo Departamento de Engenharia Civil da UFV, foi obtido a partir do processamento fotogramétrico de imagens do satélite Ikonos, apresentando resolução espacial de 5 metros e referenciado ao sistema de projeção UTM fuso 23 sul, com datum SAD-69.

A variável analisada foi a altitude presente no MDS exibido na Figura 3. Utilizou-se uma amostragem sistemática, com malha regular de 5 x 5 metros, obtendo cerca de 230 mil pontos de altitude. Esses dados foram analisados pelo software ArcGIS<sup>TM</sup> (ESRI, 2016).

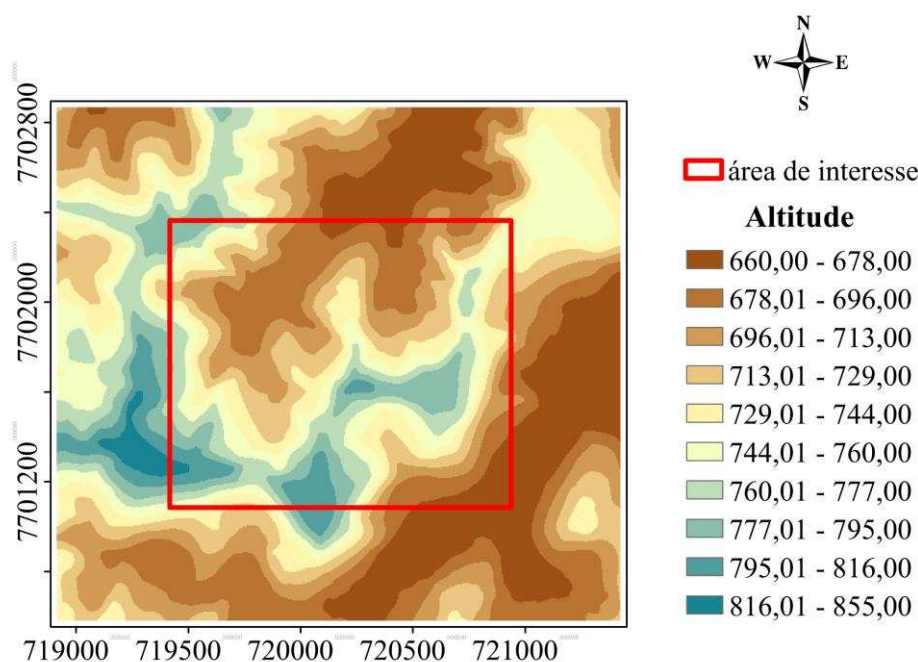


Figura 3 – Mapa da altitude da região de estudo.

A fim de estudar o efeito de borda, primeiramente selecionou-se uma área no centro da região de estudo. Essa área denominada de área de interesse (apresentada na Figura 3) contém uma extensão bem inferior a região de estudo, ocupando aproximadamente 35% de sua área. Externamente a área de interesse, foram criados quatro *buffers* com largura de 120 metros, 240 metros, 360 metros e 480 metros em relação ao limite lateral da área de interesse. Esses

*buffers* foram denominados *buffer 1*, *buffer 2*, *buffer 3* e *buffer 4*, respectivamente, conforme Figura 4.

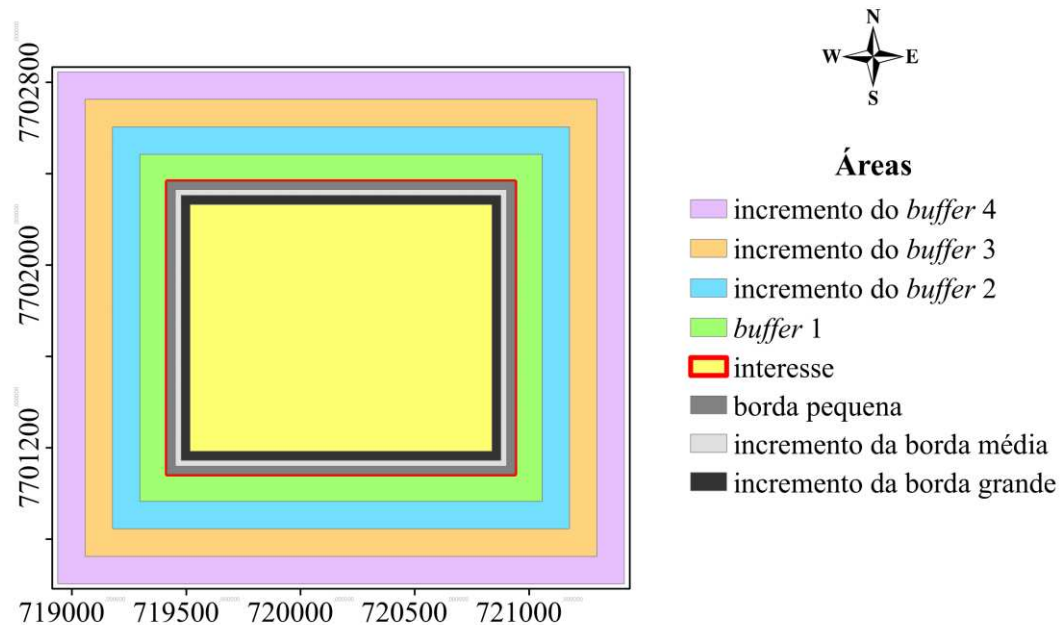


Figura 4: Representação da área de interesse do estudo, dos quatro *buffers* externos à área de interesse e das três bordas dentro da área de interesse.

Além dos *buffers*, três bordas foram criadas a partir da área de interesse. Uma borda foi definida como a área localizada nas extremidades da área de interesse, ou seja, como um *buffer* interno. A partir do limite da área de interesse foram criados três *buffers* internos (bordas) com larguras de 35 metros, 60 metros e 100 metros. Essas bordas foram denominadas borda pequena, borda média e borda grande, respectivamente, conforme Figura 4.

Após a definição das bordas, uma seleção regular e uma irregular dos pontos de altitude foram feitas. Para cada seleção, realizou-se a krigagem simples em cinco áreas: na área de interesse e na área de interesse acrescida pela área de cada um dos quatro *buffers*. Por meio dos dez mapas do erro de predição obteve-se a média e o desvio-padrão da variância de krigagem para cada borda. Ao comparar essas medidas para os diferentes tamanhos de borda, concluiu-se sobre a minimização do efeito de borda ao utilizar pontos amostrais fora de uma área de interesse para a qual se deseja predizer os valores de uma variável regionalizada.

## 4. RESULTADOS E DISCUSSÃO

Para uma melhor compreensão, os resultados obtidos foram devidamente separados entre aqueles referentes à seleção regular e à seleção irregular. Após a apresentação desses, uma comparação entre os resultados referentes à seleção regular e aqueles referentes à seleção irregular foi realizada.

### 4.1. Resultados da seleção regular

Os 3.591 pontos de altitude apresentados na Figura 5 foram selecionados a partir de uma seleção regular de 40 x 40 metros na região de estudo. Aproximadamente 36% desses, ou seja, 1.287 pontos estão dentro da área de interesse do estudo. Na Figura 5, também se pode visualizar os *buffers* externos à área de interesse que foram criados.

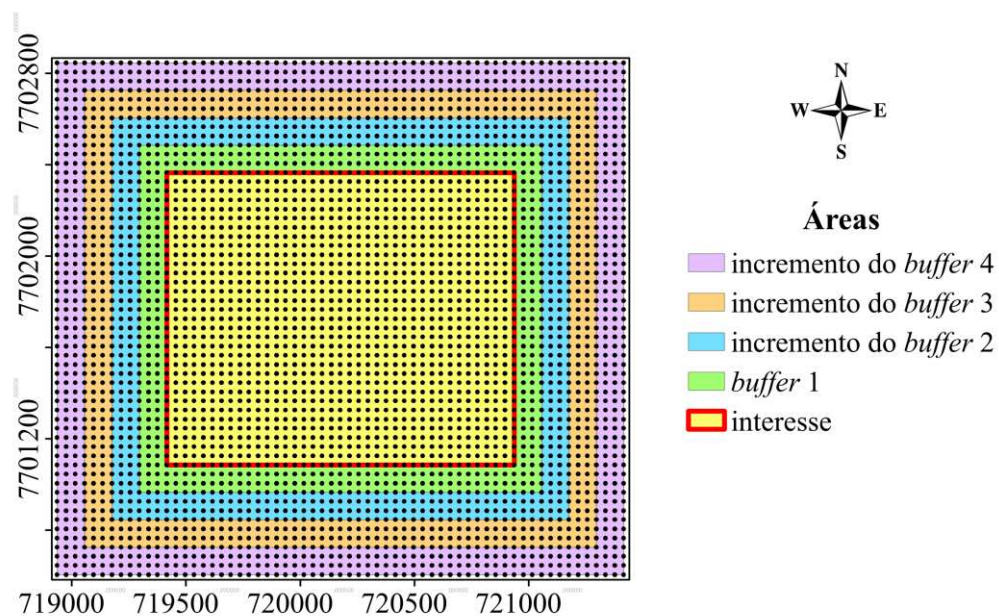


Figura 5: Representação da área de interesse do estudo, dos quatro *buffers* externos à área de interesse e os pontos amostrais em formato de malha regular quadrática.

Apresenta-se na Figura 6 apenas a área de interesse do estudo, na qual as três bordas (*buffers* internos) podem ser visualizadas. A área da borda pequena, da borda média e da borda grande abrangeram respectivamente 9,82%; 16,53% e 26,73% da área de interesse do estudo. Cerca de 10,88% dos pontos na área de interesse, isto é 140 pontos, estão contidos na borda pequena. Os 272 pontos contidos na borda média representam aproximadamente

21,13% do total de pontos pertencentes a área de interesse. Já os 396 pontos contidos na borda grande representam cerca de 30,77% dos pontos contidos na área de interesse.

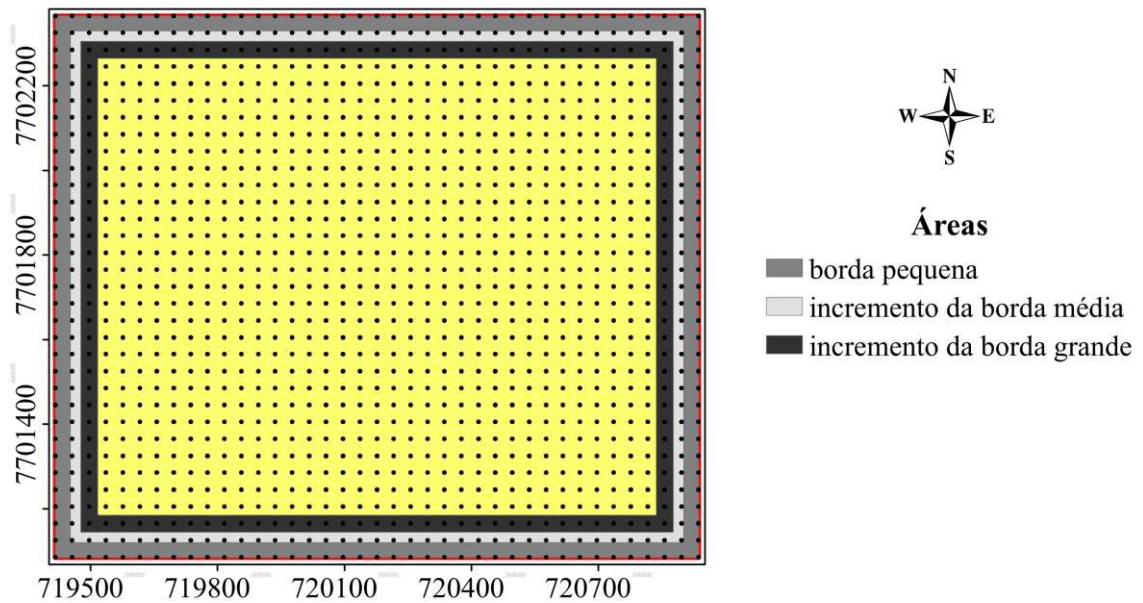


Figura 6: Representação das três bordas dentro da área de interesse e os pontos amostrais em formato de malha regular quadrática contidos na área de interesse do estudo.

Pela Figura 7, revela-se uma tendência de primeira ordem no plano XZ e uma tendência de segunda ordem no plano YZ nos dados de altimetria para a malha regular quadrática. De acordo com Vieira et al. (2000), a krigagem universal é o método de interpolação a ser utilizado quando os dados apresentam tendência. No entanto, ao comparar a krigagem simples com a krigagem universal, Santos et al. (2011) mostrou que quando se usa a krigagem simples não há necessidade de modelar a tendência separadamente.

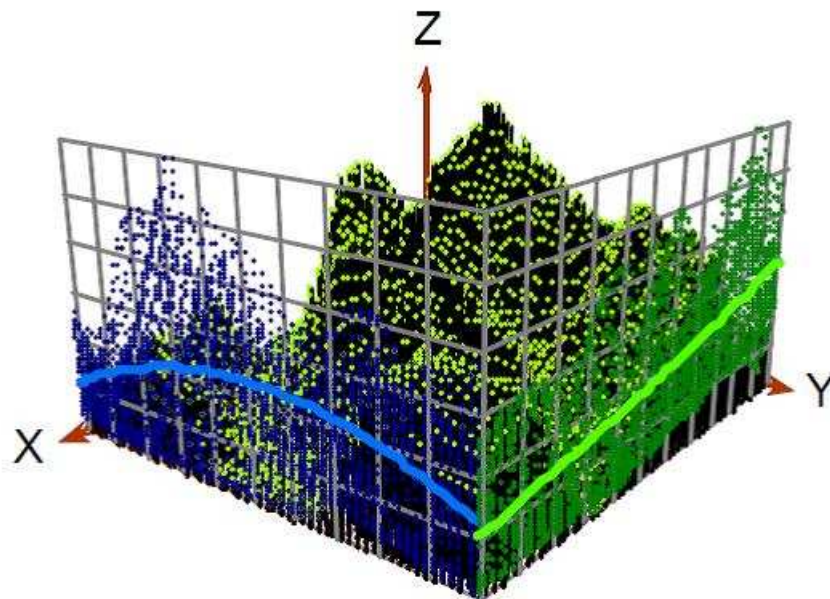


Figura 7: Gráfico de tendência da altitude para a malha regular quadrática.

Apresentam-se na Tabela 1 as principais estatísticas descritivas para as cinco áreas consideradas no estudo. Essas medidas estatísticas são: a quantidade de pontos amostrados em cada área, a área em metros quadrados de cada área, a média da variável altitude para os pontos amostrais dentro de cada área e o desvio-padrão da variável altitude para os pontos amostrais dentro de cada área. A diferença entre a quantidade de pontos amostrada entre cada área e a área anterior está sendo apresentada entre parênteses.

Tabela 1: Estatísticas descritivas da área de interesse do estudo e da área de interesse acrescida pela área de cada um dos quatro buffers para a seleção regular.

Medidas	Área				
	interesse	interesse + <i>buffer</i> de 120 m	interesse + <i>buffer</i> de 240 m	interesse + <i>buffer</i> de 360 m	interesse + <i>buffer</i> de 480 m
Pontos amostrados	1.287	1.755 (468)	2.295 (540)	2.907 (612)	3.591 (684)
Área (m <sup>2</sup> )	1.945.600	2.675.200	3.520.000	4.480.000	5.555.200
Média (m)	731,56	729,86	726,85	723,68	721,35
Desvio-padrão (m)	36,06	39,57	40,70	39,65	38,39

A Tabela 2 retorna as principais estatísticas dos modelos de semivariograma para as cinco áreas consideradas no estudo. As medidas estatísticas são: existência de anisotropia nos dados

de cada área; o modelo teórico de semivariograma utilizado considerando os pontos amostrais de cada área; o efeito pepita, a contribuição e o alcance do semivariograma utilizado considerando os pontos amostrais de cada área.

Tabela 2: Resumo geoestatístico dos modelos de semivariograma para a área de interesse do estudo e da área de interesse acrescida pela área de cada um dos quatro buffers para a seleção regular.

Medidas	Área				
	interesse	interesse + <i>buffer</i> de 120 m	interesse + <i>buffer</i> de 240 m	interesse + <i>buffer</i> de 360 m	interesse + <i>buffer</i> de 480 m
Anisotropia	Não	Não	Não	Não	Não
Modelo	Gaussiano	Gaussiano	Gaussiano	Gaussiano	Gaussiano
Efeito pepita (m <sup>2</sup> )	70,83	75,20	63,28	49,11	45,42
Contribuição (m <sup>2</sup> )	1251,12	1525,44	1598,32	1556,62	1495,50
Alcance (m)	512,28	580,84	600,55	596,11	609,84

Pela Tabela 1, percebe-se que ao aumentar a área e a quantidade de pontos considerados, houve diminuição da média da variável altitude. Pelos resultados da Tabela 2, pode-se visualizar que o alcance sofre um aumento a medida que considera-se para o modelo de semivariograma mais pontos externos à área de interesse, com exceção da área que abrange a área de interesse mais a área do *buffer* de 360 metros de distância da área de interesse. Além disso, é possível perceber da Tabela 2 que o modelo teórico de semivariograma utilizado foi o gaussiano e que todos os modelos de semivariograma foram considerados isotrópicos.

#### 4.1.1 Cálculo da variância de krigagem

Uma vez calculado os parâmetros do modelo de semivariograma para a área de interesse, apresentados na primeira coluna da Tabela 2, foi possível predizer os valores de altitude para essa área por meio da krigagem simples. Logo, apresenta-se na Figura 8 o mapa de krigagem de altitude da área de interesse dividido em dez classes com comprimento irregular.

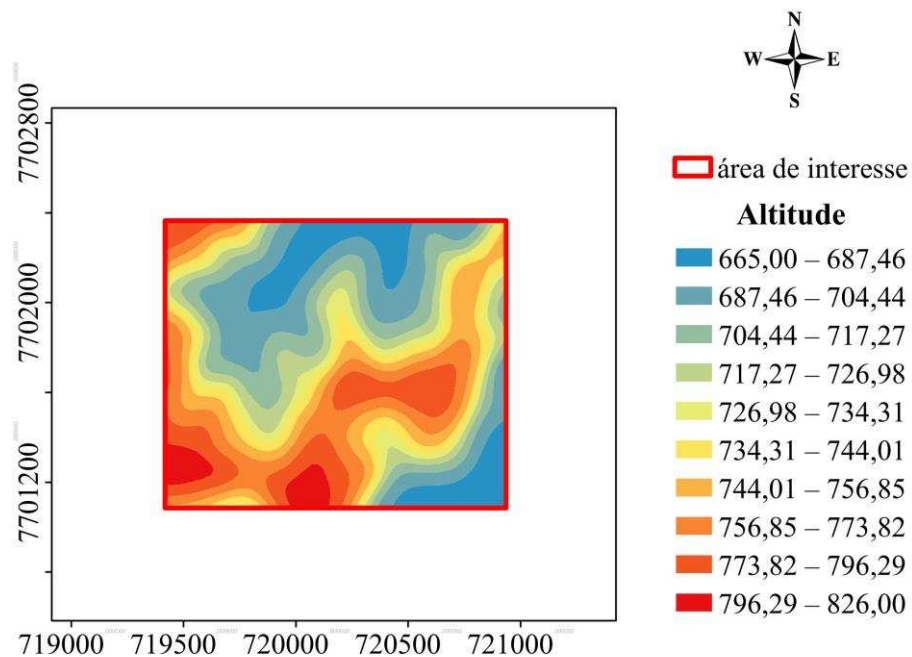


Figura 8: Mapa de krigagem da altitude da área de interesse do estudo, utilizando na interpolação os pontos amostrais em formato de malha regular quadrática contidos nessa área.

Na realização da krigagem é obtido não apenas o gráfico de predição, mas também o gráfico do erro de predição. Elevando ao quadrado o valor de cada célula do gráfico do erro de predição foi obtido o gráfico (a) da Figura 9 que apresenta a variância de krigagem para a área de interesse. Já os gráficos (b), (c) e (d) foram obtidos por meio de recorte do gráfico (a), de modo que fosse possível avaliar a variância de krigagem nas áreas das três bordas (*buffers* internos).

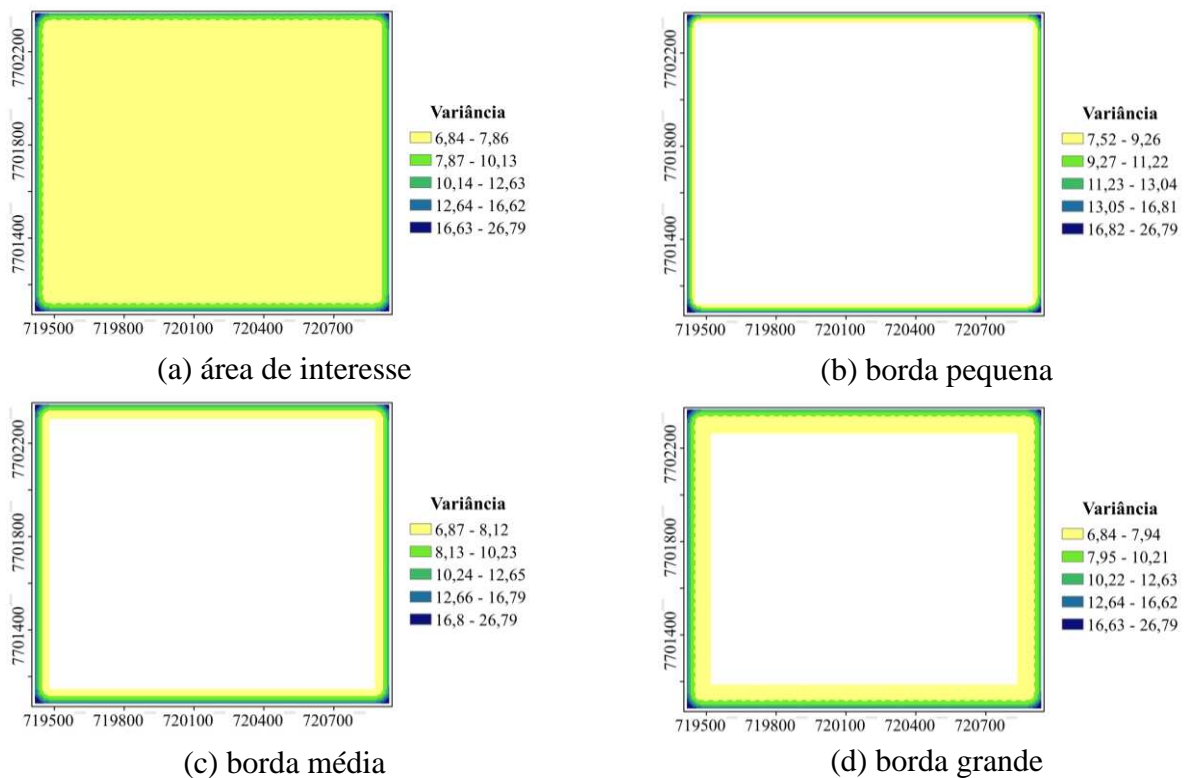


Figura 9: Mapa da variância de krigagem da área de interesse do estudo (a), da borda pequena (b), da borda média (c) e da borda grande (d), utilizando na interpolação os pontos amostrais em formato de malha regular quadrática contidos na área de interesse do estudo.

De maneira análoga, foram feitos os mapas de krigagem e da variância de krigagem para as outras áreas consideradas no estudo, referentes tanto à seleção regular quanto à seleção irregular. Esses mapas podem ser visualizados no Anexo.

#### 4.1.2 Resultados da variância de krigagem

No gráfico (a) da Figura 9, pode-se observar que a variância de krigagem aumenta a medida que aproxima-se do limite da área de interesse, ou seja, ela é bem maior nas extremidades da área. Logo, quanto mais perto da região central da área de interesse menor é a variância de krigagem. Isso pode ser melhor visualizado nos gráficos (b),(c), e (d) da Figura 9, uma vez que o menor valor de variância de krigagem na borda pequena (7,52) é maior que o menor valor de variância de krigagem na borda média (6,87), sendo esse maior que o menor valor de variância de krigagem na borda grande (6,84).

Apresenta-se nas Tabelas 3, 4, 5 e 6 um resumo estatístico descritivo dos principais resultados para a variância de krigagem na borda pequena, na borda média, na borda grande e na área de interesse, respectivamente. Nessas Tabelas é apresentado a porcentagem do valor da média da variância de krigagem, ao utilizar os pontos pertencentes a área de interesse

acrescida da área de cada um dos *buffers*, sobre o valor da média da variância de krigagem ao utilizar, na interpolação, apenas os pontos da área de interesse. Além disso, nas Tabelas 3, 4, 5 e 6 é apresentado a porcentagem da redução da média da variância de krigagem, ao utilizar os pontos pertencentes a área de interesse acrescida da área de cada um dos *buffers*, ao invés de utilizar, na interpolação, apenas os pontos da área de interesse.

Tabela 3: Resumo estatístico descritivo da variância de krigagem na borda pequena para seleção regular.

Área	Distância	Média			Desvio-padrão		
		Valor	%	Redução	Valor	%	Redução
interesse	0 m	10,01	100%	0%	2,12	100%	0%
interesse + <i>buffer</i> 1	120 m	6,51	65,03%	34,97%	$5,05 \times 10^{-2}$	2,38%	97,62%
interesse + <i>buffer</i> 2	240 m	5,60	55,94%	44,06%	$3,93 \times 10^{-2}$	1,85%	98,15%
interesse + <i>buffer</i> 3	360 m	4,69	46,85%	53,15%	$2,07 \times 10^{-2}$	0,98%	99,02%
interesse + <i>buffer</i> 4	480 m	4,27	42,66%	57,34%	$2,12 \times 10^{-2}$	1,00%	99,00%

Tabela 4: Resumo estatístico descritivo da variância de krigagem na borda média para seleção regular.

Área	Distância	Média			Desvio-padrão		
		Valor	%	Redução	Valor	%	Redução
interesse	0 m	8,84	100%	0%	2,17	100%	0%
interesse + <i>buffer</i> 1	120 m	6,51	73,64%	26,36%	$4,89 \times 10^{-2}$	2,25%	97,75%
interesse + <i>buffer</i> 2	240 m	5,60	63,35%	36,85%	$3,93 \times 10^{-2}$	1,81%	98,19%
interesse + <i>buffer</i> 3	360 m	4,69	53,05%	46,95%	$1,98 \times 10^{-2}$	0,91%	99,09%
interesse + <i>buffer</i> 4	480 m	4,26	48,19%	51,81%	$2,10 \times 10^{-2}$	0,97%	99,03%

Tabela 5: Resumo estatístico descritivo da variância de krigagem na borda grande para seleção regular.

Área	Distância	Média			Desvio-padrão		
		Valor	%	Redução	Valor	%	Redução
interesse	0 m	8,08	100%	0%	1,95	100%	0%
interesse +buffer 1	120 m	6,51	80,57%	19,43%	$4,69 \times 10^{-2}$	2,41%	97,59%
interesse +buffer 2	240 m	5,60	69,31%	30,69%	$3,75 \times 10^{-2}$	1,92%	98,08%
interesse +buffer 3	360 m	4,69	58,04%	41,96%	$1,98 \times 10^{-2}$	1,02%	98,98%
interesse +buffer 4	480 m	4,26	52,72%	47,28%	$1,93 \times 10^{-2}$	0,99%	99,01%

Tabela 6: Resumo estatístico descritivo da variância de krigagem na área de interesse para seleção regular.

Área	Distância	Média			Desvio-padrão		
		Valor	%	Redução	Valor	%	Redução
interesse	0 m	7,20	100%	0%	1,16	100%	0%
interesse +buffer 1	120 m	6,51	90,42%	9,58%	$4,84 \times 10^{-2}$	4,17%	95,83%
interesse +buffer 2	240 m	5,60	77,78%	22,22%	$3,73 \times 10^{-2}$	3,22%	96,78%
interesse +buffer 3	360 m	4,69	65,14%	34,86%	$1,98 \times 10^{-2}$	1,71%	98,29%
interesse +buffer 4	480 m	4,26	59,17%	40,83%	$2,01 \times 10^{-2}$	1,73%	98,27%

Pela Tabela 3, percebe-se que ao utilizar apenas os pontos dentro da área de interesse, na interpolação, a média da variância de krigagem foi de 10,01. Já ao considerar também os pontos pertencentes ao *buffer* de 480 metros de distância da área de interesse, a média da variância de krigagem caiu para 4,27. Logo houve um ganho de 57,34% na precisão das predições na borda pequena da área de interesse. O desvio-padrão da variância de krigagem também sofre uma grande queda, passando de 1,68 para apenas  $2,12 \times 10^{-2}$ .

Na Tabela 4, pode-se observar que a utilização dos pontos dentro dos *buffers*, na interpolação, ao invés de utilizar apenas os pontos dentro da área de interesse acarreta uma redução da média da variância de krigagem observada na borda média da área de interesse. Pois ao utilizar os pontos pertencentes ao *buffer* de 120 metros, de 240 metros, de 360 metros e de 480 metros de distância da área de interesse houve um ganho na precisão das predições de 26,36%, 36,85%, 46,95% e 51,81%, respectivamente. Ganhos na precisão similares a esses também ocorreram nas bordas pequena e grande da área de interesse, como pode ser verificado nas Tabelas 3 e 5.

É possível compreender das Tabelas 3 e 5 que a média da variância de krigagem na borda grande da área de interesse (8,08) é menor do que na borda pequena é (10,01). Uma vez que a borda grande contém uma parte maior da região central da área de interesse do que a borda pequena, reafirma-se o que pôde ser visualizado na Figura 9, quanto mais próximo ao centro da área de interesse menor a incerteza quanto as predições da altitude.

Pela Tabela 6, visualiza-se que a média e o desvio padrão da variância de krigagem em toda a área de interesse decresce ao utilizar, na interpolação, os pontos amostrais dentro dos *buffers*. Isso indica que as predições em toda a área de interesse, e não apenas nas bordas, estão sendo beneficiadas pela utilização de pontos exteriores a área de interesse na interpolação. Como por exemplo, houve uma redução de 40,50% da média da variância de krigagem e uma redução de 69,16% do desvio padrão da variância de krigagem ao utilizar os pontos dentro do *buffer* de 480 metros de distância em vez de apenas os pontos dentro da área de interesse.

De modo geral, o desvio padrão da variância de krigagem nas bordas está sendo reduzido a medida que mais pontos fora da área de interesse vão sendo utilizados na interpolação (Tabelas 3, 4 e 5). Entretanto, para a borda pequena, o desvio padrão da variância de krigagem que foi de  $2,07 \times 10^{-2}$  ao utilizar os pontos dentro do *buffer* de 360 metros de distância da área de interesse aumentou para  $2,12 \times 10^{-2}$  ao utilizar os pontos pertencentes ao *buffer* de 480 metros de distância (Tabela 4). De acordo com a Tabela 4, esse aumento no desvio-padrão também ocorreu para a borda média.

Pelas Tabelas 3, 4 e 5, percebe-se que os valores de média da variância de krigagem ao utilizar os pontos dentro do *buffer* de 360 metros de distância da área de interesse e ao utilizar os pontos pertencentes ao *buffer* de 480 metros de distância são muito similares. Esse fato acrescido do pequeno aumento percebido para o desvio-padrão da variância de krigagem nos casos citados acima, leva a crer que a partir de certa distância e certa quantidade de pontos foras da área de interesse, a melhora das predições ao utilizar pontos fora da área de interesse na interpolação se torna constante.

#### **4.2. Resultados da seleção irregular**

Os 2.232 pontos de altitude apresentados na Figura 10 resultam de uma seleção irregular de 1% dos pontos contidos inicialmente na região de estudo. Cerca de 34,23% dos pontos apresentados nessa figura, isto é 764 pontos, estão dentro da área de interesse. Na

Figura 10, também se pode visualizar os *buffers* que foram criados, que são os mesmos utilizados na análise da seleção regular.

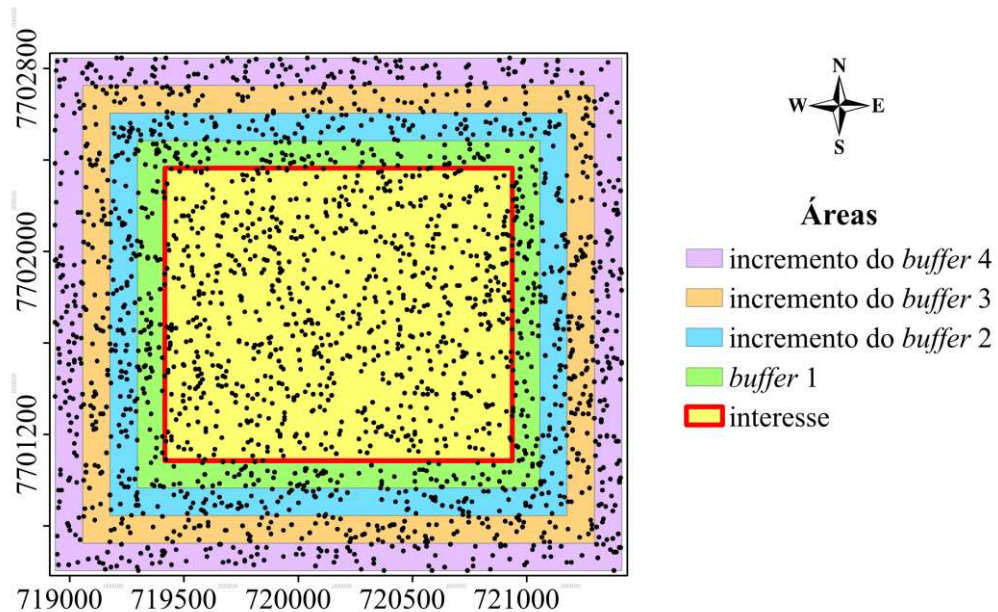


Figura 10: Representação da área de interesse do estudo, dos quatro *buffers* externos a área de interesse e os pontos amostrais em formato de malha irregular.

Apresenta-se na Figura 11 apenas a área de interesse do estudo, na qual as três bordas (*buffers* internos) dessa área podem ser visualizados. Cerca de 12,17%, isto é 93 pontos, dos pontos na área de interesse estão contidos na borda pequena. Os 143 pontos contidos na borda média representam aproximadamente 18,71% do total de pontos pertencentes a área de interesse. Já os 210 pontos contidos na borda grande representam cerca de 27,49% dos pontos contidos na área de interesse.

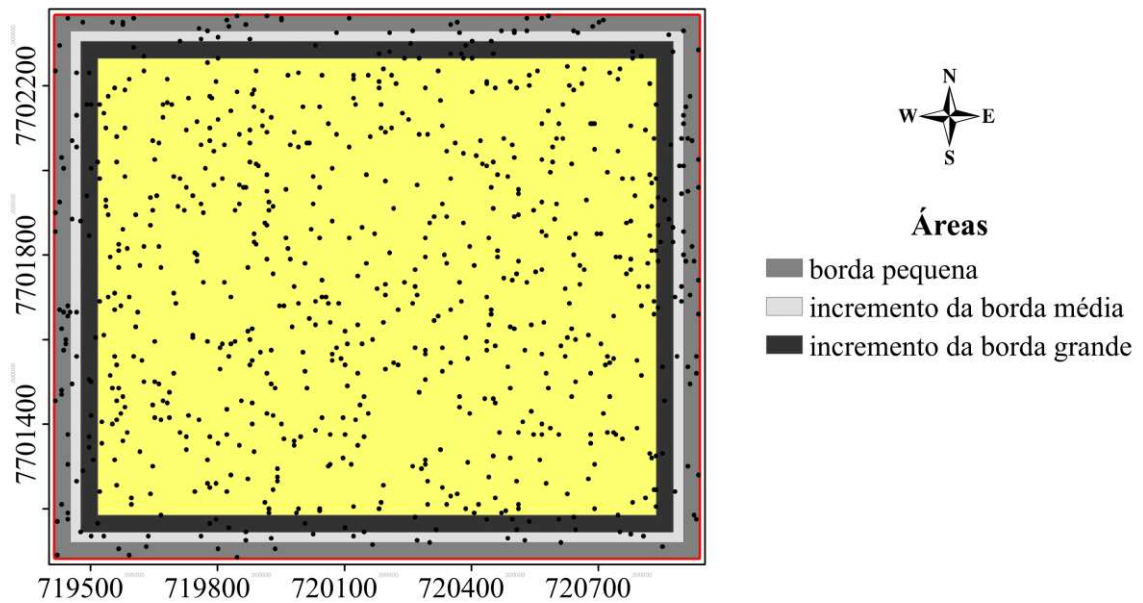


Figura 11: Representação das três bordas dentro da área de interesse e os pontos amostrais em formato de malha irregular contido na área de interesse do estudo.

Pela Figura 12, revela-se uma tendência de primeira ordem no plano XZ e uma tendência de segunda ordem no plano YZ nos dados de altimetria para a malha irregular. De acordo com Vieira et al. (2000), a krigagem universal é o método de interpolação a ser utilizado quando os dados apresentam tendência. No entanto, ao comparar a krigagem simples com a krigagem universal, Santos et al. (2011) mostrou que quando se usa a krigagem simples não há necessidade de modelar a tendência separadamente.

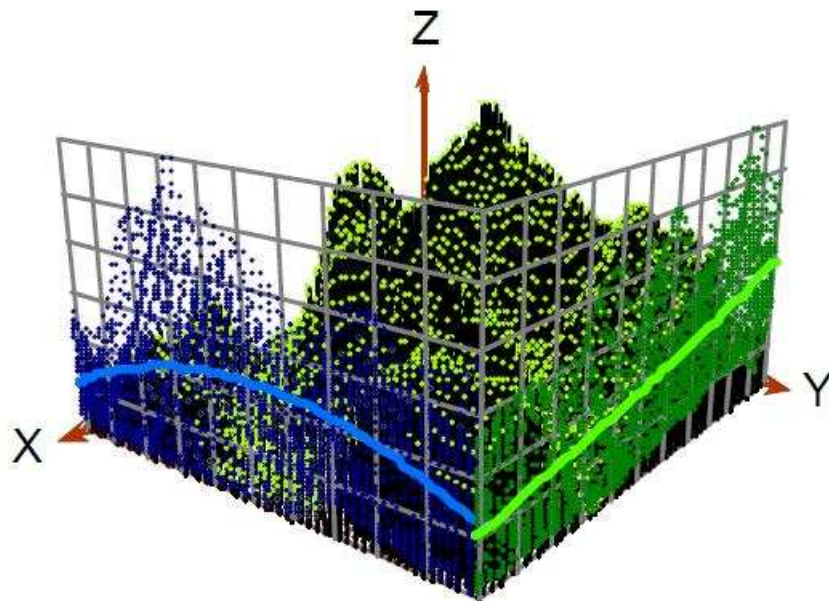


Figura 12: Gráfico de tendência da altitude para a malha irregular.

Apresentam-se na Tabela 7 as principais estatísticas descritivas para as cinco áreas consideradas no estudo para a seleção irregular. Essas medidas estatísticas são: a quantidade de pontos amostrados em cada área, a área em metros quadrados de cada área, a média da variável altitude para os pontos amostrais dentro de cada área e o desvio-padrão da variável altitude para os pontos amostrais dentro de cada área. A diferença dessas medidas entre uma área e a área anterior está sendo apresentada entre parênteses.

Tabela 7: Estatísticas descritivas da área de interesse do estudo e da área de interesse acrescida pela área de cada um dos quatro buffers para a seleção irregular.

Medidas	Área				
	interesse	interesse + <i>buffer</i> de 120 m	interesse + <i>buffer</i> de 240 m	interesse + <i>buffer</i> de 360 m	interesse + <i>buffer</i> de 480 m
Pontos amostrados	764	1.086 (322)	1.445 (359)	1.806 (361)	2.232 (426)
Área (m <sup>2</sup> )	1.945.600	2.675.200	3.520.000	4.480.000	5.555.200
Média (m)	732,04	730,73	728,04	724,16	721,77
Desvio-padrão (m)	34,91	39,14	41,05	40,73	39,56

A Tabela 8 retorna as principais estatísticas dos modelos de semivariograma para as cinco áreas consideradas no estudo para a seleção irregular. As medidas estatísticas são: existência de anisotropia nos dados de cada área; o modelo teórico de semivariograma

utilizado considerando os pontos amostrais de cada área; o efeito pepita, a contribuição e o alcance do semivariograma utilizado considerando os pontos amostrais de cada área.

Tabela 8: Resumo geoestatístico dos modelos de semivariograma para a área de interesse do estudo e da área de interesse acrescida pela área de cada um dos quatro *buffers* para a seleção irregular.

Medidas	Área				
	interesse	interesse + <i>buffer</i> de 120 m	interesse + <i>buffer</i> de 240 m	interesse + <i>buffer</i> de 360 m	interesse + <i>buffer</i> de 480 m
Anisotropia	Não	Não	Não	Não	Não
Modelo	Gaussiano	Gaussiano	Gaussiano	Gaussiano	Gaussiano
Efeito pepita (m <sup>2</sup> )	84,80	61,71	51,21	39,87	33,07
Contribuição (m <sup>2</sup> )	1216,73	1516,53	1618,69	1608,12	1540,4
Alcance (m)	551,10	589,98	608,72	608,84	613,43

Pela Tabela 7, pode-se visualizar que ao aumentar a área e a quantidade de pontos considerados, a média da variável altitude diminui. Pelos resultados da Tabela 8, percebe-se que o alcance sofre um aumento a medida que considera-se para o modelo de semivariograma mais pontos externos à área de interesse, enquanto o efeito pepita sofre uma redução. Além disso, é possível perceber da Tabela 8 que o modelo teórico de semivariograma utilizado foi o gaussiano e que todos os modelos de semivariograma foram considerados isotrópicos.

#### 4.2.1 Resultados da variância de krigagem

No gráfico (a) da Figura 13, pode-se observar que os valores da variância de krigagem distribuem-se de maneira irregular, de modo que existe valores pertencentes a segunda (11,00 – 17,01) e a terceira classe (17,02 – 27,53) de valores tanto na região central, quanto nas extremidades da área de interesse. Como era esperado, pode-se visualizar o efeito de borda nos gráficos (b),(c), e (d) da Figura 13, pois os maiores valores da variância de krigagem encontram-se nas bordas da área de interesse.

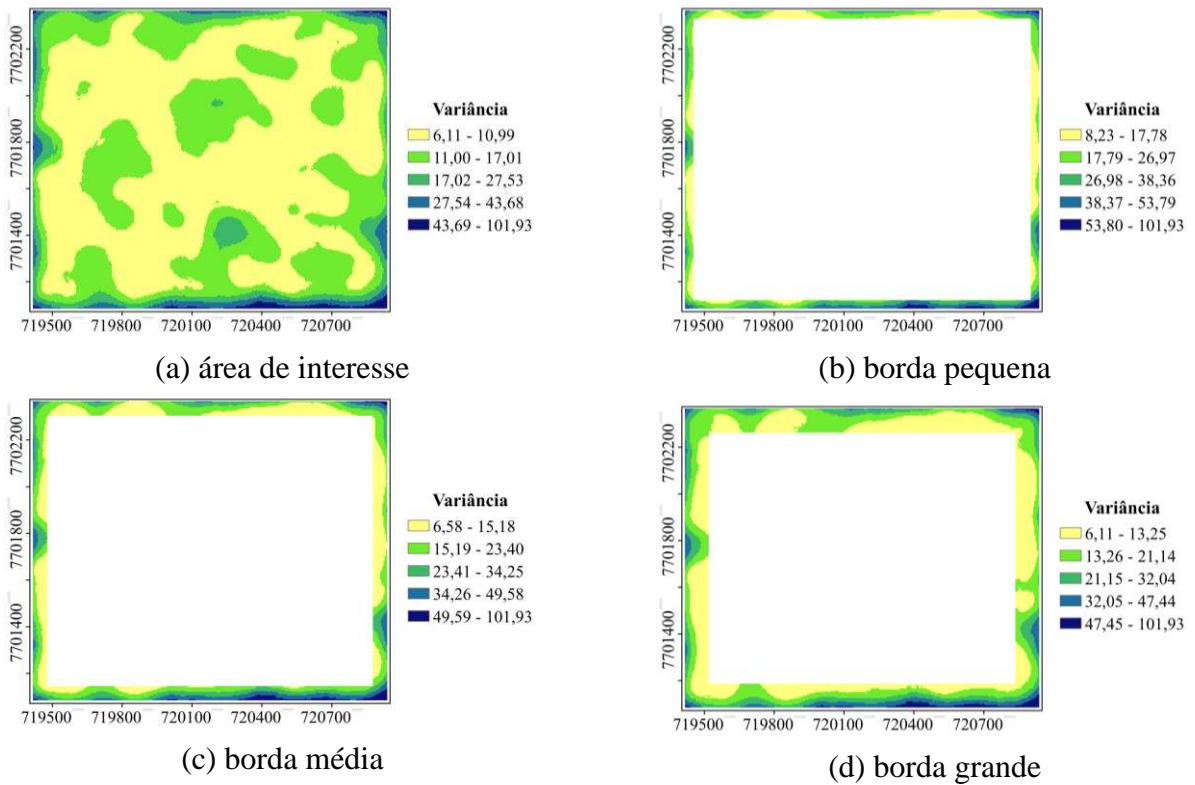


Figura 13: Mapa da variância de krigagem da área de interesse do estudo (a), da borda pequena (b), da borda média (c) e da borda grande (d), utilizando na interpolação os pontos amostrais em formato de malha irregular contidos na área de interesse do estudo.

Apresenta-se nas Tabelas 9, 10, 11 e 12 um resumo estatístico descritivo dos principais resultados para a variância de krigagem na borda pequena, na borda média, na borda grande e na área de interesse, respectivamente. Nessas Tabelas é apresentado a porcentagem do valor da média da variância de krigagem, ao utilizar os pontos pertencentes a área de interesse acrescida da área de cada um dos *buffers*, sobre o valor da média da variância de krigagem ao utilizar, na interpolação, apenas os pontos da área de interesse.

Além disso, nas Tabelas 9, 10, 11 e 12 é apresentado a porcentagem da redução da média da variância de krigagem, ao utilizar os pontos pertencentes a área de interesse acrescida da área de cada um dos *buffers*, ao invés de utilizar, na interpolação, apenas os pontos da área de interesse.

Tabela 9: Resumo estatístico descritivo da variância de krigagem na borda pequena para seleção irregular.

Área	Distância	Média			Desvio-padrão		
		Valor	%	Redução	Valor	%	Redução
interesse	0 m	24,83	100%	0%	12,00	100%	0%
interesse +buffer 1	120 m	8,34	33,59%	66,41%	2,04	17,00%	83,00%
interesse +buffer 2	240 m	7,18	28,92%	71,08%	1,81	15,08%	84,92%
interesse +buffer 3	360 m	6,07	24,45%	75,55%	1,64	13,67%	86,33%
interesse +buffer 4	480 m	5,20	20,94%	79,06%	1,45	12,08%	87,92%

Tabela 10: Resumo estatístico descritivo da variância de krigagem na borda média para seleção irregular.

Área	Distância	Média			Desvio-padrão		
		Valor	%	Redução	Valor	%	Redução
interesse	0 m	20,74	100%	0%	11,02	100%	0%
interesse +buffer 1	120 m	8,39	40,45%	59,55%	1,99	18,06%	81,94%
interesse +buffer 2	240 m	7,23	34,86%	65,14%	1,75	15,88%	84,12%
interesse +buffer 3	360 m	6,11	29,46%	70,54%	1,59	14,43%	85,57%
interesse +buffer 4	480 m	5,24	25,27%	74,73%	1,41	12,79%	87,21%

Tabela 11: Resumo estatístico descritivo da variância de krigagem na borda grande para seleção irregular.

Área	Distância	Média			Desvio-padrão		
		Valor	%	Redução	Valor	%	Redução
interesse	0 m	17,15	100%	0%	9,97	100%	0%
interesse +buffer 1	120 m	8,37	48,80%	51,20%	1,94	19,46%	80,54%
interesse +buffer 2	240 m	7,19	41,92%	58,08%	1,71	17,15%	82,85%
interesse +buffer 3	360 m	6,06	35,34%	64,66%	1,54	15,45%	84,55%
interesse +buffer 4	480 m	5,21	30,38%	69,92%	1,37	13,74%	86,26%

Tabela 12: Resumo estatístico descritivo da variância de krigagem na área de interesse para seleção irregular.

Área	Distância	Média			Desvio-padrão		
		Valor	%	Redução	Valor	%	Redução
interesse	0 m	12,28	100%	0%	6,28	100%	0%
interesse +buffer 1	120 m	8,25	67,18%	32,82%	1,82	28,98%	71,02%
interesse +buffer 2	240 m	7,10	57,82%	42,18%	1,60	25,48%	74,52%
interesse +buffer 3	360 m	5,96	48,53%	51,47%	1,41	22,45%	77,55%
interesse +buffer 4	480 m	5,10	41,53%	58,47%	1,24	21,82%	78,18%

Pela Tabela 9, percebe-se que ao utilizar apenas os pontos dentro da área de interesse, na interpolação, a média da variância de krigagem foi de 24,83, já ao considerar também os pontos pertencentes ao *buffer* de 480 metros de distância da área de interesse, a média da variância de krigagem caiu para 5,20, logo houve um ganho de 79,06% na precisão das predições na borda pequena da área de interesse. O desvio-padrão da variância de krigagem também sofre uma grande queda, passando de 12,00 para apenas 1,45.

Na Tabela 10, pode-se observar que a utilização dos pontos dentro dos buffers, na interpolação, ao invés de se utilizar apenas os pontos dentro da área de interesse acarreta uma redução da média da variância de krigagem observada na borda média da área de interesse. Pois ao utilizar os pontos pertencentes ao *buffer* de 120 metros de distância da área de interesse, de 240 metros, de 360 metros e de 480 metros houve um ganho na precisão das predições de 59,55%, 65,14%, 70,54% e 74,3%, respectivamente. Ganhos na precisão similares a esses também ocorreram nas bordas pequena e grande da área de interesse, como pode ser verificado nas Tabelas 9 e 11.

Pela Tabela 12, visualiza-se que a média e o desvio padrão da variância de krigagem em toda a área de interesse decresce ao utilizar, na interpolação, os pontos amostrais dentro dos *buffers*. Isso indica que as predições em toda a área de interesse, e não apenas nas bordas, estão sendo beneficiadas pela utilização de pontos exteriores a área de interesse na interpolação. Como por exemplo, houve uma redução de 58,47% da média da variância de krigagem e uma redução de 78,18% do desvio padrão da variância de krigagem ao utilizar os pontos dentro do *buffer* de 480 metros de distância em vez de apenas os pontos dentro da área de interesse.

Pelas Tabelas 9, 10 e 11, percebe-se que os valores do desvio-padrão da variância de krigagem ao utilizar os pontos dentro do *buffer* de 360 metros de distância da área de

interesse e ao utilizar os pontos pertencentes ao *buffer* de 480 metros de distância são muito similares. Portanto, acredita-se que a partir de certa distância e certa quantidade de pontos foras da área de interesse, os ganhos na precisão das predições ao utilizar pontos fora da área de interesse na interpolação vão tornando-se muito pequenos.

### **4.3. Comparação entre os resultados da seleção regular e da irregular**

Os gráficos da variância de krigagem apresentados no Anexo mostram que ao fazer uma seleção regular pode-se observar um padrão da variância de krigagem bem definido e comportado em toda a área de interesse. Já ao fazer uma seleção irregular, não há um padrão bem definido na distribuição dos valores da variância na área de interesse, ou seja, os valores da variância de krigagem variam muito ao longo da área de interesse. Esse resultado está de acordo com Wackernagel (2003) que afirmou que quando uma amostragem irregular é realizada, a variância de krigagem é capaz de fornecer uma análise da variação da precisão das predições causada pela disposição irregular dos pontos amostrais.

Pela primeira linha das Tabelas 3, 4, 5, 9, 10 e 11, pode-se perceber que o valor da média e do desvio-padrão da variância de krigagem nas bordas ao utilizar, na interpolação, apenas os pontos dentro da área de interesse são bem maiores ao considerar os pontos da malha irregular do que ao considerar os pontos da malha regular retangular. Portanto, a incerteza nas predições mostrou-se maior nas bordas da área de interesse para a seleção irregular do que para a seleção regular, ou seja, o efeito de borda foi maior no primeiro caso.

Esse resultado está de acordo com o que foi apontado por Walvoort, Brus e Gruijter (2010), que declarou que o efeito de borda é intensificado quando uma amostragem irregular é realizada.

No entanto, ao utilizar a estratégia sugerida por Harzheim e Warnecke (2010), ou seja, ao usar não apenas os pontos dentro da área de interesse na interpolação, mas também aqueles pertencentes aos *buffers*, o ganho na precisão das predições nas bordas foi maior para a seleção irregular do que para a regular (Tabelas 3, 4, 5, 9, 10 e 11). Isso mostra que a utilização dessa estratégia é ainda mais eficiente na minimização do efeito de borda quando uma seleção irregular é feita.

## CONCLUSÃO

Os resultados mostraram que a estratégia de coletar pontos amostrais fora da área de interesse para a qual se deseja prever é muito eficaz na minimização do efeito de borda. Ao utilizar essa estratégia na interpolação por meio da krigagem simples, houve grande redução da média e do desvio-padrão da variância de krigagem nas bordas (*buffers* internos) da área de interesse. Portanto, as previsões nas bordas da área de interesse mostraram-se mais precisas.

Além disso, a utilização dessa estratégia não só melhorou as previsões na borda da área de interesse, como era esperado, mas também em toda a extensão da área de interesse. Isso foi verificado com o aumento na precisão das previsões em toda a área de interesse ao utilizar os pontos do *buffer* de 480 metros de distância da área de interesse foi de 40,83% para a seleção regular e de 58,47% para a seleção irregular.

A seleção irregular mostrou-se mais afetada pelo efeito de borda do que a seleção regular, pois o valor da média e do desvio-padrão da variância de krigagem ao utilizar os pontos da área de interesse são maiores ao considerar os pontos da malha irregular do que ao considerar os pontos da malha regular retangular. Por isso, a necessidade de amostrar fora da área de interesse para obter melhores previsões é maior quando uma seleção irregular tem que ser feita.

Os valores de média e desvio-padrão da variância de krigagem ao utilizar os pontos do *buffer* de 480 metros de distância da área de interesse foram bem próximos aos valores obtidos ao utilizar os pontos do *buffer* de 360 metros de distância da área de interesse. Esse resultado mostrou que a partir de certa distância da área de interesse, a melhoria na precisão das previsões torna-se pequena.

De acordo com a literatura especializada, o efeito de borda é intensificado quando a região de estudo apresenta um formato irregular. No entanto, nesse trabalho o efeito de borda foi avaliado para uma região de estudo de formato regular retangular, nenhuma consideração foi feita para uma região irregular. Por isso, recomenda-se para trabalhos futuros mensurar o efeito de borda, tendo como variante o aumento do número de pontos amostrais, para uma região de estudo irregular, como a área do estado de Minas Gerais.

## REFERÊNCIAS

- BUSSAB, W. O.; MORETTIN, P. A. **Estatística básica**. 6. ed. São Paulo: Saraiva, 2010. 540 p.
- CÂMARA, G.; MEDEIROS, J.S. **Geoprocessamento para projetos ambientais**. v. 1, Online Book, 1998. São José dos Campos, Brasil. INPE.
- CAMARGO, E. C. G. In: **Geoestatística: fundamentos e aplicações**. [s.n.], 1998. cap. 5. Disponível em: <[http://dpi.inpe.br/gilberto/tutoriais/gis\\_ambiente/5geoest.pdf](http://dpi.inpe.br/gilberto/tutoriais/gis_ambiente/5geoest.pdf)>.
- CRESSIE, N. **Statistics for spatial data**. New York: J. Wiley, 1993. 900 p.
- ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE – **Esri. ArcGIS Desktop: Release 10**. Version 10.4.1. Readlands, CA: 2016.
- HARZHEIM, L.; WARNECKE, U. Robustness optimization of the position of an anti-roll bar link to avoid the toggling of a rear axle stabilizer. **Structural and Multidisciplinary Optimization**, Springer, v. 42, n. 2, p. 315–323, 2010.
- ISAAKS, E. H.; SRIVASTAVA, R. M. **An introduction to applied Geostatistics**. New York: Oxford University Press, 1989. 561 p.
- LUCAS, T. P. B.; ABREU, M. L.; PARIZZI, M. G. O efeito de borda da interpolação iqd em mapas de chuva: uma comparação entre malhas de postos pluviométricos. **Revista Geonorte**, v.3, n. 9, p. 1364–1376, 2016.
- SANTOS, G. R. et al. Krigagem simples versus krigagem universal: qual o preditor mais preciso? **Energia na Agricultura**, v. 26, n. 2, p. 49–55, 2011.
- TOBLER, W. R. A computer movie simulating urban growth in the detroit region. **Economic geography**, Taylor & Francis, v. 46, n. 2, p. 234–240, 1970.
- VIEIRA, S. R. et al. Geoestatística em estudos de variabilidade espacial do solo. **Tópicos em ciência do solo**, v. 1, p. 1–54, 2000.
- WACKERNAGEL, H. **Multivariate geostatistics: an introduction with applications**. 3. ed. Berlin: Springer, 2003. 388 p.
- WALVOORT, D. J. J.; BRUS, D. J.; GRUIJTER, J. An r package for spatial coverage sampling and random sampling from compact geographical strata by k-means. **Computers & Geosciences**, Elsevier, v. 36, n. 10, p. 1261–1267, 2010.
- WEBSTER, R.; OLIVER, M. A. **Geostatistics for environmental scientists**. 2. ed. Chichester: J. Wiley, 2007. 315 p. (Statistics in Practice).
- YAMAMOTO, J. K.; LANDIM, P. M. B. **Geoestatística: conceitos e aplicações**. São Paulo: Oficina de textos. Online Book, 2015. 215 p.

## ANEXO

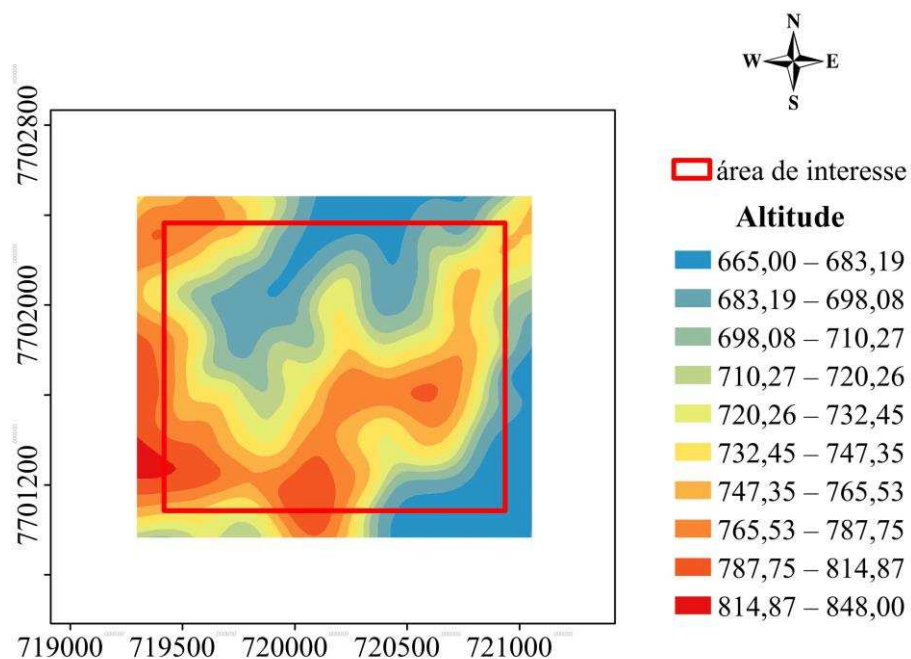


Figura 14: Mapa de krigagem da altitude da área de interesse mais área do *buffer* de 120 metros de distância da área de interesse do estudo, utilizando na interpolação os pontos amostrais em formato de malha regular quadrática contidos nessa área.

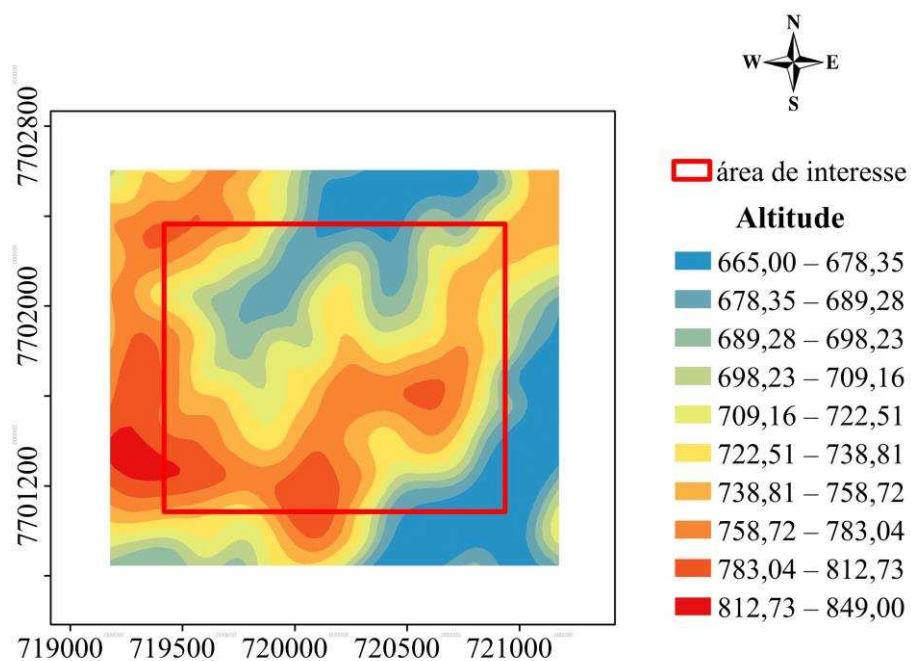


Figura 15: Mapa de krigagem da altitude da área de interesse mais área do *buffer* de 240 metros de distância da área de interesse do estudo, utilizando na interpolação os pontos amostrais em formato de malha regular quadrática contidos nessa área.

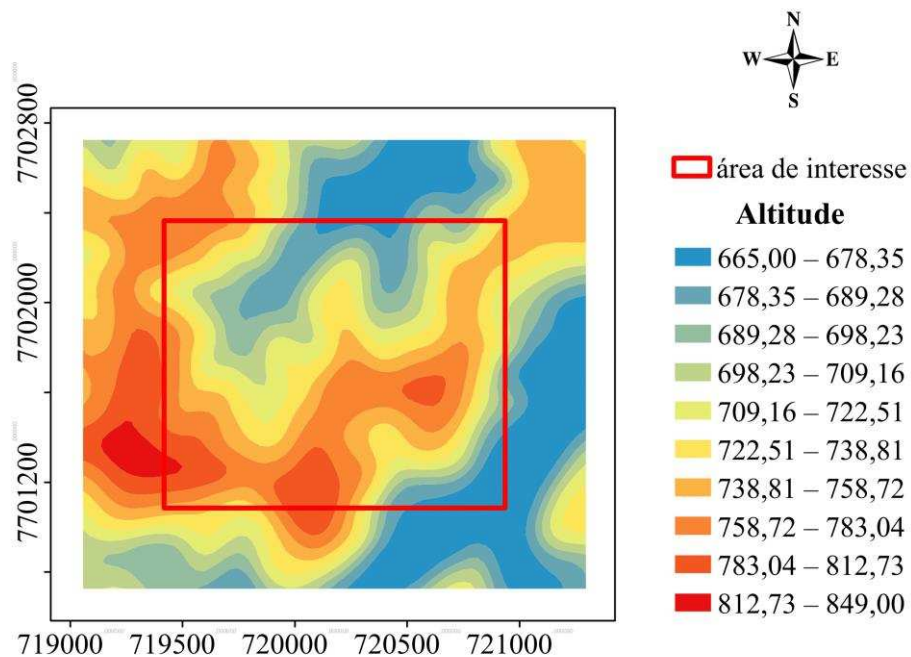


Figura 16: Mapa de krigagem da altitude da área de interesse mais área do *buffer* de 360 metros de distância da área de interesse do estudo, utilizando na interpolação os pontos amostrais em formato de malha regular quadrática contidos nessa área.

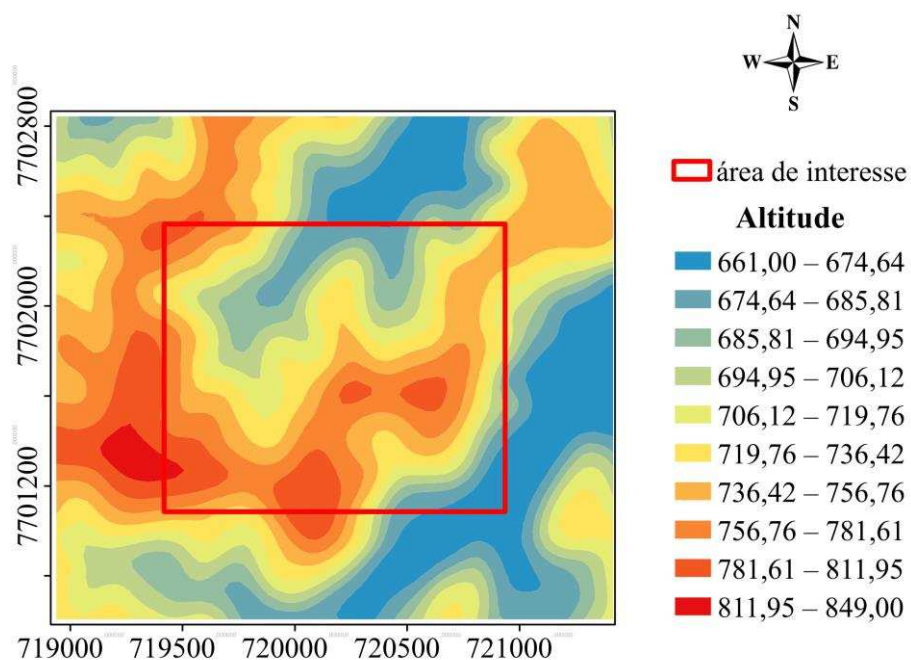


Figura 17: Mapa de krigagem da altitude da área de interesse mais área do *buffer* de 480 metros de distância da área de interesse do estudo, utilizando na interpolação os pontos amostrais em formato de malha regular quadrática contidos nessa área.

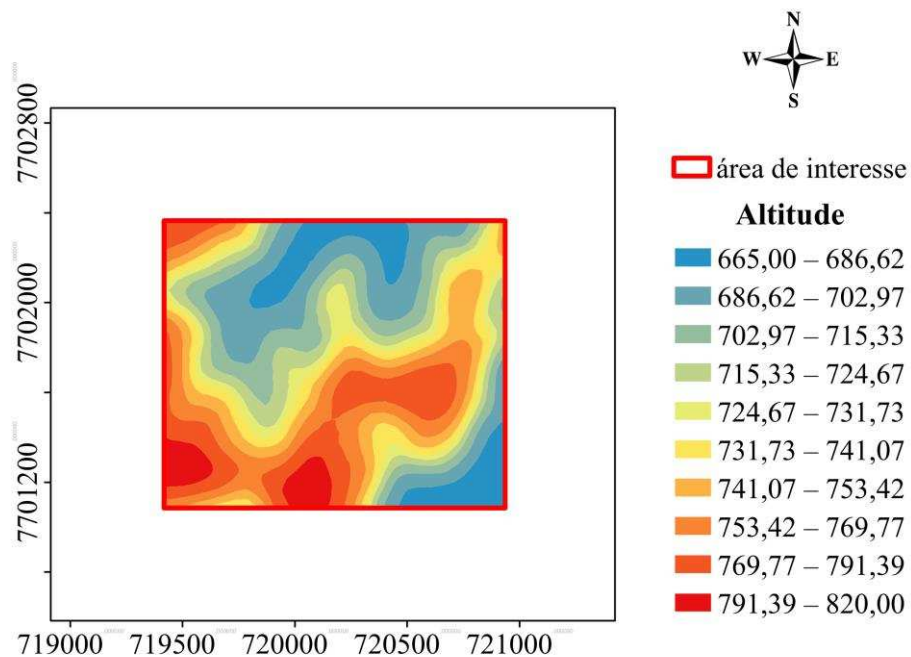


Figura 18: Mapa de krigagem da altitude da área de interesse do estudo, utilizando na interpolação os pontos amostrais em formato de malha irregular contidos nessa área.

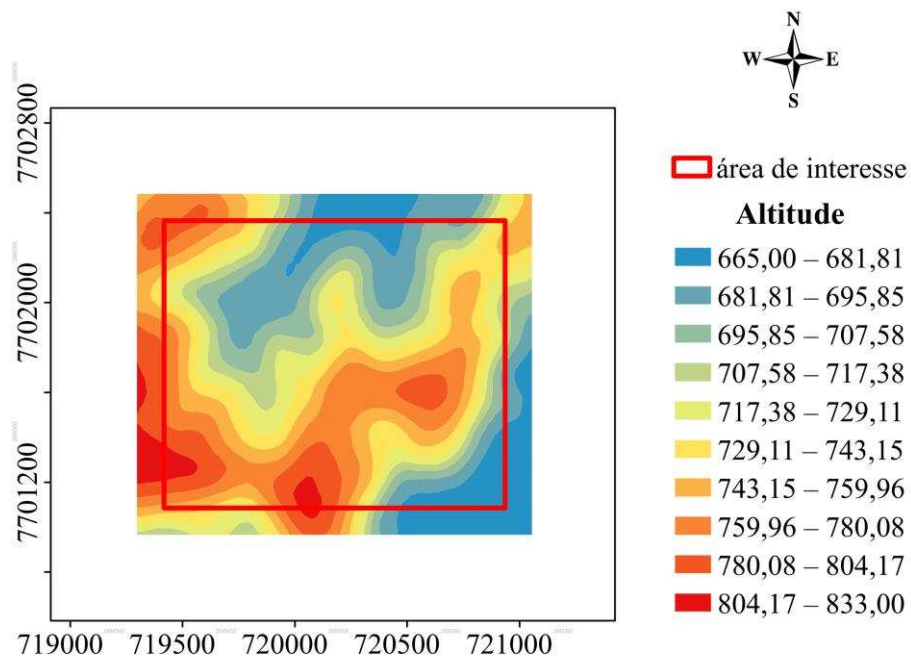


Figura 19: Mapa de krigagem da altitude da área de interesse mais área do *buffer* de 120 metros de distância da área de interesse do estudo, utilizando na interpolação os pontos amostrais em formato de malha irregular contidos nessa área.

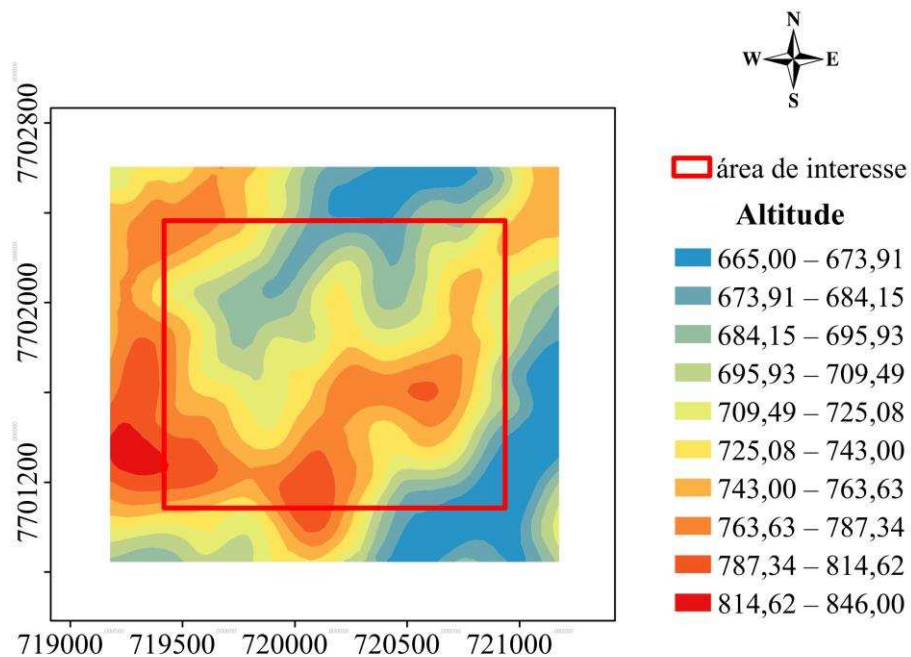


Figura 20: Mapa de krigagem da altitude da área de interesse mais área do *buffer* de 240 metros de distância da área de interesse do estudo, utilizando na interpolação os pontos amostrais em formato de malha irregular contidos nessa área.

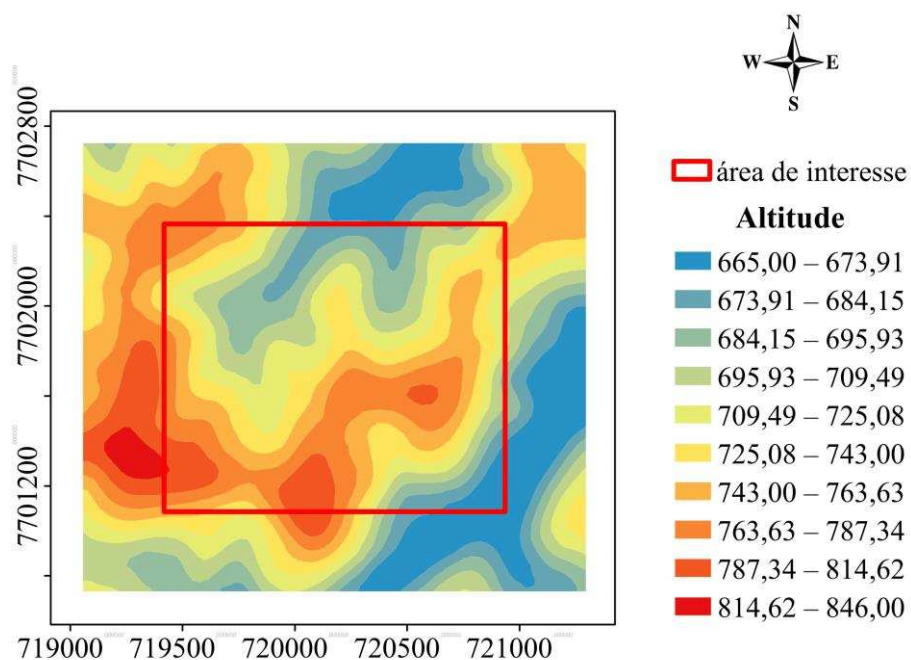


Figura 21: Mapa de krigagem da altitude da área de interesse mais área do *buffer* de 360 metros de distância da área de interesse do estudo, utilizando na interpolação os pontos amostrais em formato de malha irregular contidos nessa área.

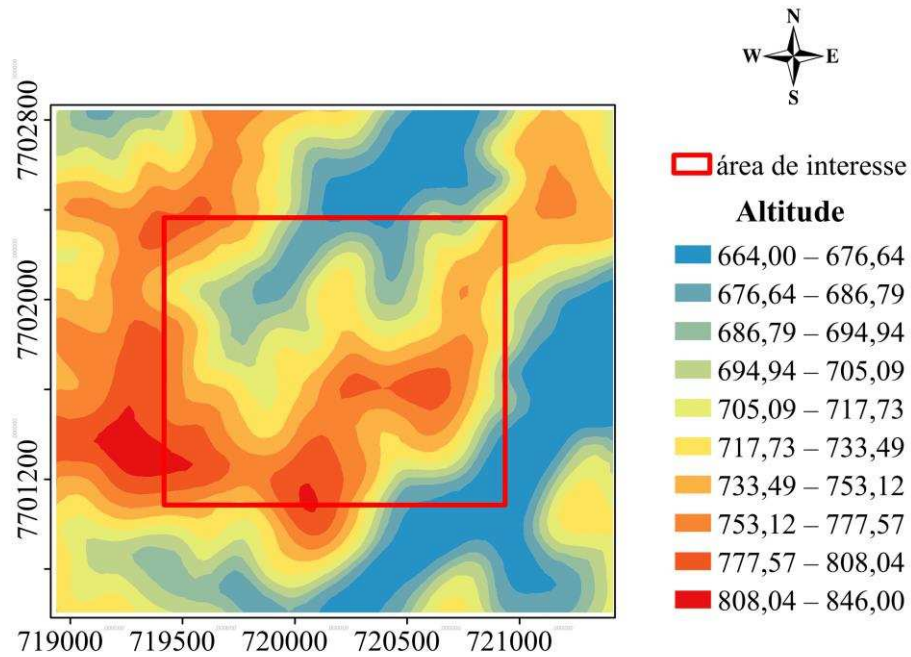


Figura 22: Mapa de krigagem da altitude da área de interesse mais área do *buffer* de 480 metros de distância da área de interesse do estudo, utilizando na interpolação os pontos amostrais em formato de malha irregular contidos nessa área.

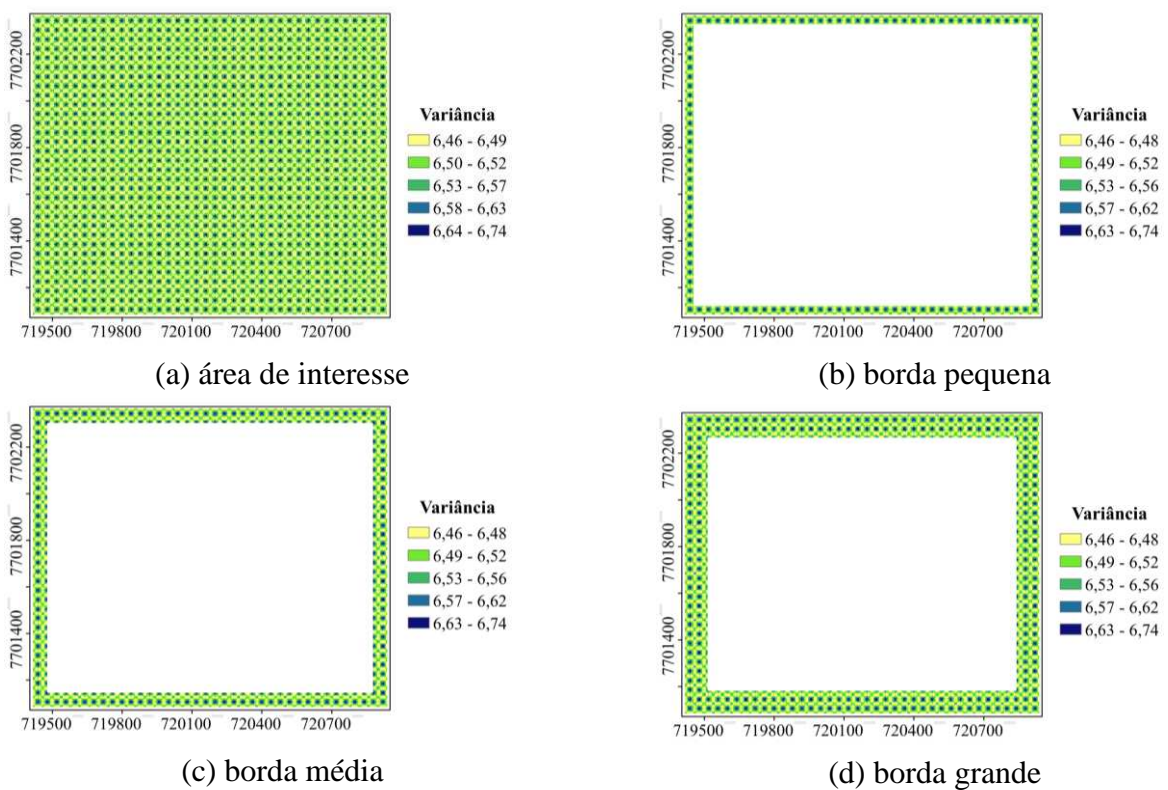


Figura 23: Mapa da variância de krigagem da área de interesse do estudo (a), da borda pequena (b), da borda média (c) e da borda grande (d), utilizando na interpolação os pontos amostrais em formato de malha regular quadrática contidos na área de interesse e na área do *buffer* de 120 metros de distância da área de interesse.

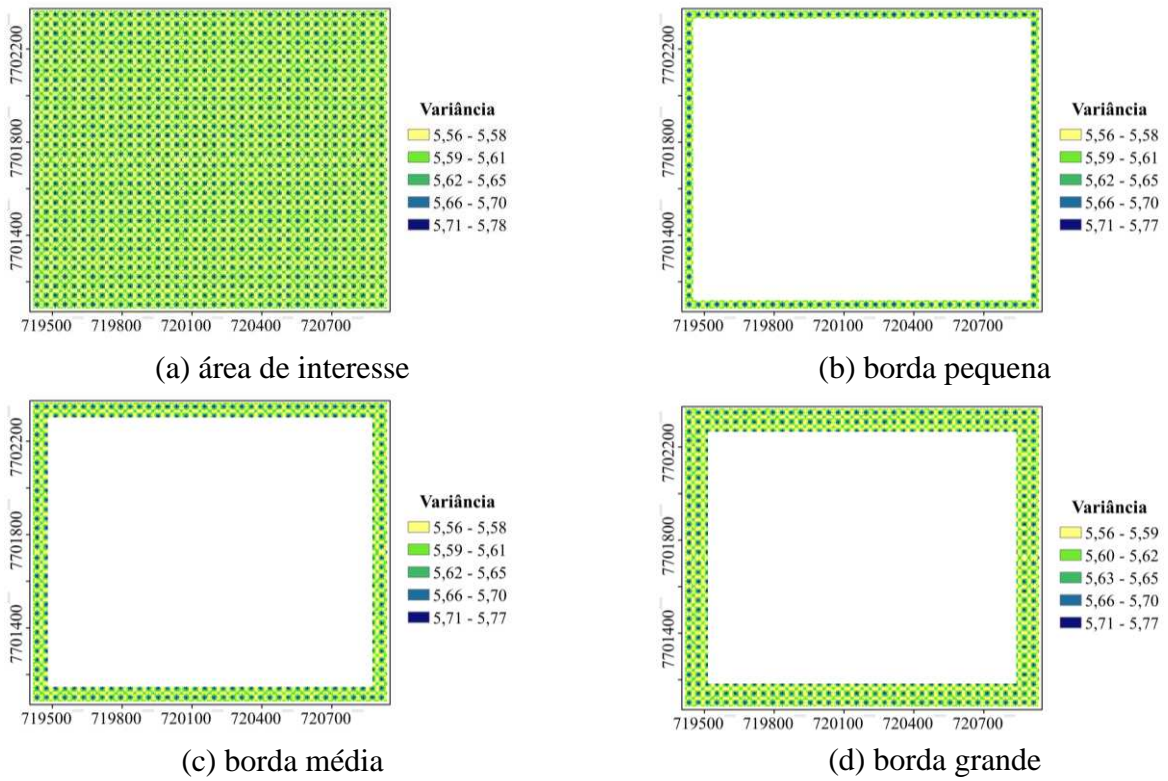
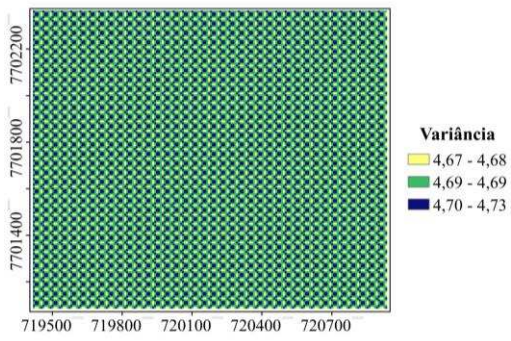
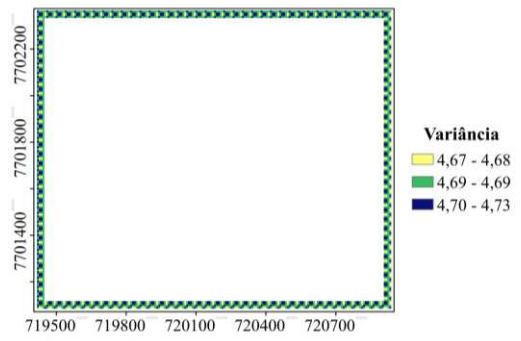


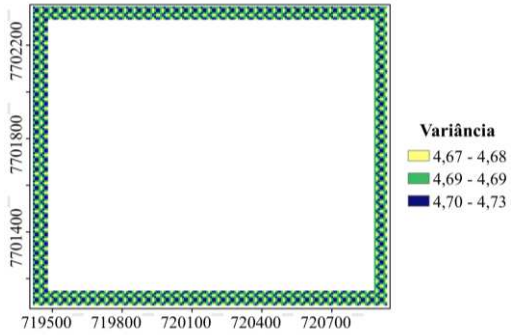
Figura 24: Mapa da variância de krigagem da área de interesse do estudo (a), da borda pequena (b), da borda média (c) e da borda grande (d), utilizando na interpolação os pontos amostrais em formato de malha regular quadrática contidos na área de interesse e na área do *buffer* de 240 metros de distância da área de interesse.



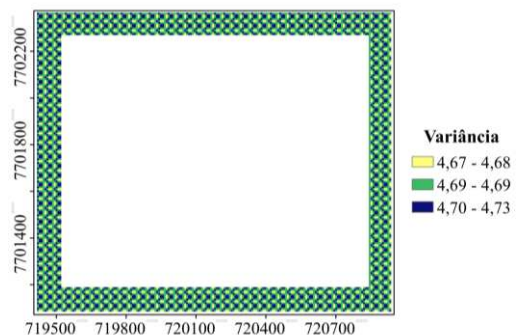
(a) área de interesse



(b) borda pequena



(c) borda média



(d) borda grande

Figura 25: Mapa da variância de krigagem da área de interesse do estudo (a), da borda pequena (b), da borda média (c) e da borda grande (d), utilizando na interpolação os pontos amostrais em formato de malha regular quadrática contidos na área de interesse e na área do *buffer* de 360 metros de distância da área de interesse.

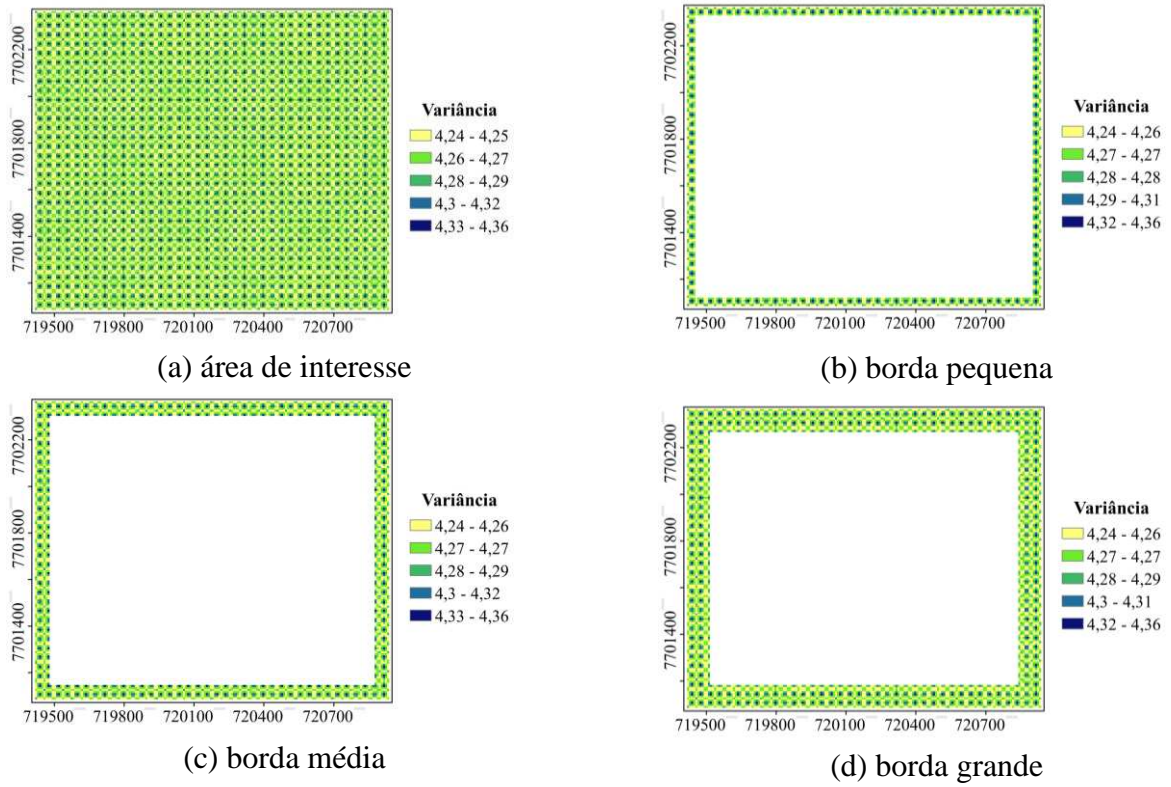
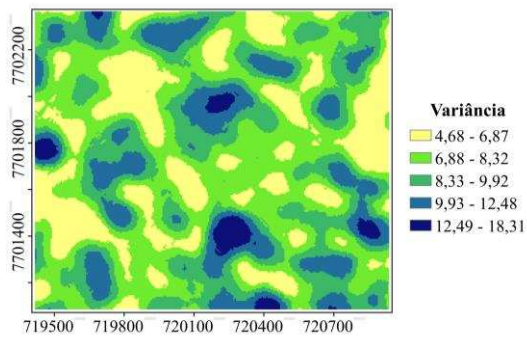
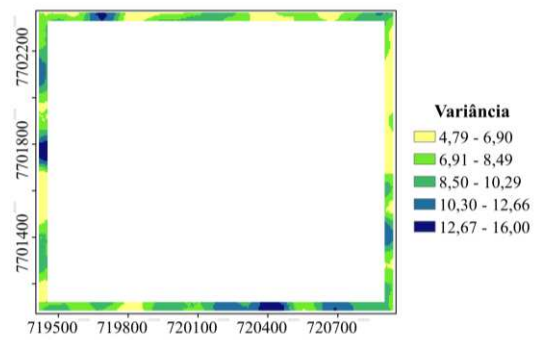


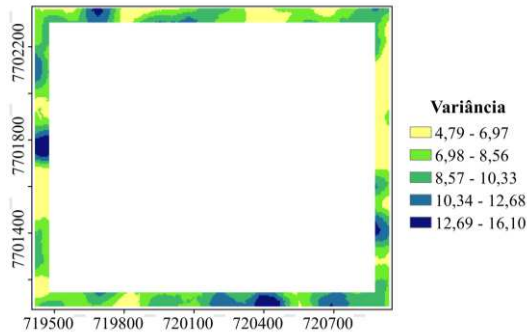
Figura 26: Mapa da variância de krigagem da área de interesse do estudo (a), da borda pequena (b), da borda média (c) e da borda grande (d), utilizando na interpolação os pontos amostrais em formato de malha regular quadrática contidos na área de interesse e na área do *buffer* de 480 metros de distância da área de interesse.



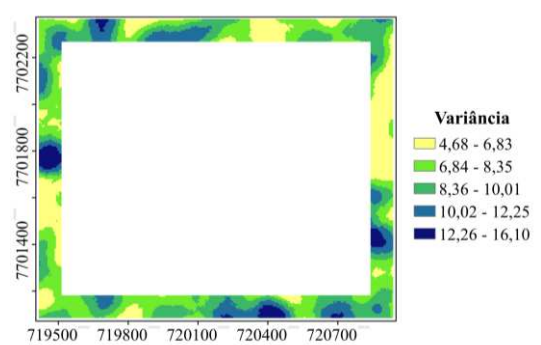
(a) área de interesse



(b) borda pequena

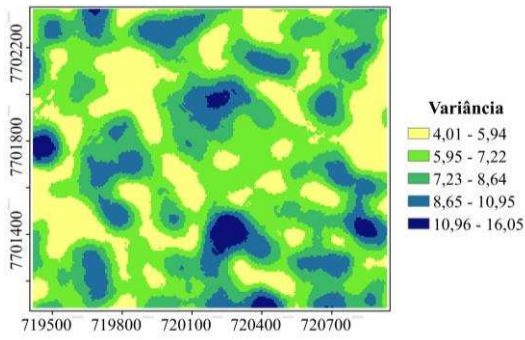


(c) borda média

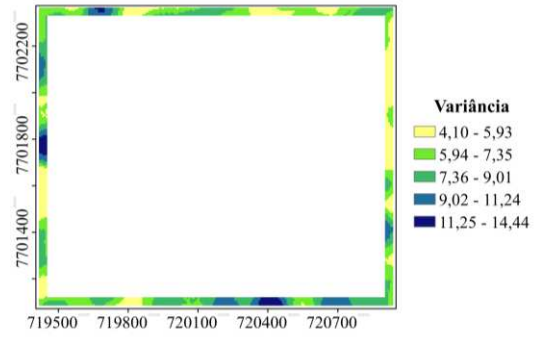


(d) borda grande

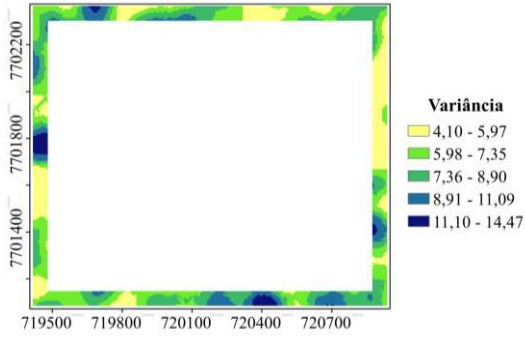
Figura 27: Mapa da variância de krigagem da área de interesse do estudo (a), da borda pequena (b), da borda média (c) e da borda grande (d), utilizando na interpolação os pontos amostrais em formato de malha irregular contidos na área de interesse e na área do *buffer* de 120 metros de distância da área de interesse.



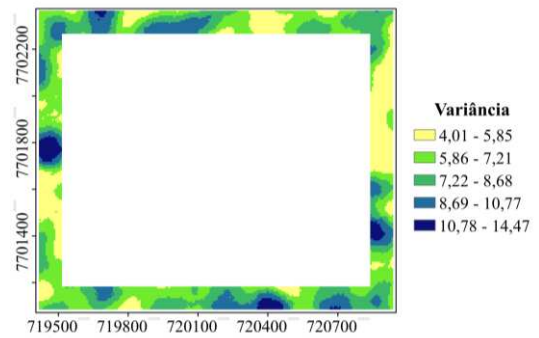
(a) área de interesse



(b) borda pequena

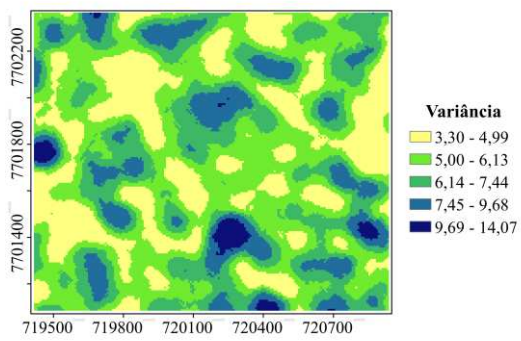


(c) borda média

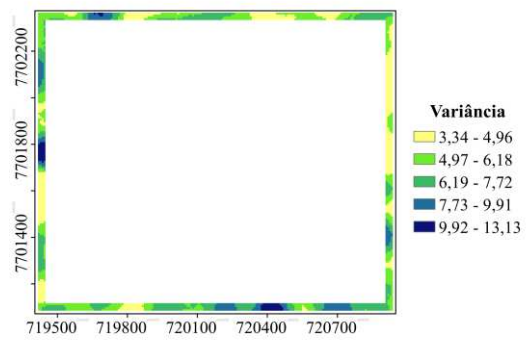


(d) borda grande

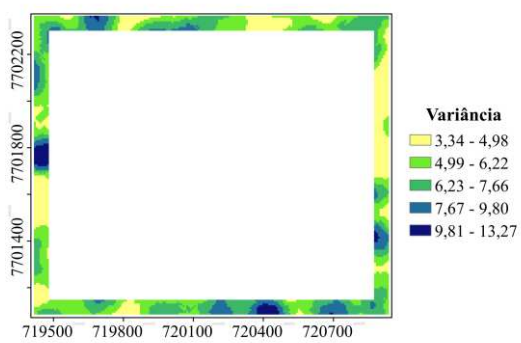
Figura 28: Mapa da variância de krigagem da área de interesse do estudo (a), da borda pequena (b), da borda média (c) e da borda grande (d), utilizando na interpolação os pontos amostrais em formato de malha irregular contidos na área de interesse e na área do *buffer* de 240 metros de distância da área de interesse.



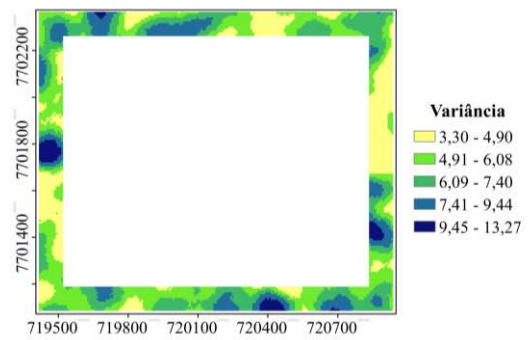
(a) área de interesse



(b) borda pequena

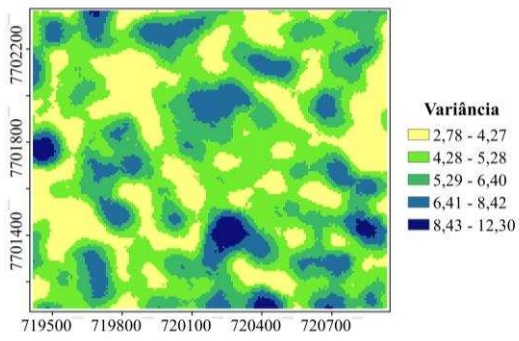


(c) borda média

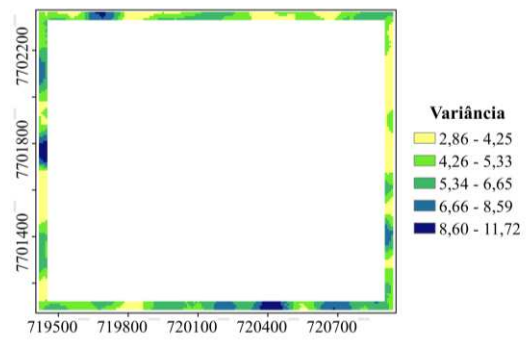


(d) borda grande

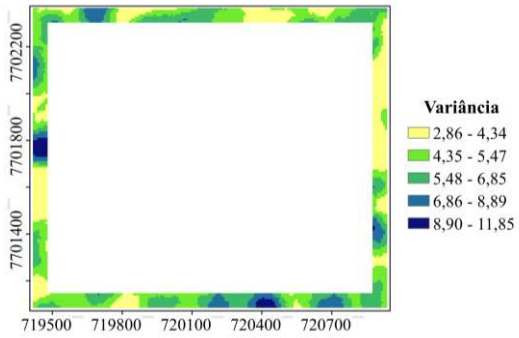
Figura 29: Mapa da variância de krigagem da área de interesse do estudo (a), da borda pequena (b), da borda média (c) e da borda grande (d), utilizando na interpolação os pontos amostrais em formato de malha irregular contidos na área de interesse e na área do *buffer* de 360 metros de distância da área de interesse.



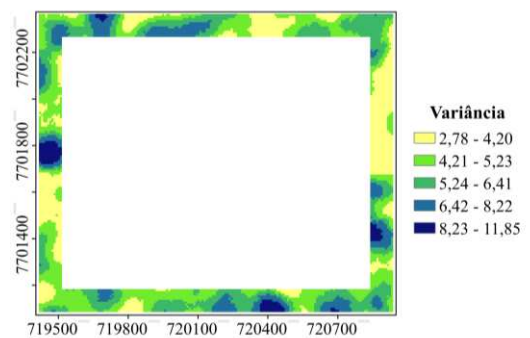
(a) área de interesse



(b) borda pequena



(c) borda média



(d) borda grande

Figura 30: Mapa da variância de krigagem da área de interesse do estudo (a), da borda pequena (b), da borda média (c) e da borda grande (d), utilizando na interpolação os pontos amostrais em formato de malha irregular contidos na área de interesse e na área do *buffer* de 480 metros de distância da área de interesse.