

FLÁVIA FERREIRA BATISTA

**INFLUÊNCIA DE MODELOS DE DEPENDÊNCIA ESPACIAL NA
DEFINIÇÃO DE MAPAS TEMÁTICOS**

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
2012

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

T

B333i
2012

Batista, Flávia Ferreira, 1985-

Influência de modelos de dependência espacial na definição de mapas temáticos / Flávia Ferreira Batista. – Viçosa, MG, 2012.

ix, 59f. : il. ; 29cm.

Inclui anexos.

Orientador: Nerilson Terra Santos

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 39-42

1. Análise espacial (Estatística). 2. Mapas de zoneamento.
3. Geologia - Métodos estatísticos. 4. Agricultura de precisão.
5. Análise discriminatória. I. Universidade Federal de Viçosa.
II. Título.

CDD 22. ed. 519.5

FLÁVIA FERREIRA BATISTA

**INFLUÊNCIA DE MODELOS DE DEPENDÊNCIA ESPACIAL NA
DEFINIÇÃO DE MAPAS TEMÁTICOS**

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: 24 de Julho de 2012

Eric Batista Ferreira

Fabyano Fonseca e Silva

Gérson Rodrigues dos Santos
(Presidente da Banca)

AGRADECIMENTOS

À Deus, o maior responsável por todas as minhas conquistas.

À meus pais e irmãs, pelo amor, incentivo, confiança e apoio incondicional, sem os quais se tornaria difícil a realização dos meus sonhos.

À todos os meus amigos e familiares, que mesmo distantes nunca deixaram de me apoiar e torceram por mim e pela minha felicidade.

À todos os amigos que conquistei na pós-graduação e em Viçosa, em especial aos que pude estudar junto e me ensinaram que estudar “brincando” é a única maneira de levar o aprendizado para a vida toda.

À todos os professores que lecionaram neste curso e em especial ao professor Nerilson Terra Santos e Gérson Rodrigues dos Santos pelas orientações e apoio.

À Universidade Federal de Viçosa, em particular aos professores do DET.

À Coordenação de Aperfeiçoamento do Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudo.

A todos, muito obrigado.

SUMÁRIO

LISTA DE FIGURAS	v
LISTA DE TABELAS	vii
RESUMO	viii
ABSTRACT	ix
1. INTRODUÇÃO	1
2. REFERENCIAL TEÓRICO	3
2.1 Agricultura de Precisão e Zonas de Manejo	3
2.2 Geoestatística	4
2.3 Semivariograma	5
2.4 Validação Cruzada	9
2.5 Índice de Dependência Espacial	10
2.6 Krigagem	11
2.7 Mapas Temáticos de Zonas de Manejo	12
2.8 Índice Kappa	13
3. MATERIAL E MÉTODOS	17
3.1 Simulações	17
3.2 Recursos Computacionais	19
3.3 Análises Estatísticas	19
4. RESULTADOS E DISCUSSÃO	23
4.1 Análise Exploratória	23
4.2 Ajuste dos Semivariogramas Experimentais	24
4.3 Validação dos Modelos	29
4.4 Mapas Temáticos de Zonas de Manejo	31
4.5 Comparação entre os Mapas Temáticos de Zonas de Manejo	35

5. CONCLUSÕES	38
6. REFERÊNCIAS BIBLIOGRÁFICAS	39
ANEXOS	43
ANEXO I.....	44
ANEXO II	49
ANEXO III.....	57

LISTA DE FIGURAS

Figura 1: Exemplo de semivariograma experimental	6
Figura 2 - Representação gráfica de modelos teóricos de semivariogramas.	9
Figura 3: Semivariogramas experimentais e modelos ajustados para os dados com forte dependência espacial em cada grid.....	25
Figura 4: Semivariogramas experimentais e modelos ajustados para os dados com moderada dependência espacial em cada Grid.....	27
Figura 5: Semivariogramas experimentais e modelos ajustados para os dados com fraca dependência espacial em cada <i>grid</i>	28
Figura 6: Mapas de zonas de manejo obtidos pela interpolação de krigagem para o <i>grid</i> G121 por diferentes estruturas de DE.	32
Figura 7: Mapas de zonas de manejo obtidos pela interpolação de krigagem para o <i>grid</i> G256 por diferentes estruturas de DE.	33
Figura 8: Mapas de zonas de manejo obtidos pela interpolação de krigagem para o <i>grid</i> G961 por diferentes estruturas de DE.	34
Figura 1II: Gráficos de validação cruzada para modelos ajustados a dados com forte DE e <i>gridG121</i>	49
Figura 2II: Gráficos de validação cruzada para modelos ajustados a dados com moderada DE e <i>gridG121</i>	50
Figura 3II: Gráficos de validação cruzada para modelos ajustados a dados com fraca DE e <i>gridG121</i>	51
Figura 4II: Gráficos de validação cruzada para modelos ajustados a dados com moderada DE e <i>gridG256</i>	52
Figura 5II: Gráficos de validação cruzada para modelos ajustados a dados com fraca DE e <i>gridG256</i>	53
Figura 7II: Gráficos de validação cruzada para modelos ajustados a dados com forte DE e <i>grid G961</i>	54

Figura 8II: Gráficos de validação cruzada para modelos ajustados a dados com moderada DE e <i>grid G961</i>	55
Figura 9II: Gráficos de validação cruzada para modelos ajustados a dados com fraca De DE e <i>grid G961</i>	56
Figura 1III: Mapas de zonas de manejo rasterizados para dados que apresentam DE Forte e <i>grid G121</i> . (a) Esférico, (b) Gaussiano e (c) Exponencial.....	57
Figura 2III: Mapas de zonas de manejo rasterizados para dados que apresentam DE moderada e <i>grid G121</i> . (a) Esférico, (b) Gaussiano e (c) Exponencial.	57
Figura 3III: Mapas de zonas de manejo rasterizados para dados que apresentam DE fraca e <i>grid G121</i> . (a) Exponencial, (b) Esférico e (c) Gaussiano.	58
Figura 4III: Mapas de zonas de manejo rasterizados para dados que apresentam DE forte e <i>grid G256</i> . (a) Gaussiano, (b) Esférico e (c) Exponencial.....	58
Figura 5B: Mapas de zonas de manejo rasterizados para dados que apresentam DE moderada e <i>grid G256</i> . (a) Esférico, (b) Gaussiano e (c) Exponencial.	58
Figura 6III: Mapas de zonas de manejo rasterizados para dados que apresentam DE Fraco e <i>grid G256</i> . (a) Exponencial, (b) Esférico e (c) Gaussiano.....	58
Figura 7III: Mapas de zonas de manejo rasterizados para dados que apresentam DE forte e <i>grid G961</i> . (a) Gaussiano, (b) Esférico e (c) Exponencial.....	59
Figura 8III: Mapas de zonas de manejo rasterizados para dados que apresentam DE moderada e <i>grid G961</i> . (a) Esférico, (b) Gaussiano e (c) Exponencial.	59
Figura 9III: Mapas de zonas de manejo rasterizados para dados que apresentam DE Fraco e <i>grid G961</i> . (a) Exponencial, (b) Esférico e (c) Gaussiano.....	59

LISTA DE TABELAS

Tabela 1: Matriz de erro de classificação.....	14
Tabela 2 - Índice Kappa e correspondente classificação do desempenho.	15
Tabela 3: Descrição dos parâmetros e modelos a utilizados nas simulações de dados.	18
Tabela 4: Intervalos de valores utilizados na definição das zonas de manejo.	20
Tabela 5: Resumo das estatísticas descritivas dos dados simulados com grade de 961 pontos (G961).	23
Tabela 6: Resumo das estatísticas descritivas dos dados simulados com grade de 256 pontos (G256).	24
Tabela 7: Resumo das estatísticas descritivas dos dados simulados com grade de 256 pontos (G121).	24
Tabela 8: Parâmetros dos modelos semivariográficos ajustados aos dados com forte DE e diferentes <i>grids</i>	25
Tabela 9: Parâmetros dos modelos semivariográficos ajustados aos dados com moderada DE e diferentes <i>grids</i>	26
Tabela 10: Parâmetros dos modelos semivariográficos ajustados aos dados com fraca DE e diferentes <i>grids</i>	28
Tabela 11: Apresentação das principais medidas da validação cruzada feita pela krigagem ordinária nos dados com forte DE e diferentes <i>grids</i>	30
Tabela 12: Apresentação das principais medidas da validação cruzada feita pela krigagem ordinária nos dados com moderada DE e diferentes <i>grids</i>	30
Tabela 13: Apresentação das principais medidas da validação cruzada feita pela krigagem ordinária nos dados com fraca DE e diferentes <i>grids</i>	31
Tabela 14: Apresentação dos resultados do Índice Kappa(k) de concordância e teste Z segundo <i>grid</i> de amostragem e estrutura de dependência espacial.	36

RESUMO

BATISTA, Flávia Ferreira, M.Sc., Universidade Federal de Viçosa, julho de 2012. **Influência de modelos de dependência espacial na definição de mapas temáticos.** Orientador: Nerilson Terra Santos. Coorientadores: Gérson Rodrigues dos Santos e Moysés Nascimento.

Nos últimos anos a Geoestatística vem sendo amplamente utilizada na área de agricultura de precisão, isso se deve ao fato de ser uma ferramenta que permite analisar a variabilidade espacial existente na área de produção agrícola, possibilitando avaliar a necessidade de criação de subáreas ou zonas onde serão realizados manejos de forma diferenciada. Um dos focos da agricultura de precisão é a identificação de zonas de manejo dentro do campo baseadas na variabilidade existente, e por meio da Geoestatística é possível produzir os mapas temáticos que auxiliam no estabelecimento das zonas de manejo através de ajustes de modelos de dependência espacial. Assim, o objetivo geral deste estudo foi analisar a influência de modelos de dependência espacial na definição de mapas temáticos de zonas de manejo, usando diferentes modelos de semivariogramas e diferentes grades de amostragem. Para a realização do estudo foram analisados dados simulados gerados no software SAS, onde se considerou diferentes estruturas de dependência espacial (DE) e diferentes grades de amostragens (grid), com diferentes tamanhos e densidades de pontos, sendo um total de 9 conjunto de dados simulados. Para a análise da estrutura de dependência espacial foram utilizados semivariogramas experimentais e ajustados três modelos teóricos ao semivariograma experimental: exponencial, esférico e gaussiano, para cada conjunto de dados em estudo, conforme os parâmetros estipulados pelo auto-ajuste do software ArcGis 9.3. Foram ajustados 27 diferentes modelos de semivariogramas através do método da krigagem ordinária e através da técnica de validação-cruzada verificados os ajustes dos modelos semivariográficos. Mapas temáticos de zonas de manejo foram gerados a partir da técnica de krigagem ordinária, e com o auxílio do software Idrisi foi realizada a comparação dos mapas através do Índice Kappa de concordância, o qual foi testado sua significância através do teste Z. A partir das análises realizadas, concluindo-se que o modelo de dependência espacial não influencia na definição de mapas temáticos de zonas de manejo.

ABSTRACT

BATISTA, Flávia Ferreira, M.Sc., Universidade Federal de Viçosa, July, 2012.
Influence of model spatial dependence in the definition of thematic maps.
Advisor: Nerilson Terra Santos. Co-Advisors: Gérson Rodrigues dos Santos and Moysés Nascimento.

In recent years Geostatistics has been widely used in the field of precision agriculture, this is due to being a tool to analyze the spatial variability exists in the area of agricultural production, enabling assess the need for creation of subareas or zones where they are handlings performed differently. One focus of precision agriculture is to identify management zones within the field based on the variability, and through Geostatistics is possible to produce thematic maps that assist in the establishment of management zones through adjustments of models of spatial dependence. The objective of this study was to analyze the influence of models of spatial dependence in the definition of thematic maps of management zones, using different models and different semivariogram sampling grids. To conduct the study, we analyzed simulated data generated in SAS software, where he held various structures of spatial dependence (DE) and different sampling grids (grid), with different sizes and densities of points, with a total of 9 data set simulated. To analyze the spatial dependence structure experimental semivariograms were used and adjusted three theoretical models to experimental semivariogram: exponential, spherical and Gaussian for each data set under study, according to the parameters set forth by autotune software ArcGis 9.3. Were adjusted 27 different semivariogram models using the method of ordinary kriging and through cross-validation technique checked the fit of the models semivariográficos. Thematic maps management zones were generated from the technique of kriging, and with the aid of software Idrisi was performed to compare the maps using Kappa Index of agreement, which was tested by testing their significance Z. From the analyzes, it was concluded that the model of spatial dependence does not influence the definition of thematic maps of management zones.

1. INTRODUÇÃO

Na agricultura convencional a prática mais frequente é de focar grandes áreas e tratá-las como homogêneas, baseando-se na necessidade média da área, assim a mesma quantidade de insumos é distribuída uniformemente na área toda, não considerando as necessidades específicas de cada parte do campo. Aplicações indevidas, superdosagem ou sobredosagem de insumos podem acarretar danos ao meio ambiente, além de elevar o custo da produção agrícola. Com isto, é cada vez mais comum, empresários agrícolas investirem em pesquisas em busca de novas tecnologias que proporcionem maior conhecimento das áreas cultivadas, tornando possível o manejo adequado de insumos nessas áreas.

Neste contexto surge a agricultura de precisão, uma tecnologia moderna para manejo adequado do solo, das culturas e insumos, que considera as variações espaciais e temporais dos fatores que afetam a produtividade. A agricultura de precisão está fundamentada na análise da variabilidade espacial. Uma forma bastante eficiente de analisar essa variabilidade é a elaboração de mapas georeferenciados. Estes mapas são importantes na determinação das zonas de manejo, pois fornecem subsídios para avaliar a necessidade de criação de subáreas ou zonas onde serão realizados manejos diferenciados dentro de uma mesma área.

Para a determinação de zonas de manejos é necessário o estudo da variabilidade espacial e temporal dos fatores que afetam a produção. Uma ferramenta eficiente de suporte à decisão de manejo é a Geoestatística, na qual o semivariograma desempenha um papel central, pois modela a dependência espacial entre pontos dentro de um campo experimental a partir do ajuste de um modelo teórico e permite estimar pontos não amostrados por meio da técnica de interpolação de dados conhecida como krigagem. Esta análise permite produzir os mapas temáticos que auxiliam na determinação das zonas de manejo.

Através da Geoestatística é possível conhecer a estrutura de dependência espacial e produzir mapas temáticos que auxiliem no manejo agrícola. Tendo em vista que o modelo da dependência espacial ajustado é o que determina a qualidade e confiabilidade dos mapas temáticos, o objetivo geral deste trabalho é analisar a

influência de modelos de dependência espacial na definição de mapas temáticos de zonas de manejo usando diferentes modelos de semivariogramas e diferentes grades de amostragem.

Os objetivos específicos deste trabalho foram:

- Ajustar diferentes modelos de dependência espacial para um fenômeno espacial conhecido;
- Determinar as zonas de manejo obtidas a partir dos modelos ajustados;
- Comparar os mapas temáticos de zonas de manejo obtidos ao utilizar diferentes modelos de dependência espacial.

2. REFERENCIAL TEÓRICO

O presente capítulo trata-se do referencial teórico, onde serão abordados conceitos de agricultura de precisão, Geoestatística e análise de concordância Kappa, com subtópicos que abrangem o conhecimento geral sobre os mesmos.

2.1 Agricultura de Precisão e Zonas de Manejo

O sistema de produção conhecido no Brasil por agricultura de precisão é denominado de *Precision Agriculture*, *Precision Farming*, *Site-Specific Crop Management* por agricultores de países com tecnologias avançadas.

Quando pequenas áreas eram cultivadas, com grande parte do trabalho feito manualmente e individualmente, as práticas de agricultura de precisão já eram realizadas. Contudo, o aumento das áreas cultivadas e o desenvolvimento da mecanização agrícola levaram os agricultores a reduzir a coleta de informações do campo e a adotar prática de tratamento uniforme. Assim a agricultura convencional baseia-se na pressuposição de uniformidade das características dos solos e das culturas, onde os agricultores com base nesta interpretação aplicam quantidades uniformes de insumos, tais como fertilizantes e defensivos. A agricultura de precisão promete reverter o quadro atual permitindo a aplicação de insumos agrícolas nos locais corretos e nas quantidades requeridas (FATORGIS, 1998).

A agricultura de precisão se desenvolveu da necessidade de se considerar a variabilidade existente em uma área de produção agrícola. Segundo MOLIN (1997) a agricultura de precisão é uma tecnologia de “ponta” tanto para a otimização de produção quanto para a tomada de decisões no manejo agrícola. A desuniformidade espacial nessas áreas de manejo sugere tratamento localizado, visando aproveitar, ao máximo, o potencial da área cultivada e da cultura implantada além de melhorar a utilização dos insumos, com reflexos positivos nos impactos ambientais provenientes do uso excessivo destes produtos. Para ser implantada com sucesso, a agricultura de precisão necessita de um maior número possível de informações da área cultivada (TSCHIEDEL & FERREIRA, 2002)

Segundo DOBERMANN & PING (2004), na agricultura de precisão o principal conceito é aplicar os insumos no local correto, no momento adequado, as quantidades de insumos necessários à produção agrícola, para áreas cada vez menores e mais homogêneas, tanto quanto a tecnologia e os custos envolvidos o permitam. Uma dos focos da agricultura de precisão é a identificação de zonas de manejo dentro do campo baseando-se na variabilidade dos fatores que limitam a produção. DOERGE (1999) definiu zona de manejo como sendo uma sub-região de uma área que expressa uma combinação homogênea de fatores limitantes à produtividade, na qual a aplicação uniforme de uma combinação específica de insumos é apropriada.

Assim, a utilização de mapas que descrevem as características e as propriedades de uma dada área é de fundamental importância na agricultura de precisão. Esses mapas representam espacialmente a variabilidade de propriedades do solo, produtividade e características observadas em campo, fornecendo assim subsídios para a tomada de decisões que envolvem o manejo localizado dessas áreas.

2.2 Geoestatística

Atualmente a aplicabilidade e a utilização da Geoestatística como metodologia de análise de dados no espaço ou no tempo esta difundida em vários ramos da ciência, envolvendo áreas de ciências humanas, biológicas e exatas. (GUIMARÃES, 2004).

O surgimento da Geoestatística se deu na década de 50, a partir de estudos práticos desenvolvidos por Daniel G. Krige, onde ao analisar dados referentes à concentração de ouro concluiu que, a variabilidade da concentração só fazia sentido se as distâncias entre as amostras fossem consideradas. No início da década de 60, G. Matheron baseando-se nas observações de Krige desenvolveu a Teoria das Variáveis Regionalizadas. Esta teoria foi definida como uma função numérica que varia espacialmente, com uma continuidade aparente possível de ser estimada pelo variograma e cuja variação não pode ser representada por uma função matemática simples.

A teoria fundamental da Geoestatística tem como princípio que as amostras próximas no espaço sejam mais similares entre si do que as que estiverem distantes (ISAAKS & SRIVASTAVA, 1989).

2.3 Semivariograma

O semivariograma desempenha o papel central nos estudos geoestatísticos, sendo capaz de descrever tanto qualitativa quanto quantitativamente a variação espacial, além de ser o ponto chave na determinação do preditor geoestatístico – krigagem (MELLO et al, 2005).

O semivariograma experimental ou empírico é um gráfico construído através da função de semivariância versus cada valor h , sendo h a distância entre dois pontos. Segundo ISAAKS & SRIVASTAVA (1989), a função de semivariância $\gamma(h)$ é definida como sendo, a metade da esperança matemática do quadrado da diferença entre os valores de pontos no espaço, separados pelo vetor h .

$$\gamma(h) = \frac{1}{2} E[Z(x_i) - Z(x_i + h)]^2$$

Em que,

$Z(x_i)$ é o valor observado da variável no ponto amostrado (x_i).

$Z(x_i + h)$ é o valor observado no ponto amostrado ($x_i + h$).

Segundo VIEIRA (2000), dentre todos os estimadores de semivariâncias o mais utilizado é o baseado no método de momentos proposto por MATHERON (1963), e dado pela seguinte equação:

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

Em que,

$\hat{\gamma}(\mathbf{h})$ é o estimador da semivariância na distância h .

$N(\mathbf{h})$ é o número de pares de pontos separados entre si por uma distância h .

O gráfico do semivariograma pode apresentar comportamento isotrópico ou anisotrópico. Sendo isotrópico quando o gráfico é idêntico para qualquer direção de h , ou seja, comportamento único em todas as direções. E anisotrópico quando existe dependência na direção, ou seja, não apresenta comportamento único em todas as direções. No caso do anisotrópico é necessário sofrer transformações antes de ser usado (VIEIRA, 2000).

Um semivariograma experimental com características muito próximas do ideal é apresentado (Figura 1). Como observações mais próximas geograficamente devem ter um comportamento mais semelhante entre si do que aquelas separadas por grandes distâncias. Dessa maneira, é esperado que as diferenças $\{Z(\mathbf{x}_i) - Z(\mathbf{x}_i + \mathbf{h})\}$ decresçam à medida que h , a distância que os separa decresce. Assim, $\hat{\gamma}(h)$ aumenta à medida que h cresce, até atingir um valor máximo no qual se estabiliza.

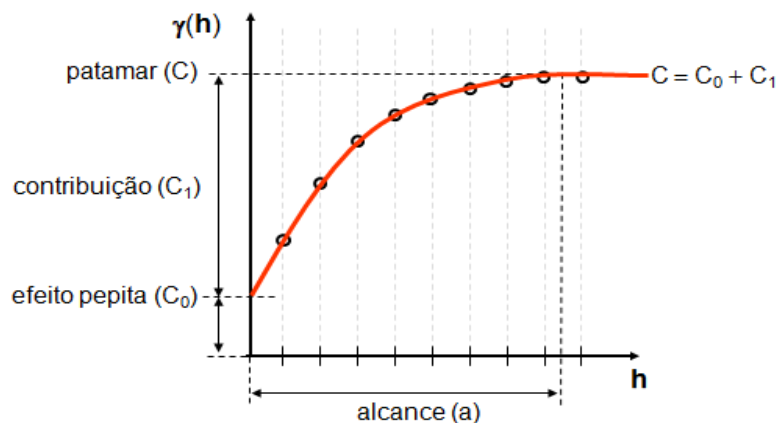


Figura 1: Exemplo de semivariograma experimental

O semivariograma possui parâmetros que auxiliam na descrição da dependência espacial, estes parâmetros são apresentados na Figura 1 e podem ser definidos da seguinte forma:

Alcance (a): É a distância dentro da qual os valores amostrais apresentam-se correlacionadas espacialmente. É o raio de dependência espacial.

Patamar (C): é o valor do semivariograma correspondente ao valor do alcance (a), ou seja, é o valor constante atingido por $\hat{\gamma}(h)$ quando a distância entre os dados, h , cresce. Deste ponto em diante, considera-se que não existe mais dependência espacial entre as amostras.

Efeito Pepita (C_0): é o valor da semivariância para a distância zero, representa a descontinuidade do semivariograma para distâncias menores do que a menor distância entre as amostras. Parte desta descontinuidade pode ser também devida a erros de medição, mas é impossível quantificar se a maior contribuição provém dos erros de medição ou da variabilidade de pequena escala, não captada pela amostragem.

Contribuição (C_1): é a diferença entre os valores do patamar (C) e do Efeito Pepita (C_0).

Conforme GUERRA (1988), o objetivo fundamental de um estudo estrutural ou estudo de semivariogramas é decidir qual o semivariograma teórico que melhor se ajusta ao semivariograma experimental, de tal modo que a partir deste modelo teórico, possam ser feitas inferências em relação ao semivariograma verdadeiro.

Diferentes modelos teóricos podem ser ajustados a um semivariograma experimental. Contudo, os modelos teóricos mais utilizados são os modelos esférico, exponencial e gaussiano (XAVIER, 2010), cujas equações e gráficos são apresentados a seguir.

- **Modelo esférico:**

É um modelo mais comumente utilizado. Seu comportamento linear para pequenos valores de h é destacado por PANNATIER (1996), o mesmo observa que a tangente à origem atinge o patamar a dois terço do alcance (Figura 2)

$$\gamma(h) = \begin{cases} 0 & , h = 0 \\ C_0 + C_1 \left[1,5 \frac{h}{a} - 0,5 \left(\frac{h}{a} \right)^3 \right] & , 0 < h < a \\ C_0 + C_1 & , h \geq a \end{cases}$$

- **Modelo exponencial**

É um modelo bastante utilizado, pressupõe pequena continuidade entre valores para pequenos intervalos. Atinge o patamar assintoticamente, com o alcance prático definido como a distância na qual o valor do modelo é 95% de C_1 (ISAACS & SRIVASTAVA, 1989).

Sua representação gráfica é apresentada na Figura 2 e obtido a partir da fórmula a seguir .

$$\gamma(h) = \begin{cases} 0 & , h = 0 \\ C_0 + C_1 \left[1 - \exp\left(-\frac{3h}{a}\right) \right] & , h \neq 0 \end{cases}$$

- **Modelo gaussiano:**

Modelo utilizado para modelar fenômenos contínuos (PANNATIER, 1996). É caracterizado pelo seu comportamento parabólico próximo a origem, sendo o único modelo que apresenta, em sua forma, um ponto de inflexão. Este modelo atinge o patamar assintoticamente, sua representação gráfica é apresentada na Figura 2, e obtido a partir da fórmula a seguir.

$$\gamma(h) = \begin{cases} 0 & , h = 0 \\ C_0 + C_1 \left[1 - \exp\left(-\frac{3h^2}{a^2}\right) \right] & , h \neq 0 \end{cases}$$

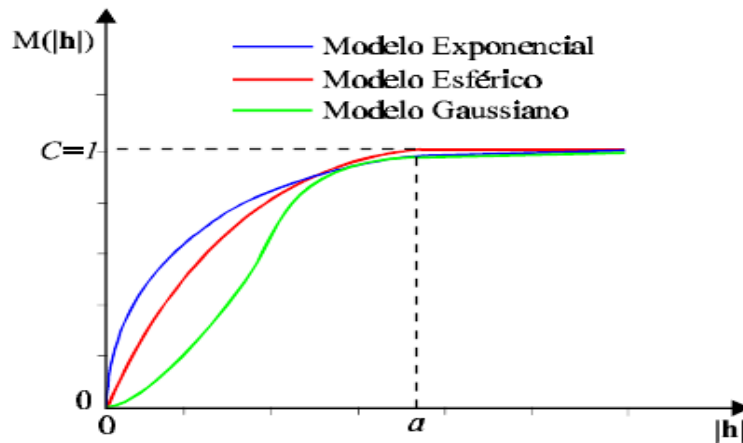


Figura 2 - Representação gráfica de modelos teóricos de semivariogramas.
 FONTE: Modificada de Isaaks e Srivastava (1989).

Todos os cálculos de Geoestatística dependem do valor do modelo do semivariograma para cada distância (h) especificada (VIEIRA et al., 1981). Sendo assim, é importante que o modelo ajustado represente a tendência do semivariograma em relação à distância. Desse modo, as estimativas da krigagem serão mais exatas e, portanto, mais confiáveis (CAMARGO et al., 2004).

2.4 Validação Cruzada

A princípio diferentes modelos teóricos podem ser ajustados a um semivariograma empírico. Contudo, espera-se que haja diferença na qualidade do ajuste destes modelos. Uma metodologia que geralmente é utilizada para avaliar a qualidade de ajuste do modelo escolhido é a validação cruzada (ISAAKS & SRIVASTAVA, 1989).

Esta técnica consiste em temporariamente eliminar um valor amostrado e prever o valor do mesmo utilizando os demais valores amostrados e o modelo do semivariograma teórico escolhido. Este procedimento deve ser realizado para cada um dos n valores amostrados, obtendo-se ao final n pares de valores observados e preditos. A diferença entre estes valores é conhecida como erro de predição.

Estatísticas são então obtidas para estes erros de predição. Análises similares devem ser realizadas para diferentes modelos teóricos de semivariograma. O modelo teórico que apresentar valores mais satisfatórios para as estatísticas destes erros de predição deve ser escolhido para representar a dependência espacial do fenômeno em estudo.

2.5 Índice de Dependência Espacial

O Índice de Dependência Espacial (IDE) pode ser usado para quantificar a dependência espacial de um fenômeno baseando-se no modelo teórico escolhido. O IDE pode ser calculado de duas maneiras diferentes. Ambas apresentam denominador igual a variância total do fenômeno, ou seja, o valor do patamar. O numerador deste índice é a diferença nas duas definições. CAMBARDELLA ET AL. (1994) definiu IDE como o percentual da variância total que é explicada pelo efeito pepita, ou seja:

$$IDE^* = \frac{C_0}{C_0 + C_1} \times 100$$

CAMBARDELLA ET AL. (1994) propuseram os seguintes intervalos para classificar a dependência espacial: $IDE^* \leq 25\%$ (forte); IDE^* entre 25% e 75% (moderada) e $IDE^* \geq 75\%$ (fraca).

ZIMBACK (2001) concluiu que a variância total ($C_0 + C_1$) é melhor explicada pela contribuição (variância estrutural), e define IDE da seguinte forma:

$$IDE = \frac{C_1}{C_0 + C_1} \times 100$$

Neste caso a dependência espacial é classificada como fraca para o $IDE \leq 25\%$; IDE entre 25% e 75%, moderada e $IDE \geq 75\%$ dependência forte.

Assim, observa-se que a determinação do IDE está diretamente ligada ao modelo teórico de semivariância e seus parâmetros, sendo importante a avaliação do

IDE a partir do modelo teórico ajustado, pois a escolha do modelo deve ser coerente com o que se espera na prática para o índice esperado do fenômeno em estudo. Supondo que o modelo escolhido apresente um IDE conforme o esperado, a fase seguinte consiste prever os valores para os pontos que não foram amostrados. Isto pode ser realizado pela krigagem.

2.6 Krigagem

Segundo VIEIRA (2000), esse método permite interpolar valores em qualquer posição no campo em estudo, sem tendência e com variância mínima, desde que exista dependência espacial entre as amostras e o semivariograma da variável seja conhecido.

Neste método utiliza-se a distância estatística a qual expressa simultaneamente a distância e a estrutura de variabilidade diferindo, portanto, dos outros métodos de interpolação que utilizam apenas a distância euclidiana. Além disso, este método considera não somente a distância dos vizinhos ao ponto a ser estimado, mas também as distâncias entre eles na influência da distribuição dos pesos (CAMARGO; 1997).

A krigagem ordinária é a versão mais comum da krigagem. Esta difere da krigagem simples por não requerer o conhecimento da média estacionária. É considerado um preditor linear, pois a predição de um ponto não-amostrado Z_k é obtida através da combinação linear dos valores observados dos pontos amostrados Z_i ($i=1,2,\dots,n$), ou seja,

$$\hat{Z}_k = \sum_{i=1}^n w_i Z_i$$

restrito a: $\sum_{i=1}^n w_i = 1$,

em que:

\hat{Z}_k é o valor predito para o local k não amostrado;

w_i são os pesos associados a cada valor z_i ;

n é o número de valores z_i das variáveis em pontos amostrados.

A krigagem ordinária para n pontos amostrais pode ser representada por meio de uma equação matricial, a qual se constitui de $n + 1$ equações, com $n + 1$ incógnitas.

$$K * \lambda = L$$

sendo:

$$K = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1n} & 1 \\ C_{21} & C_{22} & \dots & C_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix}, \quad \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix}, \quad L = \begin{bmatrix} C_{10} \\ C_{20} \\ \vdots \\ C_{n0} \\ 1 \end{bmatrix}$$

Em que K é uma matriz de semivariâncias de ordem $(n + 1) \times (n + 1)$ entre os pares de pontos amostrados; λ é o vetor de pesos; L é um vetor de semivariâncias da amostra em relação ao ponto a ser estimado Z_k ; e μ , o parâmetro de Lagrange. O vetor de parâmetros λ é obtido pelo produto do inverso da matriz K pelo vetor L , dado pela equação $\lambda = K^{-1} \cdot L$.

A krigagem é considerado um interpolador ótimo, devido a maneira como os pesos são distribuídos, e dentre todos os preditores lineares não-viesados é o que apresenta variância mínima (WEBSTER; OLIVER, 1990).

Sendo a krigagem ordinária considerada uma boa metodologia de interpolação de dados por ser um ótimo preditor, é possível a partir desta, gerar mapas temáticos que reflitam de forma precisa a variabilidade espacial existente na área em estudo.

2.7 Mapas Temáticos de Zonas de Manejo

Os mapas temáticos de zonas de manejo, também conhecido por mapas de colheita e mapas de produtividade, são amplamente utilizados em diversas áreas da

ciência, e nos últimos anos vem ganhando destaque na agricultura de precisão onde diversos estudos se baseiam na criação desses mapas a partir de análises geoestatísticas (KUZUYAKOVA, et al., 2001; BERNARDI et al, 2001; MOLIN, 2002, AMADO et al., 2007, NASCIMENTO e LANDIN, 2009).

Segundo VIEIRA AND MOLIN (2001), o conhecimento da variabilidade espacial utilizando as ferramentas estatísticas adequadas é essencial para o mapeamento e determinação das unidades de gerenciamento empregados na aplicação de técnicas de agricultura de precisão.

Assim, uma boa modelagem dos dados juntamente com a interpolação por krigagem possibilita a construção de mapas com alta precisão e a determinação das dimensões ideais do grid de amostragem (ALVES, 2005). O mapeamento permite uma visualização espacial das características da área, constituindo um aspecto relevante para seu manejo (FARIAS et al., 2004).

2.8 Índice Kappa

O Kappa é um índice que tem sido utilizado na área de estudos do solo (AMARAL et al. (2009), BAZZI et al. (2008), CARVALHO et al. (2001), COLLINGWOOD et. al (2009), MOHD AND KAMARUZAMAN (2008), KYVERYGA et. al. (2011), SUSZEK et. al. (2011), VALENTE et al. (2012)). Este índice foi introduzido por psicólogos e adaptada por CONGALTON E MEAD (1983) como uma medida de concordância (precisão) entre mapas tendo como base de calculo os elementos da matriz de confusão, também conhecida como matriz de erros (Tabela 1).

Segundo PONZINI & ALMEIDA (1996) a matriz de erros ou de confusão é constituída por igual número de linhas e de colunas, e representa o confronto entre a imagem de referência (verdade) e a imagem classificada (testada). Este confronto é realizado através de um procedimento amostral no qual certo número de pontos ou pixels pertencentes à uma dada classe (tema) são verificados em campo ou em outro mapa, considerado como referência, no sentido de comprovar ou não a atribuição correta destes à classe em questão.

Cada frequência observada na matriz corresponderá também a uma frequência esperada. A frequência observada na diagonal (x_{ii}) apresenta a concordância entre o esperado e o observado em cada categoria, ou seja, aqueles corretamente classificados. As colunas geralmente representam os dados de referência, enquanto as linhas indicam a classificação gerada a partir dos dados de sensoriamento remoto.

Uma matriz de erro é uma maneira muito eficaz para representar a precisão em que a precisão de cada categoria são claramente descritos junto com ambos os erros de inclusão (erros de comissão) e erros de exclusão (erros de omissão) presentes na classificação.

Tabela 1: Matriz de erro de classificação

		Referência				
		1	2	...	r	Total
Classificação	1	x_{11}	x_{12}	...	x_{1c}	x_{1+}
	2	x_{21}	x_{22}	...	x_{2c}	x_{2+}
	⋮	⋮	⋮	⋮	⋮	⋮
	r	x_{r1}	x_{r2}	...	x_{rr}	x_{c+}
	Total	x_{+1}	x_{+2}	...	x_{+r}	N

O Índice Kappa é calculado pelo seguinte estimador:

$$\hat{k} = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} * x_{+i})}$$

Sendo; x_{ii} o valor da linha i e coluna i da matriz de erros;

x_{i+} o total de observações da linha i;

x_{+i} o total de observações da coluna i;

N o número de observações; e r o número colunas da matriz de erros.

O maior valor assumido pelo Índice Kappa é 1 onde os valores mais elevados indica maior concordância (precisão) entre as classes dos mapas considerados, os valores próximos e até abaixo de 0, indicam nenhuma concordância, ou a

concordância foi exatamente a esperada pelo acaso. Um eventual valor de Kappa menor que zero, sugere discordância, mas seu valor não tem interpretação como intensidade de discordância. Assim, um Kappa com valor negativo, que não tem interpretação cabível.

Segundo GALPARSORO E FERNÁNDEZ (2001) o Índice Kappa de um modo geral pode ser avaliado de acordo com a tabela 2. Essa classificação sendo utilizada em diversas áreas de estudo.

Tabela 2 - Índice Kappa e correspondente classificação do desempenho.

<i>Índice Kappa</i>	<i>Desempenho</i>
$\hat{k} \leq 0,2$	Pobre
$0,2 < \hat{k} \leq 0,4$	Fraca
$0,4 < \hat{k} \leq 0,6$	Moderada
$0,6 < \hat{k} \leq 0,8$	Boa
$0,8 < \hat{k} \leq 1,0$	Muito Boa

Fonte: Adaptado de Galparsoro e Fernández (2001)

Contudo, o valor do Índice Kappa por si só, não tem a capacidade de dizer se existe uma concordância ou não entre dois mapas. É preciso avaliar a significância estatística do valor obtido para o mesmo. Esta pode ser avaliada pelo teste Z. As hipóteses para este teste são:

Hipótese Nula (H_0): Ambos os mapas apresentaram classificações semelhantes, ou seja, existe concordância entre os mapas. ($H_0 : k = 0$)

Hipótese Alternativa (H_1): Os mapas apresentaram classificações diferentes, ou seja, existe alguma divergência entre os mapas. ($H_1 : k \neq 0$)

A estatística do teste Z (CONGALTON e GREEN, 2009) é apresentada a seguir.

$$Z = \frac{\hat{k} - k}{\sqrt{\widehat{Var}(\hat{k})}} \quad Z \sim N(0,1)$$

A obtenção de variâncias é baseado no método delta (AGRESTI, 2002). A variação k , por este método, é obtida a partir da expressão:

$$\widehat{Var}(k) = \frac{1}{n} \left[\frac{d_1 (1 - d_1)}{(1 - d_1)^2} + \frac{2(1 - d_1)(2d_1d_2 - d_3)}{(1 - d_1)^3} + \frac{(1 - d_1)^2(d_4 - 4d_2^2)}{(1 - d_2)^4} \right]$$

Onde: $d_1 = \frac{\sum_{i=1}^c x_{ii}}{n}$ $d_2 = \frac{\sum_{i=1}^c x_{i+}x_{+i}}{n^2}$

$$d_3 = \sum_{i=1}^c x_{ii}(x_{i+} + x_{+i})/n^2$$

$$d_4 = \sum_{i=1}^c \sum_{j=1}^c x_{ij}(x_{j+} + x_{+i})/n^3$$

Se o valor crítico $Z_{\alpha/2} \leq Z$ a classificação é significativamente melhor do que uma classificação aleatória.

3. MATERIAL E MÉTODOS

Neste capítulo são apresentados os recursos e metodologia utilizada neste estudo. Descreve-se a forma de obtenção dos dados por meio de simulações, os softwares e as estatísticas utilizadas para a obtenção dos objetivos propostos.

3.1 Simulações

Para a realização do estudo foram geradas simulações de um fenômeno isotrópico, e utilizou-se o procedimento SIM2D do programa Statistical Analysis System (SAS, 2009).

Estas simulações foram geradas de tal forma que as grades de amostragem fossem uniformes, mas que apresentassem diferenças quanto ao número e adensamento dos pontos que compunham cada grade. Os tipos de grade de amostragem geradas são apresentadas a seguir.

- G121 - *grid* de 50 x 50 com pontos igualmente espaçados em 5 unidades, totalizando 121 pontos.
- G256 - *grid* de 150 x150 com pontos igualmente espaçados em 10 unidades, totalizando 256 pontos.
- G961 - *grid* de 150 x150 com pontos igualmente espaçados em 5 unidades, totalizando 961 pontos.

Para cada tipo de grade foram realizadas 3 diferentes simulações, sendo realizadas 18150 pontos simulados para cada *grid* G121, 38400 pontos simulados para os *grids* G256 e 144150 pontos simulados para os *grids* G961, totalizando 1467507 pontos simulados. A escolha dos dados simulados entre todas as simulações realizadas foi através da análise semivariográfica realizada a partir do comando PROC VARIOGRAM (SAS, 2009), onde escolheu-se as simulações que apresentaram parâmetros semivariográficos (alcance e patamar) iguais aos definidos na realização da simulação. Os códigos de programação utilizados nas simulações

podem ser encontrados no ANEXO I. Os valores dos dados simulados são referentes a uma variável cuja unidade de medida é expressa em ppm (parte por milhão).

Cada uma destas três grades foi gerada para três diferentes IDE: fraco, moderado e forte (ZIMBACK, 2001). Os dados gerados para estes IDE serão daqui por diante dito terem, respectivamente, estruturas de Dependência Espacial (DE) fraca, moderada e forte. Nas simulações com estrutura de DE forte foi utilizado o modelo gaussiano (*Gau*), para estrutura de DE moderada utilizou-se o modelo esférico (*Sph*) e o modelo exponencial (*Exp*) para as simulações com fraca DE.

Uma descrição de todas as simulações geradas, as respectivas estruturas de DE, modelos e parâmetros utilizados são apresentados na Tabela 3.

Tabela 3: Descrição dos parâmetros e modelos a utilizados nas simulações de dados.

<i>Grade</i>	<i>Estrutura de DE</i>	C_0	C_1	A	<i>Média</i>	<i>% IDE</i>	<i>Modelo</i>
G121	Forte	1 E-8	3	15	0	0,999	<i>Gau</i>
	Moderada	1,5	4,5	15	0	0,667	<i>Sph</i>
	Fraca	3	4	15	0	0,25	<i>Exp</i>
G256	Forte	1 E-8	3	40	0	0,999	<i>Gau</i>
	Moderada	1,5	4,5	40	0	0,667	<i>Sph</i>
	Fraca	3	4	40	0	0,25	<i>Exp</i>
G961	Forte	1 E-8	3	40	0	0,833	<i>Gau</i>
	Moderada	1,5	4,5	40	0	0,714	<i>Sph</i>
	Fraca	3	4	40	0	0,25	<i>Exp</i>

Alcance(a), Contribuição (C_1), Efeito Pepita (C_0), Índice de Dependência Espacial (*IDE*)

Ao todo, portanto, foram geradas nove grades de amostragens, resultante da combinação cruzada de três diferentes grades de adensamento com três diferentes estruturas de DE.

3.2 Recursos Computacionais

Para execução deste estudo foram utilizados os softwares SAS 9.2 (SAS, 2008), ArcGis 9.3 e o Idrisi Andes 15.0.

- O SAS 9.2: Utilizado para gerar e selecionar as melhores simulações para a realização deste estudo, como também para calcular o índice Kappa de concordância e teste Z, necessários para atingir o objetivo de comparação dos mapas.
- ArcGIS 9.3: Utilizado na realização de toda a análise geoestatística, desde a modelagem do semivariograma, obtenção de mapas de interpolação e das zonas de manejo.
- O Idrisi Andes 15.0: Utilizado para obtenção das matrizes de erros, através do comando Ermat onde é realizada a comparação dos mapas.

3.3 Análises Estatísticas

Para cada uma das nove grades de amostragens (Tabela 3) geradas a partir das simulações foram ajustados três modelos de dependência espacial: exponencial, esférico e gaussiano, conforme os parâmetros estipulados pelo auto-ajuste do software ArcGis 9.3. A partir de cada um destes modelos ajustados, foi gerado um mapa de interpolação usando a krigagem ordinária. Portanto, foram gerados um total de 27 mapas de interpolação.

Segundo BOREM (2000) uma boa medida de definição do número de classe (zonas) em um mapa é entre 3 e 5, neste estudo as zonas de manejo foram definidas em quatro classes de igual amplitude. Os intervalos de valores utilizados para a definição das zonas de manejo são apresentados na Tabela 4. Os mapas gerados a partir de um mesmo conjunto de dados apresentam intervalos idênticos.

Tabela 4: Intervalos de valores utilizados na definição das zonas de manejo.

Grade	Zona	DE		
		Forte	Moderada	Fraca
G121	Z1	-3,803 a -1,665 ppm	-5,952 a -2,843 ppm	-6,624 a -3,467 ppm
	Z2	-1,665 a 0,473 ppm	-2,843 a 0,266 ppm	-3,470 a -0,315 ppm
	Z3	0,473 a 2,611 ppm	0,266 a 3,375 ppm	-0,315 a 2,839 ppm
	Z4	2,611 a 4,749 ppm	3,375 a 6,484 ppm	2,839 a 5,993 ppm
G256	Z1	-3,337 a -1,187 ppm	-5,368 a -2,586 ppm	-4,277 a -1,552 ppm
	Z2	-1,187 a 0,963 ppm	-2,586 a 0,196 ppm	-1,552 a 1,173 ppm
	Z3	0,963 a 3,113 ppm	0,196 a 2,977 ppm	1,173 a 3,898 ppm
	Z4	3,113 a 5,263 ppm	2,977 a 5,759 ppm	3,898 a 6,623 ppm
G961	Z1	-3,339 a -1,263 ppm	-5,589 a -2,974 ppm	-4,366 a -1,805 ppm
	Z2	-1,263 a 0,813 ppm	-2,974 a -0,359 ppm	-1,805 a 0,756 ppm
	Z3	0,813 a 2,889 ppm	-0,359 a 2,26 ppm	0,756 a 3,316 ppm
	Z4	2,889 a 4,965 ppm	2,256 a 4,87 ppm	3,316 a 5,878 ppm

Para a realização da krigagem ordinária foi definida uma vizinhança circular em um setor, o número de vizinhos foi escolhido em função da análise da validação-cruzada, sendo utilizados 5, 6 e 6 vizinhos para os grids G121, G256 e G961, respectivamente.

Para avaliar a qualidade dos modelos de variograma ajustados foi utilizada a técnica de validação-cruzada (“cross-validation”), este método permite obter para cada valor observado o erro de predição a partir da diferença entre a predição e o respectivo valor observado.

A partir destes erros, a similaridade dos valores observados e preditos foi analisada através das estatísticas: erro médio (ME), erro quadrático médio padronizado (RMSS) e os coeficientes da regressão entre os valores observados e preditos descritos a seguir (VIEIRA, 2000).

- **ME** - Utilizado para determinar o grau de viés das predições. O valor próximo de zero indica que as predições são não-viesadas. Pode ser calculado pela seguinte equação:

$$ME = \frac{1}{n} \sum_{i=1}^n |\hat{Z}(x_i) - Z(x_i)|$$

em que:

n : é o número de observações

$\hat{Z}(x_i)$: São os valores preditos pela krigagem ordinária

$Z(x_i)$: São os valores observados.

- **RMSS** - Usado para avaliar se a variabilidade da predição está sendo corretamente predita. Sendo o valor ideal próximo de 1, é obtido pela seguinte fórmula:

$$RMSS = \frac{\frac{1}{n} \sqrt{\sum_{i=1}^n (\hat{Z}(x_i) - Z(x_i))^2}}{\frac{1}{n} \sum_{i=1}^n (\hat{Z}(x_i) - Z(x_i))^2} = \frac{RMS}{ASE}$$

Segundo VIEIRA (2000), uma vez que para cada um dos n locais se tem um valor medido $Z(x_i)$, se predizer um valor $\hat{Z}(x_i)$ pela krigagem, é possível fazer um gráfico dos valores pareados de $Z(x_i)$ e $\hat{Z}(x_i)$ e calcular a regressão entre eles.

Na regressão linear entre os valores preditos e estimados, coeficiente angular (a) será ótimo quando for igual à 1 e o coeficiente linear (b) quando o mesmo for nulo, SANTOS et. al. (2011). Assim, a relação entre os valores preditos e observados é alta, e conclui-se que o método de interpolação ou o modelo de dependência espacial é adequado (ISAAKS e SRIVASTAVA, 1989).

Para realizar a comparação entre os mapas de zonas de manejos de cada conjunto de dados, fez-se necessário a conversão deste em mapas do tipo raster ou mapas de pixels. Estes, foram convertidos no software ArcGis e tiveram suas classes recodificadas com valores discretos, esse passo foi necessário pois para o cálculo da matriz de erros os valores das classes devem ser discretos, definidos então como zona 1, 2, 3 e 4. No software Idrisi Andes foi realizada a comparação dos mapas raster, para esta comparação foi necessário definir qual o mapa de referência (Verdadeiro). Definiu-se como sendo mapa de referência, o mapa criado de acordo com o modelo especificado na simulação e os demais considerados como mapa de classificação, assim, sendo comparado cada mapa de zona de manejo (classificação) em relação ao mapa de referência.

Foram realizadas 18 comparações, cada um dos 9 mapas definidos como mapa de referência foi comparado com 2 mapas de zonas de manejo (classificação). Isso com intuito de verificar se mapas de zonas de manejo referentes a um determinado conjunto de dados e gerados a partir de modelos diferentes apresentam concordância (semelhança). Finalmente, foi utilizado o Índice Kappa para mensurar a concordância entre os mapas de zonas de manejo, o teste Z foi utilizado para avaliar a significância do Índice Kappa.

4. RESULTADOS E DISCUSSÃO

Apresenta-se, então, os resultados da análise descritiva dos dados simulados utilizados neste estudo e a discussão sobre os resultados obtidos pela krigagem, realizada pelo auto-ajuste do Software ArcGis 9.3.

4.1 Análise Exploratória

Antes de se realizar qualquer análise geoestatística é importante realizar uma análise descritiva e exploratória dos dados.

De acordo com GUIMARÃES (2004), o comportamento assimétrico dos dados apesar de não ser um fator limitante à análise geoestatística, pode dificultar o ajuste de modelos de semivariogramas. Segundo WEBSTER E OLIVER (2007), um coeficiente de assimetria dos dados entre os valores de 0 e 0,5 não indica a necessidade de transformação, um coeficiente entre 0,5 e 1,0 a transformação em raiz quadrada é a mais recomendável e um coeficiente maior que 1,0, se faz necessário a transformação logarítmica dos dados.

As Tabelas 5 a 7 representam o resumo da estatística descritiva dos dados simulados, observa-se que todos os conjuntos de dados não apresentam assimetria, com isso pode-se realizar a análise geoestatística pelo método de krigagem.

Tabela 5: Resumo das estatísticas descritivas dos dados simulados com grade de 961 pontos (G961).

Estatística	Estrutura de Dependência		
	Forte	Moderada	Fraca
Número	961	961	961
Média	0,183	0,285	0,093
Mediana	0,034	0,292	0,096
Mínimo	-3,803	-5,952	-6,624
Máximo	4,749	6,484	5,993
Quartil inferior	-1,069	-1,133	-1,298
Quartil superior	1,528	1,753	1,445
Desvio Padrão	1,818	2,075	1,991
Assimetria	0,179	-0,037	0,002

Tabela 6: Resumo das estatísticas descritivas dos dados simulados com grade de 256 pontos (G256).

Estatística	Estrutura de Dependência		
	Forte	Moderada	Fraca
Número	256	256	256
Média	0,938	0,072	1,109
Mediana	0,937	0,124	1,042
Mínimo	-3,337	-5,368	-4,277
Máximo	5,262	5,759	6,623
Quartil inferior	-0,47	-1,282	-0,224
Quartil superior	2,173	1,478	2,545
Desvio Padrão	1,928	2,109	1,986
Assimetria	0,085	-0,132	0,214

Tabela 7: Resumo das estatísticas descritivas dos dados simulados com grade de 121 pontos (G121).

Estatística	Estrutura de Dependência		
	Forte	Moderada	Fraca
Número	121	121	121
Média	1,171	0,0361	0,128
Mediana	1,198	-0,586	0,196
Mínimo	-3,339	-5,589	-4,366
Máximo	4,965	4,87	5,857
Quartil inferior	-0,722	-1,737	-1,097
Quartil superior	3,317	1,133	1,504
Desvio Padrão	2,313	2,148	2,013
Assimetria	-0,198	0,211	0,037

4.2 Ajuste dos Semivariogramas Experimentais

Apresenta-se abaixo os ajustes de semivariogramas e modelos estimados realizados no software ArcGis pelo método dos mínimos quadrados, onde ajustou-se 3 modelos de DE (Gaussiano, Esférico e Exponencial) para cada conjunto de dados simulados

Na Tabela 8 e Figura 6 apresenta-se respectivamente, os parâmetros efeito pepita (C_0), patamar (C_I) e alcance (a) para os modelos de semivariogramas ajustados aos dados com forte DE e *grid* de amostragem, e os modelos

semivariográficos ajustados. Para os ajustes utilizou-se lag (distância entre os pontos) definido automaticamente pelo software utilizado.

Tabela 8: Parâmetros dos modelos semivariográficos ajustados aos dados com forte DE e diferentes *grids*.

<i>grid</i>	Modelo	C_0	$C_0 + C_1$	a
G121	Simulado (<i>Gau</i>)	1 E-8	3	15
	<i>Gau</i>	0,194	4,99	29,01
	<i>Sph</i>	0	4,174	29,01
	<i>Exp</i>	0	4,028	29,01
G256	Simulado (<i>Gau</i>)	1 E-8	3	40
	<i>Gau</i>	0,184	4,583	69,61
	<i>Sph</i>	0	4,427	87,05
G961	Simulado (<i>Gau</i>)	1 E-8	3	40
	<i>Gau</i>	0,104	4,262	64,14
	<i>Sph</i>	0	4,22	84,54
	<i>Exp</i>	0	3,936	84,54

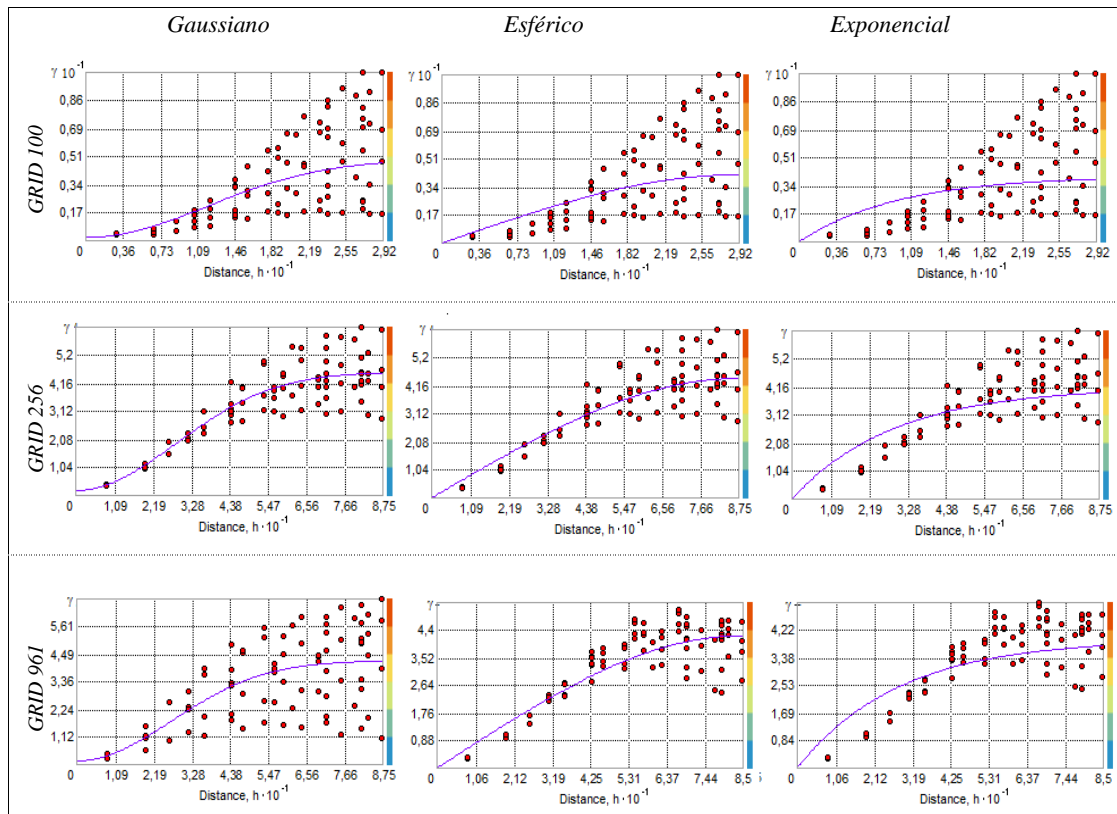


Figura 3: Semivariogramas experimentais e modelos ajustados para os dados com forte dependência espacial em cada grid.

Comparando os valores das Tabelas 5 a 7 com os valores do patamar na Tabela 8, observa-se que a variância captada pelos modelos é maior que a variância dos dados em todos os casos.

Observa-se na Tabela 8 e Figura 3 que os parâmetros dos modelos ajustados não foram próximos dos valores simulados. Na análise visual dos modelos semivariográficos, observa-se um bom ajuste do modelo gaussiano para os *grids* G961 e G256, estes apresentando patamar bem definido, e no ajuste dos demais modelos não foi alcançado um patamar.

Na Tabela 9 e Figura 4 apresenta-se respectivamente, os parâmetros efeito pepita (C_0), patamar (C_1) e alcance (a) para os modelos de semivariogramas ajustados aos dados com moderada DE e *grid* de amostragem, e os modelos semivariográficos ajustados.

Tabela 9: Parâmetros dos modelos semivariográficos ajustados aos dados com moderada DE e diferentes *grids*.

<i>grid</i>	Modelo	C_0	$C_0 + C_1$	a
G121	Simulado (<i>Sph</i>)	1,5	4,5	15
	<i>Gau</i>	2,497	4,58	13,6
	<i>Sph</i>	2,037	4,565	15,58
	<i>Exp</i>	0,569	4,693	14,89
G256	Simulado (<i>Sph</i>)	1,5	4,5	40
	<i>Gau</i>	1,95	4,539	37
	<i>Sph</i>	1,476	4,536	43,76
	<i>Exp</i>	0,854	4,728	51,42
G961	Simulado (<i>Sph</i>)	1,5	4,5	40
	<i>Gau</i>	2,0858	4,3738	29,45
	<i>Sph</i>	1,678	4,3681	34,53
	<i>Exp</i>	1,121	4,437	36,15

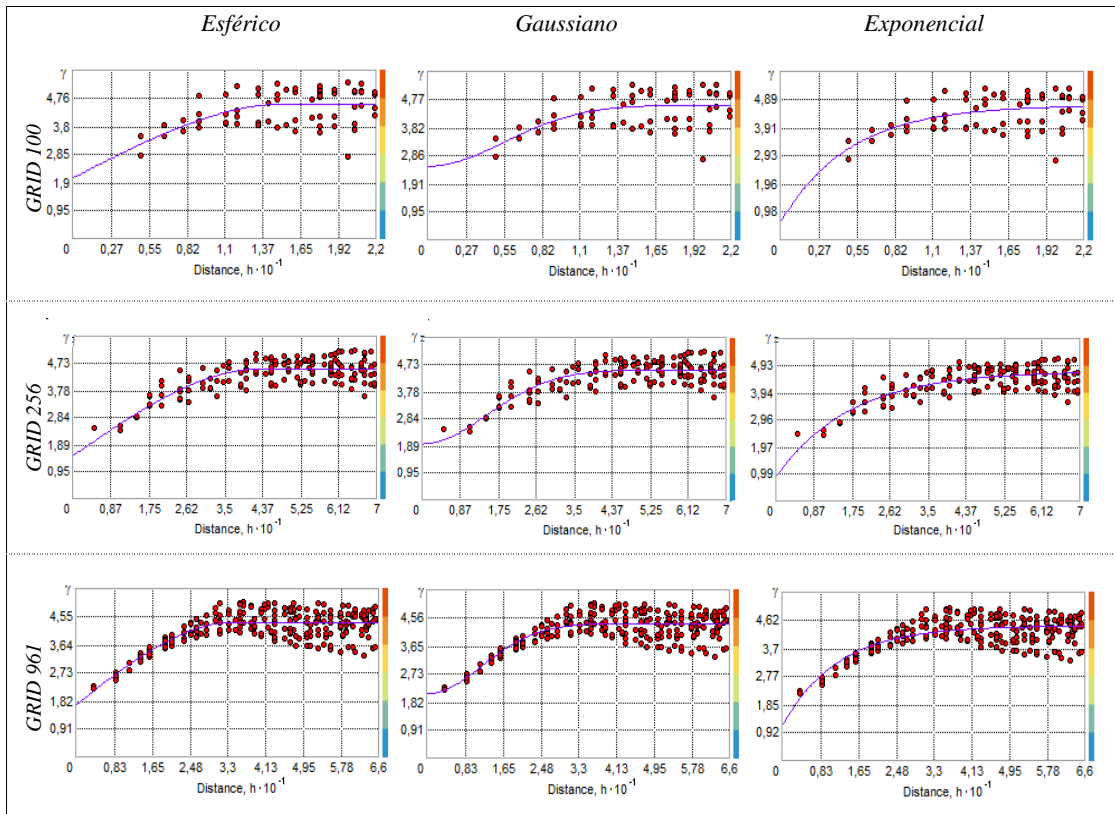


Figura 4: Semivariogramas experimentais e modelos ajustados para os dados com moderada dependência espacial em cada Grid.

Comparando os valores das Tabelas 5 a 7 com os valores do patamar na Tabela 9, observa-se que a variância captada pelos modelos é bem próxima da a variância dos dados em todos os casos.

Observa-se na Tabela 9 e Figura 4 que os parâmetros dos modelos ajustados se aproximaram dos valores das simulações, sendo os modelos esféricos os que se aproximam mais, com isso pode-se dizer que os dados simulados com moderada DE apresentaram o comportamento semelhante ao que se esperava na simulação. Todos os modelos apresentaram bom ajuste, apresentando patamar bem definido.

Na Tabela 10 e Figura 5 apresenta-se respectivamente, os parâmetros dos modelos de semivariogramas ajustados aos dados com fraca DE e *grid* de amostragem, e os modelos semivariográficos ajustados.

Tabela 10: Parâmetros dos modelos semivariográficos ajustados aos dados com fraca DE e diferentes *grids*.

<i>grid</i>	Modelo	C_0	$C_0 + C_1$	a
G121	Simulado (<i>Sph</i>)	3	4	15
	<i>Gau</i>	3,693	3,986	20,61
	<i>Sph</i>	3,604	3,977	21,46
	<i>Exp</i>	3,616	4,015	29
G256	Simulado (<i>Sph</i>)	3	4	40
	<i>Gau</i>	3,59	3,914	33,64
	<i>Sph</i>	3,561	3,707	43,18
	<i>Exp</i>	3,322	3,95	34,86
G961	Simulado (<i>Sph</i>)	3	4	40
	<i>Gau</i>	3,406	4,023	38,61
	<i>Sph</i>	3,289	4,121	44,91
	<i>Exp</i>	3,058	4,04	43,82

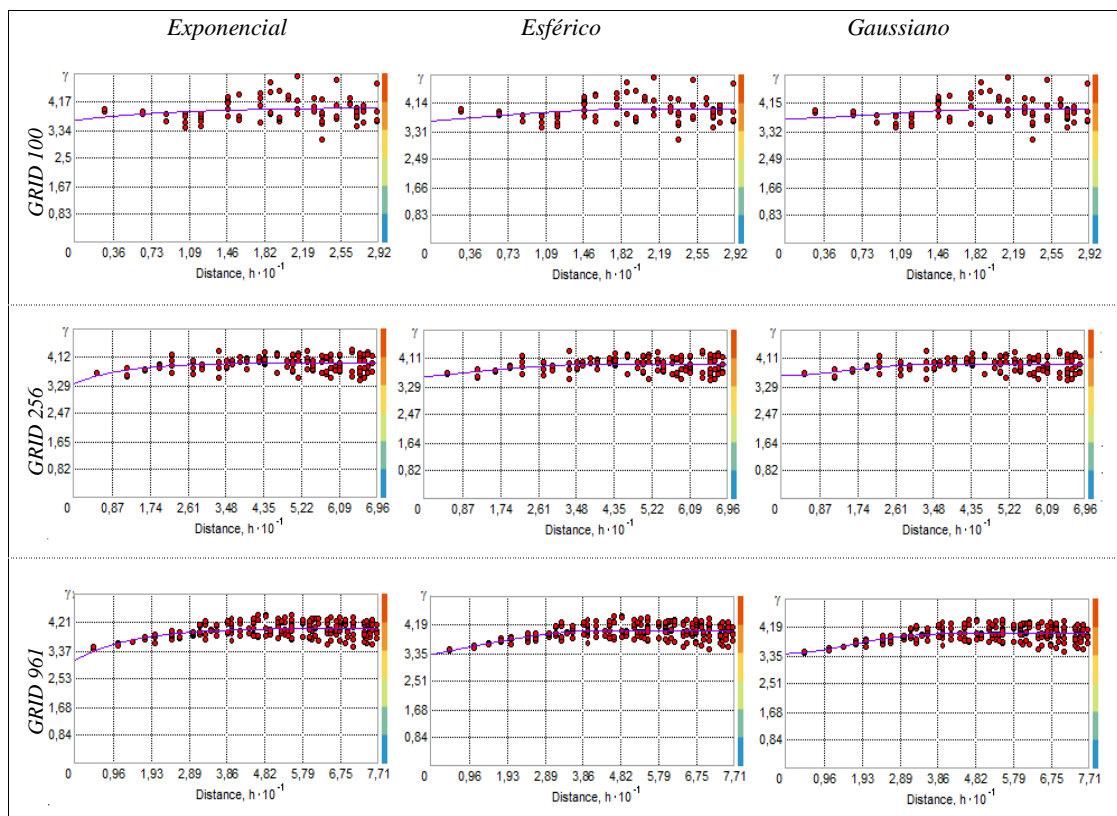


Figura 5: Semivariogramas experimentais e modelos ajustados para os dados com fraca dependência espacial em cada *grid*.

Comparando os valores das Tabelas 5 a 7 com os valores do patamar na Tabela 10, observa-se que a variância captada pelos modelos é próxima da variância dos dados em todos os casos.

Observa-se na Tabelas 8 e Figura 5 que os parâmetros efeito pepita e patamar dos modelos ajustados se aproximaram dos valores das simulações, e em alguns casos o alcance se aproxima do alcance simulado. Através da inspeção visual pode-se dizer que a maioria dos modelos foram bem ajustados.

4.3 Validação dos Modelos

Sobre os resultados da krigagem ordinária foi aplicado o procedimento de validação cruzada ou *cross-validation* (ISAAKS E SRIVASTAVA, 1989; WEBSTER E OLIVER, 2001) para avaliar o desempenho dos modelos ajustados, em termos de precisão das estimativas. Para essa avaliação utilizou-se as seguintes estatísticas: o erro médio (ME), erro quadrático médio padronizado (RMSS) e coeficientes da regressão entre os valores observados e preditos, conforme recomenda VIEIRA (2000).

Uma vez que o objetivo deste estudo foi realizar a krigagem para todos os conjuntos de dados e gerar os mapas de zonas de manejo utilizando os diferentes modelos, a etapa de validação cruzada foi realizada com o intuito de verificar as estimativas dos modelos, não sendo feita escolha por melhor modelo para a realização da krigagem.

As Tabelas 11 a 13 correspondem aos resultados obtidos na validação cruzada para os diferentes modelos ajustados em cada conjunto de dados analisados.

Tabela 11: Apresentação das principais medidas da validação cruzada feita pela krigagem ordinária nos dados com forte DE e diferentes *grids*.

<i>Grid</i>	Modelo	ME	RMSS	$\hat{\beta}_0$	$\hat{\beta}_1$
<i>G121</i>	<i>Gaussiano</i>	-0,0067	0,322	0,081	0,941
	<i>Esférico</i>	-0,0023	0,208	0,074	0,938
	<i>Exponencial</i>	0,0033	0,182	0,087	0,933
<i>G256</i>	<i>Gaussiano</i>	-0,0097	0,347	0,046	0,932
	<i>Esférico</i>	-0,0014	0,227	0,058	0,932
	<i>Exponencial</i>	0,002	0,186	0,064	0,926
<i>G961</i>	<i>Gaussiano</i>	-0,0003	0,1617	0,001	0,974
	<i>Esférico</i>	0,0002	0,0857	0,001	0,978
	<i>Exponencial</i>	0,0003	0,0647	0,001	0,978

Tabela 12: Apresentação das principais medidas da validação cruzada feita pela krigagem ordinária nos dados com moderada DE e diferentes *grids*.

<i>Grid</i>	Modelo	ME	RMSS	$\hat{\beta}_0$	$\hat{\beta}_1$
<i>G121</i>	<i>Gaussiano</i>	0,0204	0,981	-0,214	0,384
	<i>Esférico</i>	0,022	0,976	-0,214	0,38
	<i>Exponencial</i>	0,0176	0,981	-0,216	0,388
<i>G256</i>	<i>Gaussiano</i>	0,0066	0,994	0,061	0,48
	<i>Esférico</i>	0,0079	0,979	0,062	0,48
	<i>Exponencial</i>	0,011	0,967	0,049	0,498
<i>G961</i>	<i>Gaussiano</i>	0,0022	0,964	0,159	0,493
	<i>Esférico</i>	0,0014	0,971	0,166	0,5
	<i>Exponencial</i>	0,0004	0,994	0,141	0,503

Tabela 13: Apresentação das principais medidas da validação cruzada feita pela krigagem ordinária nos dados com fraça DE e diferentes *grids*.

<i>Grid</i>	Modelo	ME	RMSS	$\hat{\beta}_0$	$\hat{\beta}_1$
<i>G121</i>	<i>Gaussiano</i>	0,0209	1,044	0,155	0,036
	<i>Esférico</i>	0,0205	1,045	0,155	0,036
	<i>Exponencial</i>	0,0209	1,039	0,155	0,036
<i>G256</i>	<i>Gaussiano</i>	0,0096	0,986	0,989	0,093
	<i>Esférico</i>	0,0097	0,983	0,989	0,093
	<i>Exponencial</i>	0,0102	0,984	0,989	0,093
<i>G961</i>	<i>Gaussiano</i>	0,0044	0,989	0,128	0,149
	<i>Esférico</i>	0,004	0,994	0,13	0,149
	<i>Exponencial</i>	0,0033	1,006	0,131	0,149

Na análise das estimativas apresentadas nas Tabelas 11 a 13 observou-se que na maioria dos casos os erros de predição dos modelos ajustados foram semelhantes. Considerando como melhor ajuste aqueles que apresentaram ME mais próximo de zero e RMSS mais próximo de 1, segundo recomenda (ARCGIS, 2009). Nos os ajustes para os dados com forte DE os parâmetros da regressão apresentaram boas estimativas, não apresentando valores de RMSS próximos de 1. Os ajustes para dados com moderada DE e fraça DE apresentaram valores muito próximos de 1, e não apresentando grandes diferenças entre os modelos ajustados levando em cada *grid*.

Dessa forma, pode-se perceber que os modelos ajustados em cada *grid* não apresentam grandes diferenças entre si.

4.4 Mapas Temáticos de Zonas de Manejo

A partir da técnica de interpolação por krigagem ordinária obteve-se mapas, que neste estudo chama-se de mapas temáticos de zonas de manejo. Foram gerados

um total de 27 mapas temáticos de zonas de manejo, sendo 3 para cada conjunto de dados segundo dependência espacial DE e *grid*. Através dos mapas apresentados nas Figuras 6 a 8 pode-se fazer uma comparação visual entre os mapas gerados por diferentes modelos nos diferentes *grids* e estruturas de DE.

Na Figura 6 apresenta-se os mapas para os casos onde se tem um *grid* de amostragem pouco denso, ou seja, poucos pontos amostrados, *grid* G121, pode-se observar que os mapas gerados a partir de diferentes modelos apresentam comportamento bem semelhante. No caso onde há moderada DE é que se percebe pequenas diferenças nas zonas de manejo.

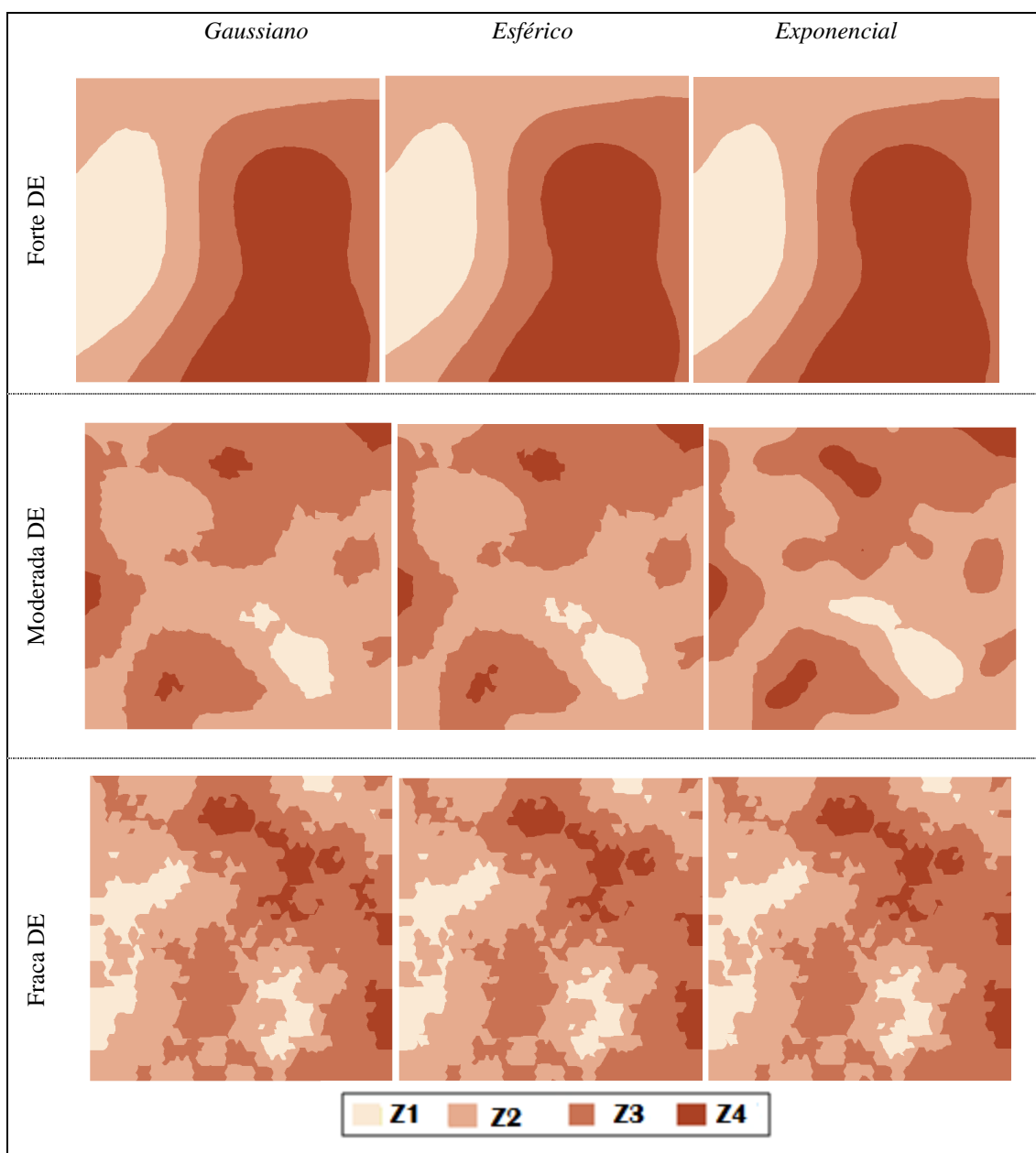


Figura 6: Mapas de zonas de manejo obtidos pela interpolação de krigagem para o *grid* G121 por diferentes estruturas de DE.

Os mapas apresentados na Figura 7 são para os casos onde temos um grid de densidade moderada, *grid* G256, neste caso é observado um mesmo comportamento apresentado para o *grid* G121. Onde visualmente não observa-se grandes diferenças entre os mapas com mesma estrutura de DE.

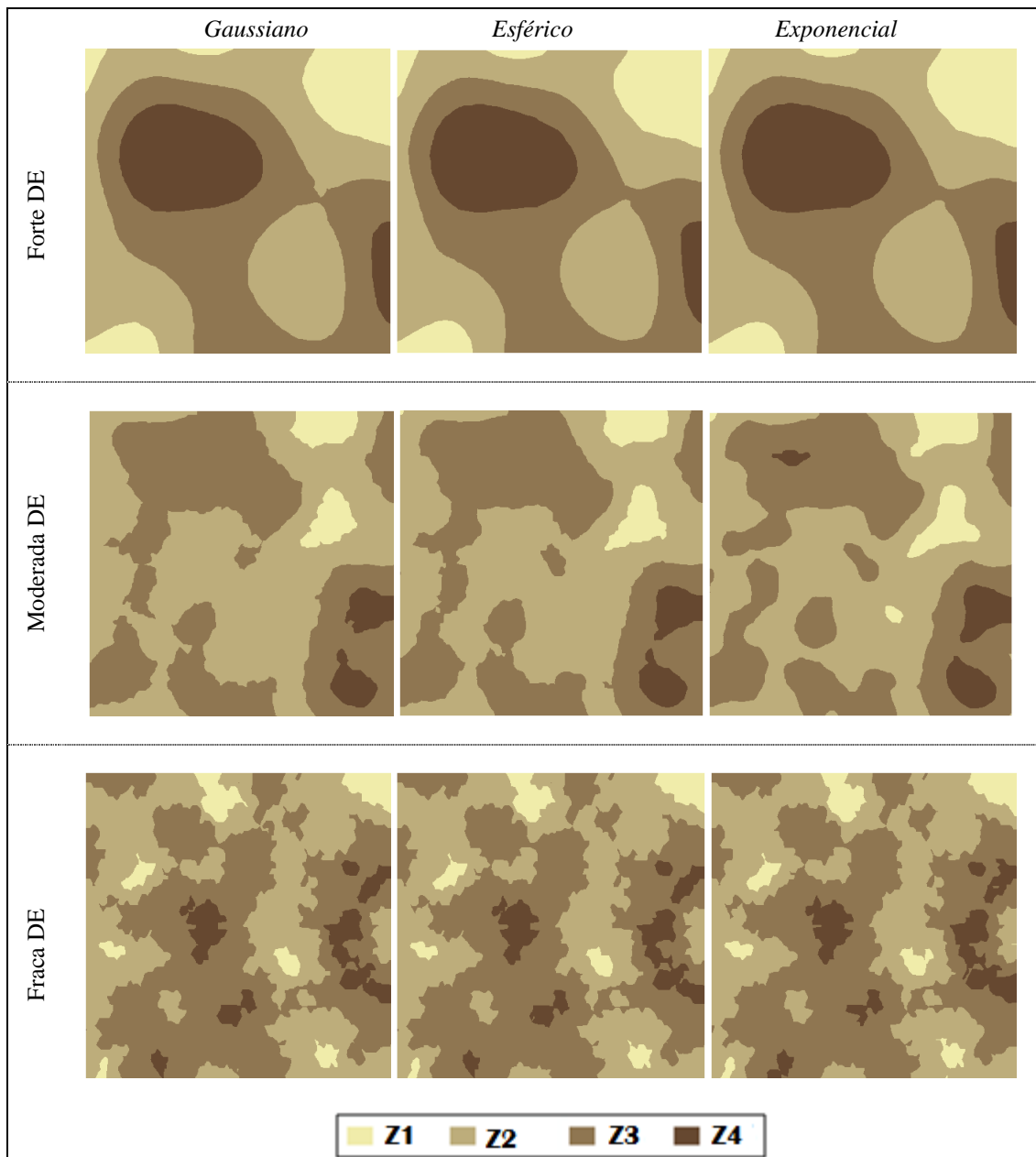


Figura 7: Mapas de zonas de manejo obtidos pela interpolação de krigagem para o *grid* G256 por diferentes estruturas de DE.

Na Figura 8 observa-se os mapas temáticos de zonas de manejos gerados para os casos onde temos *grids* de amostragens com alta densidade de pontos, o *grid* G961. Visualiza-se uma leve diferença entre os mapas gerados pelo modelo

Gaussiano e os outros dois modelos no caso de forte DE, para os casos com moderada e fraca DE a diferença é bastante sutil.

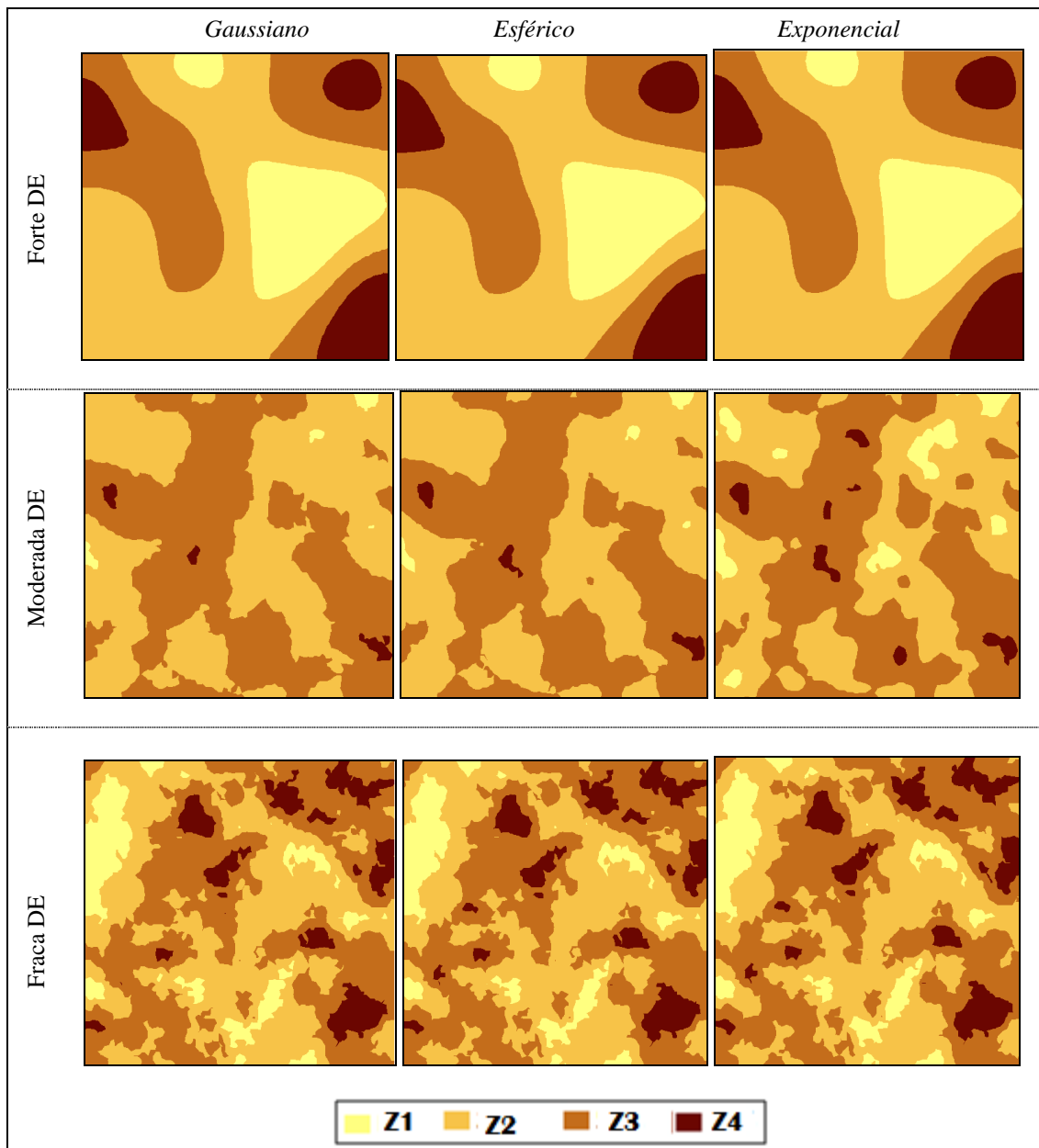


Figura 8: Mapas de zonas de manejo obtidos pela interpolação de krigagem para o *grid* G961 por diferentes estruturas de DE.

Através da análise visual dos mapas não é possível quantificar a diferença existente entre os mapas temáticos de zonas de manejo gerados por diferentes modelos de dependência espacial. Para realizar a comparação entre os mapas e quantificar esta diferença foi necessário converter os mapas temáticos de zonas de maneios em mapas tipo raster ou mapa de pixels, e através destes comparar pixel a pixel os mapas.

Após a geração dos mapas raster (Anexo III) de cada mapa temático de zona de manejo foi realizada a comparação dos mesmos através do módulo Errmat do software IDRISI, que gera uma matriz de erros e oferece vários resultados estatísticos que indicam se o resultado da classificação é o esperado ou não.

4.5 Comparação entre os Mapas Temáticos de Zonas de Manejo

Após a geração dos mapas raster (Anexo III) de cada mapa de zona de manejo foi realizada a comparação dos mesmos.

As comparações foram realizadas através do Índice Kappa (k) de concordância (COHEN, 1960), com o objetivo de verificar se existe uma concordância real entre os mapas, ou se diferem devido a alguma variação casual. Através do software IDRISI obteve-se a matriz de erros necessária para calcular este índice.

O Índice Kappa é utilizado para comparação de dois mapas, onde um é chamado de mapa de referência (Verdadeira) e o outro chamado de mapa de classificação. Neste estudo considerou-se como sendo mapa de referência aquele gerado a partir do modelo igual ao utilizado na simulação. Nos casos onde há forte DE o mapa de referência é o gerado a partir do modelo Gaussiano; quando há uma moderada DE tem-se o mapa gerado a partir do modelo Esférico e para o caso onde há fraca DE definiu-se como mapa de referência aquele gerado pelo modelo exponencial.

Visando testar estatisticamente os resultados obtidos, e verificar a significância do Índice Kappa foi realizado o teste Z a um nível de significância de 1%, calculado através do valor e variância do Índice Kappa (MA E REDMOND, 1995). Os resultados da comparação dos mapas estão apresentados na Tabela 14.

Tabela 14: Apresentação dos resultados do Índice Kappa(k) de concordância e teste Z segundo *grid* de amostragem e estrutura de dependência espacial.

Modelos Comparados	\hat{k}	95% conf.	$\sqrt{\widehat{Var}(\hat{k})}$	Z	p-valor
Grid G961					
Gau_Forte x Esf_Forte	0,9827	(0,9767; 0,9886)	0,0113	86,9646	< 0,0001
Gau_Forte x Exp_Forte	0,9832	(0,9774; 0,9891)	0,0113	87,0088	< 0,0001
Esf_Mod x Exp_Mod	0,8222	(0,8025; 0,8418)	0,0121	67,9504	< 0,0001
Esf_Mod x Gau_Mod	0,9348	(0,9223; 0,9474)	0,0122	76,6229	< 0,0001
Exp_Fraco x Esf_Fraco	0,9941	(0,9905; 0,9978)	0,0123	80,8211	< 0,0001
Exp_Fraco x Gau_Fraco	0,9776	(0,9705; 0,9847)	0,0123	79,4796	< 0,0001
Grid G256					
Gau_Forte x Esf_Forte	0,9667	(0,9596; 0,9738)	0,0114	84,7982	< 0,0001
Gau_Forte x Exp_Forte	0,9671	(0,9600; 0,9741)	0,0114	84,8333	< 0,0001
Esf_Mod x Exp_Mod	0,9089	(0,8968; 0,9210)	0,0119	76,3781	< 0,0001
Esf_Mod x Gau_Mod	0,9164	(0,9048; 0,9281)	0,012	76,3666	< 0,0001
Exp_Fraco x Esf_Fraco	0,9415	(0,9318; 0,9513)	0,0119	79,1176	< 0,0001
Exp_Fraco x Gau_Fraco	0,9371	(0,9270; 0,9472)	0,0119	78,7479	< 0,0001
Grid G121					
Gau_Forte x Esf_Forte	0,9816	(0,9763; 0,9870)	0,0145	67,6965	< 0,0001
Gau_Forte x Exp_Forte	0,9885	(0,9843; 0,9927)	0,0145	68,1724	< 0,0001
Esf_Mod x Exp_Mod	0,8515	(0,8364; 0,8666)	0,0147	57,9251	< 0,0001
Esf_Mod x Gau_Mod	0,9352	(0,9248; 0,9455)	0,0147	63,6190	< 0,0001
Exp_Fraco x Esf_Fraco	0,9721	(0,9650; 0,9792)	0,0137	70,9562	< 0,0001
Exp_Fraco x Gau_Fraco	0,9457	(0,9359; 0,9555)	0,0137	69,0292	< 0,0001

Gau_Forte: Modelo Gaussiano ajustado a dados com forte DE; Esf_Forte: Modelo Esférico ajustado a dados com forte DE; Exp_Forte: Modelo Exponencial ajustado a dados com forte DE; Gau_Mod: Modelo Gaussiano ajustado a dados com Moderada DE; Esf_Mod: Modelo Esférico ajustado a dados com Moderada DE; Exp_Mod: Modelo Exponencial ajustado a dados com Moderada DE; Gau_Fraco: Modelo Gaussiano ajustado a dados com fraca DE; Esf_Fraco: Modelo Esférico ajustado a dados com fraca DE; Exp_Fraco: Modelo Exponencial ajustado a dados com fraca DE.

Analisando os resultados apresentados na Tabela 14, observa-se que os valores dos Índices Kappa(k) na maioria dos casos foi superior a 0,80, sendo classificado assim, como uma concordância muito boa entre os mapas comparados. Na realização do teste Z, rejeitou-se a hipótese nula ($H_0: k=0$) em todos os casos, podendo-se concluir que há concordância direta e significativa entre os mapas de classificação e os de referência em todos os casos.

Este resultado é intrigante, porque mostra que não importa muito o modelo ajustado, pois as krigagens foram similares, (considerando zonas de manejo). Provavelmente, em determinadas áreas de aplicação das krigagens qualquer diferença nos mapas poderá ser significativa, mas não é o caso deste estudo.

5. CONCLUSÕES

A partir da comparação dos mapas temáticos de zonas de manejo obtidos através de diferentes modelos de dependência espacial pode-se concluir que os mesmos não influenciam na definição desse tipo de mapa. Assim sendo, não importa o modelo ajustado, pois as krigagens serão similares, (considerando zonas de manejo).

O tamanho da malha amostral neste caso não interferiu na modelagem e criação de zonas de manejo, isso não pode ser generalizado para todos os casos, por exemplo na análise espacial de células humanas, uma vez que dependendo do tipo de malha analisada qualquer diferença pode ser importante.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- AGRESTI, A., *Categorical Data Analysis*, 2nd edition. John Wiley and Sons, 2002.
- AMADO, T.J.C.; PONTELLI, C.B.; SANTI A.L.; VIANA, J.H.M.V.; SULZBACH, L.A.S. Variabilidade espacial e temporal da produtividade de culturas sob sistema plantio direto. *Pesq. agropec. bras., Brasília, v.42, n.8, p.1101-1110, ago. 2007.*
- AMARAL, M. V. F.; SOUZA, A. L.; SOARES, V. P.; SOARES, C. P. B. ET AL. Avaliação e comparação de métodos de classificação de imagens de satélites para mapeamento de estádios de sucessão florestal. *Revista Árvore*, Viçosa – MG, v.33, n.3, p.575-582, 2009.
- [ARCGIS, 2009] ArcGIS. **Performing cross-validation and validation**. Fevereiro de 2009. Disponível em: <http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=Performing_cross-validation_and_validation>. Acessado em 10 de junho de 2012.
- BAZZI, C. L.; SOUZA, E. G.; OPAZO, M. A. U. ; ET AL. Influência da distância entre passadas de colhedora equipada com monitor de colheita na precisão dos mapas de produtividade na cultura do milho. *Eng. Agríc.*, Jaboticabal, V.28, N.2, P.355-363, Abr./Jun. 2008.
- BERNARDI, J. V. E. et al. Um estudo de impacto ambiental utilizando análises estatísticas espacial e multivariada. *Holos Environment*, v.1, n.2, p.162-172, 2001.
- BOREM, A., ET AL. **Agricultura de Precisão**. Viçosa: Editora da Universidade Federal de Viçosa, 2000.
- 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 Science Society of America Journal*, n.58, p.1501-1511, 1994.
- COLLINGWOOD, A.; FRANKLIN, S.E.; GUO, X.; STENHOUSE, G. A Medium-Resolution Remote Sensing Classification of agricultural areas in alberta grizzly bear habitat. *Canadian Journal of Remote Sensing*, Vol. 35, No. 1, pp. 23-36, 2009.

CARVALHO, J. R. P.; VIEIRA, S. R.; MORAN, R. C. C. P. Como avaliar similaridade entre mapas de produtividade. **Embrapa Informática Agropecuária**, Campinas, 24p. 2001.

CONGALTON, R. G. A review of assessing the accuracy of classification of remotely sensed data. **Remote Sensing of Environment**, v. 37, n. 1, p. 35-46, 1991.

CONGALTON, R.G, GREEN, K.; Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, 2nd edition. 2009.

CONGALTON, R.G.; MEAD, R.A. A quantitative method to test for consistency and correctness in photo interpretation. **Photogrammetric Engineering & Remote Sensing**, Falls Church, v.49, n.1, p.69-74, 1983.

COHEN, J. (1960). A coefficient of agreement for nominal scales. **Educ. Psychol. Measurement** 20(1): 37-46.

DOBERMANN, A.; PING, J.L. Geostatistical integration of yield monitor data and remote sensing improves yield maps. **Agronomy Journal**, v.96, n.1, p 385-397, 2004.

DOERGE, T.A. Management zone concept. **Site-Specific Management Guidelines**, n. 2, p.1-4, 1999. Disponível em: <<http://ppi-far.org/ssmg>>. Acesso em: 10 set. 2011.

FATORGIS. **Agricultura de precisão: A tecnologia de GIS/GPS chega às fazendas**. Curitiba, 1998. Disponível na Internet. <http://www.master.inf.br/Agro2000/agripres.htm> em 19 março 2012.

GALPARSORO, L. U. FERNÁNDEZ, S. P. **Medidas de concordância: el índice Kappa**. 2001. Disponível em: <http://www.fisterra.com/mbe/investiga/kappa/kappa.htm>>. Acesso em: 15 out. 2011.

GUERRA, P.A.G. **Geoestatística operacional**. Ministério das Minas e Energia/ Departamento de Produção Mineral, Brasília, 1988. 145p. Jorge, J.A. Solo: manejo e adubação. São Paulo: Editora Polígono, 1969, 146p.

GUIMARÃES, E.C. **Geoestatística básica e aplicada**. Uberlândia: Faculdade de Matemática-Universidade Federal de Uberlândia, 2004. 77p. (Apostila).

KRIGE, D.G. "A study of the relationship between development values and recovery grades on the South African goldfields", **Journal of the South African Institution of Mining and Metallurgy**, No.61, pp.317-331, 1959.

KYVERYGA, P.M.; BLACKMER, T.M. Categorical analysis of spatial variability in economic yield response of corn to nitrogen fertilization. **Agronomy Journal**, Vol. 103, Issue 3, pp. 795-804, 2011.

KUZYAKOVA, I.F.; ROMANENKOV, V.A. & KUZYAKOV, Ya.V. Geostatistics in soil agrochemical studies. *Eur. Soil Sci.*, 34:1011-1017, 2001.

MA, Z.; REDMOND, R. L.; Tau coefficient for accuracy assessment of classification of remote sensing data. **Photogrammetric Engineering and Remote Sensing**, v.61, n .4, pag. 435- 439, apr. 1995.

MATHERON, G. **Les Variables Regionalisées et leur Estimation**. Paris : Masson, 1965.

MATHERON, G. **Principles of geostatistics**. *Economic Geology*. 1963, Vol. 58, pp. 1246-1266.

MELLO, J. M.; BATISTA, J. L. F.; Oliveira, M. S.; Ribeiro Júnior, P. J. Ajuste e seleção de modelos espaciais de semivariograma visando a estimativa volumétrica de *Eucalyptus grandis*. *Scientia Florestalis*, n. 69, p. 25-37, dez. 2005a.

MOHD HASMADI, I.; KAMARUZAMAN, J. Satellite Data Classification Accuracy Assessment Based from Reference Dataset. **International Journal of Computer and Information Engineering**, 2:6, 2008.

MOLIN, J.P. Definição de Unidades de manejo a partir de mapas de Produtividade. **Engenharia Agrícola**, Jaboticabal, v. 22 n. 1, p.83-92, 2002.

MOLIN, J.P. **Agricultura de precisão. Parte II: Diagnósticos, aplicação localizada e considerações agronômicas**. *Engenharia Agrícola*, Jaboticabal, v.17, n.2, p.108-121, 1997.

SANTOS, G. R.; OLIVEIRA, M. S.; LOUZADA, J. M.; SANTOS, A. M. R. T. Krigagem Simples Versus Krigagem Universal: Qual O Preditor Mais preciso? **Revista Energia na Agricultura**, Botucatu, v.26, n.2, p. 49-55, 2011.

ISAAKS, E.H., SRIVASTAVA, M. An introduction to applied geostatistics. **Oxford Univ. Press.**, New York, 1989.

Statistical Analysis System - SAS/STAT. 2009. SAS user's guide. Version 9.2, Cary, NC

SUSZEK, G., SOUZA, E. G.; URIBE-OPAZO, M. A., NOBREGA, L. H. P. Determination of management zones from normalized and standardized equivalent productivity maps in the soybean culture. *Engenharia Agrícola*, Jaboticabal, v.31, n.5, p.895-905, 2011.

TSCHIEDEL, M.; FERREIRA, M. F. **Introdução à agricultura de precisão: conceitos e vantagens**, Ciência Rural, Santa Maria, v.32, n.1, p.159-163, 2002.

VALENTE, D.S. M., QUEIROZ, D. M., Pinto, F.A.C.; Santos, N.T, Santos, F.L. Definition of management zones in coffee production fields based on apparent soil electrical. *Scientia Agrícola*, Vol.69, n.3, p.173-179, 2012.

VIEIRA, S. R. Geoestatística em estudos de variabilidade espacial do solo. In. NOVAES, R. F.; ALVAREZ V., V. H.; SCHAEFER, C. E G. R. **Tópicos em ciências do solo**. Viçosa, MG: Sociedade Brasileira de Ciência do Solo, 2000. v.1. p.2-54.

VIEIRA, S. R; MOLIN, J. P., **Spatial variability of soil fertility for precision agriculture**. In: European Conference on Precision Agriculture, 3., 2001, Montpellier, Anais...Montpellier: Ecole Nationale Superiure Agronomique, 2001.

XAVIER, A. C.; CECÍLIO, R. A.; LIMA, J. S. S. Módulos em *Matlab* para interpolação espacial pelo método de krigagem ordinária e do inverso da distância. **Revista Brasileira de Cartografia**, No 62/01, 2010.

WEBSTER, R., OLIVER M.A. **Geostatistics for Environmental Scientists. Second Edition.**Wiley, Chichester, 325p. 2007.

WEBSTER, R.; OLIVER, M. A. Statistical methods in soil and land resource survey. **Oxford University Press**, 1990, 316 p.

ZIMBACK, C. R. L. **Análise espacial de atributos químicos de solos para fins de mapeamento da fertilidade do solo**. 2001. 114f. Tese (Livre-Docência) –Faculdade de Ciências Agrônômicas, Unesp, Botucatu, 2001.

ANEXOS

ANEXO I

Códigos utilizados nas simulações no SAS

- Simulações para Grid com 961 pontos

```
TITLE;
TITLE1 'GAUSSIANO';
TITLE2 'IDE FORTE';

PROC SIM2D OUTSIM=IDE_FORTE;
SIMULATE
  NUMREAL=150                /*NUMBER OF REALIZATIONS*/
  SEED=123456789            /*SEED*/
  SCALE=3                    /*PARTIAL SILL*/
  RANGE=40.0                 /*RANGE*/
  NUGGET=1E-8                /*NUGGET EFFECT*/
  FORM=GAUSS                 /*MODELS: GAUSS, EXPONENTIAL E SPHERICAL */
  ;
  MEAN 0.00;                 /*MEAN*/
  GRID
    Y=0 TO 150 BY 5.0
    X=0 TO 150 BY 5.0;
RUN;

proc variogram data=IDE_FORTE outvar=sv_Forte;
  by _iter_;
  compute lagd=5 maxlag=30;
  coord xc=gxc yc=gyc;
  var SValue;
run ;

TITLE;
TITLE1 'ESFERICO';
TITLE2 'IDE MODERADO';

PROC SIM2D OUTSIM=IDE_MODERADO;
SIMULATE
  NUMREAL=150                /*NUMBER OF REALIZATIONS*/
  SEED=123456789            /*SEED*/
  SCALE=3                    /*PARTIAL SILL*/
  RANGE=40.0                 /*RANGE*/
  NUGGET=1.5                 /*NUGGET EFFECT*/
  FORM=SPHERICAL            /*MODELS: GAUSS, EXPONENTIAL E SPHERICAL
*/
  ;
  MEAN 0.00;                 /*MEAN*/
  GRID
    Y=0 TO 150 BY 5.0
    X=0 TO 150 BY 5.0;
RUN;

proc variogram data=IDE_MODERADO outvar=sv_MODERADO;
  by _iter_;
  compute lagd=5 maxlag=40;
  coord xc=gxc yc=gyc;
  var SValue;
  length SValue$3;
run ;
```

```

TITLE;
TITLE1 'EXPONENTIAL';
TITLE2 'IDE FRACO';

PROC SIM2D OUTSIM=IDE_FRACO;
  SIMULATE
    NUMREAL=150                /*NUMBER OF REALIZATIONS*/
    SEED=123456789            /*SEED*/
    SCALE=1                    /*PARTIAL SILL*/
    RANGE=40.0                 /*RANGE*/
    NUGGET=3                    /*NUGGET EFFECT*/
    FORM=EXPONENTIAL           /*MODELS: GAUSS, EXPONENTIAL E
SPHERICAL */
    ;
    MEAN 0.00;                 /*MEAN*/
    GRID
      Y=0 TO 150 BY 5.0
      X=0 TO 150 BY 5.0;
  RUN;

```

```

proc variogram data=IDE_FRACO outvar=sv_fraco;
  by _iter_;
  compute lagd=5 maxlag=40;
  coord xc=gxc yc=gyc;
  var SValue;
  run ;

```

- **Simulações para Grid com 256 pontos**

```

TITLE;
TITLE1 'GAUSSIANO';
TITLE2 'IDE FORTE';

PROC SIM2D OUTSIM=IDE_FORTE_256p;
  SIMULATE
    NUMREAL=150                /*NUMBER OF REALIZATIONS*/
    SEED=123456789            /*SEED*/
    SCALE=3                    /*PARTIAL SILL**/*antgo era 40*/
    RANGE=40.0                 /*RANGE**/*antigo 50*/
    NUGGET=1E-8                /*NUGGET EFFECT*/
    FORM=GAUSS                 /*MODELS: GAUSS, EXPONENTIAL E SPHERICAL */
    ;
    MEAN 0.00;                 /*MEAN*/
    GRID
      Y=0 TO 150 BY 10.0
      X=0 TO 150 BY 10.0;
  RUN;

```

```

proc variogram data=IDE_FORTE_256p outvar=sv_Forte_256p;
  by _iter_;
  compute lagd=10 maxlag=30;
  coord xc=gxc yc=gyc;
  var SValue;
  run ;

```

```

TITLE;

```



```

TITLE1 'ESFERICO';
TITLE2 'IDE MODERADO';

PROC SIM2D OUTSIM=IDE_MODERADO_256p;
SIMULATE
  NUMREAL=150           /*NUMBER OF REALIZATIONS*/
  SEED=123456789       /*SEED*/
  SCALE=3              /*PARTIAL SILL*/
  RANGE=40.0          /*RANGE*/
  NUGGET=1.5          /*NUGGET EFFECT*/
  FORM=SPHERICAL      /*MODELS: GAUSS, EXPONENTIAL E SPHERICAL
*/
;
MEAN 0.00;           /*MEAN*/
GRID
  Y=0 TO 150 BY 10.0
  X=0 TO 150 BY 10.0;
RUN;

proc variogram data=IDE_MODERADO_256p outvar=sv_MODERADO_256p;
  by _iter_;
  compute lagd=10 maxlag=40;
  coord xc=gxc yc=gyc;
  var SValue;
  length SValue$3;
run ;

TITLE;
TITLE1 'EXPONENTIAL';
TITLE2 'IDE FRACO';

PROC SIM2D OUTSIM=IDE_FRACO_256p;
SIMULATE
  NUMREAL=150           /*NUMBER OF REALIZATIONS*/
  SEED=123456789       /*SEED*/
  SCALE=1              /*PARTIAL SILL*/
  RANGE=40.0          /*RANGE*/
  NUGGET=3            /*NUGGET EFFECT*/
  FORM=EXPONENTIAL    /*MODELS: GAUSS, EXPONENTIAL E
SPHERICAL */
;
MEAN 0.00;           /*MEAN*/
GRID
  Y=0 TO 150 BY 10.0
  X=0 TO 150 BY 10.0;
RUN;

proc variogram data=IDE_FRACO_256p outvar=sv_fraco_256p;
  by _iter_;
  compute lagd=10 maxlag=40;
  coord xc=gxc yc=gyc;
  var SValue;
run ;

```

- **Simulações para Grid com 121 pontos**

```

TITLE;
TITLE1 'GAUSSIANO';
TITLE2 'IDE FORTE';

PROC SIM2D OUTSIM=IDE_FORTE_121p;
SIMULATE
  NUMREAL=150                /*NUMBER OF REALIZATIONS*/
  SEED=123456789            /*SEED*/
  SCALE=3                    /*PARTIAL SILL*//*antgo era 40*/
  RANGE=20.0                 /*RANGE*//*antigo 50*/
  NUGGET=1E-8                /*NUGGET EFFECT*/
  FORM=GAUSS                 /*MODELS: GAUSS, EXPONENTIAL E SPHERICAL */
  ;
  MEAN 0.00;                 /*MEAN*/
  GRID
    Y=0 TO 50 BY 5.0
    X=0 TO 50 BY 5.0;
  RUN;

proc variogram data=IDE_FORTE_121p outvar=sv_Forte_121p;
  by _iter_;
  compute lagd=5 maxlag=14;
  coord xc=gxc yc=gyc;
  var SValue;
  run ;

TITLE;
TITLE1 'ESFERICO';
TITLE2 'IDE MODERADO';

PROC SIM2D OUTSIM=IDE_MODERADO_121p;
SIMULATE
  NUMREAL=150                /*NUMBER OF REALIZATIONS*/
  SEED=123456789            /*SEED*/
  SCALE=3                    /*PARTIAL SILL*/
  RANGE=15.0                 /*RANGE*/
  NUGGET=1.5                 /*NUGGET EFFECT*/
  FORM=SPHERICAL            /*MODELS: GAUSS, EXPONENTIAL E SPHERICAL
*/
  ;
  MEAN 0.00;                 /*MEAN*/
  GRID
    Y=0 TO 50 BY 5.0
    X=0 TO 50 BY 5.0;
  RUN;

proc variogram data=IDE_MODERADO_121p outvar=sv_MODERADO_121p;
  by _iter_;
  compute lagd=5 maxlag=14;
  coord xc=gxc yc=gyc;
  var SValue;
  length SValue$3;
  run ;

TITLE;
TITLE1 'EXPONENTIAL';
TITLE2 'IDE FRACO';

PROC SIM2D OUTSIM=IDE_FRACO_121p;
SIMULATE
  NUMREAL=150                /*NUMBER OF REALIZATIONS*/
  SEED=123456789            /*SEED*/

```

```
SCALE=1                /*PARTIAL SILL*/
RANGE=15.0             /*RANGE*/
NUGGET=3              /*NUGGET EFFECT*/
FORM=EXPONENTIAL      /*MODELS: GAUSS, EXPONENTIAL E SPHERICAL */
;
MEAN 0.00;            /*MEAN*/
GRID
  Y=0 TO 50 BY 5.0
  X=0 TO 50 BY 5.0;
RUN;

proc variogram data=IDE_FRACO_121p outvar=sv_fraco_121p;
  by _iter_;
  compute lagd=5 maxlag=14;
  coord xc=gxc yc=gyc;
  var SValue;
run ;
```

ANEXO II

Gráficos de validação cruzada

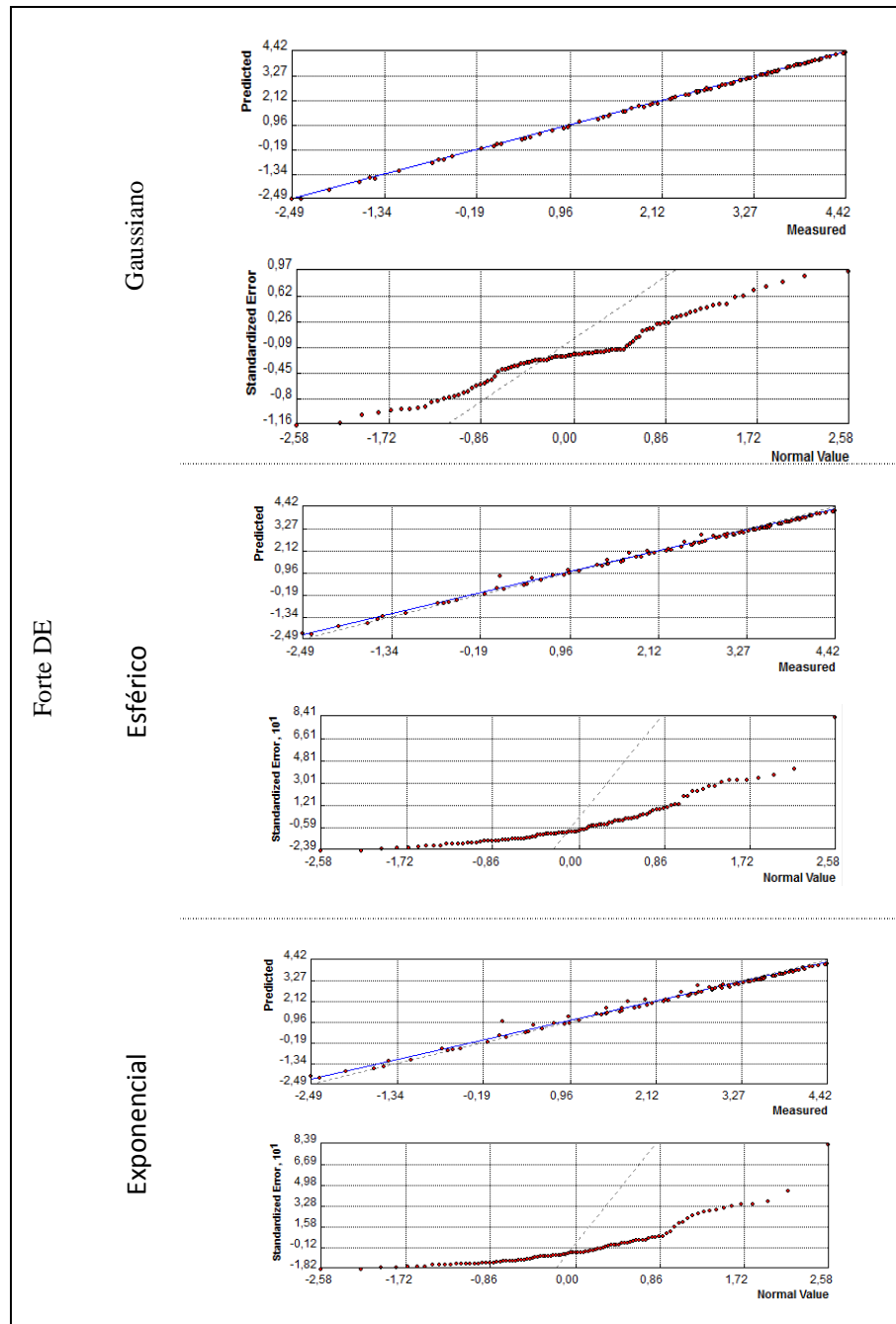


Figura III: Gráficos de validação cruzada para modelos ajustados a dados com forte DE e *gridG121*

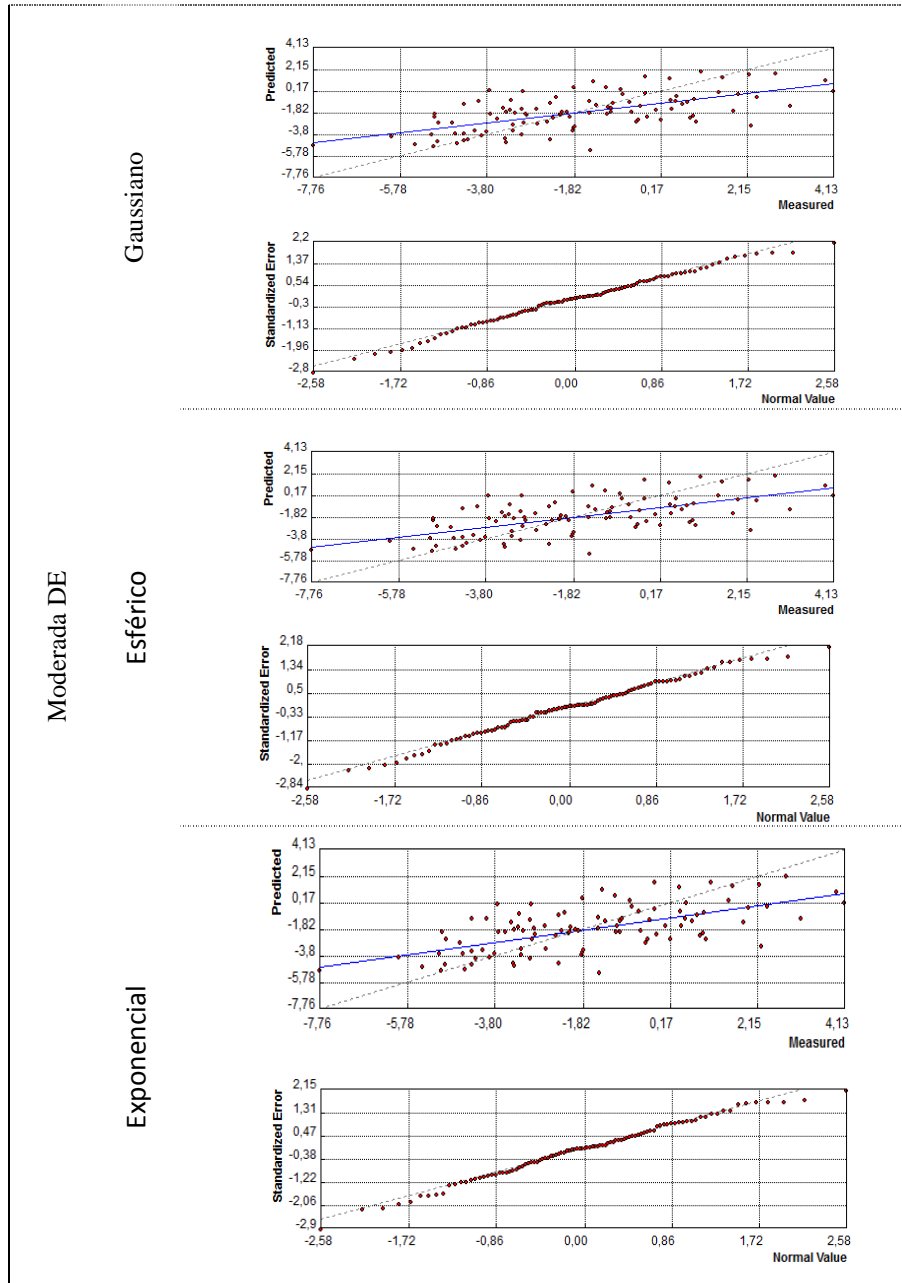


Figura 2II: Gráficos de validação cruzada para modelos ajustados a dados com moderada DE e *gridG121*

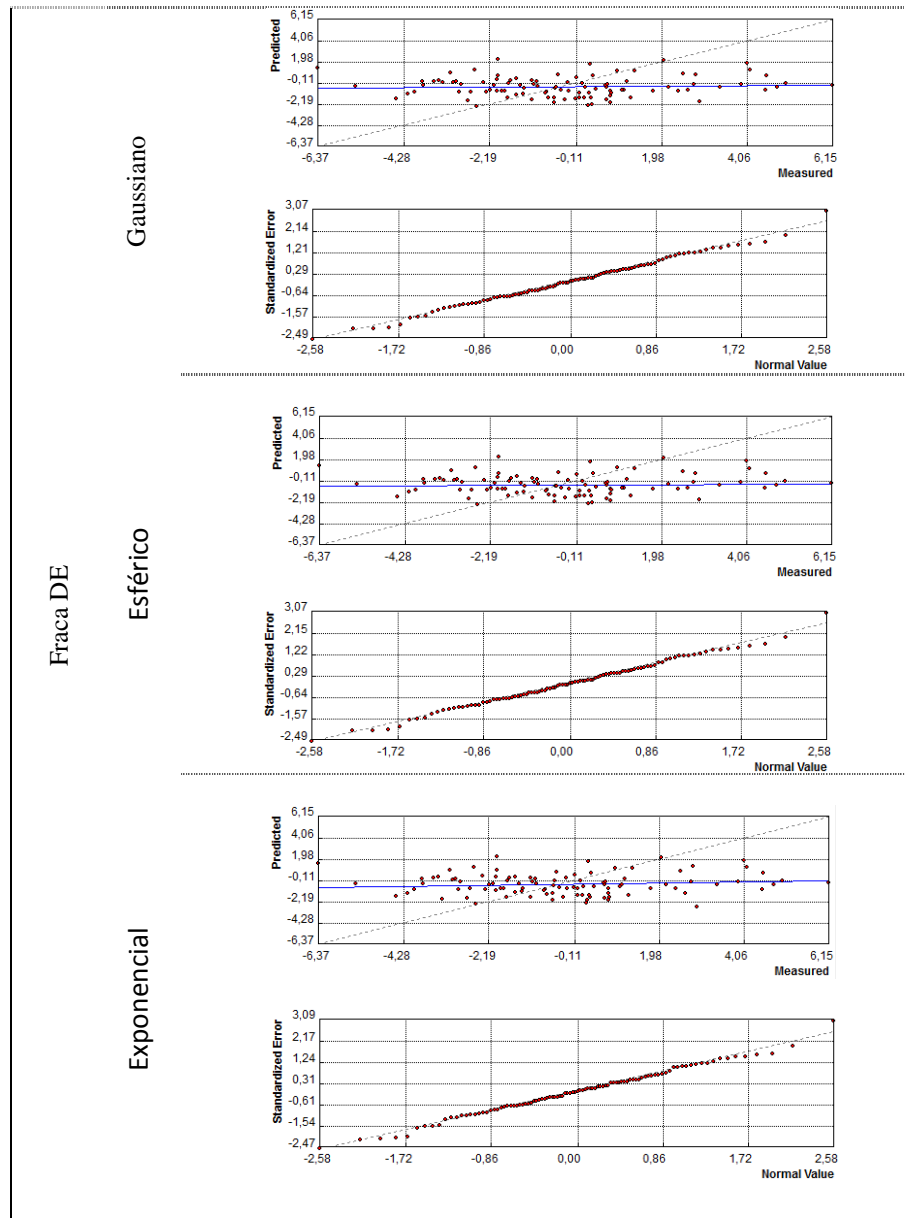


Figura 3II: Gráficos de validação cruzada para modelos ajustados a dados com fraca DE e *gridG121*

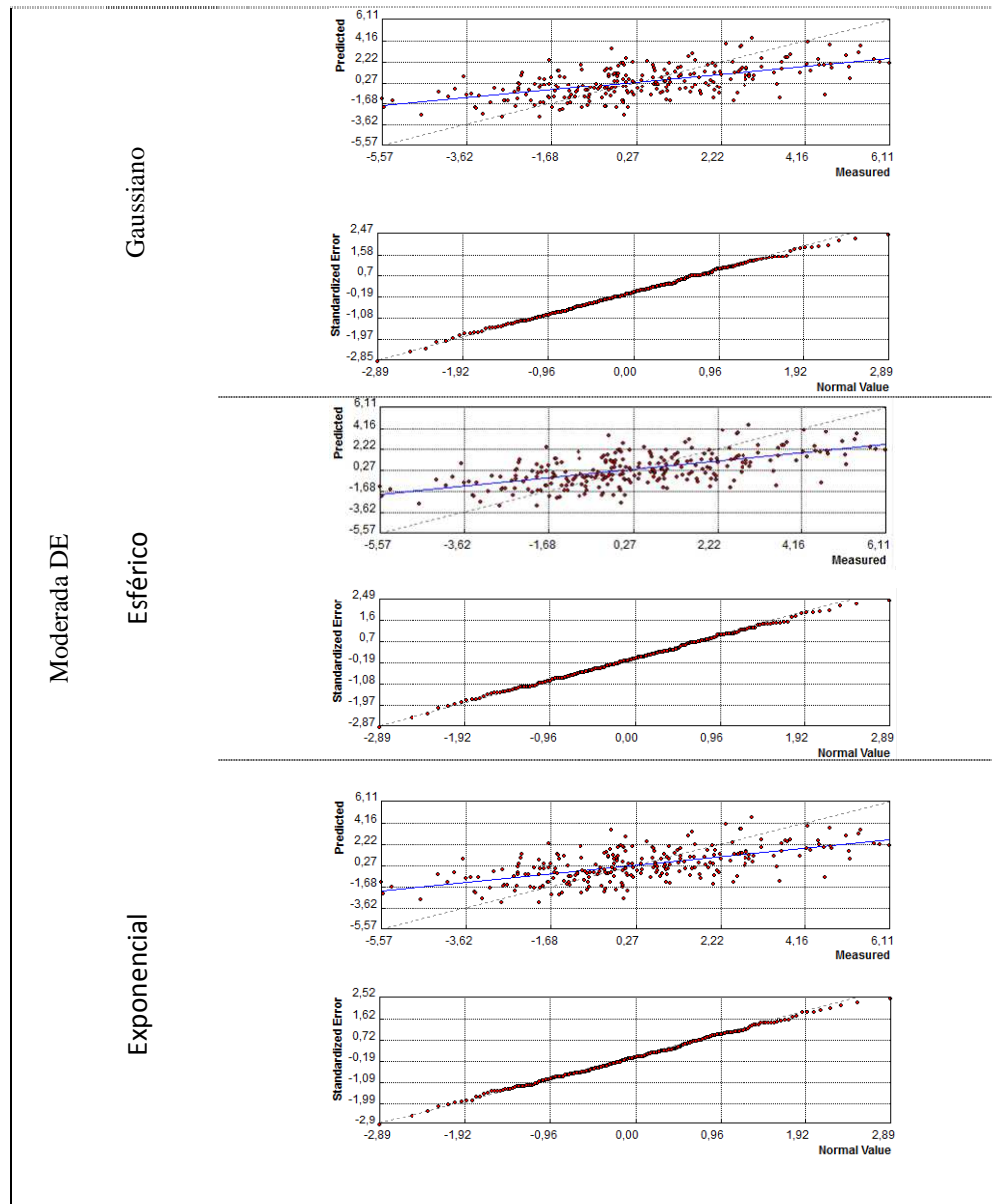


Figura 4II: Gráficos de validação cruzada para modelos ajustados a dados com moderada DE e *gridG256*

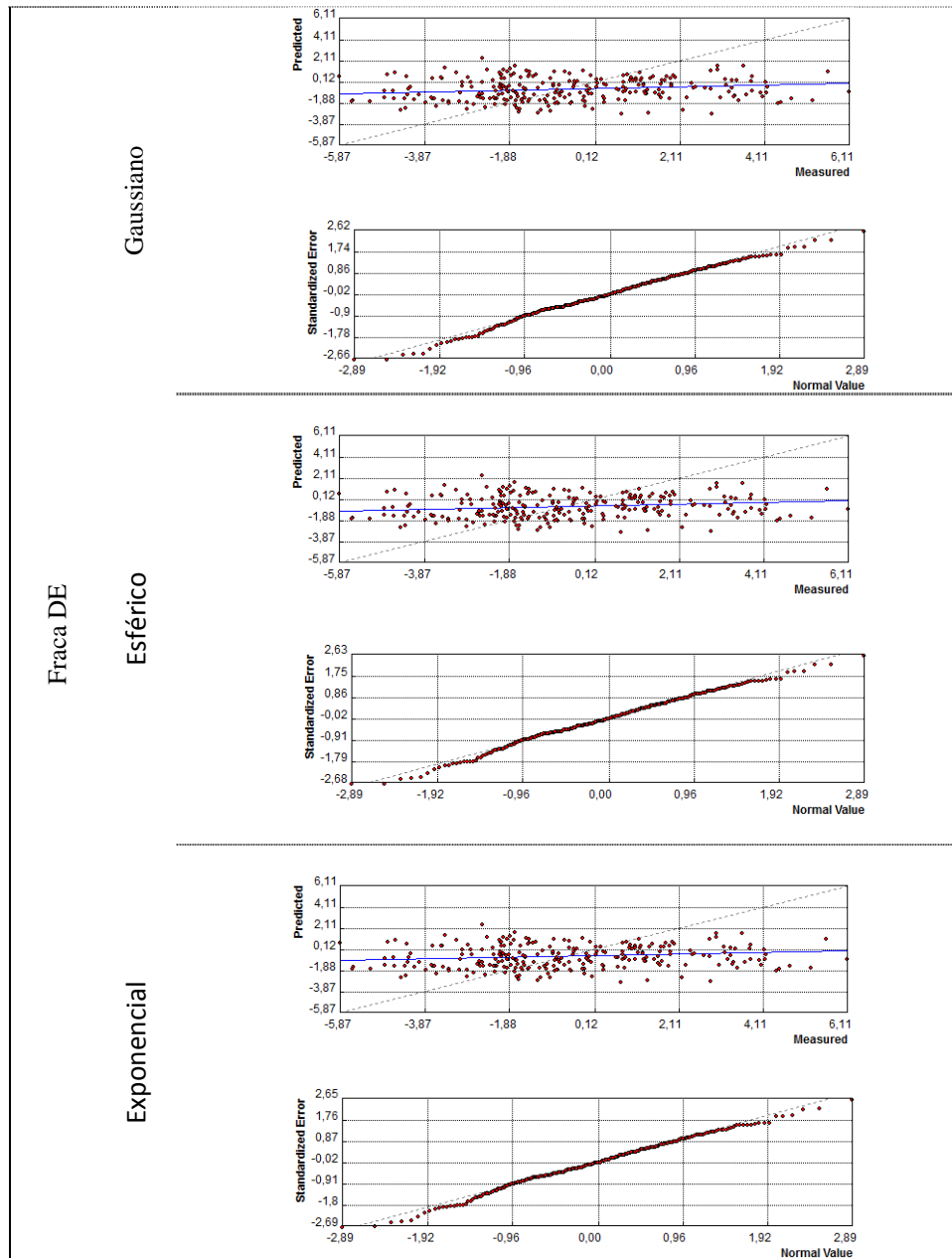


Figura 5II: Gráficos de validação cruzada para modelos ajustados a dados com fraca DE e *gridG256*

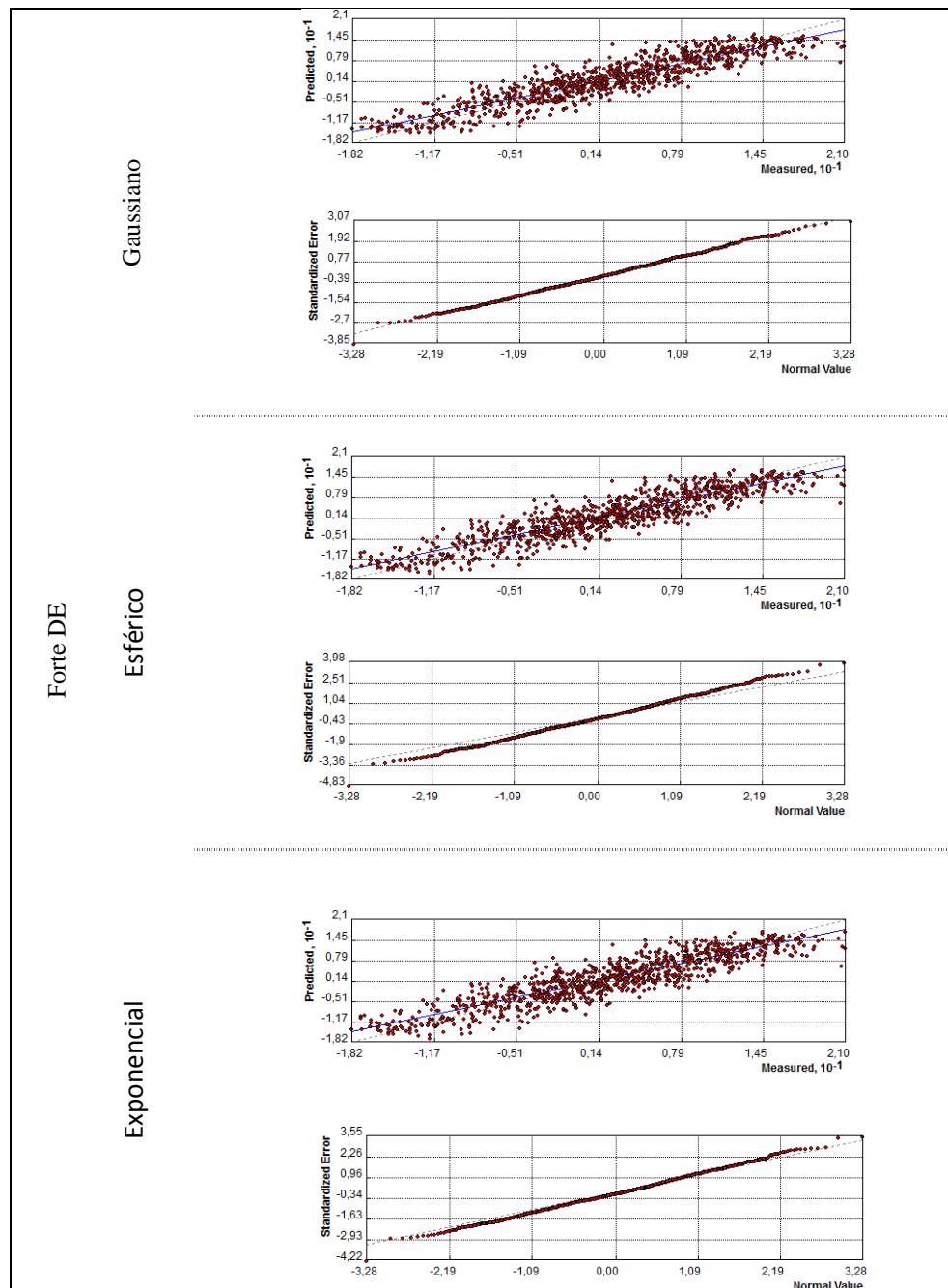


Figura 7II: Gráficos de validação cruzada para modelos ajustados a dados com forte DE e *grid G961*

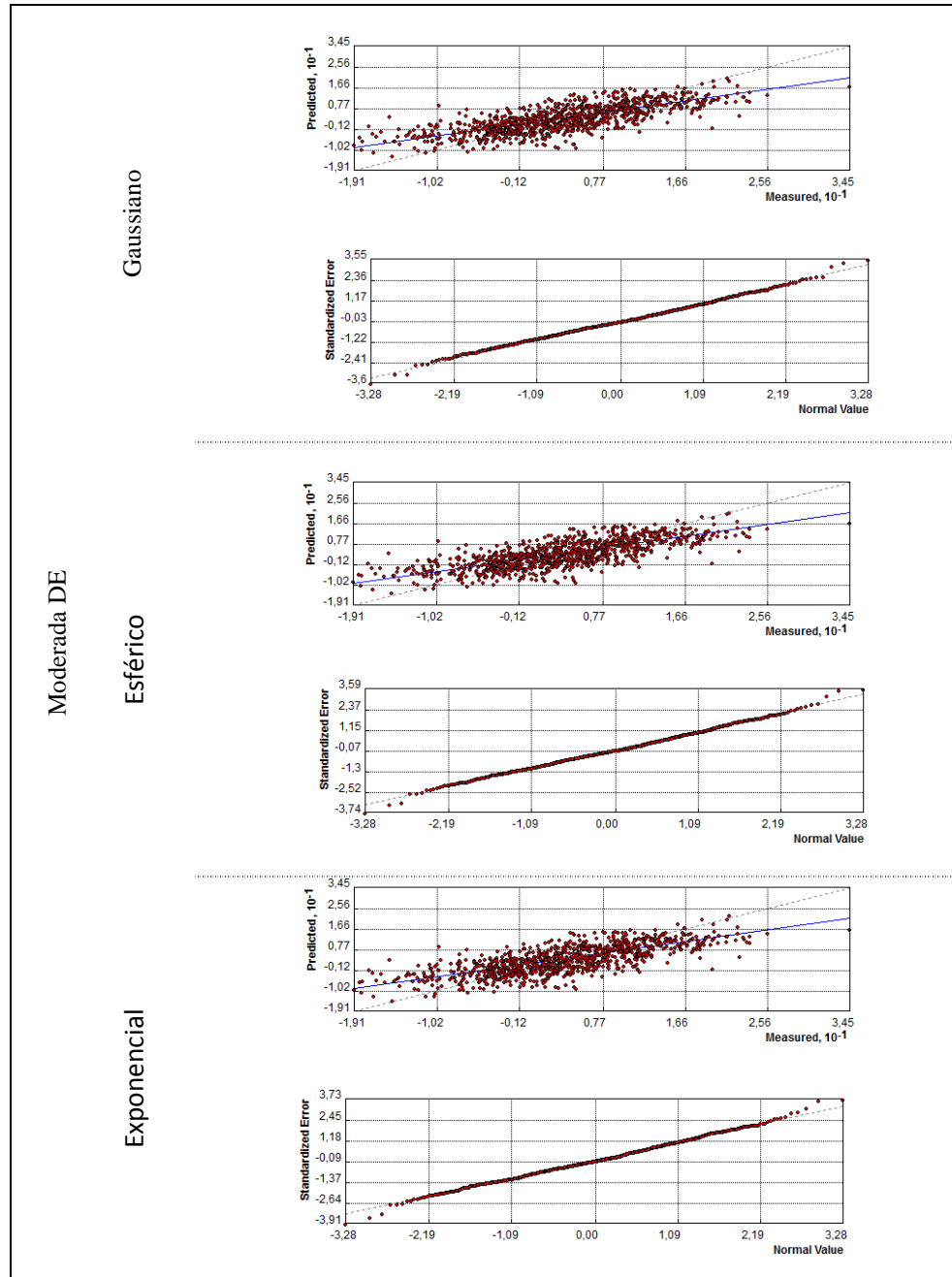


Figura 8II: Gráficos de validação cruzada para modelos ajustados a dados com moderada DE e *grid G961*

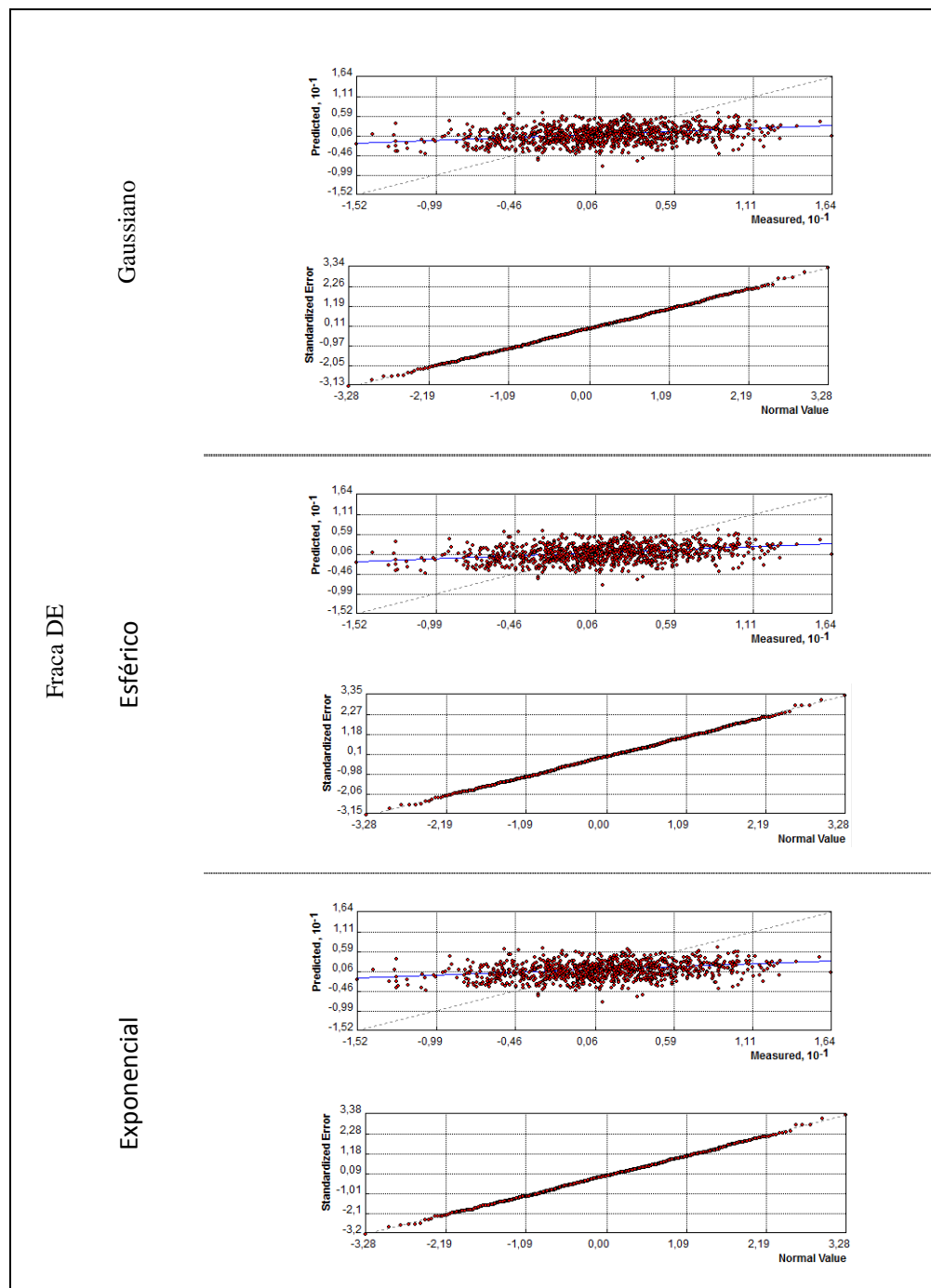


Figura 9II: Gráficos de validação cruzada para modelos ajustados a dados com fracca De DE e *grid G961*

ANEXO III

MAPAS DE ZONAS DE MANEJO DO TIPO RASTER

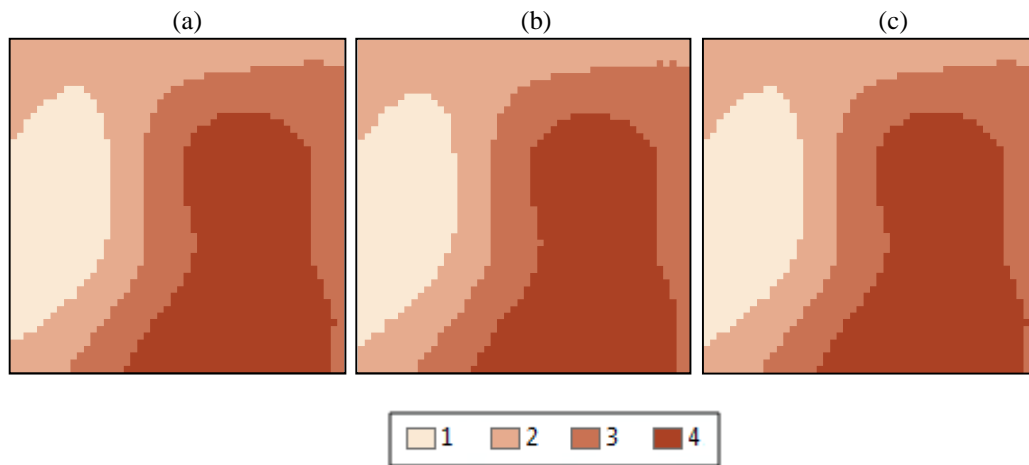


Figura 1III: Mapas de zonas de manejo rasterizados para dados que apresentam DE Forte e *grid* G121. (a) Esférico, (b) Gaussiano e (c) Exponencial.

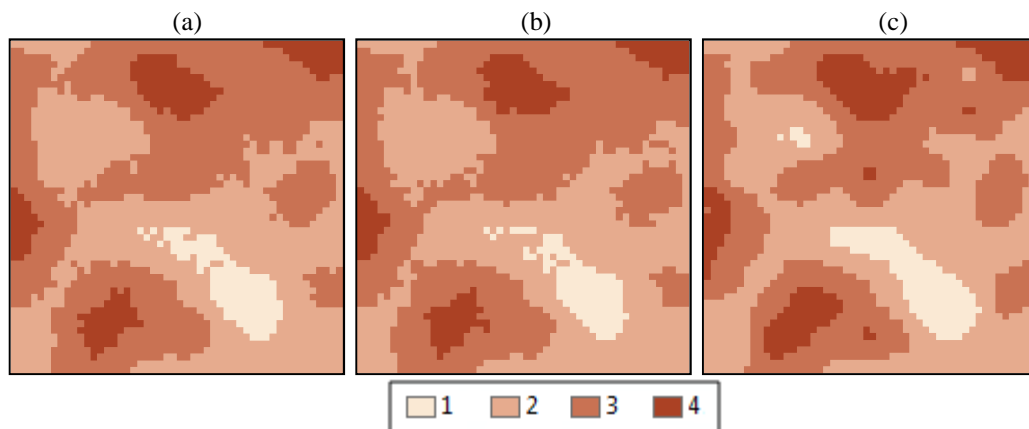


Figura 2III: Mapas de zonas de manejo rasterizados para dados que apresentam DE moderada e *grid* G121. (a) Esférico, (b) Gaussiano e (c) Exponencial.

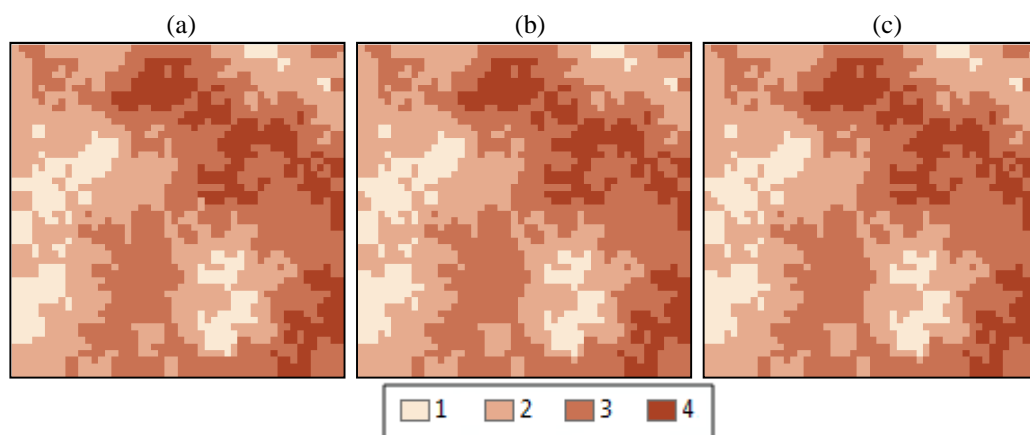


Figura 3III: Mapas de zonas de manejo rasterizados para dados que apresentam DE fraca e *grid* G121. (a) Exponencial, (b) Esférico e (c) Gaussiano.

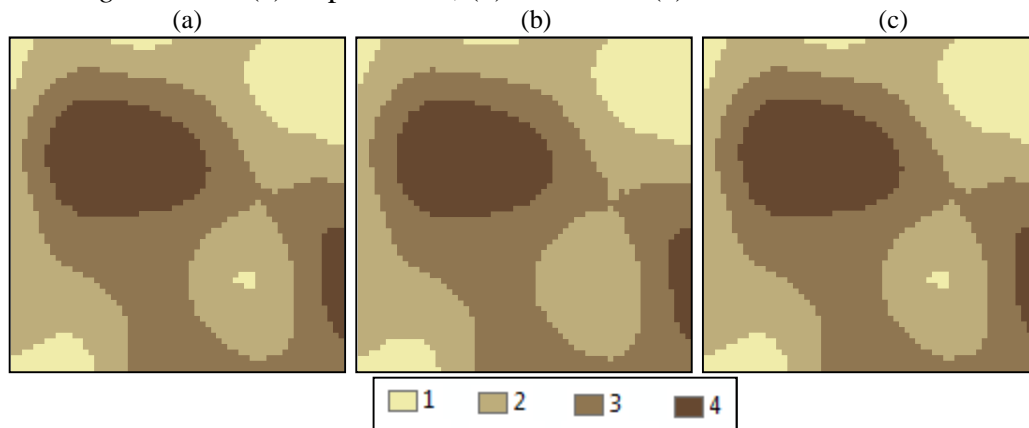


Figura 4III: Mapas de zonas de manejo rasterizados para dados que apresentam DE forte e *grid* G256. (a) Gaussiano, (b) Esférico e (c) Exponencial.

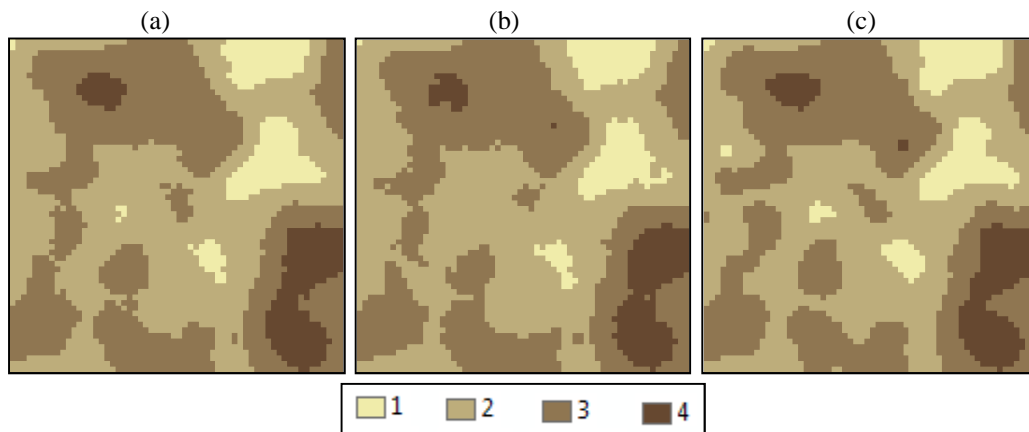


Figura 5B: Mapas de zonas de manejo rasterizados para dados que apresentam DE moderada e *grid* G256. (a) Esférico, (b) Gaussiano e (c) Exponencial.

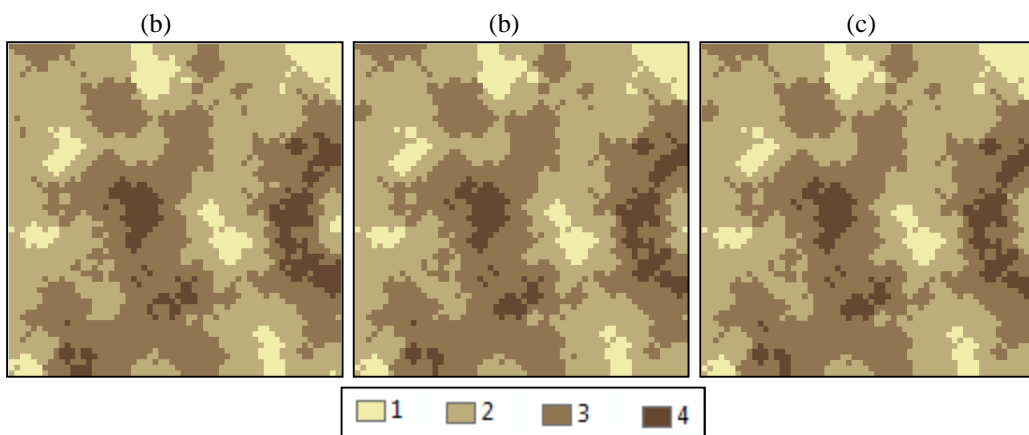


Figura 6III: Mapas de zonas de manejo rasterizados para dados que apresentam DE Fraco e *grid* G256. (a) Exponencial, (b) Esférico e (c) Gaussiano.

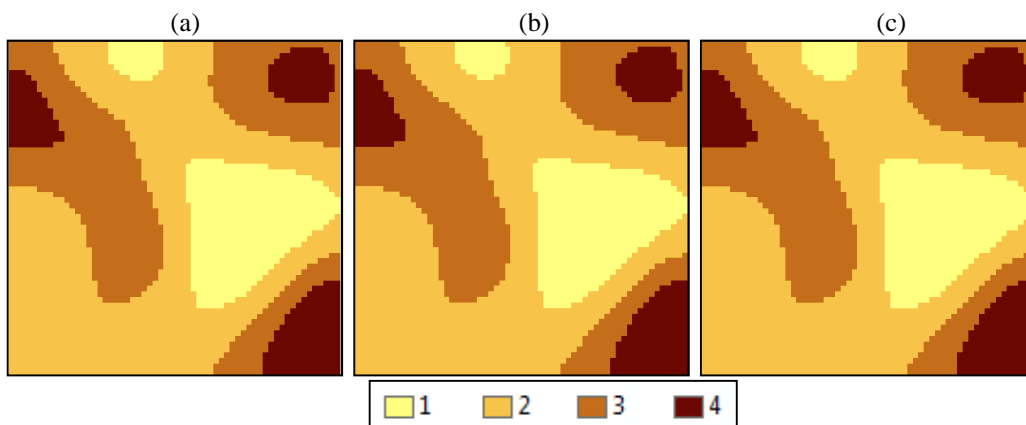


Figura 7III: Mapas de zonas de manejo rasterizados para dados que apresentam DE forte e *grid* G961. (a) Gaussiano, (b) Esférico e (c) Exponencial.

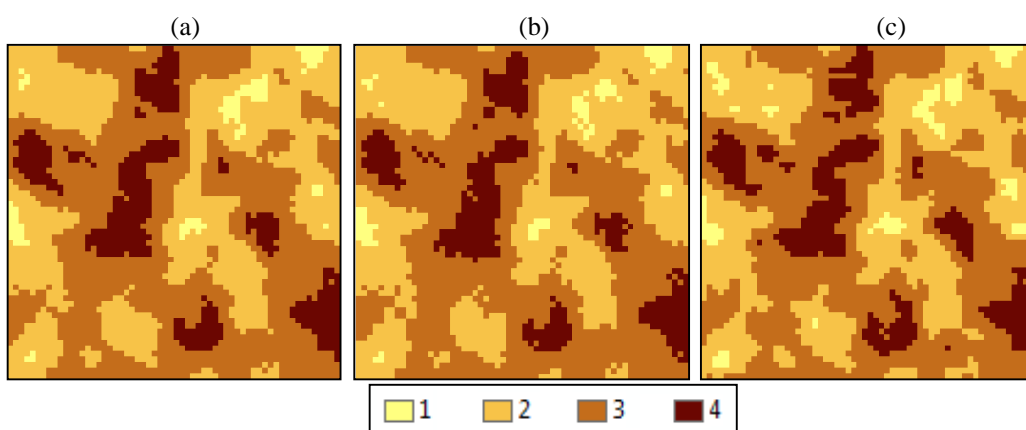


Figura 8III: Mapas de zonas de manejo rasterizados para dados que apresentam DE moderada e *grid* G961. (a) Esférico, (b) Gaussiano e (c) Exponencial.

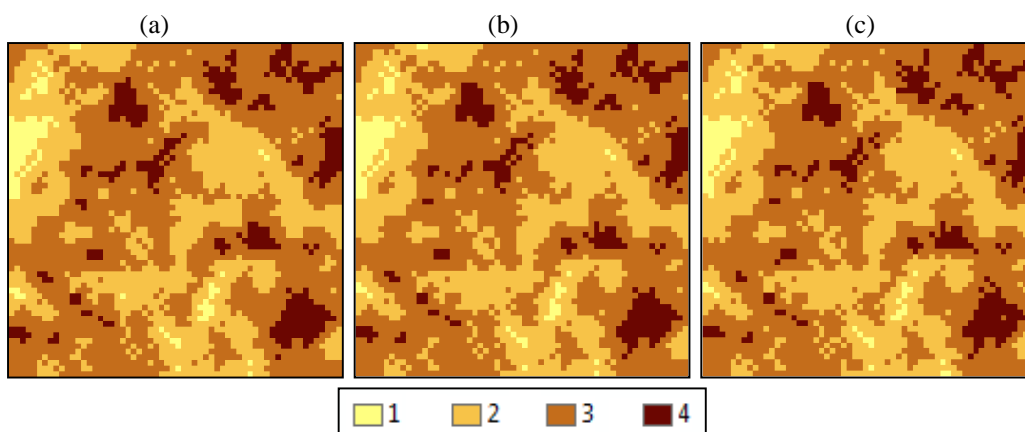


Figura 9III: Mapas de zonas de manejo rasterizados para dados que apresentam DE Fraco e *grid* G961. (a) Exponencial, (b) Esférico e (c) Gaussiano.