

CÁSSIO PINHO DOS REIS

**SIMULAÇÃO DE FATORES QUE AFETAM AS PREDIÇÕES OBTIDAS
POR KRIGAGEM ORDINÁRIA**

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
2013

**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

R375s
2013

Reis, Cássio Pinho dos, 1989-
Simulação de fatores que afetam as predições obtidas
por Krigagem Ordinária / Cássio Pinho dos Reis. – Viçosa,
MG, 2013.
x, 45f. : il. ; 29cm.

Orientador: Nerilson Terra Santos.
Dissertação (mestrado) - Universidade Federal de Viçosa.
Referências bibliográficas: f. 40-45.

1. Análise espacial (Estatística). 2. Interpolação.
3. Agricultura de precisão. I. Universidade Federal de Viçosa.
Departamento de Estatística. Programa de Pós-Graduação em
Estatística Aplicada e Biometria. II. Título.

CDD 22. ed. 519.5

CÁSSIO PINHO DOS REIS

**SIMULAÇÃO DE FATORES QUE AFETAM AS PREDIÇÕES OBTIDAS
POR KRIGAGEM ORDINÁRIA**

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: 18 de março de 2013.

Antônio Policarpo Souza Carneiro

Domingos Sárvio Magalhães Valente

Nerilson Terra Santos
(Orientador)

À todos aqueles que acreditaram.

AGRADECIMENTOS

Primeiramente a Deus, por tudo que fez na minha vida.

À toda minha família. Minha mãe Angela Pinho que sempre me incentivou em todos os momentos, me apoiando em todas as decisões tomadas. Ao meu pai Valter Reis, que sempre na medida do possível se fez presente, acreditando em mim durante todo o curso. Aos meus irmãos Sávio, Iago e Geovana pelo companheirismo mesmo estando longe. À toda Família Pinho e a Família Reis, pelo carinho que me tem, e que mesmo por alguma distância, sempre estavam lá para me apoiar.

Aos professores da graduação, Silvia Almeida, Edson Ramos e Adrilayne Araújo, da UFPA, que sempre acreditaram, apoiaram, dedicaram e me fizeram acreditar que eu sempre podia algo mais, mostrando os caminhos para vencer as barreiras.

Ao professor e orientador Nerilson Santos que me incentivou, e que mesmo estando longe, apoiou em minha pesquisa, sempre com muita paciência e boa vontade em ajudar, e ao professor e co-orientador Gerson Santos, que também ajudou a conseguir concluir esse trabalho.

À Universidade Federal de Viçosa, pela oportunidade de ingressar e concluir a pós-graduação e ao REUNI pelo apoio financeiro durante boa parte do curso.

Aos amigos do mestrado Diego, Vinícius, Fátima, Pâmela, Flávia, Nayara, Camila, Renata, Silvano e Franklin pela grande amizade que obtive em Viçosa. Aos amigos Bruno Vidigal, Wagner Pinheiro e Marcello Neiva, pelos anos de estudos e por uma amizade que certamente, não acaba por aqui. Aos amigos e companheiros de trabalho em Manaus, Marcio, Afrânio, Wal, Dani, Karine, Andréa, Klyssia, Anninson e Suzy.

E a todos aqueles que direta, ou indiretamente, contribuíram de alguma forma para o bom andamento da conclusão deste curso.

*“tudo o que eu aprendi,
é tudo que eu estou levando agora comigo”*

Soldiers of Jah Army.

SUMÁRIO

LISTA DE TABELAS	vi
LISTA DE FIGURAS	vii
RESUMO	ix
ABSTRACT.....	x
1. INTRODUÇÃO	1
2. REVISÃO BIBLIOGRÁFICA	4
2.1 Semivariograma.....	4
2.2 Krigagem	10
2.3 Comparação dos Métodos	14
2.3.1 Validação Cruzada	14
3. MATERIAL E MÉTODOS	17
4. RESULTADOS E DISCUSSÃO.....	20
4.1 Média dos Erros.....	20
4.2 Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS).....	23
4.3 Erro Padrão da Média (ASE).....	29
4.4 Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS).....	33
5. CONCLUSÕES	39
6. REFERÊNCIAS BIBLIOGRÁFICAS	40

LISTA DE TABELAS

Tabela 1: Valores dos Parâmetros Alcance, Efeito Pepita e Contribuição Escolhidos para Gerar as Simulações no Caso de Índice de Dependência Espacial Forte.....	18
Tabela 2: Valores dos Parâmetros Alcance, Efeito Pepita e Contribuição Escolhidos para Gerar as Simulações no Caso de Índice de Dependência Espacial Moderado.....	18
Tabela 3: Valores dos Parâmetros Alcance, Efeito Pepita e Contribuição Escolhidos para Gerar as Simulações no Caso de Índice de Dependência Espacial Fraco.....	18
Tabela 4: Valores da Média dos Erros de Predição das Krigagens Ordinárias de Tamanhos de grid 250x250, 300x300 e 350x250 com os IDE Forte, Moderado e Fraco, Usando Um, Dois, Quatro, Oito, Dezesesseis e Vinte Vizinhos Mais Próximos na Krigagem, para os Diferentes Modelos Estudados.	21
Tabela 5: Valores da Média Padronizada dos Erros de Predição das Krigagens Ordinárias de Tamanhos de grid 250x250, 300x300 e 350x250 com os IDE Forte, Moderado e Fraco, Usando Um, Dois, Quatro, Oito, Dezesesseis e Vinte Vizinhos Mais Próximos na Krigagem, para os Diferentes Modelos Estudados.	22
Tabela 6: Valores da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) das Krigagens Ordinárias de Tamanhos de grid 250x250, 300x300 e 350x250 com os IDE Forte, Moderado e Fraco, Usando Um, Dois, Quatro, Oito, Dezesesseis e Vinte Vizinhos Mais Próximos na Krigagem, para os Diferentes Modelos Estudados.	24
Tabela 7: Valores do Erro Padrão da Média (ASE) das Krigagens Ordinárias de Tamanhos 250x250, 300x300 e 350x250 com os IDE Forte, Moderado e Fraco, Usando Um, Dois, Quatro, Oito, Dezesesseis e Vinte Vizinhos Mais Próximos na Krigagem, para os Diferentes Modelos Estudados.	30
Tabela 8: Valores da Diferença entre a Valores da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) e o Erro Padrão da Média (ASE) das Krigagens Ordinárias de Tamanhos 250x250, 300x300 e 350x250 com os IDE Forte, Moderado e Fraco, Usando Um, Dois, Quatro, Oito, Dezesesseis e Vinte Vizinhos Mais Próximos na Krigagem, para os Diferentes Modelos Estudados.....	32
Tabela 9: Valores da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) das Krigagens Ordinárias de Tamanhos 250x250, 300x300 e 350x250 com os IDE Forte, Moderado e Fraco, Usando Um, Dois, Quatro, Oito, Dezesesseis e Vinte Vizinhos Mais Próximos na Krigagem, para os Diferentes Modelos Estudados.....	34

LISTA DE FIGURAS

Figura 1: Amostragem em Duas Dimensões de uma Variável Regionalizada, Mostrando Dois Pontos Separados por uma Distância H.	5
Figura 2: Mapa de Localização de Teores de Ferro (CLARK, 1979).....	6
Figura 3: Ajuste do Semivariograma Teórico para o Semivariograma Experimental, contendo os Parâmetros Efeito Pepita (C0), Contribuição (C1), Patamar (C) e Alcance (A).	7
Figura 4: Representação Gráfica de três Modelos Teóricos de Semivariograma: Esférico, Exponencial e Gaussiano.	8
Figura 5: Diferentes Cenários para Cada Índice de Dependência Espacial Simulado.....	17
Figura 6: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) dos Tamanhos de Grid Diferentes, das Krigagens Ordinárias em Função dos Diferentes Modelos Estudados e Número de Vizinhos, para o IDE Forte.....	26
Figura 7: Valores Médio da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) dos Tamanhos de Grid Diferentes, das Krigagens Ordinárias em Função dos Diferentes Modelos Estudados e Número de Vizinhos, para o IDE Moderado.....	26
Figura 8: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) dos Tamanhos de Grid Diferentes, das Krigagens Ordinárias em Função dos Diferentes Modelos Estudados e Número de Vizinhos, para o IDE Fraco.	27
Figura 9: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) dos Diferentes Modelos de Dependência Espacial e dos Diferentes Níveis de Dependência Espacial, das Krigagens Ordinárias em Função dos Tamanhos de Grid e Número de Vizinhos Mais Próximos na Krigagem	28
Figura 10: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) dos Diferentes Modelos de Dependência Espacial e para os Diferentes Tamanhos de Grid Amostral, das Krigagens Ordinárias em Função dos Diferentes Números de Vizinhos Mais Próximos, Para os Diferentes Índices de Dependência Espacial.	29
Figura 11: Valores Médios da Diferença entre os Valores da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) e o Erro Padrão da Média (ASE), dos Diferentes Tamanhos de Grid Amostral e dos Diferentes Números de Vizinhos, das Krigagens Ordinárias dos Diferentes Índices de Dependência Espacial, para os Modelos Estudados.....	31
Figura 12: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) dos Tamanhos de Grid Diferentes, das Krigagens Ordinárias em Função dos Diferentes Modelos Estudados e Número de Vizinhos, para o IDE Forte.	35
Figura 13: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) dos Tamanhos de Grid Diferentes, das Krigagens Ordinárias em Função dos Diferentes Modelos Estudados e Número de Vizinhos, para o IDE Moderado.	36

Figura 14: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) dos Tamanhos de Grid Diferentes, das Krigagens Ordinárias em Função dos Diferentes Modelos Estudados e Número de Vizinhos, para o IDE Fraco.....	36
Figura 15: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) dos Diferentes Modelos de Dependência Espacial e Níveis de Dependência Espacial, das Krigagens Ordinárias em Função dos Diferentes Tamanhos de Grid e Número de Vizinhos.	37
Figura 16: Valores da Média da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) dos Diferentes Modelos de Dependência Espacial e Tamanhos de Grid de Amostragem, das Krigagens Ordinárias em Função dos Diferentes Índices de Dependência Espacial e Número de Vizinhos.....	38
Figura 17: Valores da Média da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) das Diferentes Quantidades de Vizinhos e Tamanhos de Grid de Amostragem, das Krigagens Ordinárias em Função dos Diferentes Índices de Dependência Espacial e Modelos Estudados.	38

RESUMO

REIS, Cássio Pinho dos, M.Sc., Universidade Federal de Viçosa, Março de 2013. **Simulação de Fatores que Afetam as Predições Obtidas por Krigagem Ordinária**. Orientador: Nerilson Terra Santos. Coorientadores: Gerson Rodrigues dos Santos e José Ivo Ribeiro Júnior.

O objetivo deste estudo é poder identificar, como diversos fatores podem afetar a variância das predições obtidas por Krigagem Ordinária. Para tanto, foram simulados diferentes cenários para poder obter as predições de Krigagem Ordinária, usando os três modelos de semivariograma usuais, quantidades diferentes de vizinhos para realizar a predição, diferentes níveis de dependência espacial e tamanhos de amostra. Observou-se que embora todos os modelos tenderam produzir predições não viesadas, o modelo esférico tendeu a apresentar a maior variabilidade dos erros de predições na maioria dos cenários simulados, enquanto que o gaussiano obteve a menor. Verificou-se também que quando se aumentou o valor do índice de dependência espacial (IDE), o tamanho do grid e a quantidade de vizinhos mais próximos, os modelos tenderam a apresentar menores variabilidades para todos IDE estudados, sendo que o mesmo tende a se estabilizar quando se chega a uma determinada quantidade de vizinhos. Quanto à avaliação das estimativas, o modelo exponencial tendeu a ser o modelo com a pior avaliação sobre a variabilidade das predições e o esférico com a melhor avaliação na maioria dos cenários simulados, porém quando se aumenta a quantidade de vizinhos, o comportamento dos modelos se alteram de acordo com o índice de dependência espacial.

ABSTRACT

REIS, Cássio Pinho dos, M.Sc., Universidade Federal de Viçosa, March, 2013. **Simulation of Factors Affecting the Predictions Obtained by ordinary kriging.** Advisor: Nerilson Terra Santos. Co-Advisors: Gerson Rodrigues dos Santos and José Ivo Ribeiro Júnior.

The objective of this study is to identify how many factors can affect the variance of the predictions obtained by Ordinary Kriging. For this, different phenomena were simulated in order to obtain the predictions of Ordinary Kriging, using the three usual semivariogram models, different amounts of neighbors to perform the prediction, different spatial dependence indexes and sample size. It was observed that although all the models tended to produce non-biased predictions, the spherical model tended to show a greater variability of the prediction errors in most phenomena simulated, while the Gaussian got smaller. It was also found that when the value of spatial dependence indexes (FDI), the grid size and the number of nearest neighbors were increased, the models tend to have lower variability for all SDI studied, and these ones also tend to stabilize when it reaches a certain amount of neighbors. About the evaluation of the estimates, the exponential model tended to be the one with the worst rating on the variability of the predictions and the spherical the one with the best evaluation in most simulated phenomena. Nevertheless, when it increases the amount of neighbors, the models behavior changes according to the spatial dependence index.

1. INTRODUÇÃO

Compreender a distribuição espacial de dados oriundos de fenômenos ocorridos no espaço constitui hoje um grande desafio para a elucidação de questões centrais em diversas áreas do conhecimento, entre elas, saúde, meio ambiente, geologia, agronomia, entre outras (DRUCK, 2004). Na área de agricultura de precisão, por exemplo, por muito tempo o solo foi considerado uniforme para cada área de cultivo, mesmo possuindo características diferentes. Conhecer essa variabilidade espacial dos atributos é considerado um fator indispensável.

Estudar estatisticamente fenômenos espaciais requer cuidados especiais, uma vez que os métodos clássicos de análise estatística de dados geralmente supõem que, as realizações das variáveis aleatórias são independentes entre si, ou seja, que observações vizinhas não exercem influências umas sobre as outras. Porém, os fenômenos naturais apresentam frequentemente estruturação nas variações entre vizinhos. Desta forma pode-se dizer que as variações não são aleatórias e, portanto, apresentam algum grau de dependência espacial (ZIMBACK, 2003).

Portanto, as medidas descritivas da análise exploratória, como a média, variância, desvio-padrão, coeficiente de variação, valores mínimo e máximo, coeficiente de assimetria e curtose, não são suficientes para descrever a variabilidade espacial. Isto ocorre, porque estas técnicas da estatística clássica assumem que todas as amostras são aleatórias e independentes de uma distribuição de probabilidade simples, além de não considerarem as posições geográficas das amostras no campo (ORTIZ, 2002).

O avanço do conhecimento nas mais diversas áreas tem evidenciado as limitações dos métodos tradicionais da estatística, no tratamento da variabilidade espacial de diversos tipos de variáveis. Justamente pelo fato de que estas variáveis apresentarem algum grau de dependência espacial.

O começo do estudo desses fenômenos espaciais na estatística se deu na África do Sul, com o engenheiro de minas Daniel G. Krige e o estatístico H. S. Sichel, uma vez que eles desenvolveram de modo empírico uma técnica própria de

estimativa para o cálculo de reservas minerais, por meio da distribuição geográfica das amostras.

Krige estudava jazidas minerais por meio de estatísticas clássicas. Porém, ele teve a necessidade de incluir a distância entre as observações de concentração de ouro, para evitar a superestimação sistemática de reservas em mineração. Pois somente o uso da variância de suas amostras não era suficiente para explicar o fenômeno em estudo, se não levasse em consideração tais distâncias (KRIGE, 1951).

Posteriormente, George Matheron deu um tratamento formal para os trabalhos de Krige no início dos anos 60 do século passado na França, com o nome de Geoestatística. A Geoestatística trata do estudo das chamadas variáveis regionalizadas, ou seja, variáveis com condicionamento espacial, pois levava em consideração a localização geográfica e a dependência espacial dos dados, por meio da teoria das Variáveis Regionalizadas (MATHERON, 1963)

Uma variável regionalizada pode ser definida de acordo com Blais e Carlier (1968), como uma função espacial numérica, que varia de um local para outro, com uma continuidade espacial aparente e cuja variação não pode ser representada por uma função matemática simples.

De acordo com Landim (2006), o trabalho de Krige foi publicado em 1951 e os primeiros trabalhos de Matheron foram realizados após sua viagem à África do Sul quando conheceu Krige. A primeira publicação, descrevendo os resultados de Krige foi em co-autoria com Duval e L'evy em 1955. Nos anos seguintes vieram as obras mais completas de Matheron, já com o arcabouço completo da Geoestatística Linear (MATHERON, 1963).

A geoestatística é um conjunto de técnicas que permitem predizer valores de atributos numa área em estudo. Esta tem como princípio, a primeira lei da geografia: “Todas as coisas se parecem, porém as mais próximas no espaço se parecem mais”.

Vários são os métodos utilizados numa análise geoestatística geoestatísticos como por exemplo, a obtenção do semivariograma. Este gráfico é usado para investigar a magnitude da correlação entre as amostras e sua similaridade com a distância. Por meio deste, será possível ajustar um modelo teórico que será utilizado

na predição dos valores da variável em estudo para locais não amostrados, utilizando para este fim a krigagem que é um dos métodos utilizados para interpolação.

O primeiro uso de krigagem foi para solucionar problemas de mapeamentos geológicos. Posteriormente foi usado com sucesso no mapeamento de solos (BURGESS; WEBSTER, 1980), mapeamento hidrológico (KITANIDIS; VOMVORIS, 1983) e mapeamento atmosférico (LAJAUNIE, 1984).

A técnica de krigagem é preferida em detrimento a outros métodos de interpolação (polígonos, triangulação, inverso da distância) porque permite produzir predições sem tendência que apresentam variância mínima para os erros de predição.

Vários fatores relacionados à dependência espacial de um fenômeno podem influenciar na variabilidade das predições obtidas e conseqüentemente na confiabilidade do mapa de interpolação obtido.

Por conta disto, este presente trabalho, teve como objetivo principal avaliar a influência de alguns fatores na variância das predições obtidas por Krigagem Ordinária. Os objetivos específicos para este trabalho foram avaliar se o modelo do semivariograma, o número de vizinhos adotados durante a interpolação, o tamanho da grade de amostragem, e o índice de dependência espacial influenciam na variância das predições obtidas ao realizar a krigagem ordinária.

2. REVISÃO BIBLIOGRÁFICA

No estudo do comportamento das variáveis regionalizadas em questão, duas são as técnicas fundamentais dos métodos geoestatísticos: num primeiro momento o semivariograma e posteriormente a Krigagem. Para se fazer uma inferência por meio da Geoestatística, é necessário inicialmente a modelagem do semivariograma.

2.1 Semivariograma

Um dos métodos mais antigos de se estimar a dependência no espaço ou no tempo, de amostras vizinhas é a autocorrelação, que é bem útil quando se está fazendo amostragem em uma única direção (VAUCLIN et al., 1982). Porém, quando as amostras são coletadas nas duas dimensões do campo e a interpolação entre locais medidos for necessária, é preciso usar uma técnica mais adequada para medir a influencia seus vizinhos. Esta técnica é o semivariograma (VIEIRA et al., 1983).

Neste referencial cartesiano analisa-se a variabilidade espacial em função da distância numa determinada direção, e desta em relação a outras. Segundo Guerra (1988), pode ocorrer três tipos de semivariogramas: o observado ou experimental (obtido a partir das amostras colhidas no campo), o verdadeiro (real, mas desconhecido) e o teórico (de referência, que é ajustado ao experimental).

Considere duas variáveis regionalizadas, $Z(X)$ e $Z(X + H)$. Estas se referem a um mesmo atributo medido em duas posições diferentes, conforme ilustra a Figura 1, onde X denota uma posição em duas dimensões, com componentes (X_i, Y_i) , e H um vetor distância que separa os pontos (PIRES, 2010). O valor de cada ponto está relacionado com valores situados a certa distância H , sendo que a relação é tanto maior quanto menor for a distância entre os pontos.

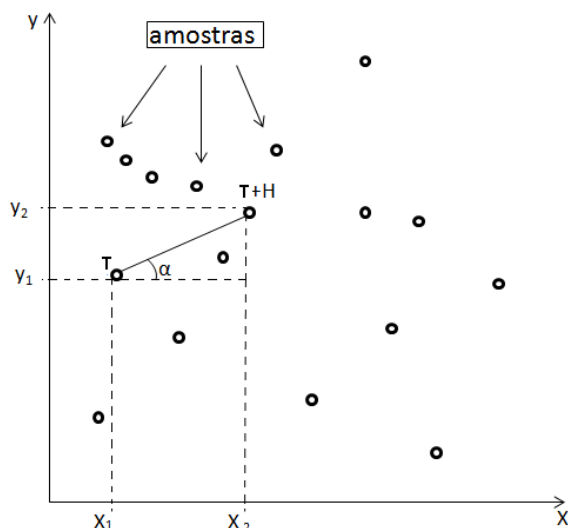


Figura 1: Amostragem em Duas Dimensões de uma Variável Regionalizada, Mostrando Dois Pontos Separados por uma Distância H .

O nível de dependência entre essas duas variáveis regionalizadas é representado pelo semivariograma, o qual é definido como a metade da esperança matemática do quadrado da diferença entre os valores de pontos no espaço, separados por um mesmo vetor H . Essa dependência espacial é estimada, segundo Isaaks e Srivastava (1989), pela equação 1

$$\gamma(H) = \frac{1}{2N(H)} \sum_{i=1}^N [Z_{(T_i)} - Z_{(T_i+H)}]^2 \quad (1)$$

onde: $\gamma(H)$ é o valor da semivariância estimada para a distância H ; $Z(T_i)$ e $Z(T_i+H)$ são valores da i -ésima observação da variável regionalizada, coletadas nos pontos T_i e T_i+H ($i=1, \dots, n$), separados por vetor H ; $N(H)$ é o número de pares de valores medidos, $Z(T_i)$ e $Z(T_i+H)$, separados por um vetor distância H .

De acordo com Landim (2006), o semivariograma mostra o grau de dependência espacial entre amostras ao longo de um suporte específico e, para sua construção, são usados as diferenças ao quadrado dos valores obtidos, assumindo-se uma estacionaridade nos incrementos.

Em Clark (1979) é apresentada a construção de um semivariograma a partir de uma rede irregular, como se pode ver na Figura 2, a qual se refere a um depósito estratiforme de ferro cujos valores são expressos em porcentagem por peso. Para o

cálculo dos semivariogramas, em diversas direções, são encontradas as somatórias dos quadrados das diferenças e posteriormente se divide por dois o número dessas diferenças.

Assim, para a direção Leste-Oeste inicia-se com o menor intervalo possível, ou seja, 30,48 metros da seguinte maneira: $(40-42)^2$, $(42-40)^2$, $(40-39)^2$, e assim por diante. Soma-se todas essas diferenças e divide-se pela quantidade de diferenças que se obteve. Para o intervalo de 60,96 metros, $(44-40)^2$, $(40-40)^2$, etc. E assim por diante, nas outras direções e para todas as distâncias.

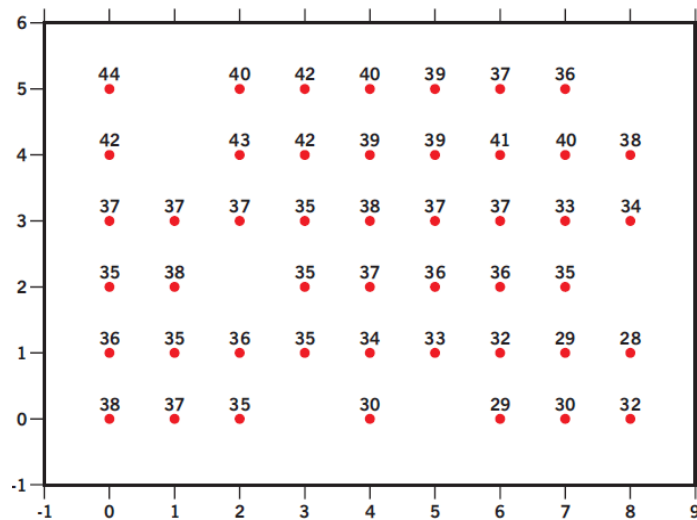


Figura 2: Mapa de Localização de Teores de Ferro (CLARK, 1979).

De posse do semivariograma experimental é necessário ajustar um modelo teórico, ou seja, a uma função que definirá os parâmetros do semivariograma: efeito pepita, alcance e patamar. A análise e o ajuste de um modelo teórico a um semivariograma experimental denomina-se Análise Estrutural, que pode ser representado pela Figura 3.

O seu padrão representa que intuitivamente, se espera que as diferenças $\{Z(T_i) - Z(T_i + H)\}$ cresçam à medida que H (a distância que os separa) também cresça, até que se estabeleça num valor máximo da semivariância.

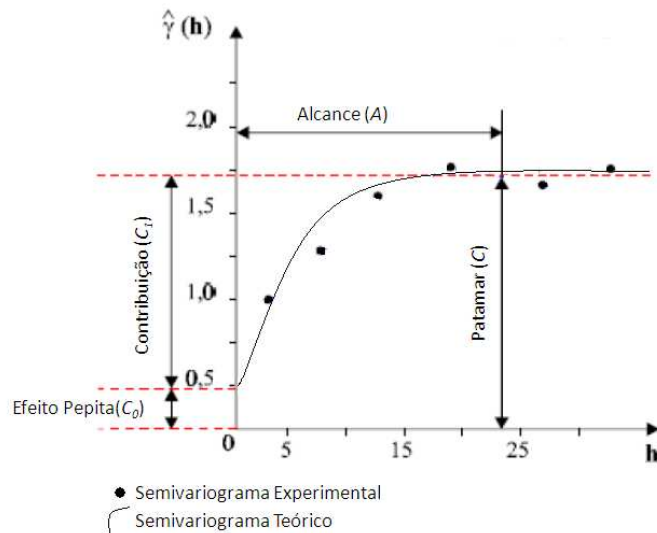


Figura 3: Ajuste do Semivariograma Teórico para o Semivariograma Experimental, contendo os Parâmetros Efeito Pepita (C_0), Contribuição (C_1), Patamar (C) e Alcance (A).

O $\gamma(H)$ é a semivariância, que permite descrever quantitativamente a variação no espaço de um fenômeno regionalizado para uma determinada distância H .

O Alcance (A) corresponde a zona de influência ou de dependência espacial de uma amostra, marcando a distância a partir da qual as amostras tornam-se independentes. No exemplo da Figura 3, o alcance ocorre próximo de $h = 25$. Após este valor, a estatística clássica por si só é capaz de estudar o fenômeno, pois não há mais dependência espacial entre as amostras.

O efeito pepita (C_0) corresponde a cota do ponto onde o semivariograma corta o eixo das ordenadas, ou seja, quando a distância é 0. De acordo com Valente (1989), este valor indica uma variabilidade não explicada (ao acaso) de um ponto para o outro, que pode ser devido a erros de medida, microestruturas não captadas pela menor escala de amostragem, erros de amostragem, de análises laboratoriais, entre outros.

Como foi visto, espera que as diferenças $\{Z(T_i) - Z(T_i + H)\}$ aumentem à medida que a distância H também aumente, até que se estabilize num valor

conhecido como Patamar (C). Este valor corresponde à distância onde toda a semivariância da amostra é devido somente a influência aleatória.

O patamar também representa uma estimativa da variância total (s^2) dos valores observados. Quando o efeito pepita for aproximadamente igual ao Patamar, o fenômeno é conhecido como Efeito Pepita Puro, o qual indica que a amostra não possui dependência espacial entre as amostras (TRANGMAR et al., 1985).

Num ajuste os modelos teóricos são superpostos aos pontos do variograma experimental. Esses modelos teóricos podem ser divididos em modelos com patamar e modelos sem patamar. Os modelos com patamar são conhecidos como modelos transitivos. Os modelos transitivos mais utilizados são o esférico, o exponencial e o gaussiano (CRESSIE 1993; DAVIS 2002). Estes modelos estão apresentados na Figura 4 com diferentes alcances (A).

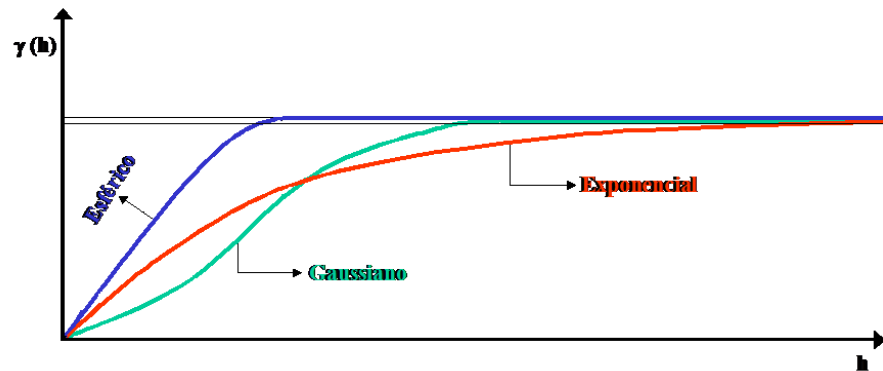


Figura 4: Representação Gráfica de três Modelos Teóricos de Semivariograma: Esférico, Exponencial e Gaussiano.

De acordo com Zimback (2003), os modelos esférico, exponencial e gaussiano são dados pelas Equações 2, 3 e 4, respectivamente.

$$Esférico : \gamma(h) = \begin{cases} 0 & , H = 0 \\ C_0 + C_1 \left[\frac{3}{2} \left(\frac{H}{A} \right) - \frac{1}{2} \left(\frac{H}{A} \right)^3 \right] & , 0 < H < a \\ C_0 + C_1 & , H \geq a \end{cases} \quad (2)$$

$$\text{Exponencial: } \gamma(h) = \begin{cases} 0 & , H = 0 \\ C_0 + C_1 \left[1 - e^{\left[-3 \left(\frac{H}{A} \right) \right]} \right] & , H \neq 0 \end{cases} \quad (3)$$

$$\text{Gaussiano: } \gamma(h) = \begin{cases} 0 & , |H| = 0 \\ C_0 + C_1 \left[1 - e^{\left[-3 \left(\frac{H}{A} \right)^2 \right]} \right] & , |H| \neq A \end{cases} \quad (4)$$

Enquanto os modelos esférico e exponencial possuem crescimento rápido na origem, o modelo gaussiano reflete mais regularidade da variável estudada, justamente por possuir uma tangente horizontal na origem.

Uma diferença fundamental entre o modelo esférico e os modelos exponencial e gaussiano é que estes atingem o patamar apenas assintoticamente, ou seja, quando tende a distância tende ao infinito, enquanto que o modelo esférico atinge o patamar no valor do alcance. De acordo com Landim (2006), o modelo esférico é o modelo mais comum, podendo-se afirmar que equivale à função de distribuição normal da estatística clássica.

Na análise estrutural do semivariograma, além do modelo teórico, do efeito pepita (C_0), do patamar ($C + C_0$) e do alcance (A), outros parâmetros podem ser utilizados para a escolha do modelo que melhor se ajusta ao semivariograma empírico.

A Soma dos Quadrados dos Resíduos (SQR) determina o ajuste do modelo teórico ao semivariograma experimental. Quanto menor o valor do SQR, melhor o ajuste de acordo com Zimmerman e Zimmerman (1991). O Coeficiente de Determinação (R^2) é uma medida da proporção da variabilidade em uma variável que é explicada pela variabilidade da outra, que no caso da Geoestatística pode indicar quantos dos pontos do semivariograma experimental encontram-se na curva do modelo teórico, embora não seja considerado um bom indicativo de ajuste, sendo que quanto maior o seu valor, melhor seria o ajuste. O critério de Akaike, que é obtido através da função de Máxima Verossimilhança, a partir dos parâmetros ajustados

para os modelos conforme os métodos, é um outro método para efeito de comparação, quanto menor este valor, melhor será o ajuste.

A modelagem, ou seja, o ajuste de um modelo teórico ao semivariograma experimental é um passo fundamental na análise variográfica, sendo um processo que envolve várias tentativas e na qual a experiência pesa muito.

Foi sugerido por vários autores (CAMBARDELLA et al. , 1994 e ZIMBACK et al., 1994), a razão entre o Efeito Pepita e o Patamar para mensurar a dependência espacial de um fenômeno.

Cambardella et al. (1994) propuseram o Índice de Dependência Espacial (IDE) expresso na Equação 5 com a seguinte interpretação para os seus valores: se $IDE \leq 25\%$ o fenômeno tem forte dependência espacial; entre 25 e 75% tem dependência espacial moderada e $\geq 75\%$ tem fraca dependência espacial.

$$IDE = \frac{C_0}{C_0 + C} \times 100 \quad . \quad (5)$$

Zimback (2001) propôs a inversão dos fatores de acordo com a Equação 6.

$$IDE = \frac{C}{C_0 + C} \times 100 \quad . \quad (6)$$

Sendo que a interpretação dos valores fica também invertida, ou seja, dependência fraca para valores $\leq 25\%$; moderada entre 25 e 75% e forte para $\geq 75\%$.

2.2 Krigagem

Para Pires (2010), na maioria das vezes o interesse da Geoestatística não se limita à obtenção de um modelo de dependência espacial, desejando-se também, prever valores em pontos não amostrados. O interesse pode ser em um ou mais pontos específicos da área ou obter uma malha de pontos interpolados que permitam visualizar o comportamento da variável na região, por meio de um mapa de superfície.

A técnica da confecção de mapas de superfície, onde são geradas estimativas de dados de pontos não amostrados a partir de pontos amostrados, denomina-se interpolação de dados. Uma maneira de estimar o valor de um ponto não amostrado é calcular uma média ponderada usando os valores observados nos pontos amostrados mais próximos ao ponto que se deseja obter tal estimativa.

Esses pesos podem ser atribuídos de diversas maneiras, como por exemplo, pontos amostrados situados em condições mais similares ao ponto de interesse, recebe um peso maior do que pontos amostrados situados em condições menos similares.

A maneira como se atribui os pesos aos pontos define o método de predição ou interpolação utilizado. Com base nisso, a Geoestatística oferece uma ampla e flexível variedade de técnicas que fornecem predições para locais não amostrados. Um dos métodos mais utilizados é a Krigagem e considerada por vários autores (BURGESS e WEBSTER, 1980; TRANGMAR et al., 1985; JOURNAL e ROSSI, 1989; OLIVIER e WEBSTER, 1991; ROBERTSON et al., 1997; BOURENNANE et al., 2003) como bastante adequada para a geração de mapas.

O termo “Krigagem” foi usado primeiramente por Matheron, em 1965, em homenagem ao engenheiro de minas sul-africano Daniel G. Krige, que primeiro formulou e implementou essa forma de interpolação, em 1951. A Krigagem pode ser usada em variáveis discretas e contínuas, e é por isso, sensível para a predição de variáveis (ROSSI et al., 1994).

Existem outros diversos métodos de interpolação de dados, como por exemplo, o método de polígono, o de triangulação, por média local amostrado e o inverso do quadrado da distância. no entanto, tais métodos não levam em consideração a continuidade existente entre valores observados em pontos amostrados próximos.

Contudo, a Krigagem leva em consideração a continuidade que existe entre os valores dos pontos amostrados e entre os pontos amostrados e o não amostrado para o qual se pretende obter uma estimativa. Adicionalmente, a Krigagem também leva em consideração a distância existente entre os pontos amostrados e entre os pontos amostrados e o não amostrado (GUMPERTZ, 2001).

De acordo com Rossi et al. (1994), três características da Krigagem a distinguem dos outros métodos de interpolação. A primeira característica é que a Krigagem pode fornecer uma predição que é maior ou menor do que os valores da amostra.

A segunda característica é que a Krigagem além de usar a distância euclidiana também usa a distância estatística ao utilizar um modelo de dependência espacial expresso pelo semivariograma.

Terceira característica, a krigagem leva em conta a minimização da variância do erro de predição, sendo que o erro de predição é definido de acordo com a Equação 7.

$$E_{(t_0)} = \hat{Z}_{(t_0)} - Z_{(t_0)} \quad . \quad (7)$$

Em que $\hat{Z}_{(t_0)}$ é a predição do ponto não amostrado t_0 obtido pela Krigagem, e $Z_{(t_0)}$ é o verdadeiro valor do mesmo ponto t_0 e $E(t_0)$ é o erro de predição.

A krigagem é dita ser o melhor interpolador linear não tendencioso (BLUP). A linearidade significa que a predição de um ponto não amostrado é função linear dos pontos amostrados. A condição de não tendência significa que, em média a diferença entre valores preditos e verdadeiros é nula. Como pode ser visto em Pires (2010), a condição de melhor preditor ocorre porque dentre todos os preditores é o que produz predições com variância mínima.

As formas mais usuais de Krigagem são a simples, disjuntiva, ordinária e a universal. Porém a Krigagem Ordinária, é uma das variações mais usada da Krigagem. Esta foi descrita por Trangmar et al. (1985), como o valor interpolado $\hat{Z}_{(t_0)}$ de uma variável regionalizada Z , num local T_0 e que pode ser determinada pela equação 8

$$\hat{Z}_{(t_0)} = \sum_{i=1}^n \left[\lambda_i \times Z_{(t_i)} \right] \quad (8)$$

onde $\hat{Z}_{(t_0)}$ é o valor predito para o local T_0 não amostrado; $Z_{(t_i)}$ o valor obtido por amostragem no campo em cada ponto amostrado T_i ; n o número de amostras vizinhas e λ_i são os pesos a cada valor observado $Z_{(t_i)}$.

A construção do preditor $\hat{Z}_{(t_0)}$, na Krigagem Ordinária, não requer o conhecimento inicial da média estacionária da amostragem, como é o caso da Krigagem Simples. Para se obter predições não viesadas é necessário que a soma dos pesos atribuídos aos valores observados em cada predição seja igual a 1. O cálculo dos pesos da Krigagem Ordinária é feito por meio da Equação 9.

$$[\lambda] = [A]^{-1}[B] \quad . \quad (9)$$

A matriz $[A]$ é constituída pela semivariância entre os pontos vizinhos amostrados $[\gamma(t_i - t_j)]$ e $[B]$ é constituída pela semivariância entre o ponto vizinho amostrado i e o ponto não amostrado $[\gamma(t_i - t_0)]$, e μ é o multiplicador de Lagrange, que é necessário para a obtenção de predições não viesadas, da forma que

$$\begin{bmatrix} \gamma(t_1 - t_1) & \gamma(t_1 - t_2) & \cdots & \gamma(t_1 - t_n) & 1 \\ \gamma(t_2 - t_1) & \gamma(t_2 - t_2) & \cdots & \gamma(t_2 - t_n) & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma(t_n - t_1) & \gamma(t_n - t_2) & \cdots & \gamma(t_n - t_n) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma(t_1 - t_0) \\ \gamma(t_2 - t_0) \\ \vdots \\ \gamma(t_n - t_0) \\ 1 \end{bmatrix} \quad .$$

De acordo com Uzumaki (1994), o sistema de Krigagem Ordinária tem solução única se o modelo de semivariância for válido. A Krigagem, além de ser um preditor não tendencioso, é um interpolador exato, isto é, se o ponto a ser predito coincidir com um dos pontos amostrados, o valor predito deverá ser igual ao valor amostrado em média.

Uma condição necessária para que um estimador seja considerado um estimador não-viesado, é que o valor esperado para o erro de predição seja nulo, como se pode ver na Equação 10.

$$E\left[\hat{Z}_{(t_i)} - Z_{(t_i)}\right] = 0 \quad . \quad (10)$$

Um dos fatores que podem influenciar a precisão do método de interpolação é o número de pontos amostrados usados para a predição (GOOVAERTS, 1997). A vizinhança de busca onde serão obtidos os pontos vizinhos, o modelo de semivariograma teórico ajustado e o nível de dependência espacial também podem influenciar a variância da predição e, conseqüentemente, uma boa validação.

2.3 Comparação dos Métodos

Com toda a subjetividade e variabilidade de resultados nos cálculos dos parâmetros do variograma, por conta de diversos fatores que podem influenciar no seu valor, é importante que se tenha um meio para checar se o modelo ajustado é satisfatório ou não (DAVID, 1988), bem como para avaliar o plano de Krigagem antes do seu uso na construção de mapas. Para um estudo espacial alguns critérios estatísticos têm sido utilizados, como por exemplo variância da Krigagem e Validação Cruzada.

2.3.1 Validação Cruzada

Nessa técnica, o valor de cada ponto amostrado ($Z_{(t_i)}$) é excluído e seu valor é predito ($\hat{Z}_{(t_i)}$) com base nos demais valores amostrados. Os pares das diferenças dos valores amostrados e preditos n de cada ponto permitem avaliar a qualidade dos ajustes da krigagem, e conseqüentemente dos semivariogramas a partir do cálculo de algumas estatísticas (JOHNSTON, 2001).

Uma destas estatísticas é a média dos erros (*Mean Prediction Errors - ME*) a qual deve apresentar valor próximo de zero, pois este valor indica que o método apresentou, na média, predições não viesadas, porém o valor desta estatística pode

sofrer efeito de escala. O valor desta estatística pode ser obtido por meio da Equação 12, segundo Silva (2006), por

$$ME = \frac{\sum_{i=1}^n [\hat{Z}_{(t_i)} - Z_{(t_i)}]}{n} . \quad (12)$$

Por conta que a ME pode ser influenciada pela escala, pode-se dividir o erro de predição pelo erro padrão de cada predição, obtendo assim, a média dos erros de predição padronizados (MS), cujo valor ideal também é próximo do valor zero, e pode ser encontrada em (Schreiber, 1997) conforme a Equação 13.

$$MS = \frac{\sum_{i=1}^n \frac{[\hat{Z}_{(t_i)} - Z_{(t_i)}]}{\hat{\sigma}_{(t_i)}}}{n} . \quad (13)$$

É desejável que o valor predito seja próximo do verdadeiro valor. A estatística Raiz Quadrada da Média do Quadrado das Diferenças entre os Valores Estimados e o Verdadeiro (*Root-Mean-Square-RMS*) permite avaliar a variabilidade do erro de predição, uma vez que menores valores de RMS, menor será a variabilidade entre valores preditos e os respectivos valores verdadeiros (Tatalovich et al., 2006). A estatística RMS é definida de acordo com a Equação 14

$$RMS = \sqrt{\frac{(\sum_{i=1}^n [\hat{Z}_{(t_i)} - Z_{(t_i)}])^2}{n}} . \quad (14)$$

Para acessar a incerteza associada às predições pode-se usar o erro padrão da média do erro de predição (*Average Standard Error-ASE*) (GEORGAKARAKOS e KITSIOU, 2008). O desejável é que o valor do ASE seja próximo ao do RMS.

Caso o valor do ASE seja superior ao RMS, a variabilidade da predição expressa pelo erro padrão ($\hat{\sigma}_{(t_i)}$) esta sendo superestimada. Por outro lado, se o valor do ASE é inferior ao RMS, a variabilidade da predição está sendo subestimada. O ASE pode ser obtido por meio da Equação 15

$$ASE = \sqrt{\frac{\sum_{i=1}^n \hat{\sigma}_{(t_i)}}{n}} . \quad (15)$$

Uma alternativa para também acessar a incerteza das predições (Szipiro, et al., 2007) é a raiz quadrada da média do quadrado das diferenças entre os valores estimados e o verdadeiro padronizado (Root-Mean-Square-Standardized – RMSS).

Se o valor de RMSS for superior a 1, a variabilidade da predição está sendo subestimada, e se for menor que 1 está sendo superestimada. Portanto, o valor ideal para o RMSS é que seja próxima de 1. Esta estimativa pode ser obtido de acordo com a Equação 16 pelo quociente entre cada erro de predição pelo respectivo erro padrão da predição, ou seja

$$RMSS = \sqrt{\frac{\left(\frac{\sum_{i=1}^n [\hat{Z}_{(t_i)} - Z_{(t_i)}]}{\hat{\sigma}_{(t_i)}} \right)^2}{n}} . \quad (16)$$

3. MATERIAL E MÉTODOS

Foram simulados 27 cenários espaciais diferentes combinando três tamanhos de grids regulares, três graus de dependência espacial e três diferentes modelos de semivariância. Os tamanhos de grids regulares simulados foram 250x250, 300x300 e 350x350. Já os modelos de semivariância foram o gaussiano, esférico e o modelo exponencial. Quanto aos níveis de dependência espacial, segundo a definição do Índice de Dependência Espacial (IDE) proposta por Zimback (2001), foram utilizados valores de 20% para fraca, 33,33% para moderada e 99,99% para forte.

Portanto, para cada nível de dependência espacial adotado, foram simulados 9 cenários diferentes, como pode se observar na Figura 5, que apresenta esses diferentes cenários.

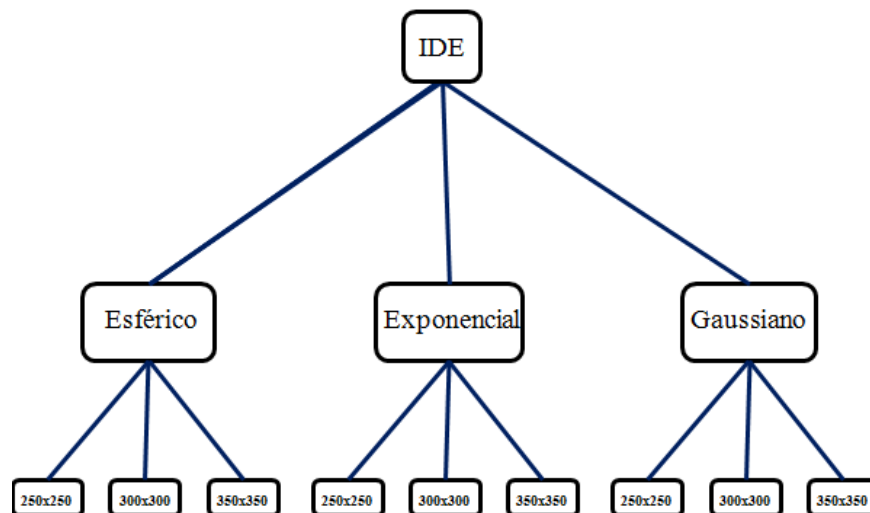


Figura 5: Diferentes Cenários para Cada Índice de Dependência Espacial Simulado.

Para a simulação desses diferentes cenários, previamente, foi necessário escolher determinados parâmetros (alcance, efeito pepita e contribuição), de modo que os valores simulados se comportassem de maneira que representasse cada cenário adotado. Todos cenários foram simulados considerando uma mesma área. O que mudou entre eles foi a quantidade de pontos observados, que foram de 250x250, 300x300 e 350x350.

Para determinar o nível de dependência espacial, o valor da contribuição foi mantido para todos os níveis, sendo alterado somente o valor do efeito pepita para cada um dos índices de dependência espacial diferentes. O valor do alcance de cada cenário foi escolhido de acordo com o proposto por Silva (2006). Ou seja, o alcance do modelo esférico representa um terço, e o do gaussiano, $1/\sqrt{3}$ do alcance teórico do modelo exponencial.

Os valores dos parâmetros (Alcance, Efeito Pepita e a Contribuição) escolhidos para gerar as simulações são apresentados nas Tabelas 1, 2 e 3 para os cenários com índices de dependência espacial Forte, Moderada e Fraca respectivamente.

Tabela 1: Valores dos Parâmetros Alcance, Efeito Pepita e Contribuição Escolhidos para Gerar as Simulações no Caso de Índice de Dependência Espacial Forte.

IDE	Modelo	Alcance	Patamar	Efeito Pepita	Contribuição
Forte	Esférico	23,09	3,000335463	0,000335463	3
Forte	Exponencial	68,28	3,000335463	0,000335463	3
Forte	Gaussiano	40,00	3,000335463	0,000335463	3

Tabela 2: Valores dos Parâmetros Alcance, Efeito Pepita e Contribuição Escolhidos para Gerar as Simulações no Caso de Índice de Dependência Espacial Moderado

IDE	Modelo	Alcance	Patamar	Efeito Pepita	Contribuição
Moderado	Esférico	23,09	6	6	3
Moderado	Exponencial	68,28	6	6	3
Moderado	Gaussiano	40,00	6	6	3

Tabela 3: Valores dos Parâmetros Alcance, Efeito Pepita e Contribuição Escolhidos para Gerar as Simulações no Caso de Índice de Dependência Espacial Fraco.

IDE	Modelo	Alcance	Patamar	Efeito Pepita	Contribuição
Fraco	Esférico	23,09	15	12	3
Fraco	Exponencial	68,28	15	12	3
Fraco	Gaussiano	40,00	15	12	3

Com o objetivo de obter um conjunto de valores que pudesse representar cada cenário, foram gerados inicialmente 150 conjuntos de valores distintos para os pontos que compõem o grid, utilizando o procedimento PROC SIM2D do software SAS (SAS Institute Inc, 1999). Para cada um destes conjuntos, foram calculados os valores das estimativas dos parâmetros alcance e patamar. Para cada cenário foram

selecionados 5 conjuntos de valores, cujas estimativas destes parâmetros mais se aproximavam dos parâmetros do modelo utilizado para a simulação.

Portanto, foram utilizados 135 conjuntos de valores, sendo 5 para cada um dos diferentes 27 cenários simulados. Para cada um desses conjuntos, foi ajustado o modelo de semivariograma correspondente ao modelo teórico utilizado na sua obtenção via a simulação. Para cada cenário foi, então, escolhido o conjunto de dados que permitiu o melhor ajuste do modelo teórico usado para a sua simulação.

O conjunto de dados de cada cenário foi utilizado para gerar as previsões por Krigagem Ordinária. Para cada cenário foi gerado uma previsão para cada um dos números de vizinhos mais próximos (1, 2, 4, 8, 16 e 20). Foram feitas no total 162 krigagens (6 krigagens para cada um dos 27 cenários diferentes).

Um estudo de validação cruzada foi feito para cada uma destas previsões. A partir da validação cruzada foram obtidas estatísticas dos erros de previsão. As estatísticas utilizadas, foram a média dos erros de previsão (ME), média padronizada dos erros de previsão (MS), raiz quadrada do erro quadrático médio do erro de previsão (RMS), erro padrão da média do erro de previsão (ASE) e a raiz quadrada do erro quadrático médio do erro de previsão padronizado (RMSS). A comparação do valor destas estatísticas obtidas para os diferentes cenários permitiram avaliar o efeito que o modelo de dependência espacial, o tamanho do grid de amostragem, o nível de dependência espacial e o número de vizinhos têm na obtenção de previsões mais acuradas ao utilizar Krigagem Ordinária.

4. RESULTADOS E DISCUSSÃO

São apresentados neste capítulo, os valores de algumas estatísticas obtidas a partir da validação cruzada. Estas estatísticas foram utilizadas para avaliar o efeito de alguns parâmetros geoestatísticos sobre a variabilidade da precisão das predições obtidas com a realização da Krigagem Ordinária, para cada um dos cenários simulados.

4.1 Média dos Erros

A Tabela 4 apresenta os valores da média dos erros de predição das Krigagens Ordinárias de tamanhos de grid de 250x250, 300x300 e 350x350 com os índices de dependência espacial forte, moderado e fraco, usando um, dois, quatro, oito, dezesseis e vinte vizinhos mais próximos na krigagem, para os diferentes modelos estudados. Os resultados indicam que todos os três modelos estudados tenderam produzir predições não viesadas, uma vez que os valores de suas respectivas médias dos erros de predição apresentaram valores próximos de zero.

Por conta que esta estatística pode sofrer efeito de escala, foi verificado a média padronizada dos erros de predição. Entretanto, essa mesma tendência de predições não viesadas pode ser verificada quando se analisa a Tabela 5. Esta tabela apresenta os valores da média padronizada dos erros de predição das Krigagens Ordinárias de tamanhos de grid de 250x250, 300x300 e 350x350 com os índices de dependência espacial forte, moderado e fraco, usando um, dois, quatro, oito, dezesseis e vinte vizinhos mais próximos na krigagem, para os diferentes modelos estudados. Pode-se observar que os valores das médias padronizadas dos erros também apresentaram valores próximos de zero. Resultados similares para estas estatísticas foram observados por Oliveira et al. (2011) e em Liao et al. (2006) para os modelos esféricos, exponencial e gaussiano.

Tabela 4: Valores da Média dos Erros de Predição das Krigagens Ordinárias de Tamanhos de grid 250x250, 300x300 e 350x250 com os IDE Forte, Moderado e Fraco, Usando Um, Dois, Quatro, Oito, Dezesesseis e Vinte Vizinhos Mais Próximos na Krigagem, para os Diferentes Modelos Estudados.

IDE	Grid	Modelo	1 Vizinho	2 Vizinhos	4 Vizinhos	8 Vizinhos	16 Vizinhos	20 Vizinhos
Forte	250x250	Esférico	0,0100185	0,0074177	0,0000686	-0,0002150	0,0004612	-0,0001181
		Exponencial	0,0055632	0,0006915	0,0000222	-0,0002547	-0,0000375	0,0000242
		Gaussiano	-0,0049240	-0,0063835	0,0010766	0,0006603	0,0001419	0,0000874
	300x300	Esférico	0,0109956	0,0049134	0,0000004	-0,0001845	-0,0004503	-0,0005137
		Exponencial	-0,0074410	-0,0045817	-0,0001200	-0,0000059	0,0006476	0,0000018
		Gaussiano	0,0045308	0,0010112	-0,0002585	-0,0000714	0,0000467	0,0001022
	350x350	Esférico	-0,0103357	-0,0101529	0,0002378	0,0008613	-0,0000343	0,0001731
		Exponencial	-0,0006109	-0,0024284	0,0000874	0,0001046	-0,0000427	0,0000085
		Gaussiano	0,0007606	0,0000811	0,0001713	-0,0000152	0,0000293	0,0000233
Moderado	250x250	Esférico	0,0499713	0,0037031	0,0044796	0,0015228	0,0008781	0,0005693
		Exponencial	0,0091887	0,0186473	0,0013750	0,0015368	0,0013732	0,0006145
		Gaussiano	-0,0005834	0,0002500	-0,0008251	0,0007594	0,0008641	0,0000981
	300x300	Esférico	0,0027041	0,0298996	0,0035095	0,0018910	0,0020442	0,0009585
		Exponencial	0,0104345	0,0026311	0,0028005	0,0017679	0,0015052	0,0013203
		Gaussiano	-0,0628442	-0,0100006	0,0004212	-0,0012290	-0,0026841	0,0009589
	350x350	Esférico	-0,0221708	-0,0120487	-0,0012418	-0,0006364	-0,0005292	-0,0004958
		Exponencial	0,0113068	-0,0072541	0,0029038	-0,0000156	0,0009884	0,0004582
		Gaussiano	-0,0270458	-0,0029026	0,0013585	0,0007252	-0,0011396	-0,0004556
Fraco	250x250	Esférico	0,0267927	0,0187375	-0,0076555	-0,0060068	-0,0085647	-0,0039074
		Exponencial	-0,1292092	0,0106927	-0,0055924	0,0018144	0,0054787	0,0033792
		Gaussiano	-0,0426782	-0,0958519	-0,0060218	0,0007400	-0,0011655	-0,0015858
	300x300	Esférico	0,0289779	0,0349931	-0,0083856	-0,0055252	-0,0030679	-0,0082981
		Exponencial	-0,0729209	-0,0050782	0,0052523	0,0016410	-0,0014091	0,0083661
		Gaussiano	0,0787392	0,0216427	-0,0071607	0,0006721	0,0040349	0,0008645
	350x350	Esférico	-0,0381092	-0,0263267	-0,0030432	-0,0020991	-0,0025892	-0,0025713
		Exponencial	-0,0798799	-0,0027604	0,0046868	0,0031319	0,0057177	0,0052905
		Gaussiano	-0,0273194	0,0098699	-0,0055077	-0,0035492	0,0029168	-0,0005692

Tabela 5: Valores da Média Padronizada dos Erros de Predição das Krigagens Ordinárias de Tamanhos de grid 250x250, 300x300 e 350x250 com os IDE Forte, Moderado e Fraco, Usando Um, Dois, Quatro, Oito, Dezesesseis e Vinte Vizinhos Mais Próximos na Krigagem, para os Diferentes Modelos Estudados.

IDE	Grid	Modelo	1 Vizinho	2 Vizinhos	4 Vizinhos	8 Vizinhos	16 Vizinhos	20 Vizinhos
Forte	250x250	Esférico	0,007152	0,006950	0,000017	-0,000278	0,000520	-0,000139
		Exponencial	0,005793	0,000865	-0,000107	0,000514	-0,000159	-0,000058
		Gaussiano	-0,010492	-0,017895	0,014506	0,006373	-0,000917	-0,002118
	300x300	Esférico	0,007920	0,005158	0,000161	-0,000067	-0,000374	-0,000434
		Exponencial	-0,007657	-0,006211	-0,000192	-0,000020	0,001008	-0,000009
		Gaussiano	0,009571	0,000068	-0,002414	-0,001138	0,000667	0,001620
	350x350	Esférico	-0,007475	-0,010277	0,000299	0,000943	-0,000051	0,000189
		Exponencial	0,000623	-0,003357	0,000113	0,000136	-0,000080	-0,000002
		Gaussiano	0,001599	0,001415	0,001550	-0,000240	0,000385	0,000329
Moderado	250x250	Esférico	0,017754	0,001473	0,001934	0,000553	0,000258	0,000116
		Exponencial	0,003799	0,009006	0,000693	0,000777	0,000672	0,000244
		Gaussiano	-0,000247	0,000116	-0,000463	0,000403	0,000449	-0,000092
	300x300	Esférico	0,000955	0,012480	0,001555	0,000787	0,000853	0,000331
		Exponencial	0,004290	0,001173	0,001420	0,000856	0,000694	0,000587
		Gaussiano	-0,026558	-0,004970	0,000225	-0,000697	-0,001528	0,000597
	350x350	Esférico	-0,007892	-0,004941	-0,000550	-0,000261	-0,000203	-0,000183
		Exponencial	0,004716	-0,003550	0,001534	-0,000042	0,000523	0,000221
		Gaussiano	-0,011413	-0,001395	0,000712	0,000371	-0,000722	-0,000327
Fraco	250x250	Esférico	0,005225	0,004230	-0,001877	-0,001507	-0,002185	-0,000959
		Exponencial	-0,026246	0,002531	-0,001433	0,000508	0,001519	0,000934
		Gaussiano	-0,008589	-0,022343	-0,001537	0,000202	-0,000316	-0,000429
	300x300	Esférico	0,005668	0,007981	-0,002060	-0,001393	-0,000721	-0,002119
		Exponencial	-0,014781	-0,001204	0,001342	0,000422	-0,000435	0,002276
		Gaussiano	0,016163	0,005126	-0,001856	0,000197	0,001152	0,000257
	350x350	Esférico	-0,007496	-0,005991	-0,000752	-0,000528	-0,000656	-0,000651
		Exponencial	-0,016111	-0,000667	0,001191	0,000822	0,001539	0,001423
		Gaussiano	-0,005612	0,002375	-0,001434	-0,000966	0,000853	-0,000136

4.2 Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS)

Na literatura geoestatística, alguns autores, usam a estatística RMS para selecionar o modelo de semivariograma, como é o caso de Oliveira et al., 2011, Montomya et al. (2011), Jamur e Freitas (2009), Lopes (2009), Silva, (2006), Liao et al. (2006) e Cerri (2005) ou para selecionar o método de interpolação (ALMEIDA et al., 2011 ; MIRANDA et al., 2009).

A Tabela 6 apresenta os valores da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) das Krigagens Ordinárias de tamanhos de grid de 250x250, 300x300 e 350x350 com os índices de dependência espacial forte, moderado e fraco, usando um, dois, quatro, oito, dezesseis e vinte vizinhos mais próximos na krigagem, para os diferentes modelos estudados. Quando se compara o valor das estatísticas RMS entre os três modelos de semivariograma, pode-se observar que para a maioria dos cenários simulados, o modelo esférico foi o que tendeu a apresentar o maior valor para aquela estatística, seguido pelos modelos exponencial e gaussiano.

Portanto, o modelo gaussiano tendeu a apresentar menor variabilidade do erro de predição entre o valor predito e o verdadeiro valor na maioria dos cenários estudados, diferentemente do estudo de Oliveira et al. (2011), uma vez que para este, o esférico foi o que apresentou o menor valor da estatística RMS, enquanto que o modelo exponencial foi o que apresentou maior valor.

Porém, quando se diminuiu a porcentagem da contribuição em relação ao patamar, aumentando o efeito pepita, ou seja, para cenários com IDE-Fraco, esta tendência não se manteve para alguns cenários simulados, como pode ser observado nas células hachuradas da Tabela 6. Por exemplo, no IDE Fraco com grid 250x250, quem tendeu a apresentar o menor valor de RMS, foi o exponencial usando 1 e 2 vizinhos, e o esférico usando 4 e 8 vizinhos. Com o tamanho de grid de 300x300, o esférico e o exponencial tenderam a apresentar os menores valores da RMS usando 4 vizinhos, respectivamente.

Já o modelo gaussiano tendeu a apresentar o maior valor usando 2, 4 e 8 vizinhos para o tamanho de grid de 250x250, e usando 4 vizinhos com 300x300. O modelo exponencial usando 1 e 2 vizinhos no 250x250, e com 1 vizinho no grid de 350x350 também tendeu a apresentar a maior variabilidade do erro entre o valor predito e o verdadeiro valor.

Tabela 6: Valores da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) das Krigagens Ordinárias de Tamanhos de grid 250x250, 300x300 e 350x250 com os IDE Forte, Moderado e Fraco, Usando Um, Dois, Quatro, Oito, Dezesesseis e Vinte Vizinhos Mais Próximos na Krigagem, para os Diferentes Modelos Estudados.

IDE	Grid	Modelo	1 Vizinho	2 Vizinhos	4 Vizinhos	8 Vizinhos	16 Vizinhos	20 Vizinhos
Forte	250x250	Esférico	1,4171690	1,1175820	0,8885174	0,8878210	0,8670692	0,8656379
		Exponencial	0,6331757	0,5058577	0,4119212	0,4116968	0,4080252	0,4075215
		Gaussiano	0,3298200	0,2500278	0,0575237	0,0360922	0,0342540	0,0344980
	300x300	Esférico	1,3728170	1,0866970	0,8781085	0,8763319	0,8640053	0,8633588
		Exponencial	0,6303080	0,5033651	0,4071564	0,4069558	0,4018343	0,4013011
		Gaussiano	0,3857286	0,2307392	0,0442935	0,0065098	0,0061446	0,0063302
	350x350	Esférico	1,3630880	1,0880760	0,8912525	0,8915817	0,8726446	0,8712619
		Exponencial	0,6406195	0,5097357	0,4088306	0,4086070	0,4041763	0,4033369
		Gaussiano	0,3594731	0,2251185	0,0425703	0,0061018	0,0055068	0,0057751
Moderado	250x250	Esférico	2,8047240	2,3668700	2,1304700	2,0635040	2,0559650	2,0545100
		Exponencial	2,5243880	2,2082160	2,0187500	1,9153620	1,8925740	1,8883590
		Gaussiano	2,4542600	2,1376620	1,9674210	1,8646020	1,8157470	1,8022990
	300x300	Esférico	2,8158860	2,4145390	2,1614880	2,0779250	2,0693010	2,0681270
		Exponencial	2,6037390	2,2358630	2,0280690	1,9287930	1,9005840	1,8964620
		Gaussiano	2,5192630	2,1804960	1,9757460	1,8653840	1,7991740	1,7909340
	350x350	Esférico	2,7769950	2,3762580	2,1538730	2,0683390	2,0605870	2,0593000
		Exponencial	2,5923000	2,2429190	2,0289680	1,9216160	1,8872050	1,8827220
		Gaussiano	2,5526640	2,2200810	1,9944160	1,8711630	1,8231270	1,8146890
Fraco	250x250	Esférico	5,1751050	4,3727150	3,9407870	3,7743000	3,7445040	3,7389200
		Exponencial	5,0727080	4,3641010	3,9771120	3,7830610	3,7040980	3,6650590
		Gaussiano	5,1414420	4,4455860	4,0284040	3,7882990	3,6945490	3,6611710
	300x300	Esférico	5,0111000	4,3486970	3,9683000	3,8218710	3,7482350	3,7381140
		Exponencial	5,0696080	4,3765020	3,9378360	3,7666070	3,6678040	3,6474450
		Gaussiano	5,0082480	4,3579360	4,0054230	3,7334730	3,6122330	3,5950850
	350x350	Esférico	5,0244980	4,3671960	4,0128870	3,7987330	3,7335590	3,7174480
		Exponencial	5,0464090	4,3262040	3,9584720	3,7582360	3,6529950	3,6407380
		Gaussiano	4,9726050	4,2864300	3,9082930	3,6838460	3,5964200	3,5779330

Os valores apresentados na Tabela 6 foram utilizados para construir as figuras apresentadas a seguir. Essas figuras têm por objetivo descrever o comportamento da estatística RMS em função de algumas combinações relevantes dos parâmetros (IDE, número de vizinhos, modelo de semivariograma e grid) usados para a simulação dos cenários.

A Figura 6 apresenta, para cada um dos modelos, o valor médio da RMS para os diferentes tamanhos de grid diferentes em função do número de vizinhos quando foi usado o IDE forte para a simulação. Por outro lado, as Figuras 7 e 8 apresentam estas combinações de parâmetros para os cenários com IDE moderado e fraco, respectivamente. Nestas combinações é possível observar que o modelo esférico foi o que tendeu a apresentar a maior variabilidade do erro entre o valor predito e o verdadeiro valor, enquanto que o gaussiano tendeu a apresentar o menor valor da RMS para todos os índices de dependência espacial.

Embora o modelo gaussiano tenha apresentado o melhor ajuste por meio da estatística RMS, alguns autores como (Corá et al., 2004; Montanari et al., 2008; Motomyia et al., 2011) obtiveram melhores ajustes usando o modelo esférico e o exponencial. Outro fator que pode explicar esse comportamento, é o nível de dispersão que as variáveis se encontravam. No estudo de Motomyia (2011) por exemplo, foram utilizados variáveis de diversas propriedades do solo cujos coeficientes de variação eram baixos e moderadas, assim como no trabalho de Corá et al. (2004).

É possível observar nas Figuras 6, 7 e 8, que a partir do momento em que se aumenta o número de vizinhos utilizado na krigagem para os três tipos de IDE, o valor da RMS tende também a ser menor. Observa-se ainda que o valor da RMS tende a se estabilizar com números diferentes de vizinhos mais próximos de acordo com o índice de dependência espacial.

Por exemplo, com o IDE forte, o valor da estatística RMS tende a se estabilizar no menor valor com 4 vizinhos. Com IDE moderado, a partir de 8 vizinhos, e com mais de 16 vizinhos, o valor da RMS tende se estabilizar no menor valor com o IDE fraco. Ou seja, quanto menor a dependência espacial do cenário, mais vizinhos são necessários para se obter o menor valor possível para a estatística RMS.

Por conta desta tendência de estabilidade do valor da estatística RMS pode-se indicar quantos vizinhos utilizar na Krigagem de acordo com o IDE apresentado pelo cenário.

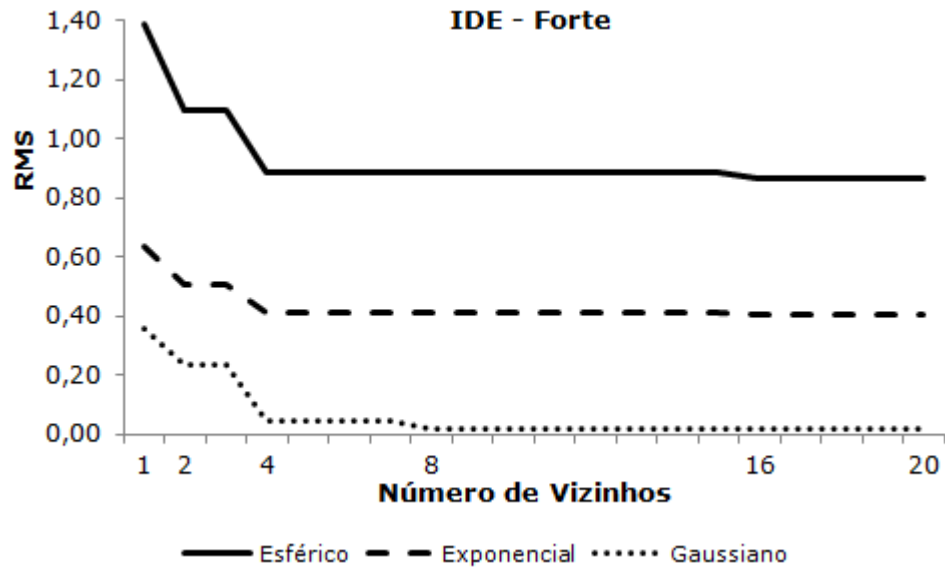


Figura 6: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) dos Tamanhos de Grid Diferentes, das Krigagens Ordinárias em Função dos Diferentes Modelos Estudados e Número de Vizinhos, para o IDE Forte.

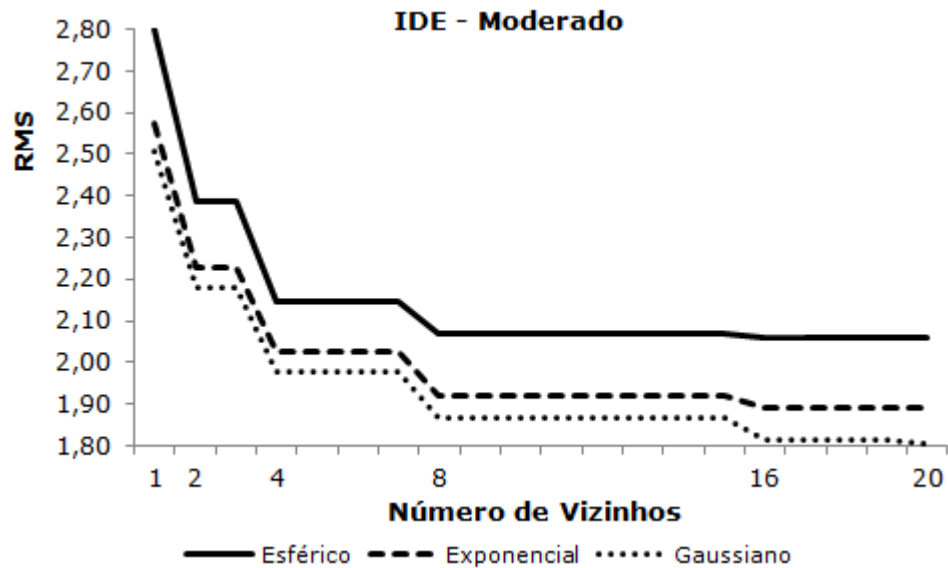


Figura 7: Valores Médio da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) dos Tamanhos de Grid Diferentes, das Krigagens Ordinárias em Função dos Diferentes Modelos Estudados e Número de Vizinhos, para o IDE Moderado.

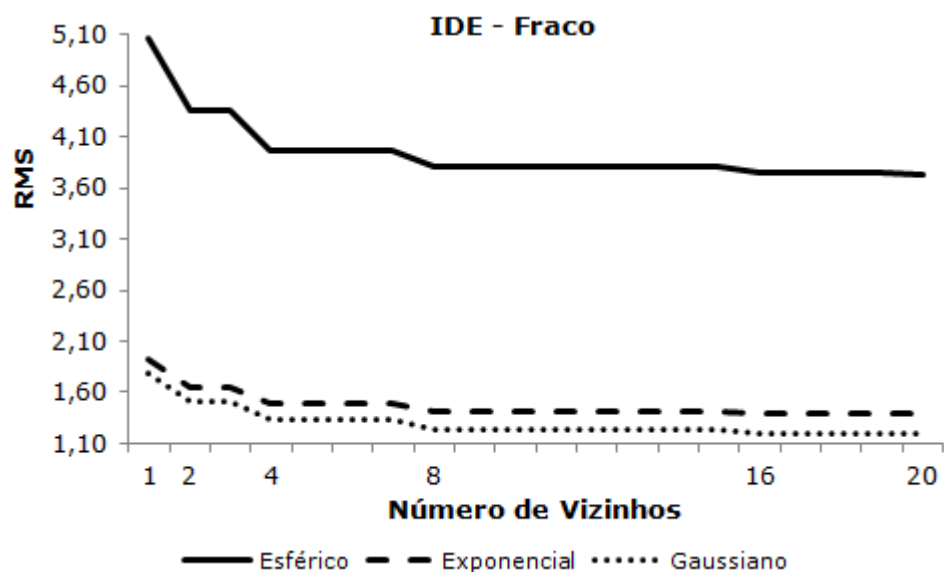


Figura 8: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) dos Tamanhos de Grid Diferentes, das Krigagens Ordinárias em Função dos Diferentes Modelos Estudados e Número de Vizinhos, para o IDE Fraco.

Além do número de vizinhos, dos modelos de dependência espacial e do IDE, outros parâmetros podem influenciar no valor da estatística RMS. Soares, et al. (2011) por exemplo, analisou o efeito do aumento da escala de investigação dos perfis de raios gama de doze poços utilizando três diferentes espaçamentos (5, 20 e 50 metros) entre pontos.

Eles observaram que quanto maior o espaçamento entre os poços estudados, maior foi o valor do RMS das Krigagens. Estudo parecido foi feito por Cerri (2005), que estudou propriedades físicas e químicas do solo em três níveis de profundidades no solo.

Porto et al. (2011) avaliaram se o valor da estatística RMS se alterava com a redução (100%, 75%, 50% e 25% dos dados) do número de amostras. A otimização da malha foi julgada, além de outros parâmetros, por meio da RMS entre cada mapa krigado com redução de dados em relação ao respectivo mapa krigado com 100% dos dados. Os autores concluíram que com a redução do número de amostras, o valor da estatística RMS tende a aumentar. Porém, esses parâmetros de redução no número de amostras e diferentes níveis de espaçamento dos pontos amostrados não serão abordados no trabalho proposto.

A Figura 9 apresenta o valor médio da estatística RMS dos diferentes modelos de dependência espacial e dos diferentes níveis de dependência espacial, em função do número de vizinhos para cada um dos tamanhos de grid usados na simulação dos cenários. Pode-se verificar que, para todos os grids, a variabilidade dos erros de predição tende a diminuir a medida que se aumenta número de vizinhos.

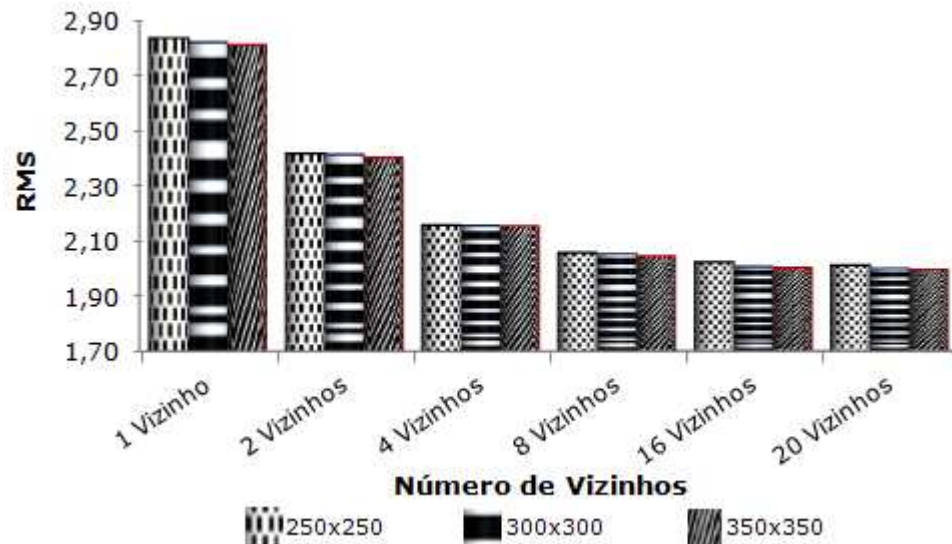


Figura 9: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) dos Diferentes Modelos de Dependência Espacial e dos Diferentes Níveis de Dependência Espacial, das Krigagens Ordinárias em Função dos Tamanhos de Grid e Número de Vizinhos Mais Próximos na Krigagem

A Figura 10 apresenta o valor médio da estatística RMS dos diferentes modelos de dependência espacial e para os diferentes tamanhos de grid amostral, em função do número de vizinhos, para os diferentes índices de dependência espacial. Nota-se que, ao usar a mesma escala para todos IDEs, pode-se perceber que a variabilidade dos erros das predições das krigagens tende a diminuir a medida que aumenta o número de vizinhos mais próximos na krigagem. Percebe-se também que quando se aumenta o percentual do IDE, o valor da estatística RMS também tende a diminuir.

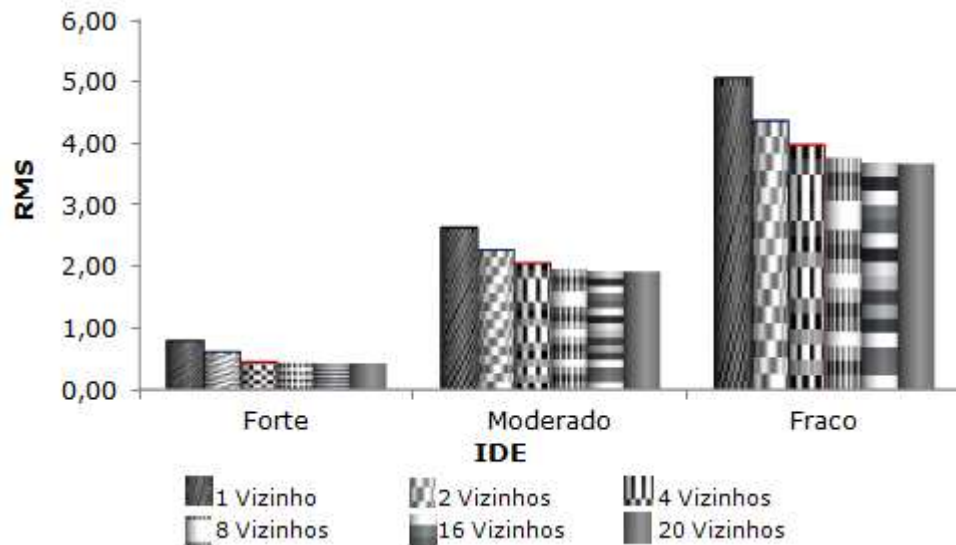


Figura 10: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) dos Diferentes Modelos de Dependência Espacial e para os Diferentes Tamanhos de Grid Amostral, das Krigagens Ordinárias em Função dos Diferentes Números de Vizinhos Mais Próximos, Para os Diferentes Índices de Dependência Espacial.

4.3 Erro Padrão da Média (ASE)

A estatística ASE é usada junto com a RMS para validar alguns modelos geoestatísticos (SILVA, 2006; OLIVEIRA et al., 2011), portanto saber como obter o menor valor do ASE e mais próximos da estatística RMS, pode ajudar na escolha do modelo geoestatístico.

A Tabela 7 apresenta os valores do erro padrão da média do erro de predição (ASE) das Krigagens Ordinárias de tamanhos de grid de 250x250, 300x300 e 350x350 com os índices de dependência espacial forte, moderado e fraco, usando um, dois, quatro, oito, dezesseis e vinte vizinhos mais próximos na Krigagem, para os diferentes modelos estudados. Nota-se que os modelos estudados se comportaram de maneira diferente de acordo com o índice de dependência espacial.

Para a maioria dos cenários usados, o modelo esférico tendeu a apresentar maiores valores do ASE, seguido do exponencial e gaussiano, porém, no IDE fraco com tamanho de 250x250, o modelo exponencial tendeu a apresentar o menor valor para ASE, como pode ser observado nas células hachuradas da Tabela 7. (Oliveira et

al. 2011), em um estudo onde foi utilizado a estatística ASE, também obtiveram para o modelo esférico, o menor valor de ASE, dentre todos os modelos estudados.

Tabela 7: Valores do Erro Padrão da Média (ASE) das Krigagens Ordinárias de Tamanhos 250x250, 300x300 e 350x250 com os IDE Forte, Moderado e Fraco, Usando Um, Dois, Quatro, Oito, Dezesesseis e Vinte Vizinhos Mais Próximos na Krigagem, para os Diferentes Modelos Estudados.

IDE	Grid	Modelo	1 Vizinho	2 Vizinhos	4 Vizinhos	8 Vizinhos	16 Vizinhos	20 Vizinhos
Forte	250x250	Esférico	1,400783	1,129735	0,910966	0,910349	0,896202	0,894935
		Exponencial	0,960419	0,779218	0,637610	0,637452	0,631211	0,630662
		Gaussiano	0,469332	0,331907	0,112389	0,072184	0,058657	0,057674
	300x300	Esférico	1,383410	1,093265	0,894220	0,893398	0,878536	0,877311
		Exponencial	0,971748	0,775820	0,644562	0,644427	0,638019	0,637463
		Gaussiano	0,473392	0,307912	0,113321	0,072819	0,058965	0,058008
	350x350	Esférico	1,382761	1,090600	0,893203	0,892434	0,877443	0,876209
		Exponencial	0,980129	0,781191	0,649712	0,649594	0,643069	0,642501
		Gaussiano	0,475629	0,306304	0,113831	0,073175	0,059123	0,058166
Moderado	250x250	Esférico	2,814639	2,397195	2,136135	2,067068	2,059034	2,058420
		Exponencial	2,418710	2,067508	1,853075	1,786677	1,772995	1,771210
		Gaussiano	2,359762	2,022060	1,822868	1,734736	1,705150	1,700446
	300x300	Esférico	2,830218	2,401760	2,148966	2,078562	2,069999	2,069331
		Exponencial	2,432376	2,074114	1,865220	1,797269	1,782627	1,780681
		Gaussiano	2,366302	2,027000	1,829424	1,740597	1,709965	1,705077
	350x350	Esférico	2,809284	2,381930	2,130633	2,061479	2,053415	2,052808
		Exponencial	2,397373	2,043773	1,838277	1,771107	1,756541	1,754585
		Gaussiano	2,369693	2,029632	1,832400	1,743254	1,712200	1,707184
Fraco	250x250	Esférico	5,127933	4,419563	4,007588	3,825000	3,758435	3,747333
		Exponencial	4,922930	4,250254	3,863728	3,680172	3,601610	3,586890
		Gaussiano	4,968800	4,294981	3,908455	3,710519	3,617655	3,599585
	300x300	Esférico	5,112380	4,401444	3,995568	3,813087	3,746079	3,734865
		Exponencial	4,933434	4,256056	3,871244	3,687468	3,608896	3,594211
		Gaussiano	4,871511	4,208183	3,831640	3,637516	3,546154	3,528411
	350x350	Esférico	5,083615	4,374283	3,969489	3,789639	3,725215	3,714533
		Exponencial	4,957970	4,278686	3,893425	3,706958	3,625462	3,610033
		Gaussiano	4,867655	4,204464	3,828299	3,634302	3,542941	3,525120

Para avaliar se a variabilidade das predições foi avaliada corretamente foi obtida a diferença dos valores das estatísticas RMS e ASE (Tabela 8) para todos os cenários simulados.

Pode-se perceber com a Figura 11, que apresenta a média da diferença entre a estatística RMS e o ASE dos diferentes tamanhos de grid amostral e dos diferentes números de vizinhos, em função dos diferentes índices de dependência espacial, para os diferentes modelos estudados, que com o modelo esférico, seguido do gaussiano, tendeu a apresentar valores mais próximos do desejável, ou seja o valor do ASE mais próximo da RMS, exceto para o IDE fraco, uma vez que o gaussiano tendeu a apresentar a pior avaliação da variabilidade da estimativa do erro de predição.

Em Oliveira et al. 2011, o modelo gaussiano que apresentava o maior efeito pepita entre os três modelos estudados (IDE fraco), também obteve o valor de ASE mais distante da estatística RMS.

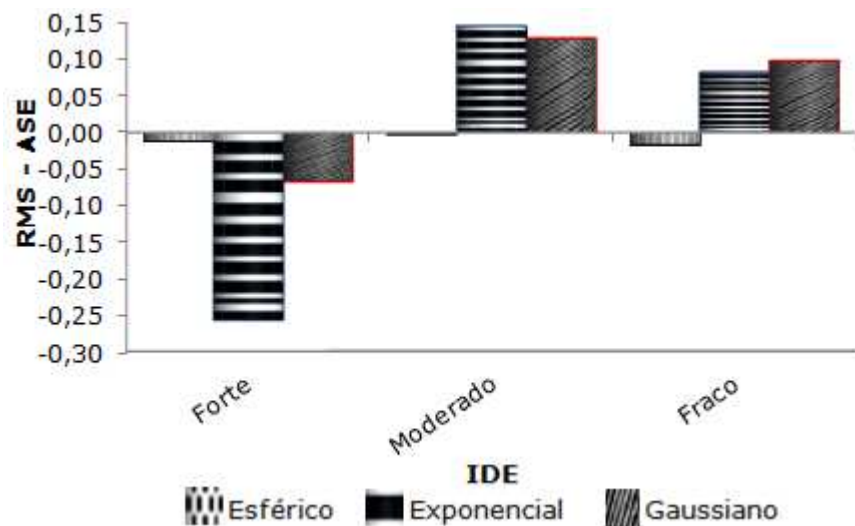


Figura 11: Valores Médios da Diferença entre os Valores da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) e o Erro Padrão da Média (ASE), dos Diferentes Tamanhos de Grid Amostral e dos Diferentes Números de Vizinhos, das Krigagens Ordinárias dos Diferentes Índices de Dependência Espacial, para os Modelos Estudados.

Porém, quando se aumenta a quantidade de vizinhos, verifica-se pela tabela 8, que o comportamento dos modelos se alterou. Por exemplo, no IDE Forte com grid 250x250 usando 16 e 20 vizinhos, o exponencial tendeu a apresentar o valor do ASE mais próximo do RMS, enquanto que no IDE Moderado de grid 250x250 usando 8 vizinhos, o gaussiano apresentou a maior distância entre os valores.

Para o IDE fraco, o modelo gaussiano foi o que tendeu a apresentar o valor do ASE mais distante da RMS para a maioria dos cenários estudados. Mas esse

comportamento se modifica quando se usa um grid de 250x250 usando 8, 16 e 20 vizinhos, e com grid de 300x300 usando 8 vizinhos, pois o modelo exponencial é quem possui o ASE mais distante da RMS. Com base nisso, diminuindo o IDE, o modelo exponencial tende a deixar de ser o modelo com a pior avaliação sobre a variabilidade das predições.

Tabela 8: Valores da Diferença entre a Valores da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição (RMS) e o Erro Padrão da Média (ASE) das Krigagens Ordinárias de Tamanhos 250x250, 300x300 e 350x250 com os IDE Forte, Moderado e Fraco, Usando Um, Dois, Quatro, Oito, Dezesesseis e Vinte Vizinhos Mais Próximos na Krigagem, para os Diferentes Modelos Estudados.

IDE	Grid	Modelo	1 Vizinho	2 Vizinhos	4 Vizinhos	8 Vizinhos	16 Vizinhos	20 Vizinhos
Forte	250x250	Esférico	0,016386	-0,012153	-0,022448	-0,022528	-0,029133	-0,029298
		Exponencial	-0,327244	-0,273360	-0,225689	-0,225756	-0,223185	-0,223141
		Gaussiano	-0,139512	-0,081879	-0,054865	-0,036092	-0,024403	-0,023176
	300x300	Esférico	-0,010593	-0,006568	-0,016112	-0,017067	-0,014531	-0,013952
		Exponencial	-0,341440	-0,272455	-0,237405	-0,237471	-0,236185	-0,236161
		Gaussiano	-0,087663	-0,077173	-0,069028	-0,066309	-0,052820	-0,051677
	350x350	Esférico	-0,019673	-0,002524	-0,001951	-0,000852	-0,004798	-0,004947
		Exponencial	-0,339510	-0,271455	-0,240881	-0,240987	-0,238893	-0,239164
		Gaussiano	-0,116156	-0,081185	-0,071260	-0,067073	-0,053617	-0,052391
Moderado	250x250	Esférico	-0,009915	-0,030325	-0,005665	-0,003564	-0,003069	-0,003910
		Exponencial	0,105678	0,140708	0,165675	0,128685	0,119579	0,117149
		Gaussiano	0,094498	0,115602	0,144553	0,129866	0,110597	0,101853
	300x300	Esférico	-0,014332	0,012779	0,012522	-0,000637	-0,000698	-0,001204
		Exponencial	0,171363	0,161749	0,162849	0,131524	0,117957	0,115781
		Gaussiano	0,152961	0,153496	0,146322	0,124787	0,089209	0,085857
	350x350	Esférico	-0,032289	-0,005672	0,023240	0,006860	0,007172	0,006492
		Exponencial	0,194927	0,199146	0,190691	0,150509	0,130664	0,128137
		Gaussiano	0,182971	0,190449	0,162016	0,127909	0,110927	0,107505
Fraco	250x250	Esférico	0,047172	-0,046848	-0,066801	-0,050700	-0,013931	-0,008413
		Exponencial	0,149778	0,113847	0,113384	0,102889	0,102488	0,078169
		Gaussiano	0,172642	0,150605	0,119949	0,077780	0,076894	0,061586
	300x300	Esférico	-0,101280	-0,052747	-0,027268	0,008784	0,002156	0,003249
		Exponencial	0,136174	0,120446	0,066592	0,079139	0,058908	0,053234
		Gaussiano	0,136737	0,149753	0,173783	0,095957	0,066079	0,066674
	350x350	Esférico	-0,059117	-0,007087	0,043398	0,009094	0,008344	0,002915
		Exponencial	0,088439	0,047518	0,065047	0,051278	0,027533	0,030705
		Gaussiano	0,104950	0,081966	0,079994	0,049544	0,053479	0,052813

4.4 Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS)

A raiz quadrada do erro quadrático médio do erro de predição padronizado (RMSS) foi usada para fazer uma avaliação das estimativas obtidas sobre a variabilidade dos erros das predições, e assim como a estatística RMS e ASE, é usado com frequência na literatura, como por exemplo em Sarangi e Madramootoo (2005), Liao et al. (2006), Silva (2006), Tatalovich et al., (2006), Oliveira et al. (2011) e Mutua (2012).

A Tabela 9 apresenta os valores de RMSS das Krigagens Ordinárias de tamanhos de grid de 250x250, 300x300 e 350x350 com os índices de dependência espacial forte, moderado e fraco, usando um, dois, quatro, oito, dezesseis e vinte vizinhos mais próximos na krigagem, para os diferentes modelos estudados. Quando se avalia o valor do RMSS por meio da Tabela 9, verifica-se que os valores dos modelos estudados tendem a se comportar de maneira diferente de acordo com o índice de dependência espacial, porém, o modelo esférico foi o que tendeu a apresentar melhor ajuste em todos os cenários estudados, ou seja, valores próximos do valor um.

Esse resultado difere dos resultados obtidos por Oliveira et al. (2011), uma vez que o modelo exponencial foi o que apresentou valor da estatística RMSS mais próximo do valor um, e para Liao et al. (2006), cujo menor valor foi para o modelo gaussiano.

Tabela 9: Valores da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) das Krigagens Ordinárias de Tamanhos 250x250, 300x300 e 350x250 com os IDE Forte, Moderado e Fraco, Usando Um, Dois, Quatro, Oito, Dezesesseis e Vinte Vizinhos Mais Próximos na Krigagem, para os Diferentes Modelos Estudados.

IDE	Grid	Modelo	1 Vizinho	2 Vizinhos	4 Vizinhos	8 Vizinhos	16 Vizinhos	20 Vizinhos
Forte	250x250	Esférico	1,011698	0,988902	0,975645	0,975543	0,967674	0,967400
		Exponencial	0,659270	0,649029	0,645738	0,645587	0,646103	0,645862
		Gaussiano	0,702743	0,744378	0,480220	0,384394	0,375033	0,386378
	300x300	Esférico	0,992343	0,993422	0,981684	0,980558	0,983014	0,983618
		Exponencial	0,648633	0,647866	0,632031	0,631852	0,630171	0,629895
		Gaussiano	0,814819	0,701786	0,390930	0,088325	0,100684	0,106975
	350x350	Esférico	0,985773	0,998311	0,997825	0,999053	0,994506	0,994328
		Exponencial	0,653607	0,651614	0,629264	0,629028	0,628550	0,627798
		Gaussiano	0,755785	0,680053	0,374128	0,083091	0,092734	0,099130
Moderado	250x250	Esférico	0,996477	0,987438	0,996900	0,997917	0,998137	0,997750
		Exponencial	1,043692	1,068028	1,087866	1,072241	1,067616	1,066308
		Gaussiano	1,040045	1,055145	1,079317	1,074843	1,064886	1,059909
	300x300	Esférico	0,994936	1,005534	1,005887	0,999697	0,999703	0,999455
		Exponencial	1,070451	1,077892	1,087162	1,072995	1,065970	1,064827
		Gaussiano	1,064641	1,075919	1,079952	1,071706	1,052185	1,050354
	350x350	Esférico	0,988506	0,997717	1,010786	1,003193	1,003356	1,003028
		Exponencial	1,081309	1,097415	1,103798	1,085073	1,074476	1,073110
		Gaussiano	1,077213	1,093939	1,088446	1,073355	1,064748	1,062916
Fraco	250x250	Esférico	1,009199	0,989393	0,983440	0,986882	0,996477	0,997955
		Exponencial	1,030425	1,026817	1,029368	1,027956	1,028467	1,021796
		Gaussiano	1,034745	1,035094	1,030709	1,020981	1,021289	1,017162
	300x300	Esférico	0,980189	0,987857	0,993111	1,002168	1,000457	1,000752
		Exponencial	1,027602	1,028262	1,017244	1,021532	1,016421	1,014927
		Gaussiano	1,028069	1,035586	1,045375	1,026402	1,018659	1,018916
	350x350	Esférico	0,988371	0,998421	1,010897	1,002353	1,002191	1,000735
		Exponencial	1,017838	1,011117	1,016723	1,013852	1,007633	1,008547
		Gaussiano	1,021561	1,019491	1,020891	1,013631	1,015119	1,015026

Os valores apresentados na Tabela 9 foram utilizados para construir as Figuras de 12 a 17. Essas figuras têm como objetivo descrever o comportamento da estatística RMSS em função de algumas combinações relevantes dos parâmetros (IDE, número de vizinhos e modelo de semivariograma) usados para a simulação dos cenários.

A Figura 12 apresenta, para cada um dos modelos, o valor médio da estatística RMSS dos tamanhos de grid diferentes, em função do número de vizinhos quando foi usado o IDE forte para a simulação. Enquanto que as Figuras 13 e 14 apresentam estas combinações de parâmetros para os cenários com IDE moderado e fraco, respectivamente.

Com o IDE Forte e Fraco, é possível observar que o modelo esférico foi o que tendeu a apresentar resultados mais próximos do desejável usando todas as quantidades de vizinhos mais próximos. Com 1 e 2 vizinhos, o exponencial tendeu a ser o modelo com a pior avaliação sobre a variabilidade dos erros das predições, porém, quando se aumentou o número de vizinhos mais próximos, o modelo gaussiano foi o que tendeu a apresentar valores mais distantes do desejável.

Já com o IDE moderado, verificou-se que o modelo esférico foi o que tendeu a apresentar valor mais próximo do desejável, seguido do gaussiano e exponencial para o IDE moderado, para todas as quantidades de vizinhos.

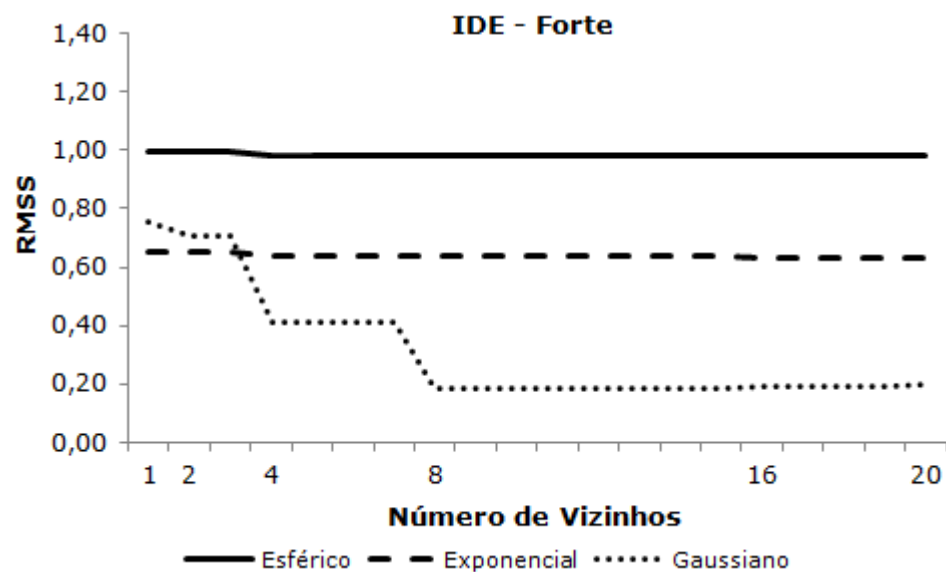


Figura 12: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) dos Tamanhos de Grid Diferentes, das Krigagens Ordinárias em Função dos Diferentes Modelos Estudados e Número de Vizinhos, para o IDE Forte.

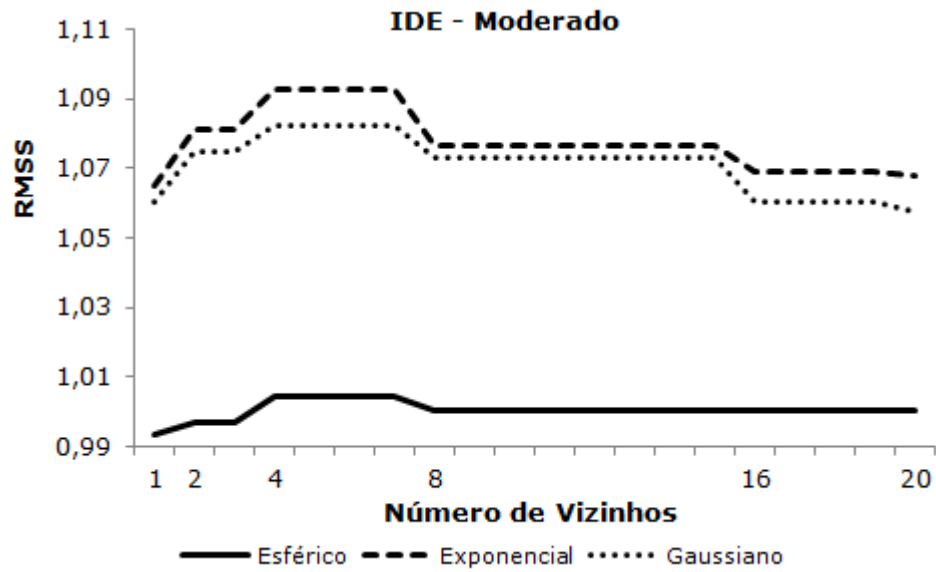


Figura 13: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) dos Tamanhos de Grid Diferentes, das Krigagens Ordinárias em Função dos Diferentes Modelos Estudados e Número de Vizinhos, para o IDE Moderado.

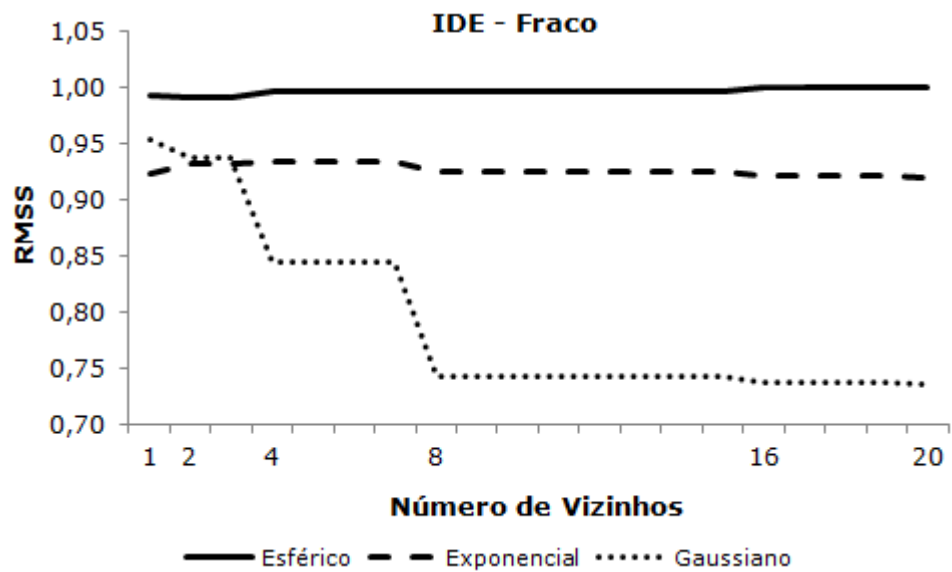


Figura 14: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) dos Tamanhos de Grid Diferentes, das Krigagens Ordinárias em Função dos Diferentes Modelos Estudados e Número de Vizinhos, para o IDE Fraco.

A Figura 15 apresenta a média da estatística RMSS dos diferentes modelos de dependência espacial e níveis de dependência espacial, em função do número de vizinhos para cada um dos tamanhos de grid usados na simulação dos cenários.

Pode-se verificar que, para todos os grids, quando se aumentou a quantidade de vizinhos, a avaliação sobre a variabilidade dos erros das previsões tendeu a piorar, ou seja, se afastando do valor um, sendo que a partir de 8 vizinhos, o valor da estatística RMSS tendeu a se estabilizar para todos os tamanhos de grid.

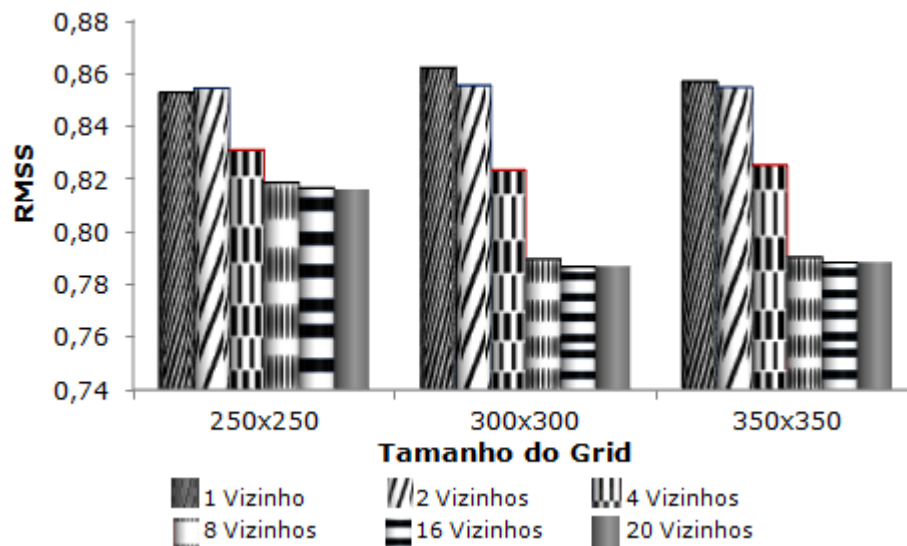


Figura 15: Valores Médios da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) dos Diferentes Modelos de Dependência Espacial e Níveis de Dependência Espacial, das Krigagens Ordinárias em Função dos Diferentes Tamanhos de Grid e Número de Vizinhos.

As Figuras 16 e 17 apresentam a média da estatística RMSS em função do diferentes índices de dependência espacial, para os diferentes números de vizinhos, e modelos estudados, respectivamente. Nota-se pela Figura 16 que, para o IDE Forte, a variabilidade dos erros entre o valor predito e o verdadeiro valor das previsões das Krigagens tende a ter piores avaliações a medida que aumenta o número de vizinhos mais próximos na Krigagem.

Quanto ao modelo estudado, verifica-se pela Figura 17, que o esférico foi o que tendeu a apresentar valores de RMSS mais próximos de um, ao contrário do IDE Moderado e Fraco, onde a avaliação do mesmo tendeu em ser constante próximo do valor desejável de 1 para todos os números de vizinhos e para todos os modelos estudados.

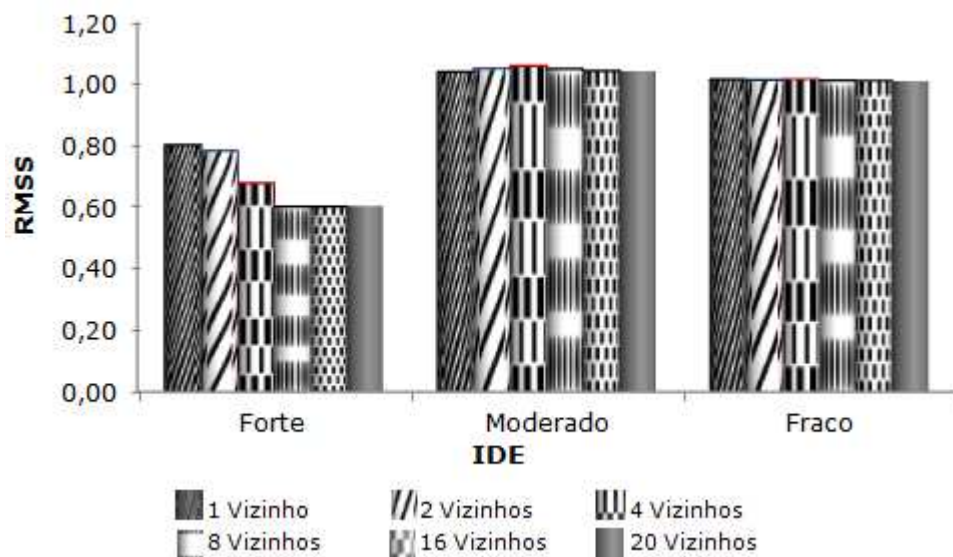


Figura 16: Valores da Média da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) dos Diferentes Modelos de Dependência Espacial e Tamanhos de Grid de Amostragem, das Krigagens Ordinárias em Função dos Diferentes Índices de Dependência Espacial e Número de Vizinhos.

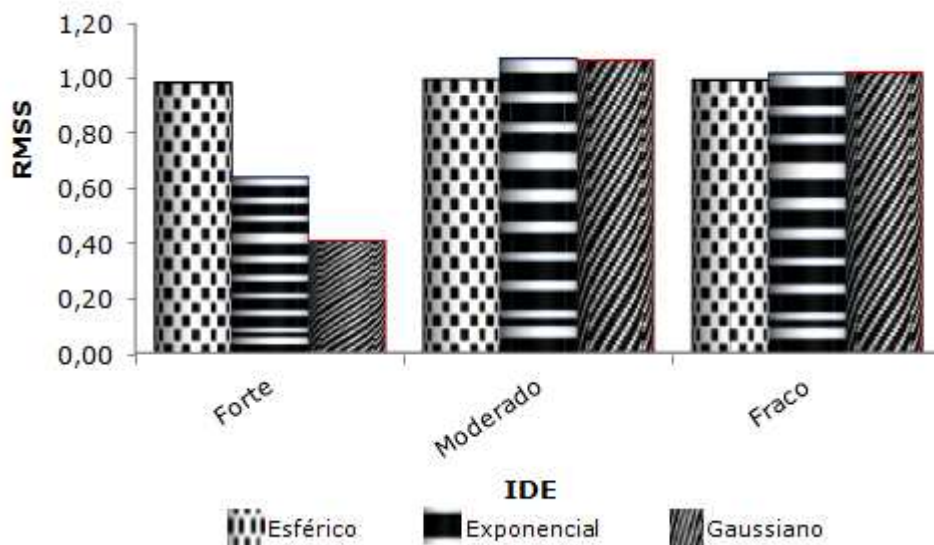


Figura 17: Valores da Média da Raiz Quadrada do Erro Quadrático Médio do Erro de Predição Padronizado (RMSS) das Diferentes Quantidades de Vizinhos e Tamanhos de Grid de Amostragem, das Krigagens Ordinárias em Função dos Diferentes Índices de Dependência Espacial e Modelos Estudados.

5. CONCLUSÕES

Para se obter uma menor variabilidade no erro da predição, recomenda-se o uso da maior grade possível de amostragem, pois o aumento do efeito pepita e a diminuição da grade de amostragem podem fazer com o que o valor da RMS aumente.

Caso o cenário estudado apresente um IDE forte, sugere-se o uso de quatro vizinhos para a realização da Krigagem, oito se for moderado o IDE, e 16 se apresentar IDE fraco, pois a partir desta quantidade, o valor da estatística RMS tende a estabilizar no menor valor de RMS, e o RMSS não tende a distanciar muito do valor ideal de um.

Dentre os modelos estudados, o modelo gaussiano apresentou menor valor de RMS, enquanto que o esférico o maior valor. O uso do modelo gaussiano também apresentou uma das melhores avaliações das estimativas, para os cenários com IDE forte ou moderado.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, A. Q.; RIBEIRO, A.; PAIVA, Y. G.; RASCON, N. J. L.; LIMA, E. P. Geoestatística no Estudo de Modelagem Temporal da Precipitação. **Rev. bras. eng. agríc. ambient.** v.15, n. 4, p. 354-358, 2011.

BLAIS, R. A.; CARLIER, P. A. Application of geostatistics in ore valuation, ore reserve estimation and grade contro.: **Canad. Inst. Min. Metall.**, v. 9, p. 48-61. 1968.

BOURENNANE, H. ; SALVADOR-BLANES, S. ; CORNU, S. ; KING, D. Scale of Spatial Dependence Between Chemical Properties of Topsoil and Subsoil Over a Geologically Contrasted Area (Massif Central, France). **Geoderma**, v.112, n.3-4, p.235-251, 2003.

BURGESS, T. M.; WEBSTER, R. Optimal Interpolation and Isarithmic Mapping of Soil Properties. II. Block Kriging. **Journal of Soil Science**, v.31, p.333-341, 1980.

CAMBARDELLA, C. A., MOORMAN, T. B., NOVAK, J. M., PARKIN, T. B., KARLEN, D. L., TURCO, R. F., KONOPKA, A. E. Field-scale variability of soil properties in Central Iowa soils. **Soil Sci. Soc. Am. J.**, v.58, p. 1501-1511, 1994.

CERRI, D. G. P. **Agricultura de Precisão em Cana-de-açúcar: Instrumentação de uma Colhedora, Mapeamento da Produtividade e de Atributos do Solo.** 2005. 156f. Tese (Doutorado em Engenharia Agrícola) - Universidade Estadual de Campinas, Campinas, SP, 2005.

CLARK, I. 1979. Practical Geostatistics. London: **Applied Science Publishers Ltd.** 129p.

CORA, J. E., ARAUJO, A. V., PEREIRA, G. T.; BERALDO, J. M. G. Variabilidade Espacial de Atributos do Solo para Adoção do Sistema de Agricultura de precisão na Cultura de Cana de Açúcar. **Revista Brasileira de Ciência do Solo**, Viçosa, v.28, n. 6, p.1013-1021, 2004.

CRESSIE, N.A.C. **Statistics for Spatial Data**. New York, John Wiley & Sons, 1993.

DAVID, M. **Handbook of applied advanced geostatistical ore reserve estimation**. New York, Elsevier, 1988. 216p.

DAVIS, J. C. **Analysis of Sequences of Data**. New York, John Wiley & Sons. 2002.

DRUCK, S. **Análise Espacial de Dados Geográficos**, EMBRAPA, Brasília, 2004.

GEORGAKARAKOS, S. ; KITSIOU, D. Mapping Abundance Distribution of Small Pelagic Species Applying Hydroacoustics and Co-Kriging Techniques. **Essential Fish Habitat Mapping in the Mediterranean Developments in Hydrobiology**. v. 203. p. 155-169, 2008.

GOOVAERTS, P. **Geostatistics for natural resources evaluation**. Oxford Univ. Press., New York, 1997. 650p.

GUERRA, P. A. G. **Geoestatística operacional**. Brasília, MME/DNPM, 1988. 145p.

GUMPERTZ, M. . **Applied Spatial Statistics**. North Carolina State University, 2001. Course Notes for ST 564.

ISAAKS, E.H., SRIVASTAVA, M. **An introduction to Applied Geostatistics**. Oxford Univ. Press., New York, 1989. 600p.

JAMUR, K. P.; FREITAS, R. G. C. Estudo Comparativo dos Modelos Gravimétricos Globais EIGEN-GL04C, EIGEN-CG03C e EIGEN-05C Utilizando um Modelo Regional e Dados gps/niv para o Estado do Paraná. **Bol. Ciênc. Geod.** v. 15, n. 4, p.527-543, out-dez, 2009.

JOHNSTON, K. ; HOEF, J. M. V. ; KRIVORUCHKO, K. ; LUCAS, N. **Using ArcGIS Geostatistical Analyst**. Redlands: ESRI Press, 2001.

JOURNEL, A. G.; ROSSI, M. E. When do You Need a Trend Model in Kriging? **Mathematical Geology**, New York, v.21, n.7, p.715-739, 1989.

KITANIDIS, P. K.; VOMVORIS, E. G. A geostatistical approach to the inverse problem in groundwater modelling (steady state) and one dimensional simulations. **Water Resources Research**, v. 19, n.3, p. 677-690, 1983.

KRIGE, D.G. A statistical approach to some basic mine evaluation problems on the Witwatersrand. **Journal of South African Institution of Mining and Metallurgy, Johannesburg**. v.52, p. 119-139, 1951.

LAJAUNIE, C. A geostatistical approach to air pollution modelling. **In: Geostatistics for Natural Resources Characterization**, ed. A. G. Journel e A. Marechal, p. 877-891, 1984.

LAMPARELLI, R. A. C.; ROCHA, J. V.; BORGHI, E. **Geoprocessamento e Agricultura de Precisão**. Ed. Agropecuária, 2001, 118p.

LANDIM, P. M. B. Sobre Geoestatística e Mapas. **Terræ Didática**. v. 2, p. 19-33, 2006.

LIAO D. ; PEUQUET, D. J. ; DUAN Y. ; WHITSEL, E. A. ; DOU, J. ; SMITH R. L. et al. GIS Approaches for the Estimation of Residential-Level Ambient PM Concentrations. **Environ Health Perspect**, v. 114, p. 1374–1380. 2006.

MATHERON, G. Principles of geostatistics. **Economic Geology**. n. 5, p. 1246-1266, 1963.

MIRANDA, F. A.; FREITAS, S. R. C.; FAGGION, P. L. Integração e Interpolação de Dados de Anomalias Ar Livre Utilizando-se a Técnica de Rna e Krigagem. **Bol. Ciênc. Geod.** v. 15, n.3, p.428-443, 2009.

MONTANARI, R. ; PEREIRA, G. T.; MARQUES JÚNIOR, J. ; SOUZA, Z. M. ; PAZETO, R. J.; CAMARGO, L. A. Variabilidade Espacial de Atributos Químicos em Latossolo e Argissolos. **Ciência Rural**, v.38, n.5, p.1266-1272, 2008.

MOTOMIYA, A. V. A. ; MOTOMIYA, W. R.; MOLIN, J. P. ; LIRA, A. ;DO OLIVEIRA, J. R. G; BISCARO, G. A. Variabilidade Espacial de Atributos Químicos do Solo e Produtividade do Algodoeiro. **Revista Agrarian**, v.4, n.11, p.01-09, 2011.

MUTUA, F. A Comparison of Spatial Rainfall Estimation Techniques: A Case Study of Nyando River Basin Kenya. **Jomo Kenyatta University of Agriculture and Technology**, v.14, n.1, p. 95-112, 2012.

OLIVEIRA, A. I. L.; FRANCA-ROCHA, W. J. S.; SILVA, A. B. Modelagem da Superfície Freática na Cabeceira da Bacia do Paraguaçu, Área do Agropolo Mucugê – Ibicoara, Estado da Bahia: Parâmetro para Estudos de Contaminação da Água Subterrânea em Aquíferos Livres. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 15. (SBSR), 2011, Curitiba. **Anais**. São José dos Campos: INPE, p. 5592-5599, 2011.

OLIVIER, M. A.; WEBSTER, R. How Geostatistics Can Help You. **Soil Use and Management**, v.7, n.4, p. 206-217, 1991.

ORTIZ, G. C. **Aplicação de métodos geostatístico para identificar a magnitude e a estrutura da variabilidade espacial de variáveis físicas do solo**. Piracicaba, SP: Escola Superior de Agricultura “Luiz de Queiroz”, 2002. Dissertação (Mestrado).

PIRES, C. A. F. **Geoestatística Aplicada**. Instituto de Geociências – Universidade Federal do Pará, Belém, 2010. 154p.

PORTO, A. L.; SOARES, J. A.; MONTEIRO, V. E. D. Otimização da Malha de Amostragem de Compostos Orgânicos Voláteis no Solo Através de Krigagem. **Águas Subterrâneas**. v.25, n.1, p.57-73, 2011.

ROBERTSON, G. P. ; KLINGENSMITH, K. M. ; KLUG, M. J. ; PAUL, E. A. ; CRUM, J. R. ; ELLIS, B. G. Soil Resources, Microbial Activity, and Primary Production Across an Agricultural Ecosystem. **Ecological Application**, v.7, n.1, p.158-170, 1997.

ROSSI, R. E., DUNGAN, J. L., BECK, L. R. Kriging in the shadows: geostatistical interpolation for remote sensing. **Remote Sens. Environ.**, v.49, p. 32-40, 1994.

SARANGI, A. ; MADRAMOOTOO, C. A. Geo-statistical Methods for Prediction of Spatial Variability of Rainfall in a Mountainous Region. **American Society of Agricultural Engineers**. v.48, p. 943-954, 2005.

SAS Institute Inc., **SAS/STAT® User's Guide**, Version 8, Cary, NC: SAS Institute Inc., 1999. Licenciado pela Universidade Federal de Viçosa em 2012.

SCHREIBER, T. Detecting and Analysing Nonstationarity in a Time Series with Nonlinear Cross-Predictions. **Physical Review Letters**. v. 78, p. 643-845, 1997.

SILVA, G. C. **Aplicação da Geoestatística Usando a Predição Espacial de Monóxido de Carbono pelo Tráfego Veicular Urbano**. 2006. Tese (Doutorado em) Universidade Federal de Santa Catarina, Florianópolis, SC, 2006.

SOARES, J. A.; MEDEIROS, A. R. S.; ALMEIDA, D. A. R.; GONZAGA, L. M. Uso da Krigagem Ordinária na Suavização de Perfis Geofísicos de Poços. **Ambiente Mineral – Revista Brasileira de Mineração e Meio Ambiente**. V. 2, 2011.

SZPIRO, A. A. ; SHEPPARD, L. ; SAMPSON, P. D. ; KIM. S. Y. Validating National Kriging Exposure Estimation. **Environ Health Perspect**. v. 115, p. 338, 2007.

TATALOVICH, Z. ; WILSON, J. P. ; COCKBURN, M. A Comparison of Thiessen Polygon, Kriging, and Spline Models of Potential UV Exposure . **Cartography and Geographic Information Science**. V. 33, p. 217-231, 2006.

TRANGMAR, B. B., YOST, R. S., UEHARA, G. Application of geostatistics to spatial studies of soil properties. **Advances in Agronomy**, v. 38, p. 45-94, 1985.

UZUMAKI, E. T. **Geoestatística multivariada: Estudo de métodos de predição**. Dissertação (Estatística – Instituto de Matemática, estatística e Ciência da Computação), UNICAMP, 1994. 104p.

VALENTE, J. M. G. P. **Geomatemática – Lições de geoestatística**. Ouro Preto, Ed. da Fundação Gorceix, 1989. 8v.

VAUCLIN, M.; VIEIRA, S.R.; BERNARD, R.; HATFIELD, J.L. Spatial variability two transects of a bare soil. **Water Resources Research**. Washington, n. 18, v. 6, p. 1677-1686, 1982.

VIEIRA, S.R.; HATFIELD, T.L.; NIELSEN, D.R.; BIGGAR, J.W. **Geostatistical theory and application to variability of some agronomical properties**. Hilgardia, Berkeley, n. 51, v. 3, p. 1-75, 1983.

ZIMBACK, C.R.L. **Análise espacial de atributos químicos de solos para fins de mapeamento da fertilidade**. Tese de Livre-Docência (Livre-Docência em Levantamento do solo e fotopedologia), FCA/UNESP, 2001. 114p.

ZIMBACK, C. R. L. **Elementos de geoestatística**. Botucatu, SP: GEPAG FCA–UNESP, 2003. 48 p.

ZIMMERMAN, D.L., ZIMMERMAN, M. B. A comparison of spatial semivariogram estimators and corresponding ordinary kriging predictors. **Technometrics**, v. 33, p. 77-91, 1991.