

**UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA AGRÍCOLA**

**QUALIDADE DO AJUSTE DE MODELOS GEOESTATÍSTICOS UTILIZADOS
NA AGRICULTURA DE PRECISÃO**

MARIO ANTONIO FARACO

CASCADEL - PR

Maio – 2006

MARIO ANTONIO FARACO

**QUALIDADE DO AJUSTE DE MODELOS GEOESTATÍSTICOS
UTILIZADOS NA AGRICULTURA DE PRECISÃO**

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Engenharia Agrícola, junto ao Programa de Pós-Graduação em Engenharia Agrícola, com área de concentração em Engenharia de Sistemas Agroindustriais, Centro de Ciências Exatas e Tecnológicas, Universidade Estadual do Oeste do Paraná.

Orientador: Prof. Dr. Miguel Angel Uribe Opazo

CASCADEL - PR

Maio - 2006

Mario Antonio Faraco

“Qualidade do ajuste de modelos Geostatísticos utilizados na agricultura de
precisão”

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Programa de Pós-Graduação “stricto sensu” em Engenharia Agrícola, da Universidade Estadual do Oeste do Paraná – UNIOESTE, pela comissão formada pelos professores:

Orientador: Prof. Dr. Miguel Angel Uribe Opazo
UNIOESTE/CCET – Cascavel - PR

Prof^ª. Dr^ª. Terezinha Aparecida Guedes
UEM – Maringá - PR

Prof. Dr. Marcio Antonio Vilas Boas
UNIOESTE/CCET – Cascavel - PR

Prof. Dr. Eduardo Godoy de Souza
UNIOESTE/CCET – Cascavel - PR

Cascavel, 03 de maio de 2006.

À Anna e Francesco, meus pais, dedico.

AGRADECIMENTOS

Ao professor Dr. Miguel Angel Uribe Opazo pelo despertar para o tema, orientação e pela confiança depositada.

Ao professor Ms. Edson Antonio Alves da Silva pelo incentivo e amizade.

Ao professor Ms. Jerry Adriani Johann, pelo apoio e orientação em campo.

Aos professores Dr. Eduardo Godoy de Souza, Dr. Luís César da Silva e Dr. Reinaldo Prandini Ricieri, pelos ensinamentos em suas disciplinas.

Aos colegas de turma, pelo companheirismo e amizade, em especial à Dione, Patrícia, Mari e Araceli.

Aos colegas do Grupo Geoestatística Aplicada: Claudinei Lunkes, Gustavo Henrique Dalposso, Joelmir Borssói, Mozart, Saymon e Taís, pelo apoio e ajuda nas pesquisas de campo.

À Universidade Estadual do Oeste do Paraná - UNIOESTE, seus professores e funcionários, pela acolhida.

À CNPq e CONAB, pelo apoio financeiro para a realização da pesquisa.

À COODETEC, pelo apoio técnico.

Ao Senhor Agassiz Linhares Neto, pelo apoio técnico para implementação dos experimentos.

À Márcia e Cíntia, minhas filhas, pela compreensão nos momentos de estudo.

A Maria Elizabeth Luczynski Silva, pelo incentivo e amizade.

A Clementina, companheira e incentivadora, com muito carinho.

A todos, muito obrigado.

SUMÁRIO

UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ.....	1
LISTA DE FIGURAS.....	ix
LISTA DE TABELAS.....	xiii
RESUMO	xv
ABSTRACT.....	xvi
<u>1 INTRODUÇÃO.....</u>	<u>1</u>
<u>2 REVISÃO DE LITERATURA.....</u>	<u>3</u>
<u>RESISTÊNCIA DO SOLO À PENETRAÇÃO.....</u>	<u>3</u>
<u>DENSIDADE DO SOLO.....</u>	<u>5</u>
<u>UMIDADE DO SOLO.....</u>	<u>7</u>
<u>PRODUTIVIDADE DA SOJA.....</u>	<u>8</u>
<u>VARIÁVEIS REGIONALIZADAS.....</u>	<u>9</u>
<u>2.1.1 Geoestatística.....</u>	<u>11</u>
<u>2.1.2 Semivariograma.....</u>	<u>13</u>
<u>2.1.3 Modelos para Ajuste de Semivariogramas.....</u>	<u>18</u>
<u>Modelo linear com patamar (Lin).....</u>	<u>18</u>
<u>Modelo circular (Cir).....</u>	<u>19</u>
<u>Modelo esférico (Sph).....</u>	<u>20</u>
<u>Modelo exponencial (Exp).....</u>	<u>21</u>
<u>Modelo gaussiano (Gau).....</u>	<u>22</u>
<u>2.1.4 Estimação de Parâmetros no Ajuste de Modelos.....</u>	<u>23</u>
<u>Método dos Mínimos Quadrados Ordinários (OLS).....</u>	<u>24</u>
<u>Métodos dos mínimos quadrados ponderados (WLS1).....</u>	<u>25</u>
<u>2.1.5 Krigagem.....</u>	<u>25</u>
<u>2.1.6 Critérios de Validação dos Ajustes de Modelos Geoestatísticos.....</u>	<u>27</u>
<u>Validação cruzada.....</u>	<u>28</u>

<u>Critério Jackknifing.....</u>	<u>30</u>
<u>Critério de Filliben.....</u>	<u>32</u>
<u>Critério de Akaike.....</u>	<u>34</u>
<u>3 MATERIAL E MÉTODOS.....</u>	<u>36</u>
<u> SIMULAÇÃO DE DADOS COM ESTRUTURA DE DEPENDÊNCIA</u>	
<u> ESPACIAL.....</u>	<u>36</u>
<u> LOCALIZAÇÃO E HISTÓRICO DA ÁREA ESTUDADA.....</u>	<u>37</u>
<u> AMOSTRAGEM.....</u>	<u>38</u>
<u> PROCEDIMENTO PARA COLETA DE DADOS.....</u>	<u>39</u>
<u> 3.1.1 Resistência do Solo à Penetração, Densidade e Umidade.....</u>	<u>39</u>
<u> 3.1.2 Produtividade da Soja.....</u>	<u>40</u>
<u> ANÁLISE DOS DADOS.....</u>	<u>41</u>
<u>4 RESULTADOS E DISCUSSÃO.....</u>	<u>43</u>
<u> SIMULAÇÃO.....</u>	<u>43</u>
<u> ANÁLISE DESCRITIVA, INFERÊNCIAS CLÁSSICAS E ANÁLISE DA</u>	
<u> VARIABILIDADE ESPACIAL.....</u>	<u>50</u>
<u> 4.1.1 Densidade do Solo.....</u>	<u>50</u>
<u> 4.1.2 Umidade do Solo.....</u>	<u>52</u>
<u> 4.1.3 Resistência do Solo à Penetração.....</u>	<u>54</u>
<u> 4.1.4 Produtividade da Soja.....</u>	<u>57</u>
<u> ANÁLISE GEOESTATÍSTICA.....</u>	<u>59</u>
<u> 4.1.5 Análise da Estacionaridade.....</u>	<u>59</u>
<u> 4.1.6 Semivariogramas Experimentais.....</u>	<u>63</u>
<u> Densidade 63</u>	
<u> Umidade 70</u>	
<u> Resistência do solo à penetração.....</u>	<u>77</u>
<u> Produtividade.....</u>	<u>84</u>
<u>5 CONCLUSÕES.....</u>	<u>89</u>
<u>REFERÊNCIAS.....</u>	<u>91</u>
<u>APÊNDICES.....</u>	<u>98</u>
<u>APÊNDICE A - GRÁFICOS DE DISPERSÃO.....</u>	<u>99</u>

APÊNDICE B - SCRIPTS DO R.....	105
APÊNDICE C - DADOS SIMULADOS.....	114
ANEXO	115
ANEXO - TABELA DO COEFICIENTE DE CORRELAÇÃO R.....	116

LISTA DE FIGURAS

Figura 1 - Semivariograma experimental com características ideais.....	16
Figura 2 - Modelo linear com patamar.....	19
Figura 3 - Modelo circular.....	20
Figura 4 - Modelo esférico.....	21
Figura 5 - Modelo exponencial.....	22
Figura 6 - Modelo gaussiano.....	23
Figura 7 - Representação da grade amostral 10x10 dos dados simulados.....	37
Figura 8 - Área em estudo de 57 ha.....	38
Figura 9 - Histograma para os dados simulados.....	44
Figura 10 - Gráfico Boxplot para os dados simulados.....	44
Figura 11 - Post-plot para os dados simulados.....	45
Figura 12 - Semivariograma experimental para os dados simulados.....	45
Figura 13 - Semivariogramas para os dados simulados.....	47
Figura 14 - Histograma para a densidade do solo nas três profundidades estudadas.....	51
Figura 15 - Gráfico Boxplot para a densidade do solo nas três profundidades estudadas.....	52
Figura 16 - Histograma para a umidade do solo nas três profundidades estudadas.....	53
Figura 17 - Gráfico Boxplot para umidade do solo nas três profundidades estudadas.....	54
Figura 18 - Histograma para a resistência do solo à penetração, nas três profundidades estudadas.....	56
Figura 19 - Gráfico Boxplot para a variável resistência do solo à penetração, nas três profundidades estudadas.....	57
Figura 20 - Histograma para a produtividade da soja no ano agrícola 2004/2005.....	58

Figura 21 - Gráfico Boxplot para a produtividade da soja no ano agrícola 2004/2005.....	59
Figura 22 - Gráficos post-plot para a densidade do solo (Mg m^{-3}), nas profundidades 0 a 10 cm (a) e 10 a 20 cm (b).....	60
Figura 23 - Gráficos post-plot para a densidade do solo (Mg m^{-3}), na profundidade 20 a 30 cm (a), e para a umidade do solo ($\text{m}^3 \text{ m}^{-3}$), na profundidade 0 a 10 cm (b).....	61
Figura 24 - Gráficos post-plot para a para a umidade do solo ($\text{m}^3 \text{ m}^{-3}$), nas profundidades 10 a 20 cm (a) e 20 a 30 cm (b).....	61
Figura 25 - Gráficos post-plot para a resistência do solo à penetração [MPa], nas profundidades 0 a 10 cm (a) e 10 a 20 cm (b).....	62
Figura 26 - Gráficos post-plot para a resistência do solo à penetração [MPa], na profundidade 20 a 30 cm (a) e para a produtividade da soja (Mg há^{-1}) (b).....	62
Figura 27 - Semivariogramas experimentais para a variável densidade nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.....	63
Figura 28 - Semivariogramas para os dados de densidade, na profundidade 0 a 10 cm.....	65
Figura 29 - Semivariogramas para os dados de densidade, na profundidade 10 a 20 cm.....	65
Figura 30 - Semivariogramas para os dados de densidade, na profundidade 20 a 30 cm.....	66
Figura 31 - Mapa para o atributo densidade na profundidade 0 a 10 cm; unidade do atributo: [Mg m^{-3}].....	69
Figura 32 - Mapa para o atributo densidade na profundidade 10 a 20 cm; unidade do atributo: [Mg m^{-3}].....	69
Figura 33 - Mapa para o atributo densidade na profundidade 20 a 30 cm; unidade do atributo: [Mg m^{-3}].....	70
Figura 34 - Semivariogramas experimentais para a variável umidade nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.....	70

Figura 35 - Semivariogramas para os dados de umidade na profundidade 0 a 10 cm.....	72
Figura 36 - Semivariogramas para os dados de umidade na profundidade 10 a 20 cm.....	72
Figura 37 - Semivariogramas para os dados de umidade na profundidade 20 a 30 cm.....	73
Figura 38 - Mapa para o atributo umidade na profundidade 0 a 10 cm; unidade do atributo: [m ³ m ⁻³].....	76
Figura 39 - Mapa para o atributo umidade na profundidade 10 a 20 cm; unidade do atributo: [m ³ m ⁻³].....	77
Figura 40 - Mapa para o atributo umidade na profundidade 20 a 30 cm; unidade do atributo: [m ³ m ⁻³].....	77
Figura 41 - Semivariogramas experimentais para a variável resistência do solo à penetração nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.	78
Figura 42 - Semivariogramas para os dados de resistência do solo à penetração na profundidade 0 a 10 cm.....	79
Figura 43 - Semivariogramas para os dados de resistência do solo à penetração na profundidade 10 a 20 cm.....	80
Figura 44 - Semivariogramas para os dados de resistência do solo à penetração na profundidade 20 a 30 cm.....	80
Figura 45 - Mapa para o atributo resistência do solo à penetração na profundidade 0 a 10 cm; unidade do atributo: [Mpa].....	83
Figura 46 - Mapa para o atributo resistência do solo à penetração na profundidade 10 a 20 cm; unidade do atributo: [Mpa].....	84
Figura 47 - Mapa para o atributo resistência do solo à penetração na profundidade 20 a 30 cm; unidade do atributo: [Mpa].....	84
Figura 48 - Semivariograma experimental para a variável produtividade.....	85
Figura 49 - Semivariogramas para os dados de produtividade.....	86
Figura 50 - Mapa para a variável produtividade; unidade do atributo: [Mg ha ⁻¹].	88

LISTA DE TABELAS

Tabela 1 - Análise exploratória e teste de normalidade para os dados simulados	43
Tabela 2 - Parâmetros dos modelos semivariográficos pelos métodos OLS e WLS1 para os dados simulados	46
Tabela 3 - Soma dos quadrados dos resíduos, AIC (\hat{A}), Filliben (r^*) e validação cruzada para os dados simulados.....	48
Tabela 4 - Análise exploratória e testes de normalidade do atributo densidade do solo nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.....	50
Tabela 5 - Análise exploratória e testes de normalidade da umidade do solo nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.....	53
Tabela 6 - Análise exploratória e testes de normalidade da resistência do solo à penetração – RSP, nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.....	55
Tabela 7 - Análise exploratória e teste de normalidade da produtividade da soja	58
Tabela 8 - Parâmetros dos modelos semivariográficos, pelos métodos OLS e WLS1, para os dados de densidade, nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.....	64
Tabela 9 - Soma dos quadrados dos resíduos, AIC (\hat{A}), Filliben (r^*) e validação cruzada para a densidade nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.....	66
Tabela 10 – Modelos escolhidos com melhor ajuste.....	68
Tabela 11 - Parâmetros dos modelos semivariográficos pelos métodos OLS e WLS1 para os dados de umidade, nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.....	71

Tabela 12 - Soma dos quadrados dos resíduos, AIC (\hat{A}), Filliben (r^*) e validação cruzada para a umidade nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.....	74
Tabela 13 - Modelos escolhidos com melhor ajuste.....	76
Tabela 14 - Parâmetros dos modelos semivariográficos pelos métodos OLS e WLS1 para os dados de resistência do solo à penetração, nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.....	79
Tabela 15 - Soma dos quadrados dos resíduos, AIC (\hat{A}), Filliben (r^*) e validação cruzada, para a resistência do solo à penetração, nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.....	81
Tabela 16 - Modelos escolhidos com melhor ajuste.....	83
Tabela 17 - Parâmetros dos modelos semivariográficos pelos métodos OLS e WLS1 para os dados de produtividade.....	85
Tabela 18 - Soma dos quadrados dos resíduos, AIC (\hat{A}), Filliben (r^*) e validação cruzada para a produtividade.....	86
Tabela 19 - Modelos escolhidos com melhor ajuste.....	87

RESUMO

Pesquisas sobre a variabilidade espacial dos atributos do solo que influenciam a produtividade são de suma importância para o desenvolvimento de novas tecnologias que beneficiam a agricultura. Para verificar a variabilidade desses atributos utilizou-se a geoestatística que disponibiliza técnicas para a obtenção de informações a respeito dessa variabilidade. Os processos de análise de dados utilizam métodos que incluem algoritmos de otimização para escolha de um modelo geoestatístico e a estimação dos parâmetros desse modelo. Foram estudados os atributos do solo: resistência do solo à penetração, densidade do solo, umidade do solo e a variável produtividade da soja. Este trabalho tem por objetivo descrever os comportamentos espaciais de dados empíricos e simulados e construir modelos de variabilidade espacial para os atributos em estudo com o objetivo principal de avaliar a qualidade dos ajustes segundo os Critérios de Akaike, Filliben, Jackknifing e Validação Cruzada. A pesquisa foi desenvolvida na região Oeste do Paraná, em uma área de 57 há, cujo solo típico é o Latossolo Vermelho Distrófico e foi utilizada uma malha com 100 parcelas georeferenciadas. Para a análise da estrutura de dependência espacial foram utilizados semivariogramas experimentais gerados a partir dos dados amostrais. Em seguida, ajustaram-se modelos teóricos aos semivariogramas experimentais e aplicaram-se as técnicas de avaliação dos ajustes aos modelos geoestatísticos. Em consequência, analisaram-se os resultados dos diversos métodos estudados, comparando-se os resultados obtidos e concluindo-se pela validação cruzada como o melhor critério de ajuste.

PALAVRAS-CHAVE: Geoestatística, estimadores, validação do ajuste.

ABSTRACT

Researches about the spatial variability of the soil attributes that influence the productivity are highly important for the development of new technologies that benefits the agriculture. To verify the variability of these attributes we used the geostatistic that offers techniques to the obtainment of information concerning this variability. The processes of data analysis use methods that include optimization algorithms to the choice a geostatistic model and the estimation of that model parameters. We studied the following soil attributes: soil resistance to penetration, soil density, soil humidity and the soybean's productivity variable. This paper has as its purpose to describe the spatial behavior of empiric and simulated data and to build models of spatial variability to the attributes in study with the main purpose of evaluating the quality of the adjustments according to the Criteria of Akaike, Filliben, Jackknifing and Cross Validation. The research was developed in the West Parana region, in a area of 54 ha where the typical soil of the region is the Red Distrofic Latosoil and a net with a 100 georeferred parcels was utilized. To the structure analyze of spatial dependency we used experimental semivariograms generated from sample data set. Afterward those theoretical models were adjusted to the experimental semivariograms and techniques of evaluation of the adjustments were applied to the geostatistic models. Consequently, the results of the several methods studied were analyzed and the gotten results were compared, concluding for the cross validation as the best adjustment criterion.

KEYWORDS: Geostatistics, estimators, adjustment validation.

1 INTRODUÇÃO

As melhores técnicas agrícolas sempre procuram alcançar o equilíbrio entre a aplicação de insumos e as práticas agrícolas com os atributos do solo e a necessidades das plantas.

Um procedimento adotado, até a algumas décadas atrás, na aplicação de fertilizantes, considerava indicadores estatísticos como a média e a variância. Embora esses procedimentos continuem a ser utilizados, pesquisadores e agricultores buscam novos métodos, pois reconhecem a variabilidade espacial dos atributos do solo. Esse sistema de gerenciamento de áreas agrícolas, que considera a posição das amostras relacionadas às propriedades do solo e da produtividade, é conhecido como Agricultura de Precisão.

Conforme BALASTREIRE (1998), a Agricultura de Precisão é definida como um conjunto de técnicas que permite o gerenciamento localizado de culturas e que se fundamenta na percepção da variabilidade espacial da produtividade e de fatores a ela relacionados.

A pesquisa da variabilidade espacial dos atributos do solo que influenciam a produtividade de uma cultura é importante para a utilização de uma agricultura de precisão, obtendo-se, assim, maior detalhamento de informações espacialmente referenciadas.

A Teoria das Variáveis Regionalizadas, sobre a qual se sustenta a Geoestatística, oferece técnicas de interpolação de dados não coletados, a partir de locais amostrados. A krigagem ordinária uma delas. Essas técnicas utilizam critérios que consideram a estrutura de dependência espacial modelada, a partir do ajuste de um modelo teórico a uma função, como a semivariância, que mede a dependência espacial em diferentes distâncias (*lag*) e cujo gráfico é conhecido como semivariograma experimental.

O objetivo do emprego da geoestatística é conhecer a estrutura de dependência espacial e produzir mapas temáticos que auxiliem no manejo agrícola.

O objetivo principal deste trabalho foi estudar a qualidade do ajuste de modelos geoestatísticos por meio dos critérios utilizados na agricultura de precisão.

Os objetivos específicos deste trabalho foram:

- Descrever os comportamentos espaciais dos dados empíricos e simulados, utilizando técnicas da geoestatística;
- Construir modelos empíricos de variabilidade espacial de dados simulados, para avaliar a qualidade dos modelos ajustados, segundo os Critérios de Akaike, Filliben e Validação Cruzada;
- Construir modelos empíricos de variabilidade espacial da produtividade da soja e de atributos físicos do solo tais como: resistência do solo à penetração, densidade e umidade e compará-los, avaliando a qualidade dos modelos ajustados, segundo os Critérios de Akaike, Filliben e Validação Cruzada;
- Construir mapas temáticos utilizando estrutura de dependência espacial, escolhida segundo os critérios utilizados para o ajuste do melhor modelo teórico.

2 REVISÃO DE LITERATURA

As práticas agrícolas vêm, de longa data, buscando um aumento de produtividade sem prescindir de um equilíbrio entre os insumos aplicados. Por outro lado, nas últimas décadas, cada vez mais se buscam procedimentos que não afetem ou que minimizem os efeitos da ação humana sobre o meio ambiente. Entre as diversas variáveis intervenientes nos processos agrícolas podem-se enumerar a necessidade de nutrientes das plantas, os insumos aplicados, os atributos do solo, as variedades de espécies de culturas e as práticas de manejo agrícolas. Como esses atributos variam no espaço e em diferentes escalas, requer-se o uso de técnicas adequadas para sua análise e representação.

Essa procura pelo aumento da produtividade, para atender cada vez mais às necessidades de produção de alimentos para uma população em contínuo crescimento, tem levado os pesquisadores a buscar soluções para os problemas de variabilidade dos atributos do solo, desde o início do século XX.. Nos últimos 50 anos, com o desenvolvimento de novas teorias, houve um avanço considerável da Estatística Aplicada às Ciências da Terra, em especial da Estatística Espacial, da qual a Geoestatística faz parte.

RESISTÊNCIA DO SOLO À PENETRAÇÃO

A compactação do solo é um processo de dispersão ou arranjo dos agregados e aproximação das partículas primárias (argila, silte e areia) do solo. A Resistência do Solo à Penetração – RSP é resultante de forças geradas pela

compactação, que é definida pela densidade do solo, conteúdo de água e textura do solo (SECCO, 2003).

Para a determinação da resistência do solo à penetração, é adequado que se tenha um solo uniforme, no estado de friabilidade, assim é possível não considerar a umidade na sua análise. Caso contrário, a umidade deve ser amostrada junto com a resistência do solo à penetração.

A resistência do solo à penetração é medida utilizando-se penetrômetros convencionais para uso agrícola ou de impacto, usados na construção civil. Os penetrômetros convencionais são pressionados contra o solo. A resistência ao avanço de penetração de sua ponta é lida ou registrada por meio de um dinamômetro. O penetrômetro convencional mede a resistência estática ou de ruptura, enquanto que o penetrômetro de impacto tem características dinâmicas de penetração. A resistência medida pelo penetrômetro difere da resistência encontrada pelas raízes das plantas, pois elas procuram locais com menor resistência, curvando-se pelo solo, enquanto que o penetrômetro mede a resistência encontrada, em linha reta, na introdução do equipamento no solo.

Determinar qual é o nível crítico de resistência do solo para o crescimento radicular tem sido um processo difícil, pois a resistência está diretamente relacionada à umidade do solo, que varia constantemente, e o nível crítico muda de acordo com o tipo de solo e a espécie cultivada. De uma forma geral, aceita-se como valores limitantes para o crescimento radicular da maioria das espécies cultivadas no solo, ou seja, solo compactado, valores de resistência à penetração acima de 2,0 MPa, na densidade de $1,45 \text{ Mg m}^{-3}$, para solos argilosos, e volume de macroporos inferior a 10% (SECCO, 2003).

SILVA, REICHERT e REINERT (2004) estudaram a variabilidade espacial da resistência do solo à penetração, em sistema plantio direto, em três tipos de solo, entre eles o Latossolo Vermelho Distrófico, dividindo os estados de compactação em três classes. Para o Latossolo Vermelho Distrófico, classificou-se a resistência do solo à penetração no menor estado de compactação, entre 0,20 MPa e 1,59 MPa; intermediário, entre 1,60 MPa e 2,10 MPa, e no maior estado de compactação, entre 2,11 MPa e 2,80 MPa.

DENSIDADE DO SOLO

A porosidade do solo é a fração *in situ* que não é ocupada pela matéria sólida e por onde circulam o ar e a água. São divididos de acordo com o diâmetro dos poros, sendo os macroporos aqueles de diâmetro maior que 0,06 mm e microporos os de diâmetro menor que 0,06 mm. Os macroporos permitem a aeração e infiltração da água no solo, enquanto que os microporos são responsáveis pelo armazenamento de água. Solos argilosos têm porosidade total em torno de 40% a 60%. Latossolos, com predominância de argilominerais e óxidos e boa estruturação, apresentam uma permeabilidade bastante rápida da água (AZEVEDO; DALMOLLIN, 2004).

A redução da porosidade, juntamente com um aumento da densidade, aumenta o potencial de inundação, pois reduz a taxa de infiltração, fazendo com que a água escorra superficialmente, aumentando a erosão do solo. As águas da erosão arrastam partículas do solo que, ao atingirem lagos e riachos, diminuem a sua qualidade, causam assoreamento, eutrofização e ainda deixam o solo infértil e degradado (AZEVEDO; DALMOLLIN, 2004).

A densidade do solo é definida como a relação entre a massa seca e um determinado volume, portanto, considera a porosidade. Assim, essa variável está relacionada com a estrutura e, conseqüentemente, com a compactação do solo. Valores comuns de densidade ficam entre 0,95 e 1,8 Mg m⁻³. A tendência é um aumento nos pontos mais profundos do solo, em função do peso das camadas superiores (AZEVEDO; DALMOLLIN, 2004). A densidade do solo também é afetada pelo manejo da área, sendo aumentada pela compactação intrínseca do solo e diminuída pela incorporação de matéria orgânica, bem como pelas práticas de preparo que alteram o espaço poroso, principalmente os macroporos, influenciando

nas características físico-hídricas, como porosidade de aeração, retenção de água, infiltração, disponibilidade de água para as plantas e resistência à penetração. Estudos indicam um maior grau de compactação no sistema plantio direto, em relação ao sistema convencional, devido ao excesso de tráfego de máquinas e adensamento natural das partículas (KLEIN, 1999).

Existem indícios de que no sistema de plantio direto ocorra uma faixa de maior compactação de 7 a 15 cm de profundidade, aparentemente indicados pela densidade e/ou resistência do solo à penetração (SECCO, 2003).

Conforme TORRES e SARAIVA (1999), cada solo apresenta uma amplitude de densidade com valores máximos e mínimos diferentes, em função de suas características mineralógicas, composição e estrutura. Em condições de campo, a amplitude de variação da densidade global para solos argilosos vai de $1,0 \text{ Mg m}^{-3}$, sob condições naturais e ricos em matéria orgânica, até a $1,45 \text{ Mg m}^{-3}$, em solos mal manejados e compactados. Para solos arenosos, a densidade varia, em média, de $1,25 \text{ Mg m}^{-3}$ a $1,7 \text{ Mg m}^{-3}$. Essas diferenças de densidade e de suas amplitudes entre os solos propiciam que outros parâmetros, obtidos indiretamente a partir da densidade, possam ser utilizados para se quantificar melhor o efeito da compactação sobre a produtividade das culturas.

MILANI (2005) encontrou, para a profundidade 0 a 10 cm, uma densidade média de $1,17 \text{ Mg m}^{-3}$, com coeficiente de variação de 5,52 %, e para a profundidade 10 a 20 cm, uma densidade média de $1,19 \text{ Mg m}^{-3}$, com coeficiente de variação de 4,9%, considerando uma densidade amostral de $9,2 \times 7,2 \text{ m}$, em 256 pontos amostrais.

UMIDADE DO SOLO

A água é de suma importância no desempenho dos ecossistemas terrestres, inclusive nos agroecossistemas. O teor de água no solo depende de fatores climáticos e das características do solo como textura, estrutura e porosidade. O armazenamento de água, disponibilidade para as plantas e taxa de infiltração podem variar no tempo e no espaço.

Em estudos sobre a variabilidade temporal da umidade do solo, verificou-se que ela não possui distribuição aleatória na área, porém, possui dependência espacial bem definida em caso de Argissolos Vermelhos (GONÇALVES; FOLEGATTI, 1995).

A determinação quantitativa da água do solo é feita à base de peso u e à base do volume Q , dependendo da finalidade da medida. A umidade é adimensional (kg kg^{-1}), mas suas unidades devem ser mantidas para não serem confundidas com as unidades da umidade à base de volume, que também é adimensional, mas numericamente diferente. A umidade u também é, com frequência, apresentada em percentagem. Sua medida é bastante simples: a amostra é pesada úmida e em seguida mantida em estufa a 105°C , até peso constante. A diferença entre essas massas, corresponde à massa de água. A amostra pode ter qualquer tamanho e pode ter a estrutura deformada. O ideal é que tenha entre 0,05 e 0,50 kg, tendo-se o cuidado de não deixar a água evaporar antes da pesagem úmida. A umidade à base do volume Q também é adimensional ($\text{m}^3 \text{m}^{-3}$) e, com frequência, é apresentada em percentagem. Sua medida é mais difícil, pois envolve a medida do volume e, por isso, a amostra não pode ser deformada. A técnica mais comum é a do uso de anéis volumétricos, idênticos aos utilizados para a medida da densidade do solo. REICHARDT e TIMM (2004) demonstraram que a umidade à base do volume é igual ao produto entre a umidade à base de peso e a densidade do solo.

As variações de umidade do solo são de grande importância para a análise do comportamento de uma cultura e, conseqüentemente, do armazenamento de água. Essas variações são reflexo das taxas de evapotranspiração, precipitação pluvial, irrigação e movimentos de água no perfil de solo.

PRODUTIVIDADE DA SOJA

A previsão de safras no Brasil é feita por meio da estimativa anual da produção agrícola, utilizando-se questionários aplicados aos produtores e/ou às entidades relacionadas à atividade agrícola em cada região. Com isso, a obtenção de dados confiáveis por meio dessa metodologia é difícil, lenta e onerosa, tendo como conseqüência um alto grau de subjetividade nos resultados da estimativa da produção (FONTANA et al., 2001).

As geotecnologias espaciais são técnicas que podem otimizar a estruturação de sistemas mais eficientes e dinâmicos para estimativa da produção agrícola, em nível regional e nacional. Essas tecnologias permitem a obtenção de informações precisas, em tempo hábil e com baixo custo, sobre a extensão, as condições de desenvolvimento e o potencial de produção das culturas. Essas informações são de grande importância para a economia de um país, pois permitem um planejamento adequado da sua economia agrícola, minimizando o problema de escassez ou de excesso de produtos (FONTANA et al., 2001; MOTTA; FONTANA; WEBER, 2001).

Segundo KILPP (2000), pelos mapas temáticos de produtividade pode-se facilmente identificar os diferentes índices de produtividade de uma área. Entre eles, o mapa de isolinhas de produtividade é o mais comum. A variação de produtividade numa mesma área é muito maior do que geralmente se espera. Os

mapas da produtividade apresentam essa variação e com eles pode-se identificar as áreas com baixo ou alto rendimento. A ação seguinte é a identificação das causas que afetam a produtividade, como os atributos físicos do solo e os fatores que afetam a fertilidade.

A produtividade é diretamente afetada pelas características físico-mecânicas do solo, como umidade, aeração, temperatura e resistência do solo à penetração (SECCO, 2003). Esta última, por sua vez, é resultante das forças provocadas pela compactação, teor de água e textura do solo.

Segundo ANTUNES (2004), o plantio otimizado da soja pode resultar em rendimentos de até 7 Mg ha⁻¹. Contudo, a produtividade média nas lavouras brasileiras mantém-se em 2,8 Mg ha⁻¹. Nas últimas quatro décadas, as técnicas de melhoramento da cultura têm resultado num crescimento de 0,025 Mg ha⁻¹ ao ano, rendimento que depende tanto de pesquisas científicas quanto do ambiente. O autor ressalta que o principal papel da pesquisa é desenvolver cultivares com boa produtividade de grãos, de ampla adaptação e que apresentem estabilidade na produção.

Segundo dados da CONAB (2006), na safra 2004/05, o Brasil teve uma área de 23 301 000 hectares plantados com a cultura da soja, sendo 8 588 500 hectares só na região Sul. No Estado do Paraná, a área plantada com essa cultura foi de 4 148 400 hectares, o que corresponde a pouco mais de 48% da região Sul. A produção média estadual, para essa safra, foi de 2,3 Mg ha⁻¹ e a média nacional foi de 2,19 Mg ha⁻¹.

VARIÁVEIS REGIONALIZADAS

KRIGE (1951) iniciou estudos pragmáticos a partir de um imenso arquivo de dados que representavam as minas de ouro sul-africanas. Esse

pesquisador considerou a existência de correlações espaciais e, trabalhando com dados de concentração de ouro, concluiu que não conseguia encontrar sentido nas variâncias, se considerasse a distância entre as amostras. MATHERON (1963), baseado nessas observações, desenvolveu uma teoria, que chamou de Teoria das Variáveis Regionalizadas, contendo os fundamentos da Geoestatística. Ele define variável regionalizada como uma função espacial numérica, que varia de um local para outro, com uma continuidade aparente e cuja variação não pode ser representada por uma função matemática simples. São variáveis cujos valores são relacionados, de algum modo, com a posição espacial que ocupam. Esse tipo de variável tem uma dupla característica:

- É aleatória, pois os valores numéricos observados podem variar casualmente de um ponto a outro do espaço;
- É espacialmente correlacionada, pois, mesmo variando muito no espaço, não é inteiramente independente.

As Variáveis Regionalizadas possuem características qualitativas estreitamente ligadas à estrutura do fenômeno natural que elas representam e, segundo OLEA (1975, 1977), as suas principais características são:

- Localização: uma variável regionalizada é numericamente definida por um valor, o qual está associado a uma amostra de tamanho, forma e orientação específicos. Essas características geométricas da amostra são denominadas suporte geométrico, o qual não necessariamente compreende volumes, podendo se referir também a áreas e linhas. Entretanto a teoria não restringe a R^2 ou R^3 o domínio da metodologia. Quando o suporte geométrico tende a zero, tem-se um ponto ou amostra pontual. Em suma, a teoria das variáveis regionalizadas considera a geometria espacial dos pontos amostrados distintamente da estatística clássica, na qual a forma e a orientação dos pontos amostrais não são considerados;
- Continuidade: dependendo do fenômeno observado, a variação espacial pode ser grande ou pequena. Apesar da complexidade das flutuações,

uma continuidade média geralmente está presente em pequenas distâncias;

- Anisotropia: algumas variáveis regionalizadas são anisotrópicas, isto é, apresentam variações graduais numa direção e rápidas ou irregulares em outra. A anisotropia é uma característica esperada. Quando as variações são similares em várias direções, o que a experiência demonstra ser menos freqüente, o fenômeno é considerado isotrópico.

2.1.1 Geoestatística

A Geoestatística apresenta várias técnicas usadas para análise e inferência de uma variável espacialmente georreferenciada (variáveis regionalizadas). Considera a localização geográfica e a dependência (ou auto-correlação) espacial entre amostras. Assim, os mapas temáticos, construídos para as variáveis regionalizadas, são de grande importância para avaliação do comportamento da variável na área em estudo. O estudo geoestatístico possui um grande potencial, pois com o conhecimento da amplitude da dependência entre as amostras é possível planejar de maneira mais eficaz a implantação de zoneamento de áreas.

O processo de análise de dados espacialmente referenciados incluem, mas não se limitam, a métodos gráficos e métodos exploratórios para investigar algum padrão nos dados. Os processos de análise utilizam métodos que incluem algoritmos de otimização para escolha de um modelo geoestatístico e a estimação dos parâmetros desse modelo.

Para RIBEIRO JR. (1995), o gráfico chamado de *post-plot* é um método para a avaliação da estacionaridade ou da tendência direcional dos dados. É uma importante ferramenta para esse controle. Nesse tipo de gráfico, cada parcela do experimento está classificada segundo os valores correspondentes aos seus inter-quartis (mínimo até 1º quartil; 1º quartil até mediana; mediana até 3º

quartil; 3º quartil até valor máximo). Desse modo, são definidos quatro intervalos em que aparecem concentrações de tonalidades iguais, indicando a existência de sub-regiões no mapa.

Um dos objetivos da Geoestatística é extrair da distribuição dos dados da variável regionalizada, uma imagem da sua variabilidade e uma medida da correlação existente entre os valores tomados em pontos distintos do espaço. Esse método é conhecido como análise da estrutura espacial dos dados. A Geoestatística possibilita também o estudo da estrutura de dependência espacial utilizando, dentre outros, os métodos analíticos como o da máxima verossimilhança e o bayesiano (DIGGLE; RIBEIRO JR., 2000).

O outro objetivo da Geoestatística é o de proporcionar uma medida da precisão das previsões ou estimativas feitas, por meio de um processo de interpolação, a partir de dados amostrados, de modo a dar suporte a uma teoria de estimativas de valores em pontos não amostrados.

Em resumo, os passos, num estudo empregando técnicas geoestatísticas, incluem análise exploratória dos dados, análise estrutural (modelagem do semivariograma) e realização de inferências.

Segundo BURROUGH (1987), a variação espacial regionalizada pode ser expressa pela soma de três componentes. Se s representa uma posição em uma, duas ou três dimensões, então o valor da função aleatória, ou variável regionalizada Z , em s , é dada por:

$$Z(s) = m(s) + \varepsilon'(s) + \varepsilon'' \quad (1)$$

em que:

- $m(s)$ é uma função determinística que descreve a componente estrutural de Z em s , associada a um valor médio constante ou a uma tendência constante;

- $\varepsilon'(s)$ é um termo estocástico, que varia localmente e depende espacialmente de $m(s)$;

- ε'' é um ruído aleatório ou erro residual não correlacionado, com distribuição normal, com média zero e variância σ^2 .

- $s \in S$, sendo S um Espaço Euclidiano \mathbf{R} , \mathbf{R}^2 ou \mathbf{R}^3 .

2.1.2 Semivariograma

O semivariograma, definido como o gráfico da função semivariância *versus* a distância h , é uma técnica usada para medir a dependência entre pontos amostrais, distribuídos segundo um sistema espacial de referência e para a interpolação de valores necessários à construção de mapas de isolinhas (ISAACS; SRIVASTAVA, 1989). A semivariância é uma função da distância h , que é estimada somente em um conjunto discreto de distâncias (*lags*). A partir dessa estimativa, ajusta-se um modelo que depende de parâmetros desconhecidos a serem determinados e com características espaciais (CRESSIE, 1985). Em termos práticos, é a técnica estatística que permite estudar a dispersão natural das variáveis regionalizadas. Cada semivariograma construído para uma direção fixa é chamado de semivariograma direcional.

Segundo CRESSIE (1989), o semivariograma é a técnica central da geoestatística. A análise variográfica/covariográfica é uma arte, no melhor sentido do termo. Artes requerem bons instrumentos como, neste caso, um bom programa interativo; mas, também, experiência e habilidade para sintetizar e, às vezes, ir além dos dados.

RIBEIRO JR. e DIGGLE (2001) relatam que a técnica mais usada para descrever a dependência espacial na geoestatística é o semivariograma, o qual descreve a associação espacial dos pontos amostrais em função das distâncias entre eles. É uma técnica exploratória que auxilia os métodos geoestatísticos, contribuindo na modelagem da dependência espacial, determinando os parâmetros básicos do modelo: alcance, efeito pepita e patamar. Os principais estimadores utilizados na construção de semivariogramas são os de Matheron (MATHERON, 1963), Cressie e Hawkins (CRESSIE; HAWKINS, 1980), Pairwise (ISAACS; SRIVASTAVA, 1989) e New1 e New2 (LI; LAKE, 1994).

Segundo GUERRA (1988), existem três tipos de semivariograma:

- Semivariograma observado ou experimental, que é aquele obtido a partir do conjunto de dados amostrais, portanto o único conhecido;
- Semivariograma verdadeiro, que é o semivariograma real, e é sempre desconhecido;
- Semivariograma teórico, que é um semivariograma de referência.

O semivariograma teórico é aquele que tem suas propriedades conhecidas, as quais são aproveitadas na análise do semivariograma. Dentre os distintos modelos teóricos existentes sobressaem-se: o semivariograma linear, o semivariograma circular, o semivariograma esférico, o semivariograma exponencial e o semivariograma gaussiano.

Segundo ISAACS e SRIVASTAVA (1989), a função $\gamma(h)$ é a semivariância dada pela metade da esperança matemática do quadrado da diferença entre os valores de pontos no espaço, separados pelo vetor h :

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

em que:

- $Z(s)$: é o valor da semivariância medida;
- s : é a localização;
- h : a distância que separa duas amostras, e
- $s \in S \subset \mathbf{R}, \mathbf{R}^2$ ou \mathbf{R}^3 .

Para que seja possível construir um semivariograma experimental, é necessária a utilização de um estimador para a semivariância expressa na equação (2). Segundo CRESSIE (1993), um estimador natural, baseado no método dos momentos e proposto por MATHERON (1963) é dado por:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(s_i) - Z(s_i + h)]^2 \quad (3)$$

em que:

- $N(h)$: é o número de pares de pontos amostrais separados pelo vetor h ;

- $Z(s_i)$ e $Z(s_i + h)$: são os valores amostrados nas posições s_i e $s_i + h$, respectivamente.

O estimador dado na equação (3) é conhecido como estimador clássico ou estimador de Matheron, com uso muito difundido na literatura e muito utilizado quando os dados amostrais apresentam distribuição normal de probabilidades.

O estimador de Matheron é pouco resistente e muito afetado por observações atípicas (*outliers*), sobretudo aquelas que produzem distribuições de caudas pesadas (assimétricas, leptocúrticas). Nestes casos, é adequado o uso de um estimador que não seja afetado por pontos discrepantes, que não possam ser eliminados ou substituídos (RIBEIRO JR., 1995).

CRESSIE e HAWKINS (1980) propuseram o uso do seguinte estimador, para dados que não têm distribuição normal de probabilidades:

$$\hat{\gamma}(h) = \frac{1}{2} \frac{\left[\frac{1}{N(h)} \sum_{i=1}^{N(h)} \left| Z(s_i + h) - Z(s_i) \right|^{1/2} \right]^4}{0,457 + \frac{0,494}{N(h)}} \quad (4)$$

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.

VIEIRA (1998) destaca que o ajuste de um modelo teórico ao semivariograma experimental é um dos aspectos mais importantes da Teoria das Variáveis Regionalizadas e pode ser uma das maiores fontes de ambigüidade e polêmicas nessas aplicações.

O estudo da estrutura espacial dada pela análise da função de semivariância, utilizando o semivariograma, não constitui o objetivo final da análise espacial. É necessário estimar os valores em locais não amostrados, para se obter o conhecimento da distribuição espacial da variável regionalizada. Dessa

forma, a análise da estrutura de dependência espacial deve ser vista como um passo fundamental que precede as técnicas de estimação de qualquer valor em qualquer posição da área de estudo (VIEIRA; CARVALHO, 2001).

A semivariância é uma função crescente de h , visto que os valores tomados entre dois pontos diferentes são, em média, tanto mais diferentes quanto mais afastados estejam um do outro. Desse modo, o semivariograma dá sentido preciso à noção tradicional de zona de influência das amostras.

Na Figura 1 é ilustrado um semivariograma que apresenta características ideais. Como medições mais próximas umas das outras devem ser mais parecidas do que aquelas separadas por grandes distâncias, $\gamma(h)$ aumenta à medida que h cresce, até atingir um valor máximo no qual se estabiliza. O valor de h no qual $\gamma(h)$ se estabiliza é chamado alcance (a) que é o raio de dependência espacial. Embora, por definição, $\gamma(h)$ seja igual a zero para h igual a zero, na prática é comum que, à medida que h tenda para zero, $\gamma(h)$ se aproxime de um valor positivo. Tal valor positivo, usualmente denotado por C_0 , é conhecido como efeito pepita (*nugget effect*), que representa erros de medição ou de variabilidade em pequena escala (CRESSIE, 1993).

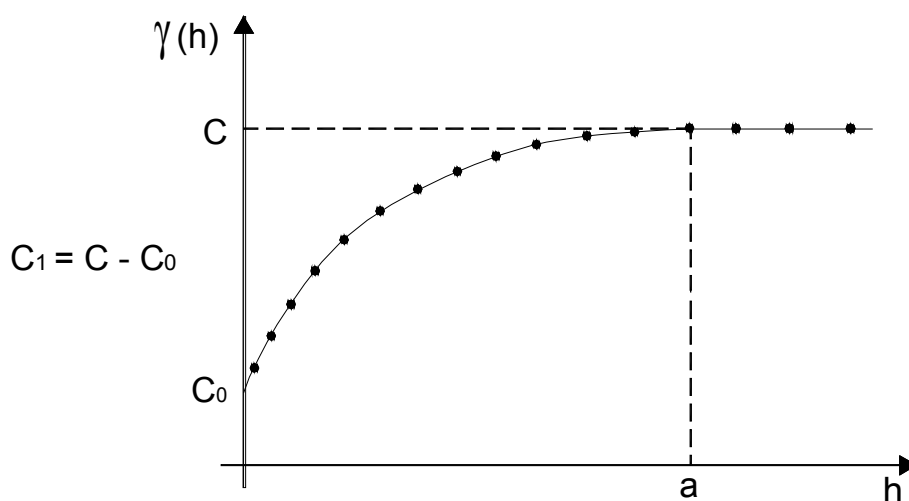


Figura 1 - Semivariograma experimental com características ideais.

O patamar (C), o alcance (a) e o efeito pepita (C_0) são os parâmetros do semivariograma, isto é, são os parâmetros, por meio dos quais, em conjunto com o modelo ajustado, busca-se quantificar a dependência espacial da variável regionalizada em estudo. É usual denotar por C_1 como a contribuição, em que C_1 é a diferença $C - C_0$.

O patamar (C) representa o valor máximo no qual $\gamma(h)$ se estabiliza, tornando-se independente à variação de h . Esse valor de h é representado pelo alcance (a) e é um parâmetro importante no estudo da variabilidade espacial, pois é a distância limite na escolha entre utilizar a geoestatística e a estatística clássica. Nas distâncias menores que o alcance, as observações amostrais são espacialmente dependentes e usa-se a geoestatística. Para as distâncias maiores que o alcance pode-se, então, utilizar a estatística clássica.

Quando o alcance é menor que a distância entre as amostragens, tem-se o chamado efeito pepita puro e uma distribuição espacial completamente ao acaso, deixando-se de aplicar os princípios da geoestatística.

É necessário que se escolha um ponto de corte até onde será estudado o semivariograma (*cutoff*), ou seja, a distância máxima a ser utilizada no cálculo da semivariância. Em geral, usa-se 25%, 33% ou 50% da distância máxima da área em estudo, levando-se em conta o bom senso (CLARK, 1979). Também é necessária a escolha dos intervalos de h (*lag*), para os quais as semivariâncias são calculadas. Quando são determinados os semivariogramas direcionais para o estudo da anisotropia, os ângulos mais usados são 0° , 45° , 90° e 135° . Uma primeira estimativa dos parâmetros do semivariograma teórico pode ser obtida graficamente a partir do semivariograma experimental.

O grau de dependência espacial pode ser medido pela relação entre o efeito pepita e o patamar, chamado coeficiente de efeito pepita, dado por:

$$\varepsilon = \frac{C_0}{C_0 + C_1}. \quad (5)$$

Segundo os valores encontrados para este coeficiente, pode-se classificar o grau de dependência espacial, do atributo em estudo, como forte, se for menor

que 0,25; moderada, se os valores estiverem entre 0,25 e 0,75; e fraca dependência espacial, se o valor for maior que 0,75 (CAMBARDELLA *et al.*, 1994).

2.1.3 Modelos para Ajuste de Semivariogramas

Como apresentados na Seção 2.5.2, o semivariograma experimental é constituído de valores estimados de $\gamma(h)$ correspondentes a valores discretos de h . A interpolação por krigagem irá requerer a estimativa da semivariância para outros valores da distância. Portanto, uma curva contínua, que corresponda a um modelo teórico, deve ser ajustada aos pontos do semivariograma empírico (ISAAKS; SRIVASTAVA, 1989).

Dentre os distintos modelos teóricos de semivariogramas, relatados na literatura, são analisadas, a seguir, as características do que se considera os principais modelos.

Modelo linear com patamar (Lin)

O modelo linear com patamar é dado pela função definida na equação (6):

$$\gamma(h) = \begin{cases} C_0 + \frac{C_1}{a}h, & \text{se } 0 \leq h \leq a \\ C_0 + C_1, & \text{se } h > a \end{cases} \quad (6)$$

sendo, $\frac{C_1}{a}$ o coeficiente angular para $0 \leq h \leq a$.

A Figura 2 representa um gráfico do modelo linear obtido da equação (6).

Conforme CRESSIE (1993), o modelo linear com patamar é um modelo válido apenas em R (na reta) não sendo válido para dimensões maiores.

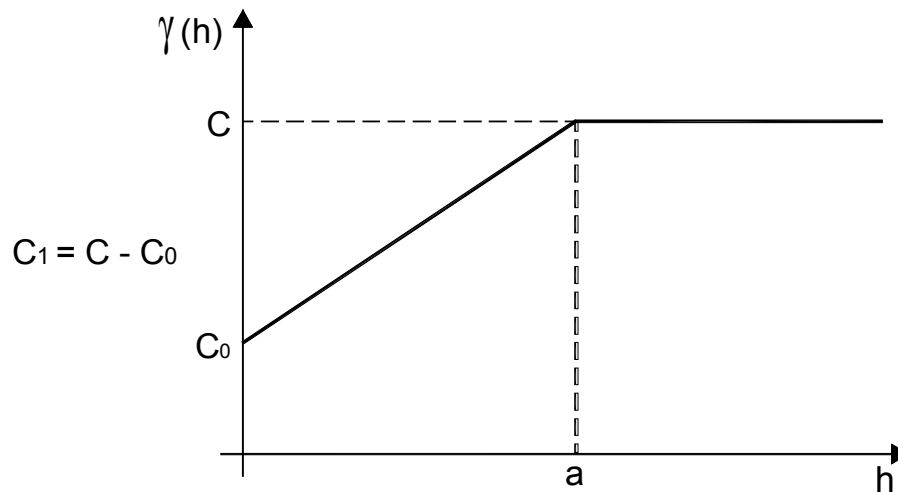


Figura 2 - Modelo linear com patamar.

Modelo circular (Cir)

O modelo circular de semivariograma é dado pela função definida na equação (7):

$$\gamma(h) = \begin{cases} C_0 + C_1 \left[1 - \frac{2}{\pi} \cos^{-1} \left(\frac{h}{a} \right) + \frac{2h}{\pi a} \sqrt{1 - \frac{h^2}{a^2}} \right], & \text{se } 0 \leq h \leq a \\ C_0 + C_1, & \text{se } h > a \end{cases} \quad (7)$$

Segundo McBRATNEY e WEBSTER (1986), o modelo circular é válido em R e em R^2 , mas não em R^3 . O modelo circular, apresentado na equação (7), é representado graficamente na Figura 3.

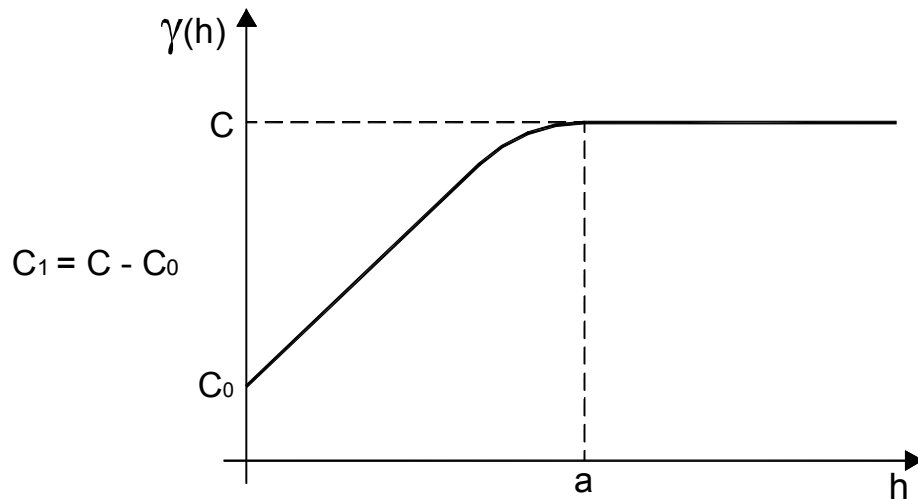


Figura 3 - Modelo circular.

Modelo esférico (Sph)

O modelo esférico de semivariograma é dado pela função apresentada na equação (8):

$$\gamma(h) = \begin{cases} 0, & \text{se } h = 0 \\ C_0 + C_1 \left[1,5 \left(\frac{h}{a} \right) - 0,5 \left(\frac{h}{a} \right)^3 \right], & \text{se } 0 < h \leq a \\ C_0 + C_1, & \text{se } h > a \end{cases} \quad (8)$$

A representação gráfica do modelo esférico é apresentada na Figura 4.

PANNATIER (1996) também destaca o comportamento linear desse modelo de semivariograma, para pequenos valores de h , e observa que a tangente à origem atinge o patamar a dois terços do alcance.

Os modelos esféricos são válidos nos espaços R , R^2 e R^3 (CRESSIE 1993).

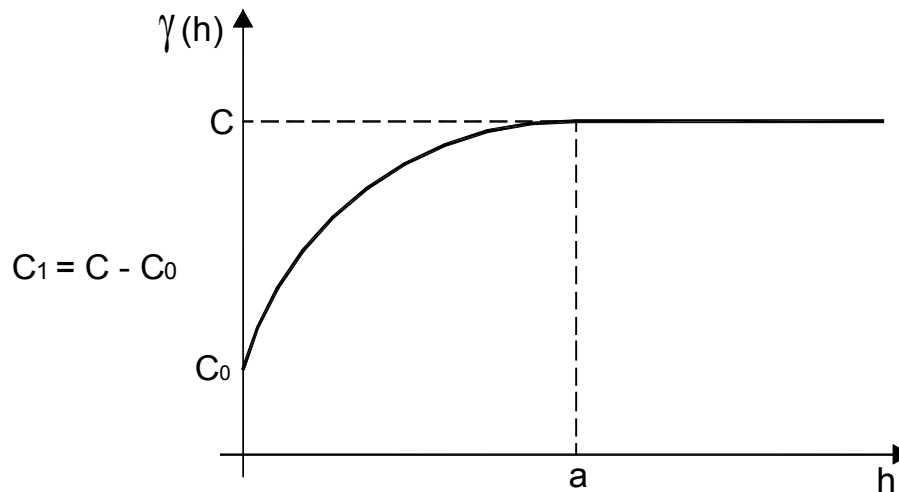


Figura 4 - Modelo esférico.

Modelo exponencial (Exp)

O modelo exponencial, apresentado na equação (9) e representado graficamente na Figura 5, é outro modelo muito utilizado no ajuste de semivariâncias experimentais (PANNATIER, 1996).

$$\gamma(h) = C_0 + C_1 \left[1 - e^{-\frac{3h}{a}} \right], \text{ se } h \geq 0 \quad (9)$$

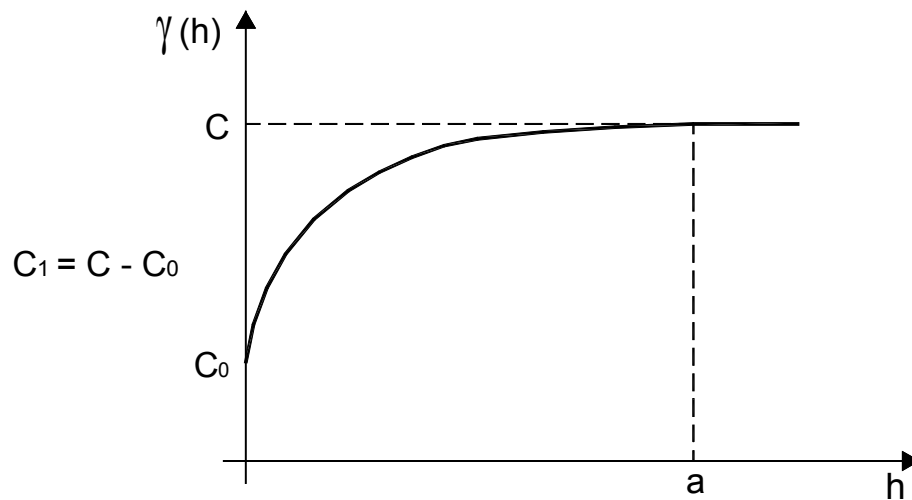


Figura 5 - Modelo exponencial.

Esse modelo atinge o patamar assintoticamente, com o alcance prático definido como a distância na qual o valor do modelo é 95% de C_1 (ISAAKS; SRIVASTAVA, 1989). Segundo CRESSIE (1993), os modelos exponenciais são válidos nos espaços R , R^2 e R^3 .

Modelo gaussiano (Gau)

O modelo gaussiano que, segundo PANNATIER (1996), é utilizado para modelar fenômenos contínuos, é dado pela equação (10).

$$\gamma(h) = C_0 + C_1 \left[1 - e^{-3\left(\frac{h}{a}\right)^2} \right], \text{ se } h \geq 0 \quad (10)$$

O que caracteriza esse modelo é seu comportamento parabólico próximo à origem e é o único modelo que apresenta, em sua forma, um ponto de inflexão (ISAAKS; SRIVASTAVA, 1989).

A representação gráfica do modelo gaussiano é apresentada na Figura 6 e é válido nos espaços R , R^2 e R^3 .

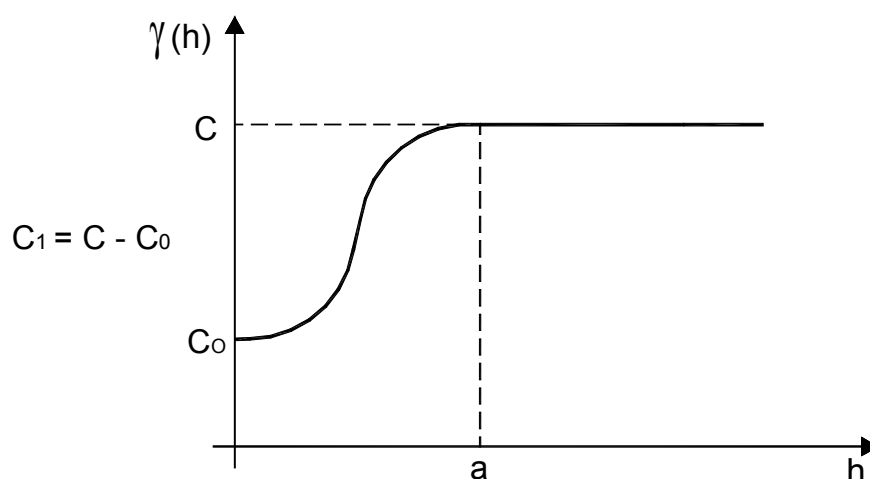


Figura 6 - Modelo gaussiano.

2.1.4 Estimação de Parâmetros no Ajuste de Modelos

Os solos variam continuamente no espaço, ao menos em escalas práticas. Semivariogramas de atributos do solo são, portanto, funções contínuas. Os semivariogramas experimentais, entretanto, são constituídos por um conjunto discreto de valores, que são estimativas sujeitas a erros (WEBSTER, 1985).

Escolhido, dentre os modelos definidos na Seção 2.5.3, um modelo que pareça adequado, é necessário ajustá-lo ao semivariograma experimental, estimando-se os parâmetros efeito pepita C_0 , o patamar C e alcance a .

Muitos pesquisadores optam por um ajuste visual. BARNES (1991) comenta que uma diretriz, que muitas vezes orienta esse tipo de ajuste, e que nem sempre é adequada, é a que estima o patamar como a variância dos dados amostrados. O autor demonstra que, se o patamar é claramente presente no gráfico do semivariograma, seu valor pode ser usado como uma estimativa da variância populacional. Porém, a variância amostral não pode ser usada como uma estimativa do patamar do semivariograma.

Segundo McBRATNEY e WEBSTER (1986), um procedimento baseado em critérios estatísticos deve ser preferido a um ajuste visual, na estimação do vetor de parâmetros desconhecidos de um determinado modelo geoestatístico. A escolha de um bom modelo ajustado não é um procedimento automático. Requer um bom julgamento, baseado na experiência e compreensão das limitações da função.

Os métodos de estimação de parâmetros encontrados na literatura são: Mínimos Quadrados Ordinários, Mínimos Quadrados Ponderados, Máxima Verossimilhança e Métodos Bayesianos. Descrevem-se, a seguir, os métodos que fazem parte deste trabalho.

Método dos Mínimos Quadrados Ordinários (OLS)

Seja $\theta = (C_0, C_1, a)^T$ o vetor de parâmetros desconhecidos a serem estimados.

Segundo o Método dos Mínimos Quadrados Ordinários, o vetor θ a ser escolhido é o que minimiza a expressão (11):

$$\sum_{j=1}^k [\gamma^*(h_j) - \gamma(h_j, \theta)]^2 \quad (11)$$

em que, $\gamma^*(h_j)$ é o valor estimado da semivariância correspondente ao lag h_j para $j = 1, 2, \dots, k$, utilizando o estimador dado na equação (3) ou (4), e $\gamma(h_j, \theta)$ é a semivariância correspondente a h_j , dada pelo modelo teórico escolhido para ser ajustado, dependente do vetor de parâmetro θ .

Segundo CRESSIE (1985), o método dos mínimos quadrados ordinários é apenas um método numérico e não um método estatístico, entretanto, é o que mais é citado pela literatura.

Métodos dos mínimos quadrados ponderados (WLS1)

O método WLS1 (CRESSIE, 1993) é um método de mínimos quadrados ponderados, com pesos diretamente proporcionais ao número de pares de pontos amostrais, que contribuem para a semivariância estimada a cada *lag*. Segundo o método, o vetor de parâmetros θ a ser estimado é o que minimiza a expressão:

$$\sum_{j=1}^k N(h_j) [\gamma^*(h_j) - \gamma(h_j, \theta)]^2. \quad (12)$$

Na literatura há indicações de outros métodos de estimação de mínimos quadrados como, por exemplo, os propostos por CRESSIE (1993) e PANNATIER (1996).

2.1.5 Krigagem

A krigagem é um método de interpolação geoestatístico que utiliza a estrutura de dependência espacial, determinada pelo ajuste de um modelo teórico ao semivariograma experimental. Sendo assim, ela é um método usado com o objetivo de estimar valores de variáveis, para locais em que não foram efetuadas medidas, a partir de valores adjacentes interdependentes conhecidos.

Dentre os diferentes métodos, a krigagem ordinária é a que utiliza um estimador linear não-viciado, com mínima variância (*BLUE - Best Linear Unbiased Estimator*), para interpolação dos pontos em posições não-amostradas (ISAAKS; SRIVASTAVA, 1989). É linear, porque suas estimativas são feitas por combinações lineares; *Unbiased* (sem viés), porque o erro de estimativa esperado é nulo; e *Best* (melhor), porque é o melhor estimador que minimiza a variância destes erros de estimativa. Como foi dito, interpolação por krigagem é uma combinação linear, isto é, uma média móvel de vizinhos, que considera a

estrutura de variabilidade encontrada na variável regionalizada em estudo. Pontos próximos da posição a ser interpolada levam maiores pesos que os mais distantes.

A krigagem ordinária é interpretada como o valor interpolado de uma variável regionalizada Z , num local s_0 , podendo ser determinada pela equação (13), KRIGE (1951):

$$\hat{Z}(s_0) = \sum_{i=1}^n \lambda_i Z(s_i) \quad (13)$$

restrito a:

$$\sum_{i=1}^n \lambda_i = 1,$$

em que:

- $\hat{Z}(s_0)$: valor estimado para local s_0 não amostrado;

- n : número de valores medidos $Z(s_i)$ das variáveis em pontos amostrados;

- λ_i : pesos associados a cada valor $Z(s_i)$ medido.

A krigagem ordinária para n pontos amostrais constitui-se de $n + 1$ equações, com $n + 1$ incógnitas, podendo ser representada por meio de uma equação matricial dada em (14):

$$K \cdot \lambda = L \quad (14)$$

sendo:

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

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 s_0 ; e μ , o

parâmetro de Lagrange. O vetor de parâmetros λ é obtido pelo produto do inverso da matriz K pelo vetor L , dado na equação (15):

$$\lambda = K^{-1} \cdot L \quad (15)$$

De acordo com RIBEIRO JR. (1995) e CAMARGO (1997), o processo de krigagem se diferencia dos outros métodos de interpolação pela forma de atribuição dos pesos, pois nesse método não se utiliza a distância euclidiana entre os pontos, mas uma distância estatística que expressa a distância e a estrutura de variabilidade (semivariância ou covariância). Não apenas a distância dos vizinhos ao ponto a ser estimado é considerada, mas, também, as distâncias entre eles na influência da distribuição dos pesos. Assim, os vizinhos agrupados têm importância individual relativamente menor do que aqueles isolados.

No método da krigagem, os pesos são atribuídos de acordo com a variabilidade espacial expressa no semivariograma (VIEIRA, 1998). No entanto, o que torna a krigagem um interpolador ótimo é a maneira como os pesos são distribuídos, não sendo tendencioso, tendo variância mínima e possibilitando que se conheça a variância da estimativa (WEBSTER; OLIVER, 1990).

2.1.6 Critérios de Validação dos Ajustes de Modelos Geoestatísticos

Entre os métodos de validação de modelos, os que vêm sendo utilizados são aqueles cujas técnicas são baseadas na comparação entre valores teóricos de modelos geoestatísticos e valores empíricos. Essas técnicas podem ajudar a escolher um, entre os diferentes modelos de semivariogramas baseados na análise dos erros de estimação.

Os resultados de validação são comumente utilizados para comparar a distribuição da estimação de erros, ou resíduos, dos diferentes procedimentos. Tal comparação, geralmente, não indica qual alternativa é a melhor. Os resíduos da validação têm informações importantes e um estudo cuidadoso da sua

distribuição pode prover indícios de problemas em um procedimento de estimação. Desde que tais indícios possam conduzir ao melhoramento de procedimentos de estimação de casos específicos, a validação é o passo preliminar útil antes que as estimações finais sejam calculadas (ISAAKS; SRIVASTAVA, 1989).

Outra metodologia para validação de ajustes são os cálculos de regressão linear aplicados aos pares de dados, referentes ao valor amostrado e ao valor estimado pelo modelo.

A seguir, apresentam-se os critérios de Validação Cruzada, *Jackknifing*, Filliben e Akaike para validação dos ajustes dos modelos teóricos.

Validação cruzada

Ao ajustar um modelo teórico aos pontos de um semivariograma experimental pode haver dúvida se tal ajuste é o melhor possível. Faz-se, então, vários ajustes, baseados nos mínimos quadrados, e decide-se, sob algum critério, qual o melhor modelo a adotar.

Um critério razoável é a avaliação dos resíduos obtidos, quando se compara o valor do semivariograma teórico, ajustado com o valor correspondente do semivariograma experimental, em uma dada distância h .

A validação cruzada é uma técnica de avaliação de erros de estimativa que permite comparar valores, estimados e amostrados, usando-se somente a informação disponível em nossa amostra de dados. Tais comparações são úteis para ajudar a escolher entre os diferentes modelos de estimação. Em várias situações práticas, é necessário verificar os resultados de diferentes modelos e escolher o que mais se aproxima do semivariograma experimental (ISAAKS; SRIVASTAVA, 1989).

Segundo DAVIS (1987), a validação cruzada é usada para encontrar o melhor modelo entre os estudados. DELFINER (1976) a apresenta, nesse

enfoque, quando a usa para buscar por uma função de covariância generalizada, para ser usada no seu procedimento de estimação entre um número finito de modelos.

Na aplicação da validação cruzada, o modelo é testado nos locais das amostras existentes. O valor da amostra, em certa localização, é temporariamente descartado do conjunto de dados da amostra; o valor na mesma localização é, então, estimado por krigagem, usando-se as amostras restantes. Esse procedimento pode ser visto como um experimento no qual se imita o processo de estimação ao supor que nunca se toma uma amostra naquela localização. Com a estimação calculada, pode-se compará-la ao valor da amostra que foi inicialmente removida do conjunto de dados amostrais. Esse procedimento é repetido para todas as amostras disponíveis, e é designado como o método de “deixar um fora” (ISAAKS; SRIVASTAVA, 1989).

Conhecendo-se os valores amostrados e os valores estimados, pode-se conhecer também a variância total da estimativa, sendo, assim, possível avaliar a qualidade ou precisão do processo.

Existem limitações à validação cruzada que devem ser lembradas ao se analisar os resultados de um estudo para escolher um modelo teórico. Por exemplo, ela gera pares de valores amostrados e estimados somente no local da amostra. Seus resultados normalmente não refletem exatamente a real performance de um método de estimação, porque estimação no local das amostras é tipicamente não-representativa da estimação de todos os locais não amostrados (ISAAKS; SRIVASTAVA, 1989).

Espera-se, segundo GONÇALVES (1997), que os erros de estimação, dados por $e = Z_i - \hat{Z}_{(i)}$, em que $\hat{Z}_{(i)}$ é o valor estimado sem a i -ésima observação, tenham média nula, variância unitária e distribuição normal de probabilidades. As características do erro de estimação indicam não somente a eficácia do ajuste dos diferentes modelos teóricos ao semivariograma experimental e a modelagem do processo em questão, mas, também, a avaliação da estacionariedade e do peso ou importância da presença de dados atípicos.

Segundo ISAACS e SRIVASTAVA (1989), a validação cruzada também permite avaliações qualitativas de como os métodos de estimação funcionam. As características espaciais dos resíduos devem ser vistas com atenção, quando a validação cruzada é usada como uma técnica puramente quantitativa para a escolha entre os procedimentos de estimação.

Espera-se que as estimativas sejam condicionalmente não tendenciosas em relação a qualquer escala de valores; também se espera que sejam condicionalmente não tendenciosas com respeito à localização. O mapa de contorno dos resíduos pode revelar áreas nas quais as estimativas estão consistentemente tendenciosas (ISAACS; SRIVASTAVA, 1989).

A disponibilidade de pares de valores estimados e amostrados permite examinar questões que estão especificamente relacionadas com os objetivos do processo de estimação. Na maioria dos estudos práticos de geoestatística existem critérios econômicos ou técnicos que são mais relevantes que o critério estatístico que se está usando para avaliar um conjunto de estimativas (ISAACS; SRIVASTAVA, 1989).

Critério Jackknifing

Segundo DAVIS (1987), logo a partir da primeira menção na literatura geoestatística, ver DAVID (1976) e, especialmente, DELFINER (1976), confusão e má aplicação cercaram o uso da validação cruzada em geoestatística. Aparentemente, por este último ter usado tanto a validação cruzada e uma técnica de redução de tendências, chamada *jackknifing*, em seu trabalho, tem existido uma tendência por parte de outros autores, como PARKER, JOURNAL e DIXON (1979), de referir-se à validação cruzada como *jackknifing*. Infelizmente, essa confusão ainda existe. Vários *softwares* disponíveis comercialmente contêm rotinas para *jackknifing* o semivariograma. Conclui-se que o termo validação cruzada significa coisas diferentes para pessoas diferentes.

Jackknife é um estimador introduzido por QUENOUILLE (1956) para reduzir tendências direcionais. TUKEY (1958) estendeu o uso utilizando o *jackknife* para construir intervalos de confiança aproximados. Desde então, muitos trabalhos apareceram sobre *jackknifing*.

A literatura, desde o aparecimento dessa técnica, apresenta não só diferentes maneiras de redigir a palavra (*jacknife*, *jack-knife*, *jack-knifing*, *jack knifing*), como até entendimentos diferentes para o processo. Alguns pesquisadores utilizam a expressão *jackknifing* como se fosse um substantivo. Neste trabalho, a palavra *jackknife* significa o estimador. E a palavra *jackknifing*, o método de obtenção do estimador.

VIEIRA, CARVALHO e MORAES (2004) apresentam o procedimento de *jackknifing* como se fosse a validação cruzada relatada por ISAACS e SRIVASTAVA (1989). Segundo eles, em todo o processo do estudo da estrutura espacial e construção de mapas por krigagem, existe um grau de incerteza sobre as hipóteses assumidas ou sobre os parâmetros ajustados aos modelos. Essa incerteza é o erro da estimação, o qual pode ser avaliado usando-se o procedimento de autovalidação, comumente chamado de *jackknifing*. Esse procedimento envolve a estimativa de cada ponto medido, fazendo-se de conta que ele não existe, porque, senão, a solução do sistema de krigagem forneceria o peso associado a ele, com valor unitário e todos os outros pesos iguais a zero. A razão para isso está em que a krigagem é um interpolador exato, passando exatamente pelo ponto medido, quando este é usado no cálculo. Porém, quando se considerar que o valor não existe, ele será estimado normalmente como se fosse ponto perdido, levando-se em conta a variabilidade espacial local, expressa nas primeiras distâncias, no semivariograma (VIEIRA; CARVALHO; MORAES, 2004). Esse entendimento dos pesquisadores fundamenta-se em JOURNEL e HUIJBREGTS (1978).

Portanto, conforme DAVIS (1987), o *jackknifing* é um processo de redução de tendências direcionais e não um critério para escolha de um melhor modelo geoestatístico.

Critério de Filliben

Ajustar modelos teóricos, como os apresentados na Seção 2.5.3, aos semivariogramas experimentais, é um procedimento subjetivo. A qualidade de ajuste pode ser verificada por meio da técnica de validação cruzada. Como, para cada local, tem-se um valor medido e pode-se estimar um outro valor, utilizando-se a krigagem, então, pode-se calcular a regressão linear entre esses pares de dados e calcular o coeficiente linear α , o coeficiente angular β , a correlação entre os pares r^* e o erro reduzido com sua média e variância (VIEIRA, *et al.*, 1983; VIEIRA, 1997). O melhor ajuste é alcançado quando os valores obtidos se aproximam dos seguintes valores ideais: $\alpha = 0$, $\beta = 1$, $r^* = 1$, com média do erro reduzido igual a zero e variância do erro reduzido igual a um.

Definido o modelo, os interpoladores de krigagem são usados e os pontos são estimados (VIEIRA *et al.*, 1983; CARVALHO, SILVEIRA; VIEIRA, 2002; CARVALHO; ASSAD, 2003) para a construção de mapas de isolinhas, ou mapas de superfícies tridimensionais, úteis na verificação e interpretação da variabilidade espacial. O teste de Filliben para resíduos ortonormais, a ser definido a seguir, pode ser de grande ajuda na determinação de qual o melhor modelo geoestatístico.

Considerando um vetor de dados observados, de tamanho n , pode-se aleatorizar a ordem dos elementos do vetor e, assim, obter-se uma nova amostra. Para a k -ésima posição ($1 \leq k \leq n$) de um dado, pode-se estimar, por krigagem ordinária, o valor de $\hat{Z}(s_k)$, usando-se somente os $k-1 = p$ valores anteriores de dados e normalizados pelo desvio padrão do erro da krigagem.

Os erros normalizados (LEE, 1994) para $n-p$ variáveis são definidos como:

$$\varepsilon_k = \frac{Z(s_k) - \hat{Z}(s_k)}{\sigma_k} \quad (16)$$

para $k = p + 1, \dots, n$, em que $Z(s_k)$ são os valores amostrados, $\hat{Z}(s_k)$ são os valores estimados pelo interpolador de krigagem ordinária e σ_k é o desvio padrão da estimativa da krigagem. Os $n - p$ resíduos obtidos pela equação (16) são chamados resíduos ortonormais, isto é, são não correlacionados, linearmente independentes e têm variância unitária.

No teste de Filliben (FILLIBEN, 1975), dado pela equação (18), calcula-se o coeficiente de correlação r^* , definido como o produto dos momentos entre as observações ordenadas X_i e a estatística das medianas ordenadas M_i , de distribuição normal $N(0, 1)$. Quanto mais próximo r^* é de 1, mais normal é a distribuição dos dados. Se os valores tabelados de r (Anexo) para um determinado tamanho de amostra são maiores do que r^* , a hipótese de que os dados seguem uma distribuição normal é rejeitada para certo nível de significância. Praticamente, o que se testa é se os resíduos ortonormais seguem uma distribuição normal a um determinado nível de significância.

Se Y representa o vetor de $n - p$, resíduos ortonormais e X uma amostra ordenada de Y , pode-se gerar estatísticas das medianas ordenadas m_i . A partir de uma população com distribuição uniforme $N(0, 1)$, por meio do seguinte algoritmo:

$$\begin{aligned} m_i &= 1 - m_n && \text{para } i = 1, \\ m_i &= (i - 0,375)/(n + 0,365) && \text{para } i = 2, 3, \dots, n - 1, \\ m_i &= \sqrt[n]{(0,5)} && \text{para } i = n. \end{aligned}$$

A seguir, obtém-se a estatística mediana ordenada M_i de uma população com distribuição normal a partir de Tabelas de Distribuição Normal Padronizada.

$$m_i = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{M_i} \exp(-t^2/2) dt \quad (17)$$

Assim, pode-se calcular:

$$r^* = \frac{\sum_{i=1}^n (X_i - \bar{X})(M_i - \bar{M})}{\sqrt{\left[\sum_{i=1}^n (X_i - \bar{X})^2 \right] \left[\sum_{i=1}^n (M_i - \bar{M})^2 \right]}} \quad (18)$$

em que:

- X_i : resíduos ordenados;
- \bar{X}_i : média dos resíduos;
- M_i : estatística de uma distribuição Normal (0,1);
- \bar{M} : média da estatística de uma distribuição Normal (0,1).

Critério de Akaike

Procura-se uma solução satisfatória entre o bom ajuste e o princípio da parcimônia, aplicando-se o chamado Critério de Informação de Akaike (*Akaike's Information Criterion - AIC*).

AKAIKE (1973) e depois, SAKAMOTO, ISHIGURO e KITAGAWA (1986), desenvolveram estudos visando conhecer como os modelos são usados para fazer previsões. O Teorema de Akaike busca demonstrar que se dois modelos representam dados igualmente satisfatórios, então, com o modelo mais simples, pode-se esperar um melhor desempenho para a previsão de novos dados. Portanto, o Critério de Akaike impõe uma penalidade para a complexidade.

SOBER (2002) pondera sobre a seleção de modelos em termos de aproximação das previsões e afirma que os critérios de seleção de modelos têm implicações filosóficas importantes. Modelos, em primeiro lugar, são expressões que contêm parâmetros ajustáveis.

Para a tomada de decisão sobre o melhor modelo geoestatístico, quando do ajuste do semivariograma, pode ser usado um índice de desempenho como a

Soma dos Quadrados dos Resíduos. Esse critério é muito utilizado para a escolha do melhor modelo, porém não pondera sobre o número de componentes usados para o modelo estatístico estimado. Para um conjunto de dados, a parte variável do *AIC* é estimada por:

$$\hat{A} = n \ln(R) + 2p \quad (19)$$

em que:

- n : é o número de pontos experimentais do semivariograma;
- R : é a soma dos quadrados dos resíduos;
- p : é o número de parâmetros no modelo.

A decisão para escolha entre os modelos estudados recai sobre aquele que apresentar o menor valor de \hat{A} .

3 MATERIAL E MÉTODOS

SIMULAÇÃO DE DADOS COM ESTRUTURA DE DEPENDÊNCIA ESPACIAL

Segundo CRESSIE (1993), dados são importantes para o desenvolvimento de métodos estatísticos e se dados reais não são disponíveis, pode-se utilizar dados simulados. Validar um procedimento sobre dados dos quais se conhecem os verdadeiros parâmetros, pode ser uma parte essencial para reconhecer sua importância.

Um requisito para qualquer processo de simulação é o de que todos os parâmetros do processo sejam especificados. A fonte de valores aleatórios é, usualmente, um gerador de pseudonúmeros randômicos (CRESSIE, 1993).

Por meio do *software R* (R DEVELOPMENT CORE TEAM, 2005), utilizou-se uma função geradora de números aleatórios que tem como argumento uma “semente” fornecida pelo usuário. Essa função, denominada *grf*, utiliza o método de decomposição de Cholesky, descrito por CRESSIE (1993). A geração dos dados seguiu uma distribuição normal de probabilidades e foi utilizado um modelo de dependência espacial esférico. O vetor de parâmetros θ para o modelo de dependência é $\theta = (0; 15; 60)$, em que $C_0 = 0$, $C_1 = 15$ e $a = 60$ significam, respectivamente, efeito pepita, contribuição e alcance.

Os dados simulados foram alocados a uma malha quadrada de 100 pontos, com lado igual a 100 unidades, que pode ser visualizada na Figura 7.

Portanto, o espaçamento entre os pontos no sentido do eixo dos x e dos y ficou estabelecido como sendo de 10 unidades.

Conhecidos os valores gerados, passou-se à análise estatística e geoestatística, construiu-se o semivariograma, realizou-se a krigagem ordinária e estudaram-se os critérios de validação dos modelos.

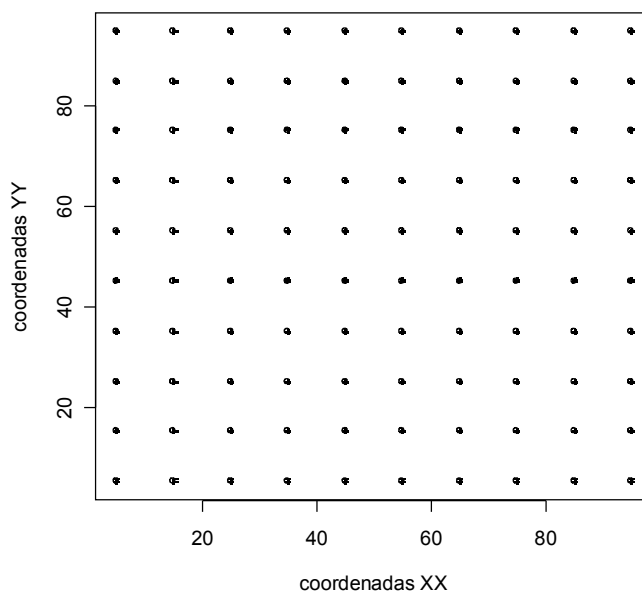


Figura 7 - Representação da grade amostral 10x10 dos dados simulados.

LOCALIZAÇÃO E HISTÓRICO DA ÁREA ESTUDADA

A pesquisa aplicada desenvolveu-se na região Oeste do Estado do Paraná, em uma área de produção de grãos de 57 ha, com Latossolo Vermelho Distrófico, numa propriedade localizada no município de Cascavel. Os dados provêm do experimento realizado no ano agrícola 2004/2005, pelo Grupo de Geoestatística Aplicada – GGEA, da Universidade Estadual do Oeste do Paraná. A variedade da soja semeada na área em estudo foi a COODETEC 216 (CD216).

O levantamento topográfico e o posicionamento dos locais de amostragem foram feitos por meio de receptores GPS, pelo método estático, com correção diferencial pós-processada.

O clima da região apresenta-se como temperado mesotérmico e superúmido, tipo climático Cfa (Köppen) e com temperatura anual média de 21°C.

AMOSTRAGEM

Na área experimental foram demarcadas as 100 parcelas com auxílio de teodolito e trena. A área foi georreferenciada com auxílio de um aparelho GPS, visando a sua correta localização em um sistema de coordenadas geográficas (latitude/longitude).

Foram usados pontos georreferenciados, com espaçamento médio de 75 m entre parcelas. A área em estudo pode ser visualizada na Figura 8. Os pontos correspondem às parcelas amostradas.

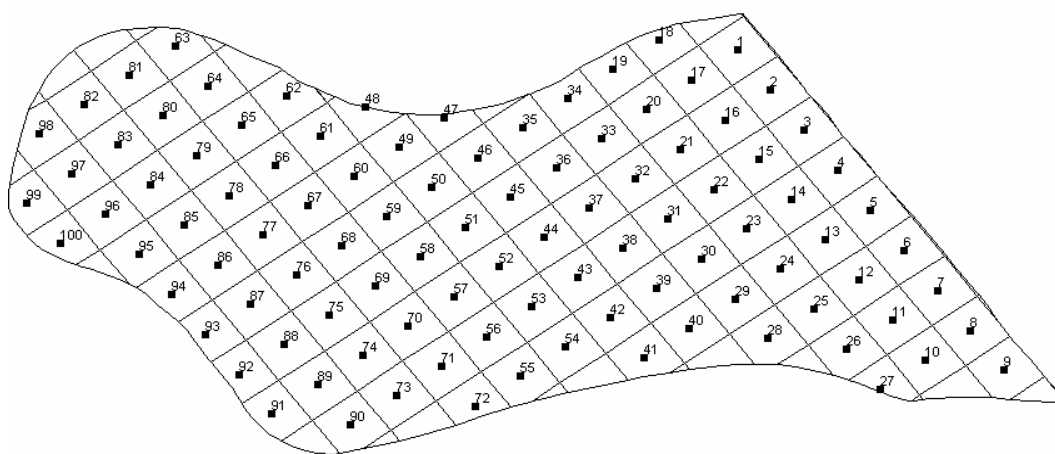


Figura 8 - Área em estudo de 57 ha.

O sistema de coordenadas geográficas utilizado foi o Universal Transverse Mercatur (UTM), que utiliza coordenadas métricas. Os valores mínimos para as coordenadas, em metros, são: $X = 239\,683,8$ e $Y = 7\,237\,073,1$ e os valores máximos são: $X = 240\,993,8$ e $Y = 7\,237\,660,1$. A distância mínima entre dois pontos amostrados é de 74,8 m e a distância máxima é de 1337,6 m.

As parcelas 16, 21, 67 e 68, visualizáveis na Figura 8, estavam cultivadas por milho, motivo pelo qual não se têm amostras de produtividade de soja nesses locais.

PROCEDIMENTO PARA COLETA DE DADOS

3.1.1 Resistência do Solo à Penetração, Densidade e Umidade

Para os dados da resistência do solo à penetração usou-se o penetrômetro, aparelho com uma haste de 600 mm, para avaliação de resistência do solo à penetração, desenvolvido no laboratório de Mecanização Agrícola da UNIOESTE (TIEPPO, 2004). Esse aparelho fornece resultados a cada 0,25 segundos, em kgf, à medida que vai sendo introduzido no solo. Assim, após a transformação dos dados para MPa, determinou-se o valor da resistência do solo à penetração nas camadas 0-10, 10-20 e 20-30 cm, calculando-se a média das resistências medidas nesses intervalos. Foram realizadas coletas para diferentes profundidades, procurando-se manter a velocidade de penetração em 3 cm seg^{-1} . As leituras que não apresentaram consistência foram eliminadas.

Os dados de densidade do solo foram coletados em três camadas, sendo a primeira entre 0 e 10 cm, a segunda entre 10 e 20 cm e a terceira entre 20 e 30 cm. A obtenção dos dados seguiu os procedimentos recomendados por EMBRAPA (1997), também denominado de método do anel volumétrico. Em campo, um anel de volume V , de bordos cortantes, foi introduzido no solo até ficar completamente cheio. Depois de medida sua massa em balança de precisão, a amostra foi colocada em estufa a uma temperatura de 105°C , por 24 horas. Após o resfriamento da amostra, foi realizada nova medição da massa. Determinou-se a densidade do solo (D_s) pela razão entre o peso de solo seco e o volume total de solo coletado no cilindro; sua unidade dada em Mg m^{-3} .

Para a determinação da umidade do solo utilizaram-se as mesmas amostras coletadas para a determinação da densidade. Portanto, a umidade foi determinada nas profundidades entre 0 e 10 cm, entre 10 e 20 cm e entre 20 e 30 cm. Determinou-se, a seguir, a umidade à base de peso, ou gravimétrica, pela razão entre a massa de água e a massa de solo seco e a unidade, dada em kg kg^{-1} . A umidade à base de peso é adimensional. Em seguida, multiplicou-se pela densidade média, nesse intervalo, para obtenção da umidade volumétrica, dada em $\text{m}^3 \text{ m}^{-3}$.

3.1.2 Produtividade da Soja

Para estabelecer a produtividade, em cada ponto amostral, foi colhida a soja de uma área de 1m^2 , pesada, e a produtividade transformada para Mg ha^{-1} . Também foi verificada a umidade para cada amostra. A umidade de correção para a produtividade considerada foi de 13%.

ANÁLISE DOS DADOS

O estudo dos dados iniciou-se com a análise exploratória para avaliação do comportamento geral e identificação de pontos discrepantes.

A análise preliminar da tendência direcional foi realizada pelos gráficos *post-plot* e gráficos de dispersão, para os dados *versus* linhas e dados *versus* colunas. Uma análise de tendência dos dados, por meio de um exame gráfico, que represente a área de coleta, é denominado *post-plot* e é constituído de pontos marcados por coordenadas que representam os locais amostrados. O tamanho dos pontos é proporcional aos intervalos separados pelos quartís. Portanto, aparecem, nesse gráfico, quatro tamanhos de pontos que indicam as representações dos dados correspondentes a quatro intervalos. Definiu-se esses intervalos pela letra P, seguida de números de 1 a 4 e tendo a seguinte descrição:

- P-1: intervalo entre o valor mínimo observado e o 1º quartil;
- P-2: intervalo entre o 1º quartil e a mediana;
- P-3: intervalo entre a mediana e o 3º quartil;
- P-4: intervalo entre o 3º quartil e o valor máximo observado.

Para a análise da estrutura de dependência espacial utilizaram-se semivariogramas experimentais construídos utilizando-se o estimador de Matheron (MATHERON, 1963), para os dados que apresentaram distribuição normal, e o estimador de Cressie e Hawkins (CRESSIE; HAWKINS, 1980), quando os dados não apresentaram normalidade. Com a finalidade de se aplicarem os critérios de validação ora em estudo, ajustaram-se três modelos teóricos ao semivariograma experimental: exponencial, esférico e gaussiano, considerados adequados aos dados em análise. Para a estimação dos parâmetros, utilizaram-se dois métodos: o dos mínimos quadrados ordinários (OLS) e o dos mínimos quadrados ponderados (WLS1).

Para a estimação e aplicação de ajuste de modelos e avaliação de critérios de ajuste foi utilizado o *software R* (R DEVELOPMENT CORE TEAM, 2005) e o *software Surfer 6.0*, para confecção dos mapas temáticos.

Com a comparação entre os resultados alcançados, foram analisados os comportamentos dos critérios de Validação Cruzada, Akaike e Filliben. Também se procurou conhecer semelhanças e divergências entre os diversos critérios de validação.

Finalmente, foram construídos mapas temáticos da produtividade da soja, referente ao ano agrícola 2004/2005, e dos atributos físicos, do solo em estudo, que tiveram melhor ajuste utilizando-se os critérios de validação e *software* gráfico.

4 RESULTADOS E DISCUSSÃO

SIMULAÇÃO

A geração dos dados seguiu uma distribuição normal de probabilidades e foi utilizado um modelo de dependência espacial esférico, com vetor de parâmetros $\theta = (0; 15; 60)$, ou seja, efeito pepita $C_0 = 0$, contribuição $C_1 = 15$ e alcance $a = 60$, na construção da matriz de covariância. Os dados gerados por meio da simulação estão apresentados no Apêndice C e os resultados da estatística descritiva dos dados simulados são apresentados na Tabela 1, sendo o valor mínimo de 0,30 e o valor máximo de 19,26. O coeficiente de variação calculado para os dados simulados é de 40,35 %, considerado muito alto (GOMES, 2000). Conclui-se que os dados apresentam grande dispersão em relação à sua média.

Tabela 1 - Análise exploratória e teste de normalidade para os dados simulados

DADOS	N	MÉDIA	MÍNIMO	1° QUARTIL	MEDIANA	3° QUARTIL	MÁXIMO	DESVIO PADRÃO
SIMULADOS	100	9,447	0,30	6,86	8,81	11,86	19,26	3,81

TESTES NORMALIDADE (p-valor)						
DADOS	CV (%)	SW	AD	KS	As	K
SIMULADOS	40,35	0,043	0,034	0,05 *	0,530	0,226

Notas: N: quantidade de dados; CV: coeficiente de variação; SW: Shapiro-Wilk; AD: Anderson-Darling; KS: Kolmogorov-Smirnov; As: assimetria; K: curtose; * não rejeita a normalidade ao nível de 5 % de probabilidade.

A forma da distribuição de freqüências pode ser analisada no histograma apresentado na Figura 9. Sua forma sugere uma distribuição normal de probabilidades. Os testes Kolmogorov-Smirnov de normalidade, aplicados aos dados, indicaram uma distribuição normal de probabilidades, ao nível de 5% de significância.

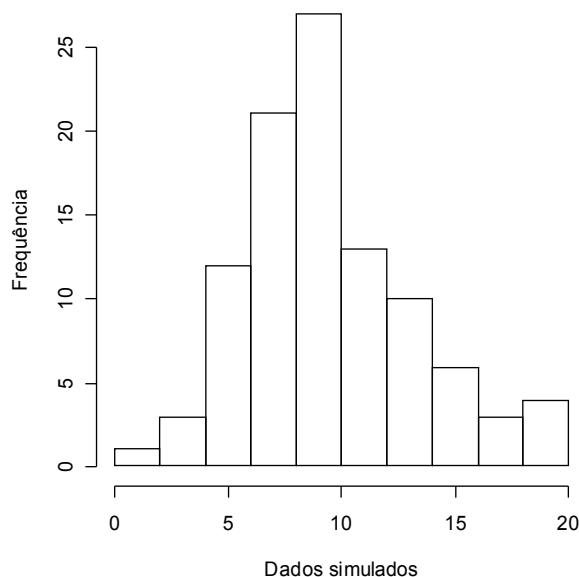


Figura 9 - Histograma para os dados simulados.

No gráfico *Boxplot*, apresentado na Figura 10, observa-se que os valores dos dados simulados que se encontram entre o 1º quartil e o 3º quartil, e que constituem 50 % dos valores amostrados, estão entre 6,86 e 11,86.

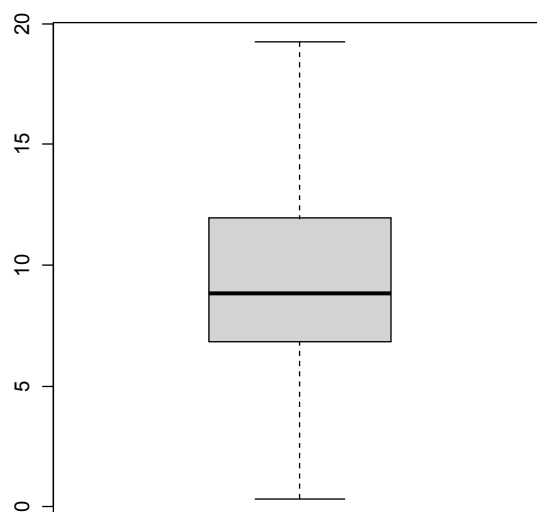


Figura 10 - Gráfico *Boxplot* para os dados simulados.

Pela análise do gráfico *post-plot*, apresentado na Figura 11, verifica-se que os dados simulados não apresentam tendências direcionais.

O semivariograma experimental, apresentado na Figura 12, foi calculado pelo estimador de Matheron, em razão dos dados simulados apresentarem distribuição normal de probabilidades.

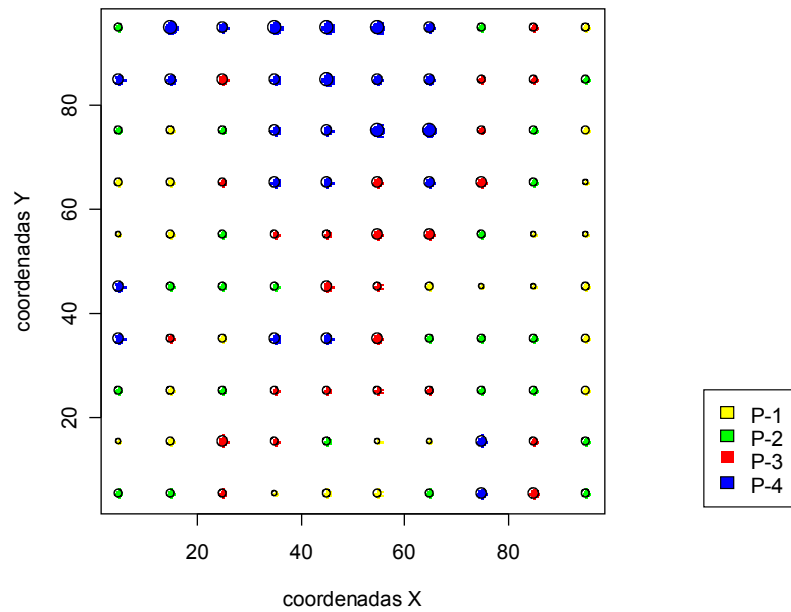


Figura 11 - *Post-plot* para os dados simulados.

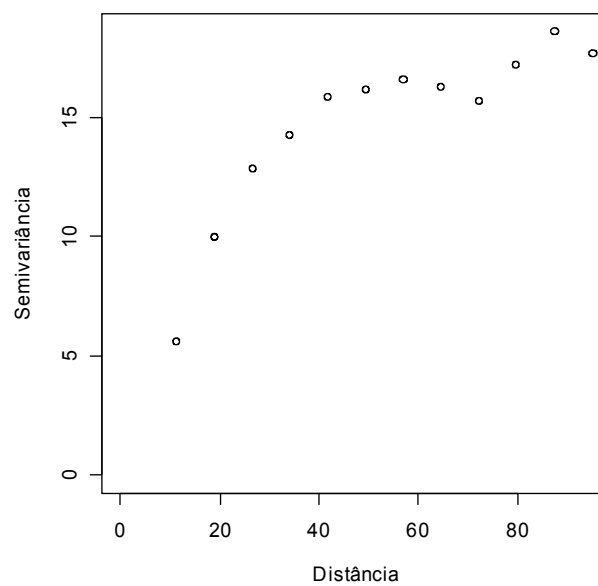


Figura 12 - Semivariograma experimental para os dados simulados.

Com a finalidade de se aplicar os critérios de validação, ora em estudo, ajustaram-se três modelos teóricos ao semivariograma experimental: exponencial, esférico e gaussiano, considerados adequados aos dados em análise. Para a estimação dos parâmetros utilizaram-se dois métodos: o dos mínimos quadrados ordinários (OLS) e o dos mínimos quadrados ponderados (WLS1). O método de máxima verossimilhança, por ser um método analítico, não é utilizado.

A Tabela 2 apresenta os parâmetros efeito pepita (C_0), patamar ($C_0 + C_1$) e o alcance (a) para os modelos ajustados. Na Figura 13 mostram-se os seis semivariogramas obtidos. Verifica-se, pela Tabela 2, que o modelo exponencial ajustado pelo método WLS1 é o que melhor se aproxima dos parâmetros verdadeiros.

Tabela 2 - Parâmetros dos modelos semivariográficos pelos métodos OLS e WLS1 para os dados simulados $\theta = (0;15;60)$,

Variáveis estudadas	Modelos	(C_0)		$(C_0 + C_1)$		(a)	
		OLS	WLS1	OLS	WLS1	OLS	WLS1
Dados simulados	Exponencial	0,0000	0,0000	17,9568	17,8498	70	68
	Esférico	0,6100	0,7749	16,8221	16,6272	50	50
	Gaussiano	3,5306	3,5618	16,8769	16,6854	44	43

Notas: OLS: Mínimos quadrados ordinários; WLS1: Mínimos quadrados ponderados.

Com a finalidade de confirmar os critérios de avaliação da qualidade de ajuste de modelos geoestatísticos, parte-se da suposição de que não se conhece o modelo de origem dos dados obtidos por simulação. Embora sejam obtidos modelos gaussianos para ajuste aos dados simulados, os parâmetros encontrados, neste caso, não se podem considerar satisfatórios em função de que o parâmetro C_0 é diferente de zero e também não tende para zero. Entretanto, prosseguiu-se no objetivo de se aplicar os critérios de avaliação de qualidade do ajuste, com a finalidade de se conhecer seus resultados.

Para análise dos resultados da validação de ajustes de modelos geoestatísticos é necessário aplicar os diversos métodos e proceder-se à comparação dos valores obtidos.

Os resultados apurados encontram-se na Tabela 3.

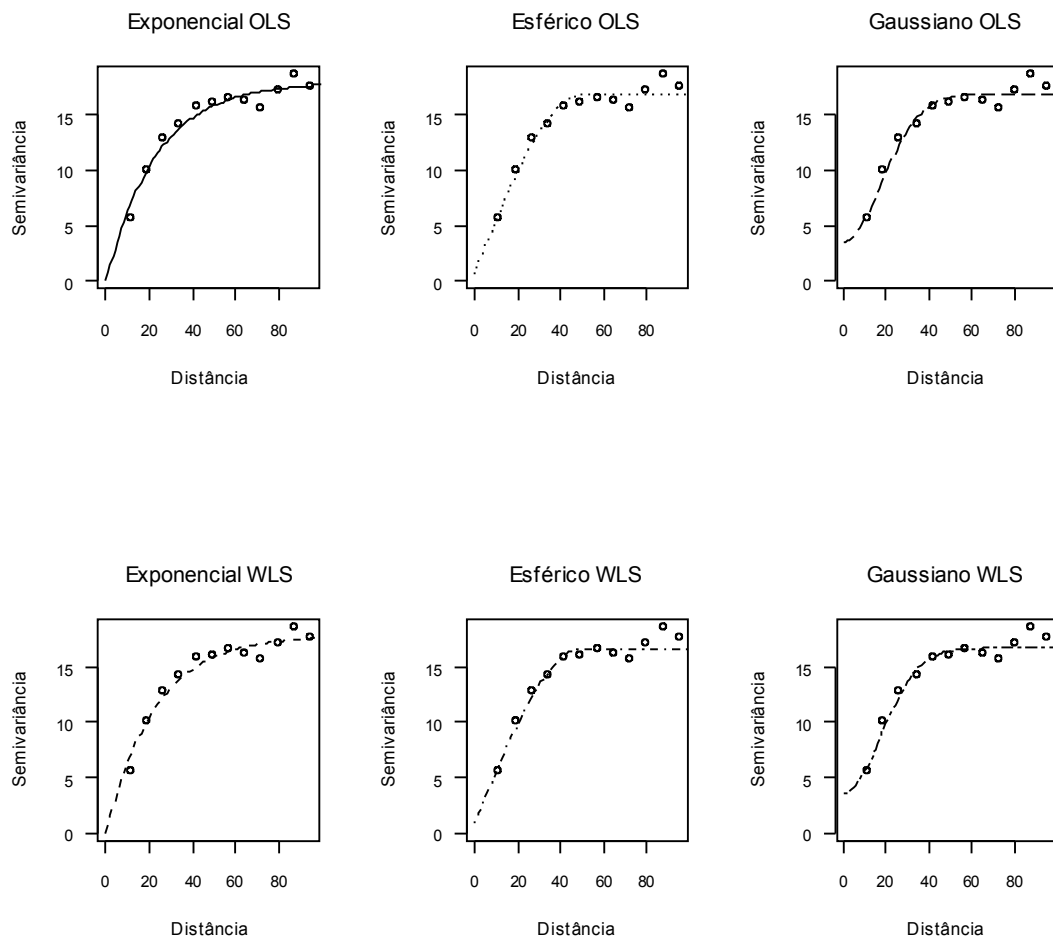


Figura 13 - Semivariogramas para os dados simulados.

Tabela 3 - Soma dos quadrados dos resíduos, AIC (\hat{A}), Filliben (r^*) e validação cruzada para os dados simulados

MODELOS	N	Soma quadrados residuais		\hat{A}		r^*	
		OLS	WLS1	OLS	WLS1	OLS	WLS1
Exponencial	100	6,97E+00	2,63E+03	200,1	793,6	0,9969802	0,9969802
Esférico	100	7,28E+00	2,24E+03	204,6	777,4	0,9976573	0,9974548
Gaussiano	100	6,79E+00	2,11E+03	197,5	771,3	0,9977535	0,9977742

VALIDAÇÃO CRUZADA : ANÁLISE DOS RESÍDUOS						
MODELOS	MIN	MEDIANA	MÉDIA	MAX	D PADRÃO	
Exponencial OLS	-4,80E+00	1,31E-01	-9,54E-03	4,75E+00	1,97E+00	
Exponencial WLS1	-4,80E+00	1,31E-01	-9,54E-03	4,75E+00	1,97E+00	
Esférico OLS	-4,97E+00	9,23E-02	-9,41E-03	4,90E+00	2,07E+00	
Esférico WLS1	-4,92E+00	8,01E-02	-9,01E-03	4,95E+00	2,07E+00	
Gaussiano OLS	-5,19E+00	9,33E-02	-9,37E-03	5,39E+00	2,10E+00	
Gaussiano WLS1	-5,17E+00	9,34E-02	-9,42E-03	5,34E+00	2,09E+00	

Notas: Em negrito a melhor escolha; N: número de pontos amostrados; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados; \hat{A} : critério de Akaike; r^* : teste de Filliben, compara-se com $r = 0,979$ (Anexo).

O Critério de Informação de Akaike (*Akaike's Information Criterion - AIC*) busca determinar valores estimados pela equação (20), apresentada na Seção 2.5.6.4. O modelo que apresentar o menor valor para \hat{A} será o escolhido. No caso dos dados simulados, é aquele correspondente ao valor 197,5 que, conforme a Tabela 3, aponta para o modelo gaussiano, com parâmetros estimados pelo método OLS.

O critério de Filliben testa se os resíduos ortonormais seguem uma distribuição normal a um determinado nível de significância. Para o caso dos dados simulados, que tem 100 valores, e para um nível de significância de 5 %, o valor de r obtido da tabela, apresentada no Anexo, é de 0,979. Comparando-se com os valores r^* , apresentados na Tabela 3, nas colunas Filliben OLS e WLS1, verifica-se que todos os valores r^* são maiores do que r e conclui-se que a hipótese de que os erros seguem uma distribuição normal não é rejeitada para um nível de significância de 5 % de probabilidade. Por esse critério, todos os modelos de semivariogramas teóricos podem ser aceitos.

O critério de validação cruzada, aplicado aos modelos em estudo, sugere que o menor valor da média dos erros de estimação indica o modelo com melhor

ajuste. Para os dados simulados em estudo, a escolha aponta para o modelo esférico, com parâmetros estimados pelo método WLS1.

Para se calcular o valor de \hat{A} do Critério de Informação de Akaike, utiliza-se o valor R , que é a soma dos quadrados dos resíduos, o número de pontos amostrados e o número de parâmetros, que no caso dos semivariogramas teóricos é constante. Portanto, no caso em estudo, o valor de \hat{A} não será influenciado pelo número de parâmetros.

Assim, além dos critérios apresentados, para fins de comparação, optou-se por calcular a soma de quadrados residuais e constatou-se, no presente estudo, que ocorreu coincidência das indicações de melhor modelo usando-se o Critério de Informação de Akaike e a soma dos quadrados dos resíduos. Verifica-se que, pelo critério da soma dos quadrados dos resíduos, ao se procurar por um modelo que apresente o menor valor, observando-se a Tabela 3, a escolha aponta para o modelo gaussiano, com parâmetros estimados pelo método OLS.

Como se partiu de um modelo de dependência espacial esférico e obtiveram-se resultados diversos, ao se aplicarem critérios de avaliação de qualidade de ajustes distintos, pode-se notar que não houve uma convergência dos critérios.

Para o caso da simulação em estudo, o critério da validação cruzada alcançou, com precisão, o modelo de dependência espacial proposto. O critério de Filliben, por outro lado, não foi restritivo, apontando todos os modelos ajustados como válidos. O critério de Akaike foi o menos preciso para este caso, pois apontou um modelo exponencial como o melhor, divergindo do modelo original.

Em resumo, a validação cruzada aponta como melhor modelo o esférico, que coincide com o modelo escolhido para gerar os dados simulados.

ANÁLISE DESCRITIVA, INFERÊNCIAS CLÁSSICAS E ANÁLISE DA VARIABILIDADE ESPACIAL

4.1.1 Densidade do Solo

Na Tabela 4, apresentam-se as estatísticas descritivas da densidade do solo e pode-se observar que os valores não apresentaram variações acentuadas, estando a média dentro de valores esperados para solos na região estudada, sendo da ordem de $1,0 \text{ Mg m}^{-3}$ a $1,45 \text{ Mg m}^{-3}$, conforme apresentado na Seção 2.2.

Tabela 4 - Análise exploratória e testes de normalidade do atributo densidade do solo nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm

DENSIDADE	N	MÉDIA	MÍNIMO	1º QUARTIL	MEDIANA	3º QUARTIL	MÁXIMO	DESVIO PADRÃO
0-10 cm	83	1,101	0,770	1,050	1,100	1,160	1,310	0,080
10-20 cm	88	1,118	0,980	1,050	1,110	1,160	1,550	0,090
20-30 cm	91	1,077	0,950	1,030	1,070	1,120	1,300	0,070

TESTES NORMALIDADE (p-valor)						
DENSIDADE	CV (%)	SW	AD	KS	As	K
0-10 cm	7,59	0,031	0,267 *	>0,150 *	-0,046	2,030
10-20 cm	7,69	0,000	0,002	>0,150 *	1,686	6,225
20-30 cm	6,42	0,013	0,078 *	>0,150 *	0,737	0,790

Notas: Unidade do atributo: $[\text{Mg m}^{-3}]$; N: número de pontos amostrados; CV: coeficiente de variação; SW: Shapiro-Wilk; AD: Anderson-Darling; KS: Kolmogorov-Smirnov; As: assimetria; K: curtose; * não rejeita a normalidade ao nível de 5 % de probabilidade. As diferenças para os valores no número de pontos ocorreram devido a erro experimental.

Observa-se também que a densidade do solo apresenta um coeficiente de variação de 7,59 %, para a profundidade 0 a 10 cm; 7,69 %, na profundidade 10 a 20 cm; e 6,42 %, na profundidade 20 a 30 cm, todas indicando homogeneidade, pois, segundo GOMES (2000), para coeficientes de variação inferiores a 10 %, a homogeneidade será considerada alta.

Pode-se observar, ainda, que a variável tem uma amplitude entre os valores mínimo e máximo de $0,54 \text{ Mg m}^{-3}$, na profundidade 0 a 10 cm. Para a profundidade 10 a 20 cm, a média e mediana são semelhantes às densidades encontradas para a profundidade 0 a 10 cm, que apresenta uma amplitude entre os valores mínimo e máximo de $0,57 \text{ Mg m}^{-3}$. Já para a profundidade 20 a 30 cm, a média e a mediana da densidade do solo foram pouco inferiores às outras duas profundidades estudadas e sua amplitude foi de $0,35 \text{ Mg m}^{-3}$.

O Teste de Normalidade de Kolmogorov-Smirnov (1965) revelou que os dados da densidade do solo, nas três profundidades, possuem distribuição normal de probabilidade, considerando-se um nível de 5 % de significância.

A Figura 14 apresenta os gráficos histogramas para esse atributo nas três profundidades estudadas.

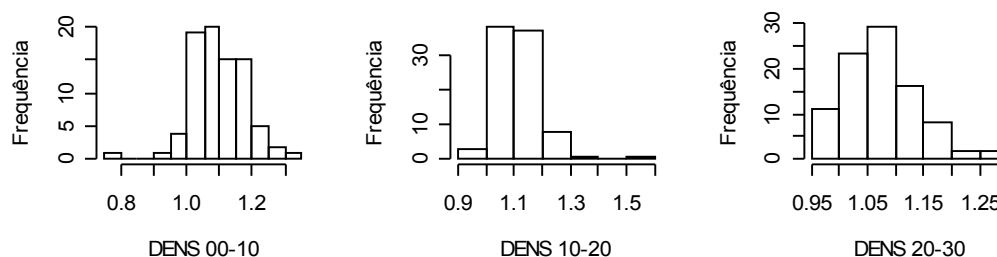


Figura 14 - Histograma para a densidade do solo nas três profundidades estudadas.

Os gráficos *Boxplot* da Figura 15 apresentam um ponto discrepante com valor mínimo para a profundidade 0 a 10 cm, um ponto discrepante máximo para a profundidade 10 a 20 cm e dois pontos discrepantes superiores para a profundidade 20 a 30 cm. Pontos discrepantes unilaterais podem provocar uma assimetria na distribuição dos dados. Entretanto, os mesmos foram mantidos, pois não houve erro na coleta dos dados.

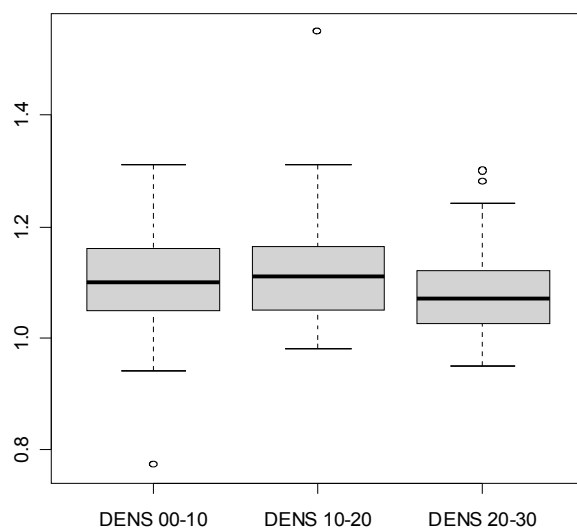


Figura 15 - Gráfico *Boxplot* para a densidade do solo nas três profundidades estudadas.

4.1.2 Umidade do Solo

Na Tabela 5 apresentam-se as estatísticas descritivas para a umidade do solo nas três profundidades estudadas. A umidade média encontrada foi de $0,369 \text{ m}^3 \text{ m}^{-3}$, com desvio padrão de 0,06 e coeficiente de variação de 15,56 %, para a profundidade 0 a 10 cm. Para a profundidade 10 a 20 cm, a umidade média foi de $0,391 \text{ m}^3 \text{ m}^{-3}$, com desvio padrão de 0,06 e coeficiente de variação de 16,17 %. Na profundidade 20 a 30 cm foi encontrada a umidade média de $0,406 \text{ m}^3 \text{ m}^{-3}$, desvio padrão de 0,07 e coeficiente de variação de 16,38 %. Esse atributo nas três profundidades estudadas possui média homogênea para a área em estudo (GOMES, 2000).

Pode-se observar, na Tabela 5, que os dados têm uma amplitude entre os valores mínimo e máximo de $0,36 \text{ m}^3 \text{ m}^{-3}$, para a profundidade 0 a 10 cm. Para a profundidade 10 a 20 cm, as estatísticas média e mediana são um pouco superiores às encontradas para a profundidade 0 a 10 cm, com uma amplitude

entre os valores mínimo e máximo $0,38 \text{ m}^3 \text{ m}^{-3}$. Para a profundidade 20 a 30 cm, as estatísticas média e mediana foram superiores às outras duas profundidades estudadas e sua amplitude foi de $0,43 \text{ m}^3 \text{ m}^{-3}$.

Tabela 5 - Análise exploratória e testes de normalidade da umidade do solo nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm

UMIDADE	N	MÉDIA	MÍNIMO	1º QUARTIL	MEDIANA	3º QUARTIL	MÁXIMO	DESVIO PADRÃO
0-10 cm	84	0,369	0,230	0,348	0,375	0,403	0,590	0,060
10-20 cm	88	0,391	0,220	0,360	0,400	0,423	0,600	0,060
20-30 cm	91	0,406	0,210	0,390	0,420	0,440	0,640	0,070

UMIDADE	CV (%)	TESTES NORMALIDADE (p-valor)				
		SW	AD	KS	As	K
0-10 cm	15,56	0,000	0,001	0,046	-0,031	2,213
10-20 cm	16,17	0,000	0,000	0,122 *	0,038	2,406
20-30 cm	16,38	0,000	0,000	<0,010	-0,670	2,204

Notas: Unidade do atributo: $[\text{m}^3 \text{ m}^{-3}]$; N: número de pontos amostrados; CV: coeficiente de variação; SW: Shapiro-Wilk; AD: Anderson-Darling; KS: Kolmogorov-Smirnov; As: assimetria; K: curtose; * não rejeita a normalidade ao nível de 5 % de probabilidade. As diferenças para os valores no número de pontos ocorreram devido a erro experimental.

Só o teste de normalidade de Kolmogorov-Smirnov, ao nível de 5 % de significância, indicou que os valores amostrados para a umidade, na profundidade 10 a 20 cm possuem distribuição normal de probabilidade. Nas outras duas profundidades não possuem normalidade. A Figura 16 mostra os histogramas para este atributo nas três profundidades estudadas.

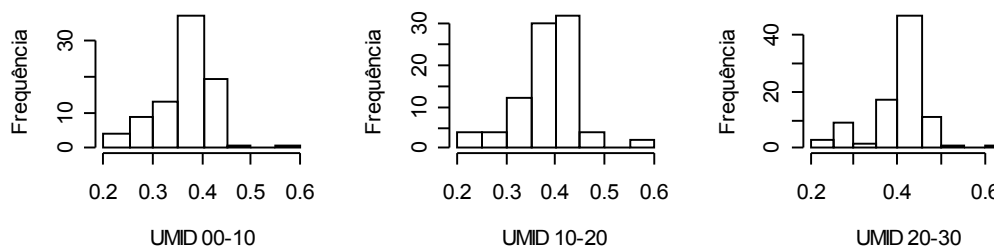


Figura 16 - Histograma para a umidade do solo nas três profundidades estudadas.

Para os dados da umidade do solo foram constatados pontos discrepantes, para as três profundidades estudadas, que são apresentados nos gráficos *Boxplot* da Figura 17. Um ponto discrepante com valor máximo e dois pontos discrepantes inferiores, para a profundidade 0 a 10 cm, um ponto discrepante máximo e três pontos discrepantes inferiores para a profundidade 10 a 20 cm, dois pontos discrepantes superiores e oito pontos discrepantes inferiores, para a profundidade 20 a 30 cm. Os pontos discrepantes foram considerados nesta análise e influenciaram nos testes de normalidade realizados.

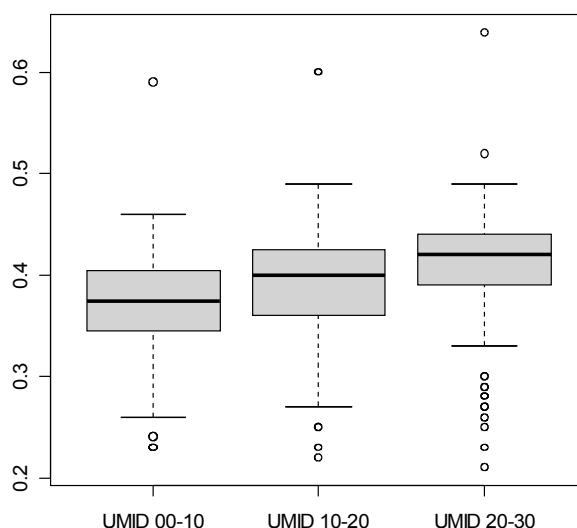


Figura 17 - Gráfico Boxplot para umidade do solo nas três profundidades estudadas.

4.1.3 Resistência do Solo à Penetração

Na Tabela 6 são apresentados os resultados das estatísticas descritivas da resistência do solo à penetração – RSP, em que se observam, na camada 0 a 10 cm, estatísticas de posição superiores às outras profundidades, notadamente a

média e a mediana. Nas camadas 0 a 10 cm e 10 a 20 cm, mais de 75 % dos dados foram superiores a 2,0 Mpa. Segundo SECCO (2003), quando 25 % dos dados encontram-se acima de 2,0 Mpa, podem ser considerados como uma resistência limitante ao crescimento radicular.

Assim, observa-se que houve a ocorrência de valores de resistência limitante ao crescimento radicular nas camadas 0 a 10 cm e 10 a 20 cm. De uma forma geral, os valores de resistência do solo à penetração foram altos.

Tabela 6 - Análise exploratória e testes de normalidade da resistência do solo à penetração – RSP, nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm

RSP	N	MÉDIA	MÍNIMO	1° QUARTIL	MEDIANA	3° QUARTIL	MÁXIMO	DESVIO PADRÃO
0-10 cm	93	2,729	0,671	2,374	2,692	3,125	4,269	0,670
10-20 cm	90	2,427	0,875	2,131	2,330	2,771	3,709	0,510
20-30 cm	88	2,101	0,770	1,767	2,031	2,291	4,958	0,530

RSP	CV (%)	TESTES NORMALIDADE (p-valor)				
		SW	AD	KS	As	K
0-10 cm	24,50	0,335 *	0,354 *	>0,150 *	-0,170	0,692
10-20 cm	21,04	0,199 *	0,05 *	0,028	0,111	0,444
20-30 cm	25,38	0	0	<0,010	2,040	9,310

Notas: RSP: resistência do solo à penetração; unidade do atributo: [Mpa]; N: número de pontos amostrados; CV: coeficiente de variação; SW: Shapiro-Wilk; AD: Anderson-Darling; KS: Kolmogorov-Smirnov; As: assimetria; K: curtose; * não rejeita a normalidade ao nível de 5% de probabilidade. As diferenças para os valores no número de pontos ocorreram devido a erro experimental.

Os valores de resistência do solo à penetração, nas três profundidades estudadas, possuem alta dispersão e baixa homogeneidade dos dados, para a área em estudo, conforme demonstram os valores encontrados no desvio padrão e no coeficiente de variação, o que é classificado como alta heterogeneidade, quando o CV % se encontra entre 20 e 30 % (GOMES, 2000).

Observa-se, na Tabela 6, que os valores da resistência do solo à penetração, na camada 0 a 10 cm, encontram-se entre 0,671 e 4,269 MPa. Os valores mais elevados ocorrem nas parcelas 10, 92, 96 e 97. Os valores da resistência do solo à penetração, na camada 10 a 20 cm, encontram-se entre 0,875

e 3,709 Mpa. Nas parcelas 13, 34 e 25 ocorrem os valores mais elevados. Observa-se ainda, na Tabela 6, que os valores para a profundidade 20 a 30 cm se encontram entre 0,770 e 4,958 Mpa. Os valores mais elevados ocorrem nas parcelas 48, 11, 7 e 64.

No histograma apresentado na Figura 18 observa-se a forma da distribuição de freqüências da resistência do solo à penetração.

Verificou-se, por meio dos testes de Shapiro-Wilk, Anderson-Darling e Kolmogorov-Smirnov, ao nível de 5 % de significância, que os dados da resistência do solo à penetração, para as profundidades 0 a 10 cm e 10 a 20 cm, possuem distribuição normal de probabilidade e na camada 20 a 30 cm os dados não possuem distribuição normal de probabilidade.

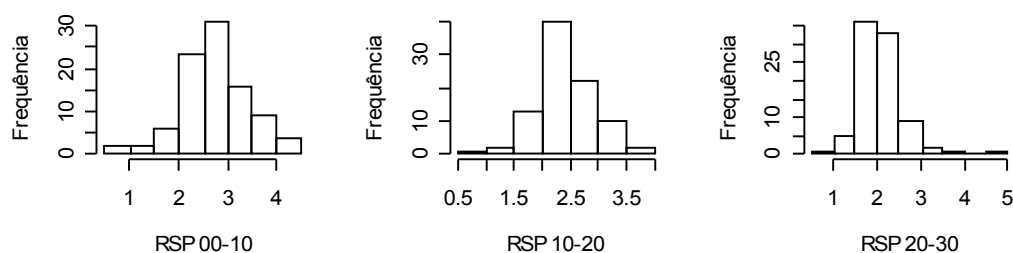


Figura 18 - Histograma para a resistência do solo à penetração, nas três profundidades estudadas.

No gráfico *Boxplot*, apresentado na Figura 19, observa-se, entre os valores da resistência do solo à penetração na área estudada, relativos à camada 0 a 10 cm, a existência de três pontos discrepantes. Os valores discrepantes foram mantidos, pois não se verificaram erros de leitura nos registros feitos pelo penetrômetro. Na mesma Figura 19, o gráfico *Boxplot*, relativo aos valores na camada 10 a 20 cm, mostra que existe um ponto discrepante de 0,875 MPa, com característica que não justifica sua eliminação. Observa-se, na camada 20 a 30 cm, pontos discrepantes apresentados pelo gráfico *Boxplot*, que também serão

mantidos, por não se terem verificado erros de leitura nos registros feitos pelo penetrômetro.

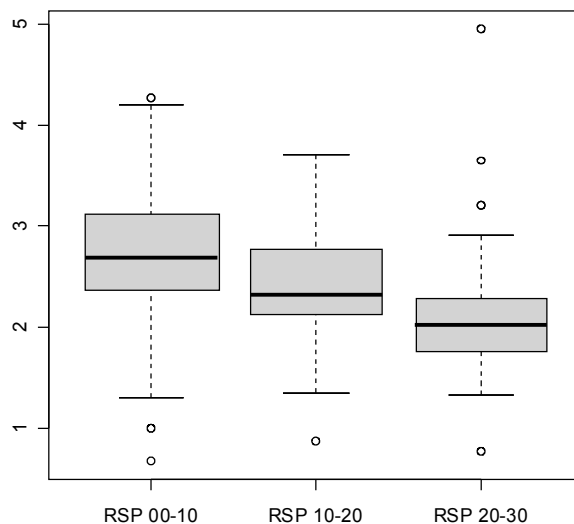


Figura 19 - Gráfico Boxplot para a variável resistência do solo à penetração, nas três profundidades estudadas.

4.1.4 Produtividade da Soja

Na Tabela 7 é apresentada a análise exploratória para a produtividade da soja, na qual foi obtida a média de produtividade de 3,22 Mg ha⁻¹, com desvio padrão de 0,38. O valor mínimo encontrado foi de 2,09 Mg ha⁻¹ e o máximo de 4,09 Mg ha⁻¹. Observa-se, portanto, que, em média, a produtividade para a área em estudo foi superior às médias do Estado do Paraná e do país, nesse ano agrícola.

O coeficiente de variação (CV) encontrado para a produtividade da soja foi de 11,71 %, existindo, portanto, média homogeneidade (GOMES, 2000)

A forma da distribuição de freqüências, observada na Figura 20, sugere uma distribuição normal. Os testes de normalidade de Shapiro-Wilk, Anderson-Darling e Kolmogorov-Smirnov revelaram que os valores amostrados para a produtividade possuem distribuição normal de probabilidade, ao nível de 5 % de significância.

O *Boxplot* para essa variável regionalizada revelou a presença de um ponto discrepante com valor de 2,09 Mg ha⁻¹ (Figura 21).

Tabela 7 - Análise exploratória e teste de normalidade da produtividade da soja

VARIÁVEL	N	MÉDIA	MÍNIMO	1° QUARTIL	MEDIANA	3° QUARTIL	MÁXIMO	DESVIO PADRÃO
PRODUTIVIDADE E	66	3,217	2,09	2,96	3,19	3,48	4,09	0,38

VARIÁVEL	CV (%)	TESTES NORMALIDADE (p-valor)				
		SW	AD	KS	As	K
PRODUTIVIDADE E	11,71	0,420	0,519	>0,150 *	-0,180	0,696

Notas: Unidade do atributo: [Mg ha⁻¹]; N: número de pontos amostrados; CV: coeficiente de variação; SW: Shapiro-Wilk; AD: Anderson-Darling; KS: Kolmogorov-Smirnov; As: assimetria; K: curtose; * não rejeita a normalidade ao nível de 5 % de probabilidade. As diferenças para os valores no número de pontos ocorreram devido a erro experimental.

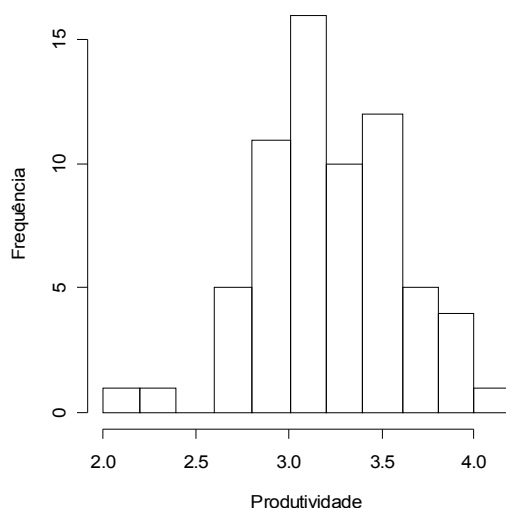


Figura 20 - Histograma para a produtividade da soja no ano agrícola 2004/2005.

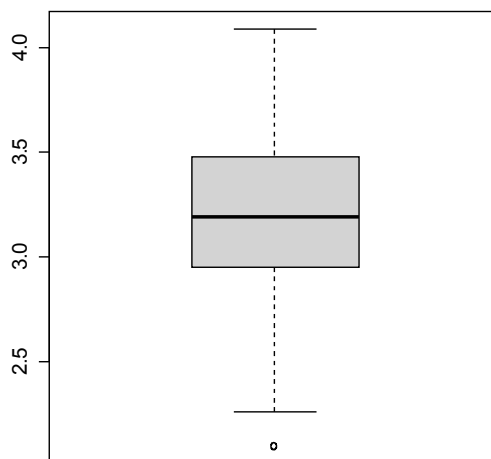


Figura 21 - Gráfico *Boxplot* para a produtividade da soja no ano agrícola 2004/2005.

ANÁLISE GEOESTATÍSTICA

4.1.5 Análise da Estacionaridade

Com o uso do gráfico *post-plot*, que é um diagrama em que se mostra a posição dos pontos associada às suas coordenadas, pode-se avaliar a tendência direcional dos dados. Nele, cada parcela do experimento está classificada segundo os valores correspondentes aos seus quartis (Seção 3.5).

As Figuras 22 a 26 mostram os *post-plot* para os atributos em estudo. Para a densidade do solo, observa-se, nas Figuras 22 (a), 22 (b) e 23 (a), a inexistência de tendências direcionais, indicando que $E(Z(s)) = m$, sendo m um valor constante. Para a umidade do solo, observa-se, nas Figuras 23 (b), 24 (a) e

24 (b), também inexistência de tendências direcionais. Para a resistência do solo à penetração, conforme Figuras 25 (a), 25 (b) e 26 (a), também não foi detectada a existência de tendência direcional. A produtividade da soja apresentou indício de tendência direcional, pois se observa na Figura 26 (b) uma concentração maior para os dados pertencentes ao intervalo P-4, definido na Seção 3.5, em determinada região, que poderia ser justificada como sendo uma região na qual o solo é mais fértil.

Outra forma de verificação da tendência direcional para os atributos em estudo é a utilização dos gráficos de dispersão para os dados *versus* linha e dados *versus* coluna. Esses gráficos podem ser observados no Apêndice A, nas Figuras 1A a 11A. Para todos os atributos em estudo não houve indícios da existência de tendência direcional.

Então, segundo esses dois tipos de gráficos, pode-se concluir que as variáveis regionalizadas em estudo não possuem tendência direcional e, portanto, são estacionárias.

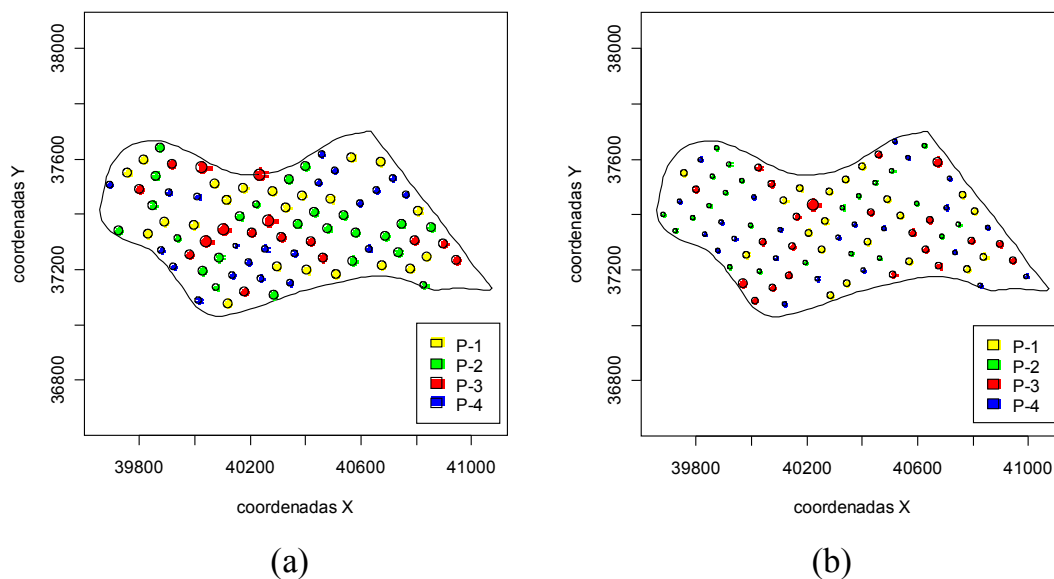


Figura 22 - Gráficos *post-plot* para a densidade do solo (Mg m^{-3}), nas profundidades 0 a 10 cm (a) e 10 a 20 cm (b).

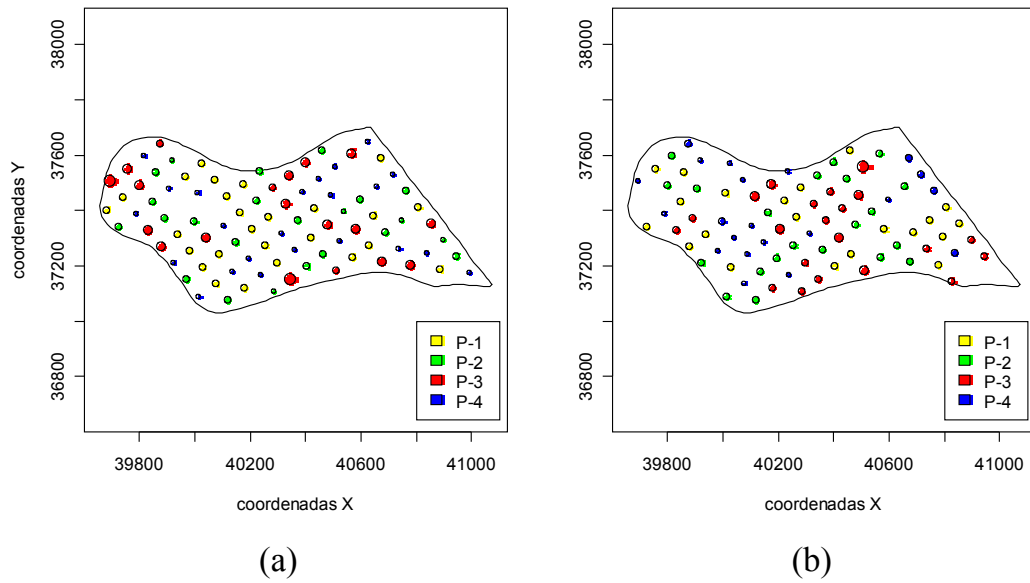


Figura 23 - Gráficos *post-plot* para a densidade do solo (Mg m^{-3}), na profundidade 20 a 30 cm (a), e para a umidade do solo ($\text{m}^3 \text{m}^{-3}$), na profundidade 0 a 10 cm (b).

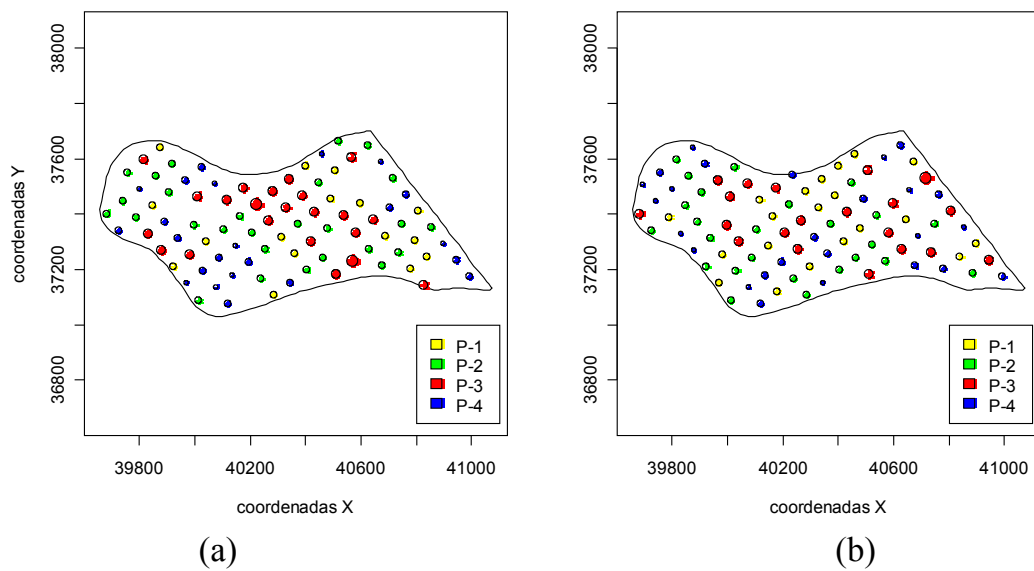


Figura 24 - Gráficos *post-plot* para a para a umidade do solo ($\text{m}^3 \text{m}^{-3}$), nas profundidades 10 a 20 cm (a) e 20 a 30 cm (b).

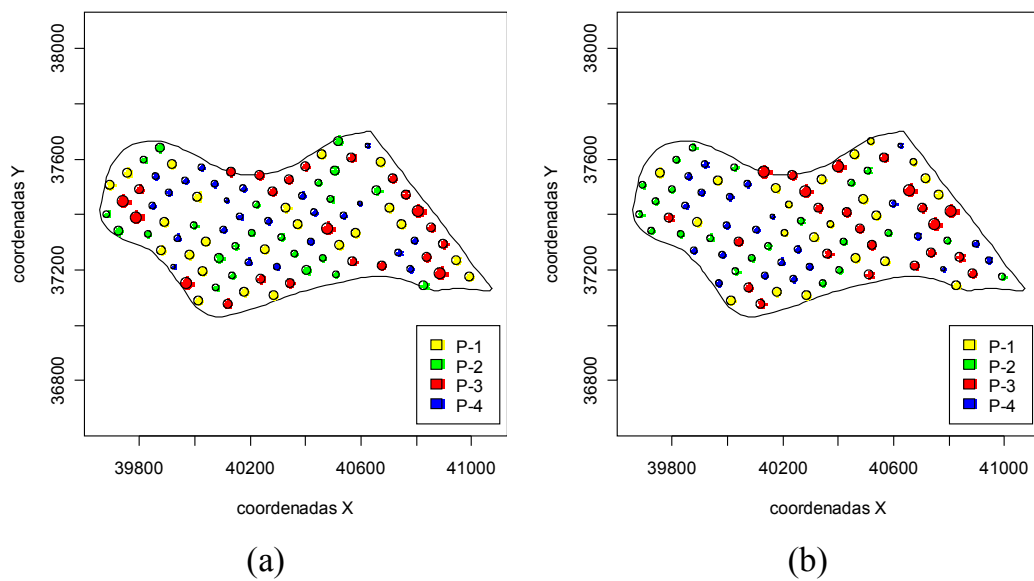


Figura 25 - Gráficos *post-plot* para a resistência do solo à penetração [MPa], nas profundidades 0 a 10 cm (a) e 10 a 20 cm (b).

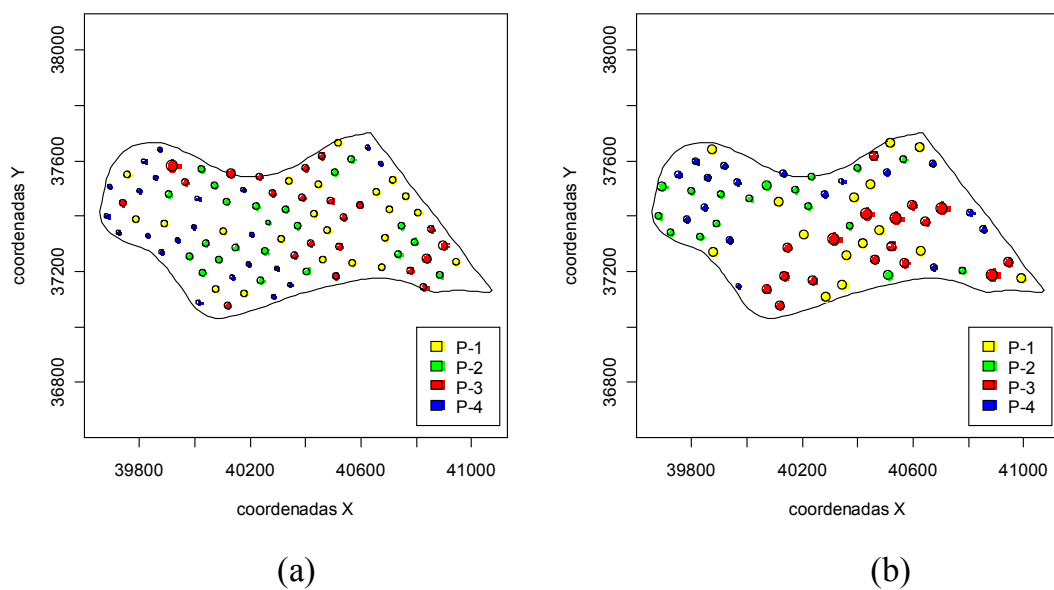


Figura 26 - Gráficos *post-plot* para a resistência do solo à penetração [MPa], na profundidade 20 a 30 cm (a) e para a produtividade da soja (Mg há^{-1}) (b).

4.1.6 Semivariogramas Experimentais

Os semivariogramas experimentais para as variáveis regionalizadas em estudo foram calculados pelo estimador de Matheron, para os casos nos quais os dados apresentaram distribuição normal de probabilidade. Para as variáveis regionalizadas que não apresentaram distribuição normal de probabilidade foi utilizado o estimador de Cressie e Hawkins, como a umidade nas profundidades 0 a 10 cm e 20 a 30 cm e a resistência do solo à penetração na profundidade 20 a 30 cm.

Densidade

A Figura 27 ilustra os semivariogramas experimentais do atributo densidade na profundidade 0 a 10 cm, 10 a 20 cm e 20 a 30 cm, que foram calculados pelo estimador de Matheron, pois os dados apresentaram normalidade.

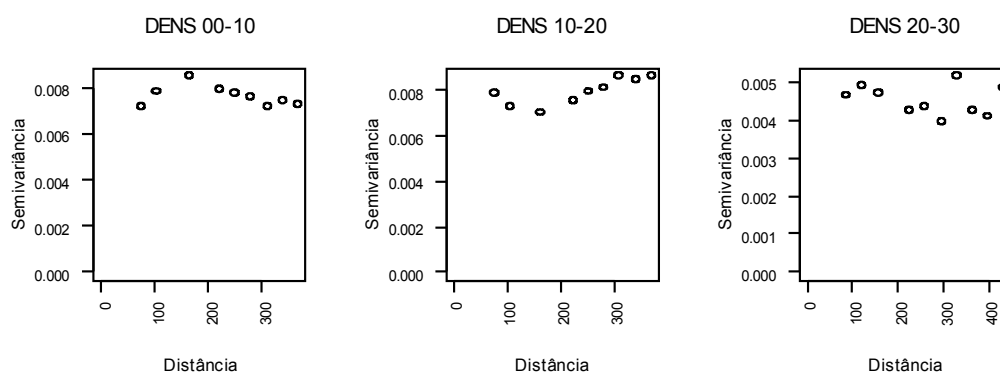


Figura 27 - Semivariogramas experimentais para a variável densidade nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.

Os modelos teóricos foram ajustados pelos métodos de estimação mínimos quadrados ordinários - OLS e mínimos quadrados ponderados - WLS1.

A Tabela 8 apresenta os parâmetros efeito pepita (C_0), patamar ($C_0 + C_1$) e alcance (a), para os modelos ajustados. Nas Figuras 28, 29 e 30 mostram-se os semivariogramas ajustados obtidos, sobrepostos ao semivariograma experimental.

Tabela 8 - Parâmetros dos modelos semivariográficos, pelos métodos OLS e WLS1, para os dados de densidade, nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm

Variáveis estudadas	Modelos	(C_0)		$(C_0 + C_1)$		(a)	
		OLS	WLS1	OLS	WLS1	OLS	WLS1
Densidade 00-10	Exponencial	0,0050	0,0050	0,0077	0,0077	95	141
	Esférico	0,0060	0,0055	0,0077	0,0077	126	140
	Gaussiano	0,0035	0,0050	0,0077	0,0077	87	97
Densidade 10-20	Exponencial	0,0050	0,0050	0,0083	0,0080	239	120
	Esférico	0,0060	0,0050	0,0080	0,0080	150	120
	Gaussiano	0,0050	0,0050	0,0079	0,0080	73	78
Densidade 20-30	Exponencial	0,0035	0,0030	0,0045	0,0045	11	141
	Esférico	0,0035	0,0030	0,0045	0,0045	78	140
	Gaussiano	0,0030	0,0025	0,0045	0,0045	30	139

Notas: OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados.

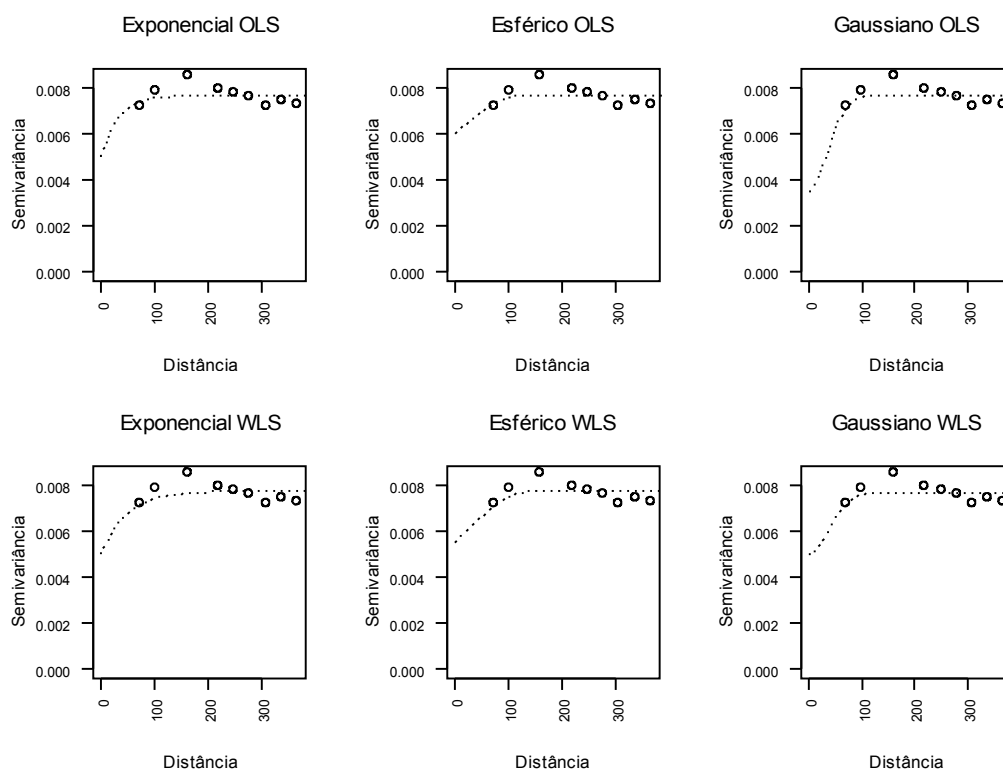


Figura 28 - Semivariogramas para os dados de densidade, na profundidade 0 a 10 cm.

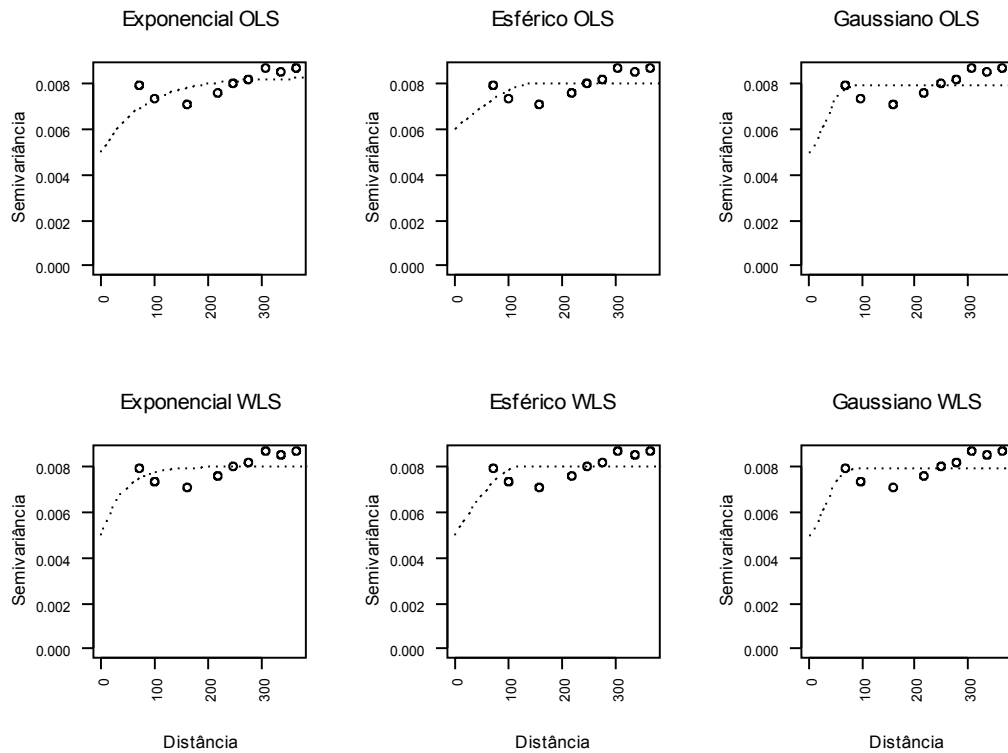


Figura 29 - Semivariogramas para os dados de densidade, na profundidade 10 a 20 cm.

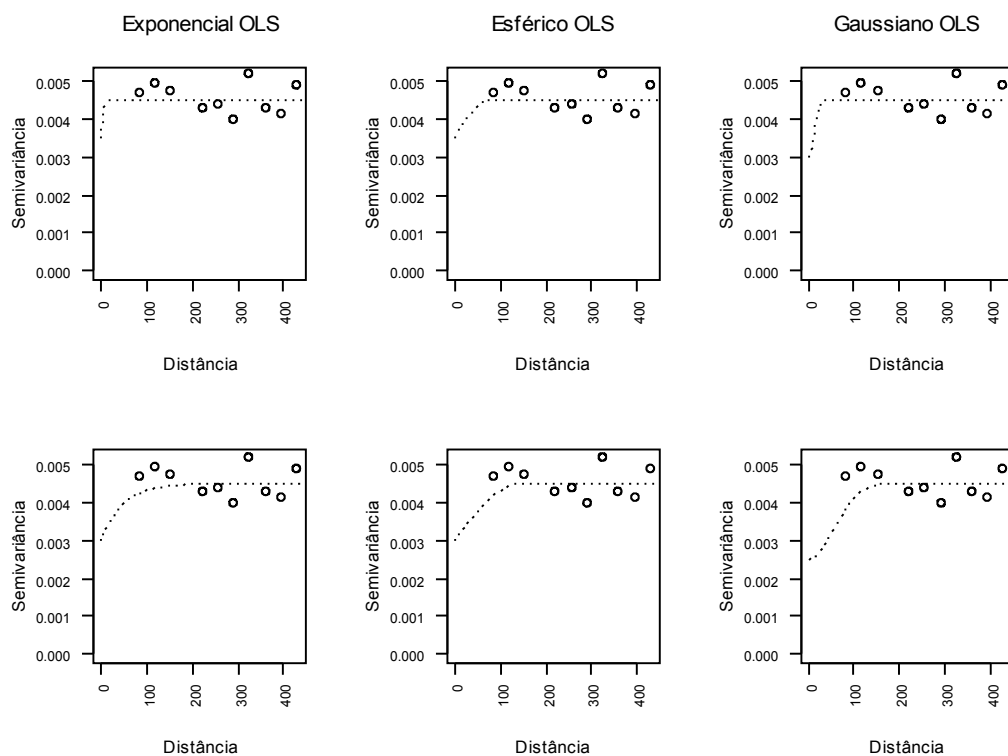


Figura 30 - Semivariogramas para os dados de densidade, na profundidade 20 a 30 cm.

Para aplicação das técnicas de validação de ajustes de modelos geoestatísticos foi necessário analisar os resultados dos diversos métodos e proceder-se à comparação de valores. Os resultados encontram-se na Tabela 9.

Tabela 9 - Soma dos quadrados dos resíduos, AIC (\hat{A}), Filliben (r^*) e validação cruzada para a densidade nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm

Variáveis estudadas	Modelos	N	Soma quadrados residuais		\hat{A}		r^*	
			OLS	WLS1	OLS	WLS1	OLS	WLS1
Dens 00-10	Exponencial	83	1,45E-06	4,15E-04	-1109,8	-640,34	0,9779950	0,9760846
	Esférico	83	1,36E-06	3,64E-04	-1115,2	-651,22	0,9773484	0,9754163
	Gaussiano	83	1,33E-06	3,57E-04	-1117,0	-652,84	0,9766188	0,9766824
Dens 10-20	Exponencial	88	2,30E-06	6,36E-04	-1136,5	-641,71	0,9439069	0,9409327
	Esférico	88	2,78E-06	7,03E-04	-1119,8	-632,89	0,9420241	0,9412659
	Gaussiano	88	2,81E-06	6,84E-04	-1118,8	-635,31	0,9399533	0,9399444
Dens 20-30	Exponencial	91	1,44E-06	3,51E-04	-1217,8	-717,88	0,9817626	0,9851343
	Esférico	91	1,44E-06	3,42E-04	-1217,8	-720,24	0,9817958	0,9856011
	Gaussiano	91	1,44E-06	4,27E-04	-1217,8	-700,04	0,9817626	0,9890379

VALIDAÇÃO CRUZADA		ANÁLISE DOS RESÍDUOS				
MODELOS		MIN	MEDIANA	MÉDIA	MAX	D PADRÃO
Dens 00-10	Exponencial OLS	-3,40E-01	-2,79E-03	-3,92E-06	2,08E-01	8,45E-02
	Exponencial WLS1	-3,45E-01	-5,17E-03	-1,07E-05	2,07E-01	8,48E-02
	Esférico OLS	-3,42E-01	-3,17E-03	-5,40E-06	2,00E-01	8,44E-02
	Esférico WLS1	-3,45E-01	-6,45E-03	1,14E-05	1,94E-01	8,45E-02
	Gaussiano OLS	-3,44E-01	-3,57E-03	-7,06E-06	1,94E-01	8,45E-02
	Gaussiano WLS1	-3,44E-01	-3,72E-03	-7,15E-06	1,96E-01	8,45E-02
Dens 10-20	Exponencial OLS	-1,36E-01	-1,33E-02	-1,13E-04	4,21E-01	8,86E-02
	Exponencial WLS1	-1,23E-01	-7,84E-03	-3,44E-05	4,30E-01	8,71E-02
	Esférico OLS	-1,28E-01	-1,26E-02	-7,09E-05	4,26E-01	8,75E-02
	Esférico WLS1	-1,26E-01	-1,07E-02	-7,10E-05	4,27E-01	8,72E-02
	Gaussiano OLS	-1,37E-01	-7,61E-03	-6,17E-06	4,35E-01	8,69E-02
	Gaussiano WLS1	-1,37E-01	-7,61E-03	-6,07E-05	4,35E-01	8,69E-02
Dens 20-30	Exponencial OLS	-1,29E-01	-7,56E-03	-1,82E-14	2,25E-01	7,00E-02
	Exponencial WLS1	-1,29E-01	-2,19E-03	1,76E-05	2,31E-01	7,10E-02
	Esférico OLS	-1,29E-01	-7,49E-03	1,31E-08	2,25E-01	7,00E-02
	Esférico WLS1	-1,25E-01	-3,47E-03	1,66E-05	2,30E-01	7,10E-02
	Gaussiano OLS	-1,29E-01	-7,56E-03	1,37E-13	2,25E-01	7,00E-02
	Gaussiano WLS1	-1,32E-01	1,75E-03	-5,87E-05	2,34E-01	7,26E-02

Notas: Em negrito a melhor escolha; N: número de pontos amostrados; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados; \hat{A} : critério de Akaike; r^* , teste de Filliben, compara-se com $r = 0,9736$ (Dens 0-10), $0,9752$ (Dens 10-20) e $0,9762$ (Dens 20-30), (Anexo).

O Critério de Informação de Akaike determinou que o modelo que apresentasse o menor valor calculado para \hat{A} seria o escolhido. Conforme a Tabela 9, o menor valor de \hat{A} , para a densidade na profundidade 0 a 10 cm, é - 1117,0, apontando para o modelo gaussiano estimado pelo método OLS. Para a densidade do solo na profundidade 10 a 20 cm, o menor valor de \hat{A} é -1136,5, que indica o modelo exponencial estimado pelo método OLS e para a densidade do solo na profundidade 20 a 30 cm, o menor valor calculado de \hat{A} é -1217,8, que aponta para os três modelos estimados pelo método OLS, exponencial, esférico e gaussiano.

O critério de Filliben testou se os resíduos ortonormais seguiam uma distribuição normal a um determinado nível de significância. Para o caso dos dados da densidade nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm, que têm 83, 88 e 91 valores amostrados, respectivamente, apresentados na Tabela 9, e para um nível de significância de 5 %, os valores tabulares de r , obtidos da tabela apresentada no Anexo, foram de 0,9736, 0,9752 e 0,9762, respectivamente. Comparando-se com os valores r^* , calculados e apresentados na Tabela 9, nas colunas r^* , OLS e WLS1, verificou-se que os valores r^* , que são maiores do que r , indicaram que a hipótese de que os dados seguem uma distribuição normal não foi rejeitada para um nível de significância de 5 %. Por esse critério, os seguintes modelos de semivariogramas teóricos puderam ser aceitos: para a densidade do solo na profundidade 0 a 10 cm todos os modelos puderam ser aceitos, para a densidade do solo na profundidade 10 a 20 cm nenhum modelo foi aceito e para a densidade do solo na profundidade 20 a 30 cm todos os modelos foram aceitos.

O critério de validação cruzada, aplicado aos modelos em estudo, sugeriu que o menor valor da média dos erros de estimação indica o modelo com melhor ajuste. Para os dados de densidade na profundidade 0 a 10 cm, em estudo, apresentados na Tabela 9, a escolha apontou para o modelo exponencial com parâmetros estimados pelo método OLS. Para os dados de densidade, na profundidade 10 a 20 cm, a escolha indicou o modelo esférico com parâmetros estimados pelo método OLS. Para os dados de densidade, na profundidade

20 a 30 cm, a escolha apontou para o modelo exponencial com parâmetros estimados pelo método OLS.

Além dos critérios apresentados, para fins de comparação, optou-se por calcular a soma dos mínimos quadrados relativos aos erros de estimação buscando-se aquele que apresenta o menor valor. Verifica-se na Tabela 9 que, para a densidade do solo na profundidade 0 a 10 cm, o menor valor apontou para o modelo gaussiano, com parâmetros estimados pelo método OLS; para a densidade do solo na profundidade 10 a 20 cm, o menor valor indicou o modelo exponencial estimado pelo método OLS; e para a densidade do solo na profundidade 20 a 30 cm, apontou para os três modelos estimados pelo método OLS, exponencial, esférico e gaussiano. Apresenta-se na Tabela 10, um resumo com os melhores modelos escolhidos, usando-se cada um dos métodos de ajuste.

Tabela 10 – Modelos escolhidos com melhor ajuste

ATRIBUTO	AIC	FILLIBEN	VALIDAÇÃO CRUZADA
Densidade 00-10	GAU OLS	TODOS	EXP OLS
Densidade 10-20	EXP OLS	NENHUM	GAU OLS
Densidade 20-30	EXP OLS ESF OLS GAU OLS	TODOS	EXP OLS

Notas: AIC: critério de Akaike; ESF: modelo esférico; EXP: modelo exponencial; GAU: modelo gaussiano; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados.

Para obter-se uma idéia visual do comportamento do atributo na área estudada, utilizou-se uma representação gráfica onde as coordenadas dos pontos e os valores da variável são representados em um sistema. Foi utilizado, para confecção dos mapas temáticos, o melhor modelo indicado pela validação cruzada.

A Figura 31 mostra o mapa temático referente ao atributo densidade do solo, na profundidade 0 a 10 cm, ajustado por um modelo de semivariograma teórico exponencial, com parâmetros estimados pelo método OLS.

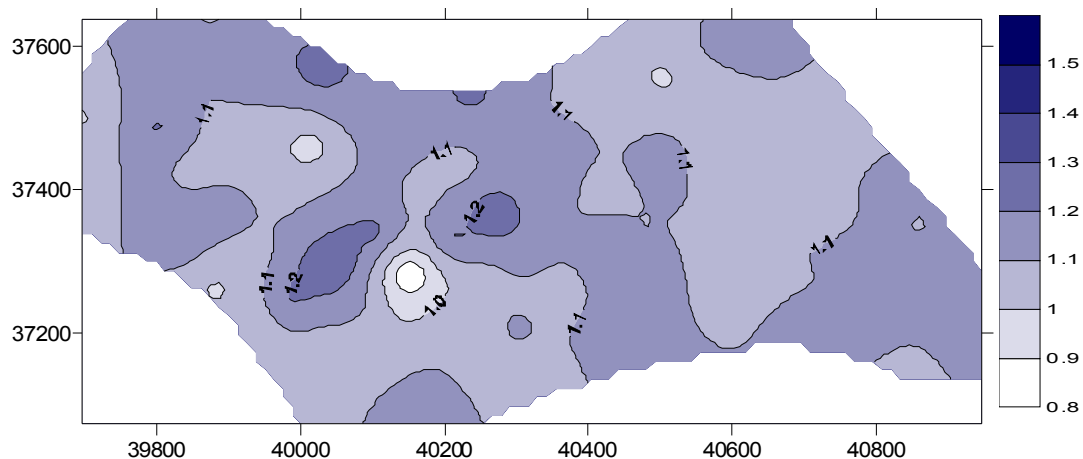


Figura 31 - Mapa para o atributo densidade na profundidade 0 a 10 cm; unidade do atributo: $[\text{Mg m}^{-3}]$.

A Figura 32 refere-se ao mapa temático do atributo densidade do solo na profundidade 10 a 20 cm, ajustado por um modelo de semivariograma teórico gaussiano, com parâmetros estimados pelo método OLS.

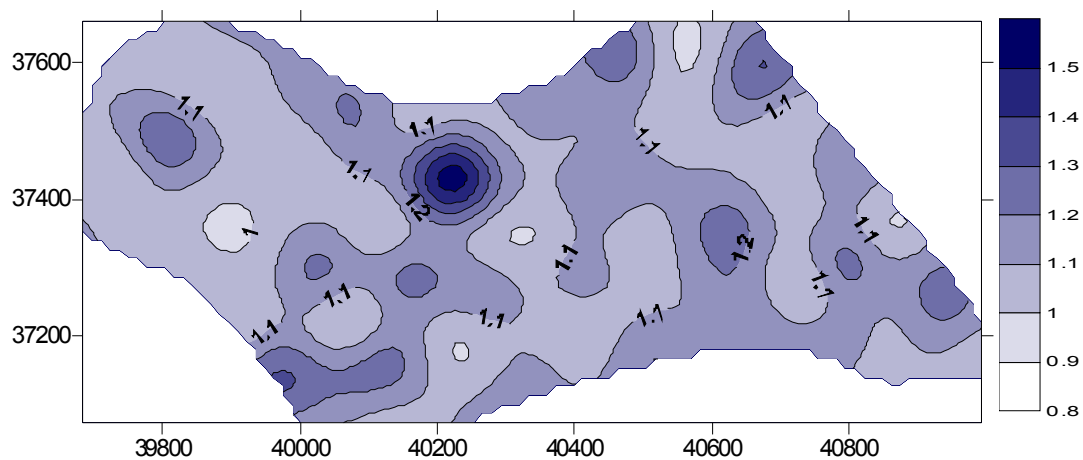


Figura 32 - Mapa para o atributo densidade na profundidade 10 a 20 cm; unidade do atributo: $[\text{Mg m}^{-3}]$.

A Figura 33 apresenta o mapa referente ao atributo densidade do solo na profundidade 20 a 30 cm, ajustado por um modelo de semivariograma teórico exponencial, com parâmetros estimados pelo método OLS.

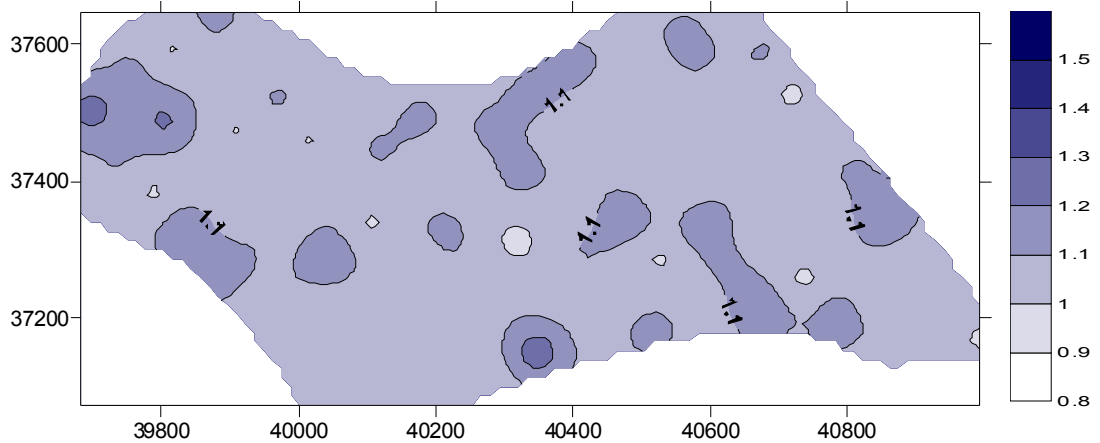


Figura 33 - Mapa para o atributo densidade na profundidade 20 a 30 cm; unidade do atributo: $[\text{Mg m}^{-3}]$.

Umidade

A Figura 34 ilustra os semivariogramas experimentais do atributo umidade nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm. O primeiro e o último foram calculados pelo estimador de Cressie e Hawkins, pois os dados não apresentaram normalidade. O segundo foi calculado pelo estimador de Matheron, pois os dados a ele referentes apresentaram normalidade.

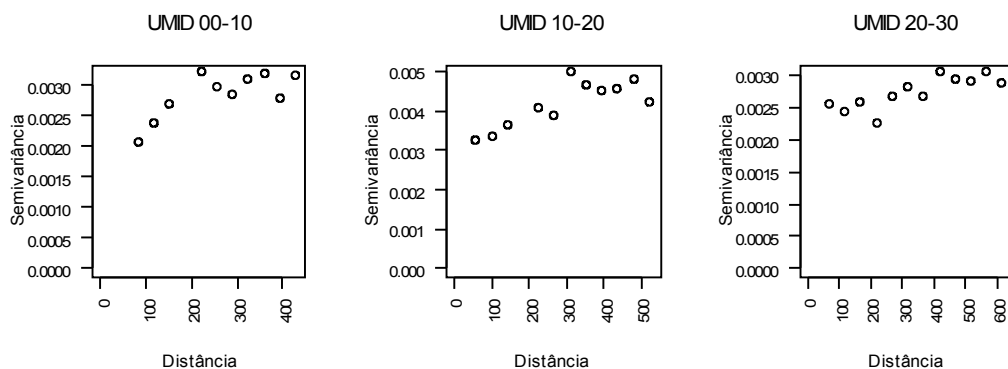


Figura 34 - Semivariogramas experimentais para a variável umidade nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.

Objetivando a aplicação dos critérios de validação ajustaram-se três modelos teóricos ao semivariograma experimental: exponencial, esférico e gaussiano, considerados adequados aos dados em análise. Para a estimação dos parâmetros, utilizaram-se dois métodos: o dos mínimos quadrados ordinários (OLS) e o dos mínimos quadrados ponderados (WLS1).

A Tabela 11 apresenta os parâmetros efeito pepita (C_0), patamar ($C_0 + C_1$) e o alcance (a) para os modelos ajustados.

Tabela 11 - Parâmetros dos modelos semivariográficos pelos métodos OLS e WLS1 para os dados de umidade, nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm

Variáveis estudadas	Modelos	(C_0)		$(C_0 + C_1)$		(a)	
		OLS	WLS1	OLS	WLS1	OLS	WLS1
Umidade 00-10	Exponencial	0,0000	0,0000	0,0031	0,0031	233	240
	Esférico	0,0009	0,0000	0,0030	0,0031	234	250
	Gaussiano	0,0013	0,0004	0,0030	0,0032	208	242
Umidade 10-20	Exponencial	0,0023	0,0000	0,0047	0,0048	512	360
	Esférico	0,0026	0,0000	0,0045	0,0048	406	350
	Gaussiano	0,0030	0,0029	0,0046	0,0046	372	346
Umidade 20-30	Exponencial	0,0010	0,0010	0,0028	0,0029	204	390
	Esférico	0,0015	0,0015	0,0028	0,0029	314	400
	Gaussiano	0,0024	0,0023	0,0031	0,0028	674	398

Notas: OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados.

Nas Figuras 35, 36 e 37 mostram-se os semivariogramas obtidos.

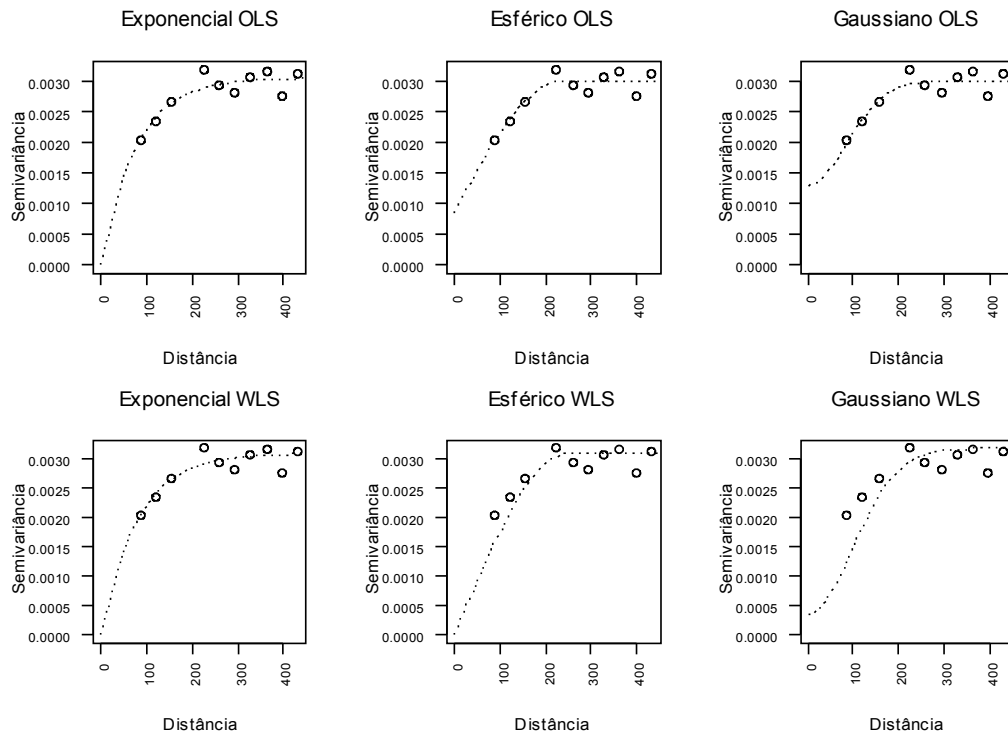


Figura 35 - Semivariogramas para os dados de umidade na profundidade 0 a 10 cm.

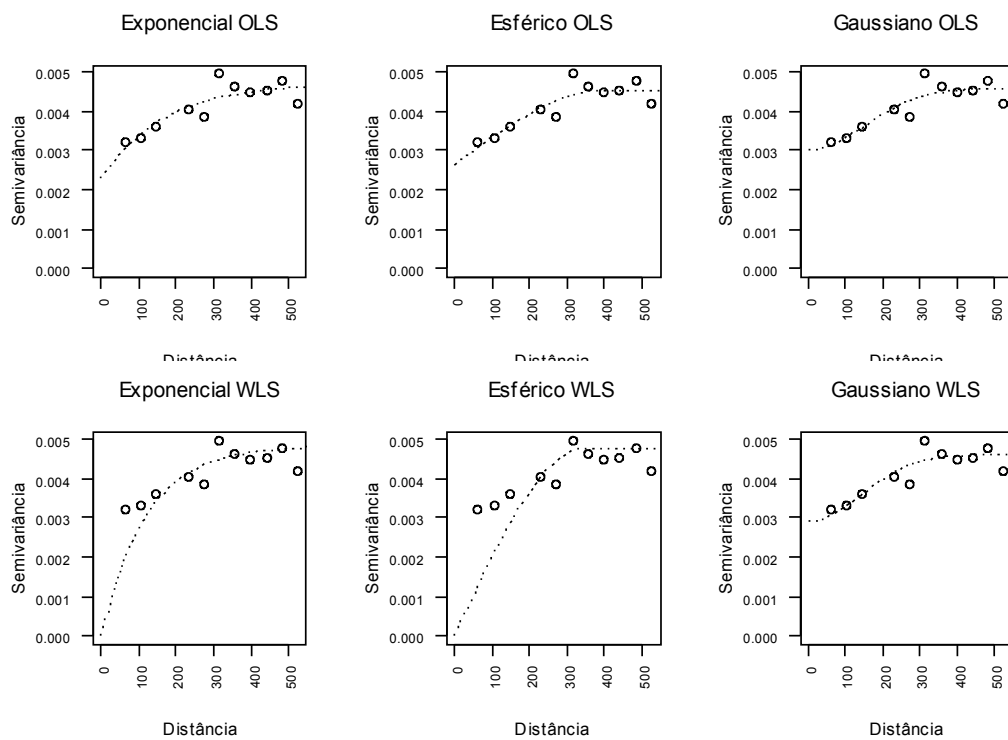


Figura 36 - Semivariogramas para os dados de umidade na profundidade 10 a 20 cm.

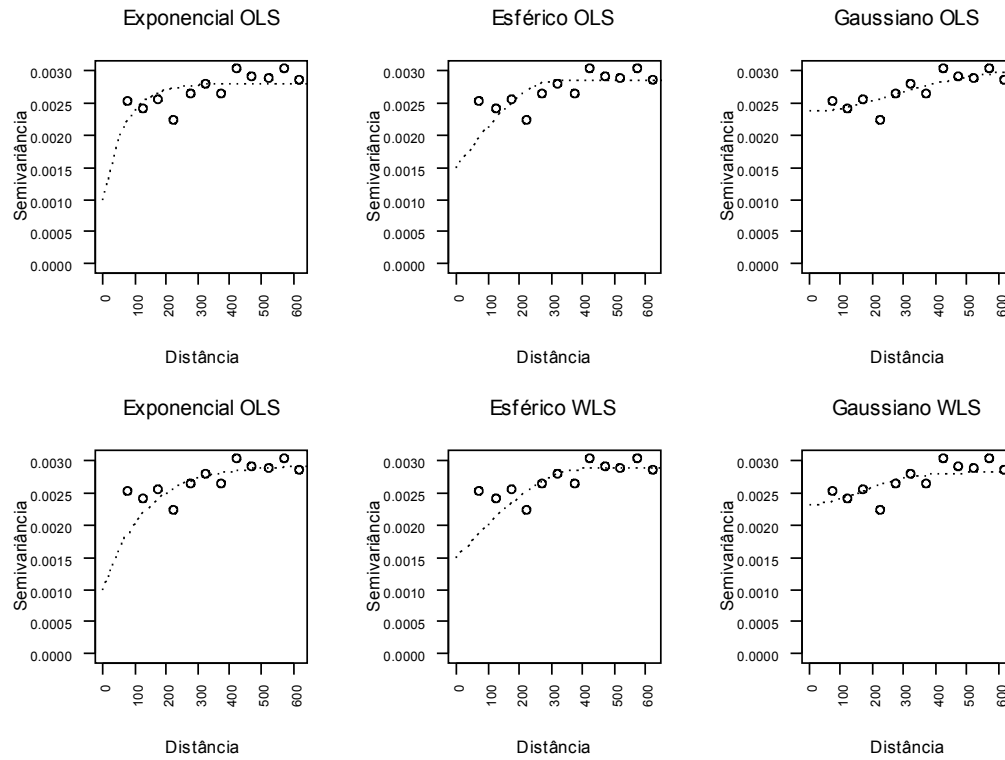


Figura 37 - Semivariogramas para os dados de umidade na profundidade 20 a 30 cm.

Para utilizar as técnicas de validação de ajustes de modelos geoestatísticos procedeu-se à análise dos resultados dos diversos métodos e compararam-se os valores. Os resultados encontram-se na Tabela 12.

Tabela 12 - Soma dos quadrados dos resíduos, AIC (\hat{A}), Filliben (r^*) e validação cruzada para a umidade nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm

Variáveis estudadas	Modelos	N	Soma quadrados residuais		\hat{A}		r^*	
			OLS	WLS1	OLS	WLS1	OLS	WLS1
Umid 00-10	Exponencial	84	2,47E-07	6,03E-05	-1272,0	-810,2	0,9779910	0,9782075
	Esférico	84	1,95E-07	5,24E-05	-1291,7	-821,9	0,9758822	0,9770408
	Gaussiano	84	2,14E-07	1,61E-04	-1284,1	-727,8	0,9777746	0,9859640
Umid 10-20	Exponencial	88	1,51E-06	3,70E-04	-1173,4	-689,5	0,9620641	0,9715090
	Esférico	88	1,88E-06	5,06E-04	-1154,0	-661,9	0,9618506	0,9752866
	Gaussiano	88	1,85E-06	6,28E-04	-1155,5	-642,8	0,9622150	0,9622963
Umid 20-30	Exponencial	91	5,51E-07	1,49E-04	-1305,4	-795,6	0,9485327	0,9552083
	Esférico	91	6,94E-07	1,47E-04	-1284,4	-796,9	0,9539306	0,9514900
	Gaussiano	91	2,21E-07	9,66E-05	-1388,6	-835,2	0,9416540	0,9404600

VALIDAÇÃO CRUZADA : ANÁLISE DOS RESÍDUOS						
MODELOS	MIN	MEDIANA	MÉDIA	MAX	D PADRÃO	
Umid 00-10	Exponencial OLS	-1,51E-01	5,85E-03	-8,77E-05	2,23E-01	5,93E-02
	Exponencial WLS1	-1,51E-01	5,73E-03	-8,48E-05	2,23E-01	5,94E-02
	Esférico OLS	-1,58E-01	7,44E-03	-7,94E-05	2,26E-01	5,93E-02
	Esférico WLS1	-1,55E-01	7,35E-03	3,75E-05	2,37E-01	6,25E-02
	Gaussiano OLS	-1,59E-01	5,97E-03	-6,20E-05	2,23E-01	6,00E-02
	Gaussiano WLS1	-1,65E-01	9,86E-03	6,15E-04	2,28E-01	6,88E-02
Umid 10-20	Exponencial OLS	-1,59E-01	9,41E-03	9,61E-06	1,97E-01	6,05E-02
	Exponencial WLS1	-1,59E-01	9,82E-03	4,10E-04	1,98E-01	6,23E-02
	Esférico OLS	-1,59E-01	8,72E-03	3,57E-06	1,97E-01	6,10E-02
	Esférico WLS1	-1,53E-01	8,64E-03	6,49E-04	2,08E-01	6,16E-02
	Gaussiano OLS	-1,60E-01	9,14E-03	-2,67E-05	1,97E-01	6,10E-02
	Gaussiano WLS1	-1,59E-01	9,63E-03	-1,14E-05	1,98E-01	6,10E-02
Umid 20-30	Exponencial OLS	-1,81E-01	1,29E-02	-1,61E-04	2,54E-01	6,80E-02
	Exponencial WLS1	-1,83E-01	1,18E-02	-1,92E-04	2,57E-01	6,95E-02
	Esférico OLS	-1,83E-01	1,10E-02	-1,46E-04	2,55E-01	6,92E-02
	Esférico WLS1	-1,82E-01	1,34E-02	-1,05E-04	2,59E-01	6,92E-02
	Gaussiano OLS	-1,83E-01	1,39E-02	-9,05E-05	2,48E-01	6,67E-02
	Gaussiano WLS1	-1,81E-01	1,34E-02	-6,38E-05	2,50E-01	6,70E-02

Notas: Em negrito a melhor escolha; N: número de pontos amostrados; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados; \hat{A} : critério de Akaike; r^* , teste de Filliben, compara-se com $r = 0,9734$ (Umid 0-10), $0,9752$ (Umid 10-20) e $0,9762$ (Umid 20-30), (Anexo).

No caso dos dados da umidade nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm, os valores \hat{A} do critério de Akaike resultaram em -1291,7, -1173,4 e -1388,6, respectivamente, que, conforme a Tabela 12, apontaram para os modelos: esférico com parâmetro estimado pelo método OLS; exponencial, com parâmetros estimados pelo método OLS; e gaussiano, com parâmetros estimados pelo método OLS.

O critério de Filliben, no caso dos dados da umidade nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm, que têm 84, 88 e 91 valores amostrados,

respectivamente, e para um nível de significância de 5 %, o valor de r obtido, da tabela apresentada no Anexo, foi de 0,9734, 0,9752 e 0,9762, na devida ordem. Comparando-se com os valores r^* , apresentados na Tabela 12, nas colunas r^* OLS e WLS1, verificou-se que os valores r^* , que são maiores do que r , indicaram que a hipótese de que os dados seguem uma distribuição normal não foi rejeitada, para um nível de significância de 5 %. Por esse critério, os seguintes modelos de semivariogramas teóricos puderam ser aceitos: para o atributo umidade de 0 a 10 cm, todos os modelos foram aceitos; para o atributo umidade de 10 a 20 cm, a escolha apontou para o modelo esférico com parâmetros estimados pelo método WLS1; e para a densidade de 20 a 30 cm, nenhum modelo foi aceito.

Para os dados de umidade na profundidade 0 a 10 cm, em estudo, a escolha apontou, segundo o critério da validação cruzada, para o modelo esférico com parâmetros estimados pelo método WLS1. Para os dados de umidade na profundidade 10 a 20 cm, a escolha apontou para o modelo esférico com parâmetros estimados pelo método OLS. Para os dados de umidade na profundidade 20 a 30 cm, a escolha apontou para o modelo gaussiano com parâmetros estimados pelo método WLS1.

Calculou-se a soma de mínimos quadrados, relativos aos erros de estimação, com o objetivo de buscar aquele que apresentasse o menor valor. Observando-se a Tabela 12, a escolha apontou para os modelos: esférico, com parâmetros estimados pelo método OLS; exponencial, com parâmetros estimados pelo método OLS; e gaussiano, com parâmetros estimados pelo método OLS, respectivamente, para os dados da densidade na profundidade 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.

Um resumo com os melhores modelos escolhidos por meio dos métodos de ajuste apresenta-se na Tabela 13.

Tabela 13 - Modelos escolhidos com melhor ajuste

ATRIBUTO	AIC	FILLIBEN	VALIDAÇÃO CRUZADA
Umidade 00-10	ESF OLS	TODOS	ESF WLS1
Umidade 10-20	EXP OLS	ESF WLS1	ESF OLS
Umidade 20-30	GAU OLS	NENHUM	GAU WLS1

Notas: AIC: critério de Akaike; ESF: modelo esférico; EXP: modelo exponencial; GAU: modelo gaussiano; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados.

O melhor modelo indicado pela validação cruzada foi utilizado para confecção dos mapas temáticos. A Figura 38 mostra o mapa temático referente ao atributo umidade do solo, na profundidade 0 a 10 cm, ajustado por um modelo de semivariograma teórico esférico, com parâmetros estimados pelo método WLS1.

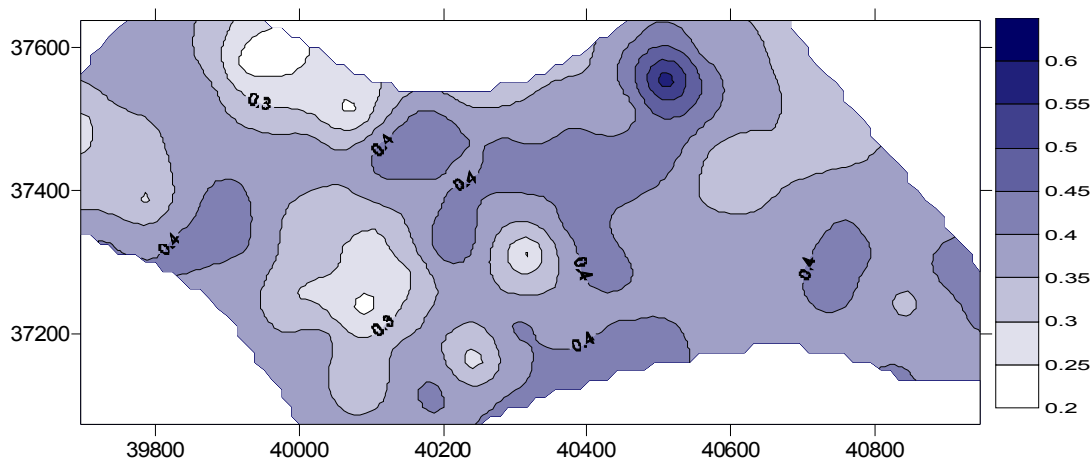


Figura 38 - Mapa para o atributo umidade na profundidade 0 a 10 cm; unidade do atributo: $[m^3 m^{-3}]$.

A Figura 39 refere-se ao mapa temático do atributo umidade do solo, na profundidade 10 a 20 cm, ajustado por um modelo de semivariograma teórico esférico, com parâmetros estimados pelo método OLS. O mapa referente ao atributo umidade do solo, na profundidade 20 a 30 cm, ajustado por um modelo de semivariograma teórico gaussiano, com parâmetros estimados pelo método WLS1, é apresentado na Figura 40.

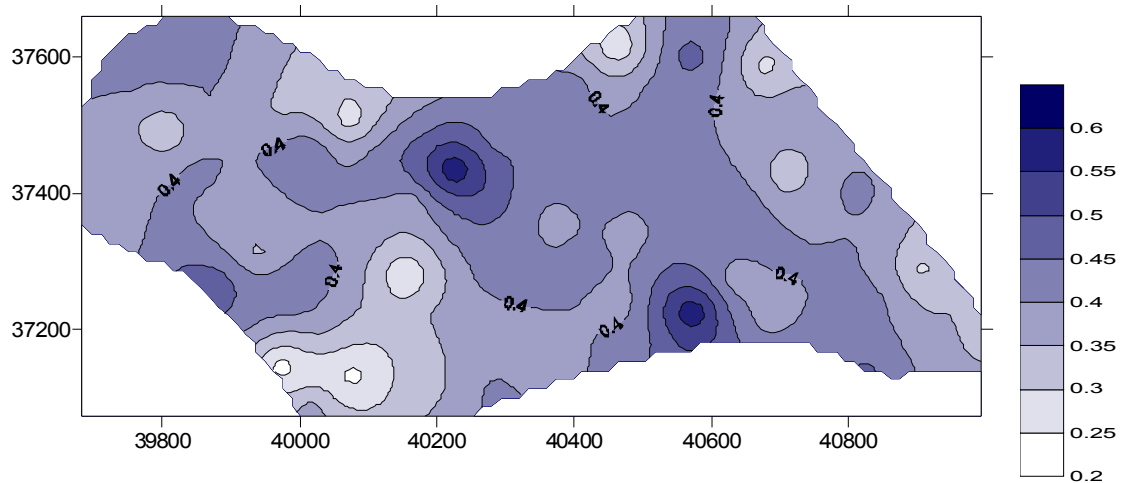


Figura 39 - Mapa para o atributo umidade na profundidade 10 a 20 cm; unidade do atributo: $[m^3 m^{-3}]$.

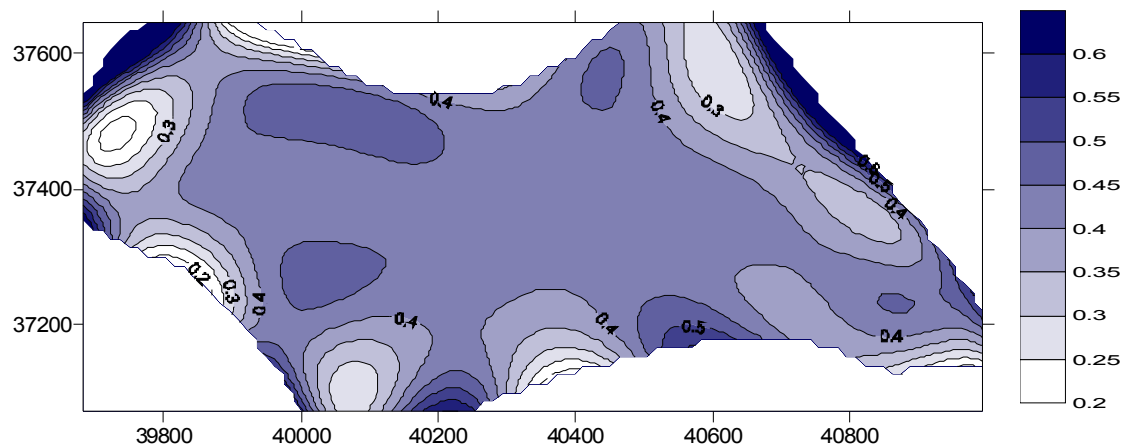


Figura 40 - Mapa para o atributo umidade na profundidade 20 a 30 cm; unidade do atributo: $[m^3 m^{-3}]$.

Resistência do solo à penetração

Os semivariogramas experimentais para a variável regionalizada resistência do solo à penetração, nas profundidades 0 a 10 cm e 10 a 20 cm, foram calculados pelo estimador de Matheron por apresentarem normalidade. O semivariograma experimental para a variável resistência do solo à penetração, na

profundidade 20 a 30 cm, foi calculado pelo estimador de Cressie e Hawkins, pois os dados não apresentaram normalidade.

A Figura 41 ilustra os semivariogramas experimentais do atributo resistência do solo à penetração nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.

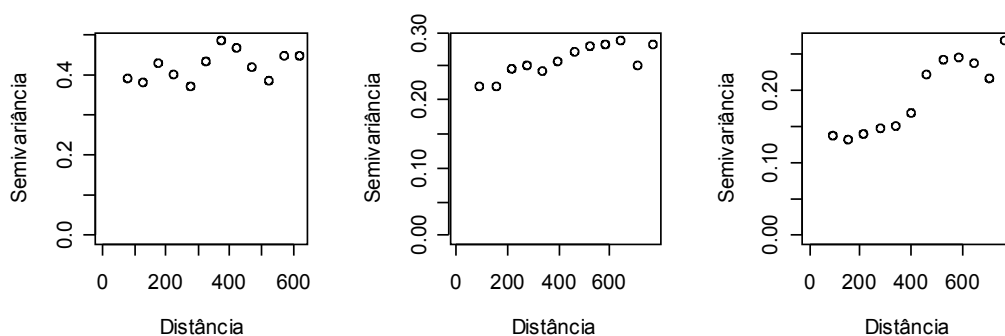


Figura 41 - Semivariogramas experimentais para a variável resistência do solo à penetração nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm.

Objetivando a aplicação dos critérios de validação ora em estudo, foram ajustados três modelos teóricos ao semivariograma experimental: exponencial, esférico e gaussiano, por serem adequados aos dados em análise. Para a estimação dos parâmetros utilizaram-se dois métodos: o dos mínimos quadrados ordinários (OLS) e o método dos mínimos quadrados ponderados (WLS1).

A Tabela 14 apresenta os parâmetros efeito pepita (C_0), patamar ($C_0 + C_1$) o alcance (a), para os modelos ajustados. Nas Figuras 42, 43 e 44 são apresentados os semivariogramas obtidos.

Tabela 14 - Parâmetros dos modelos semivariográficos pelos métodos OLS e WLS1 para os dados de resistência do solo à penetração, nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm

Variáveis estudadas	Modelos	(C_0)		$(C_0 + C_1)$		(a)	
		OLS	WLS1	OLS	WLS1	OLS	WLS1
RSP 00-10	Exponencial	0,3594	0,3518	0,4464	0,4399	638	450
	Esférico	0,3649	0,3700	0,4382	0,4378	445	440
	Gaussiano	0,3800	0,3800	0,4400	0,4375	426	369
RSP 10-20	Exponencial	0,1894	0,1699	0,2870	0,2782	901	600
	Esférico	0,2006	0,1969	0,2765	0,2752	647	600
	Gaussiano	0,2144	0,2123	0,2778	0,2771	581	554
RSP 20-30	Exponencial	0,0000	0,0000	0,2580	0,2607	784	824
	Esférico	0,0000	0,0000	0,2397	0,2401	591	609
	Gaussiano	0,1184	0,1175	0,2749	0,2864	901	946

Notas: OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados.

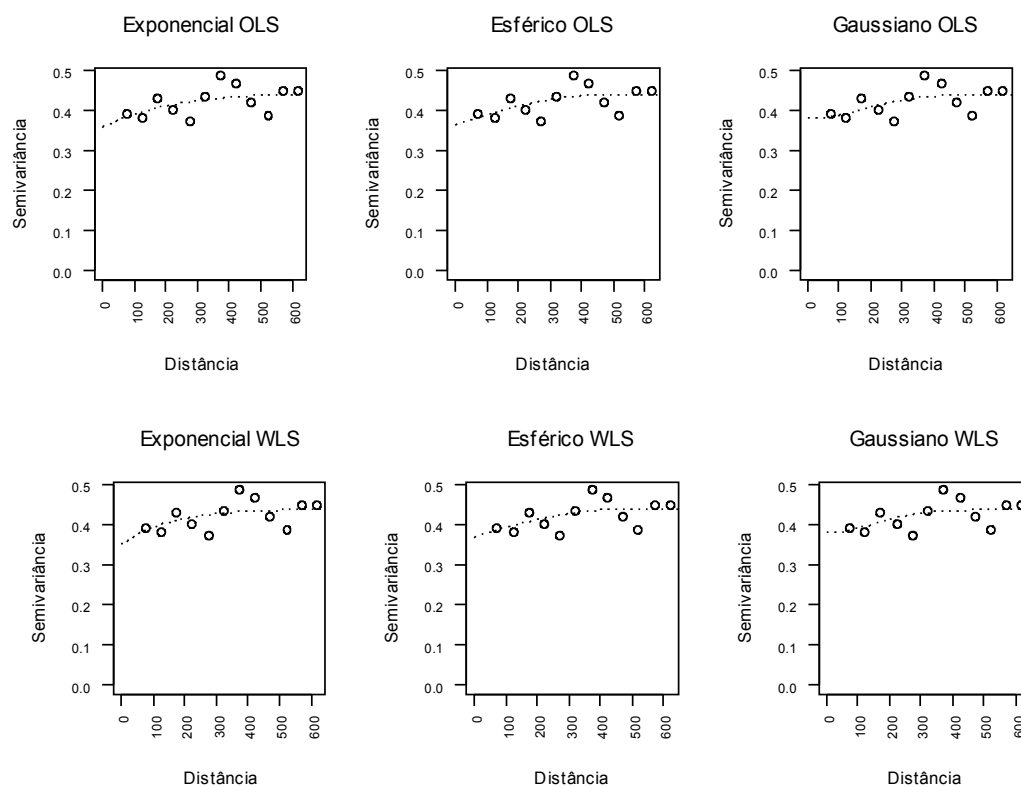


Figura 42 - Semivariogramas para os dados de resistência do solo à penetração na profundidade 0 a 10 cm.

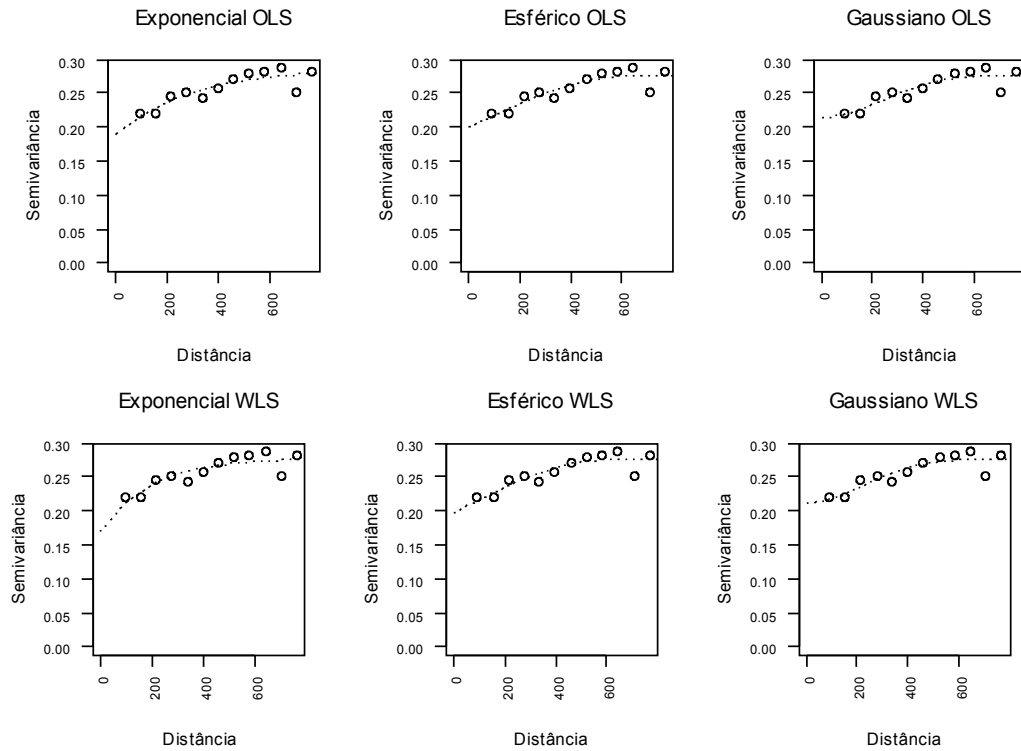


Figura 43 - Semivariogramas para os dados de resistência do solo à penetração na profundidade 10 a 20 cm.

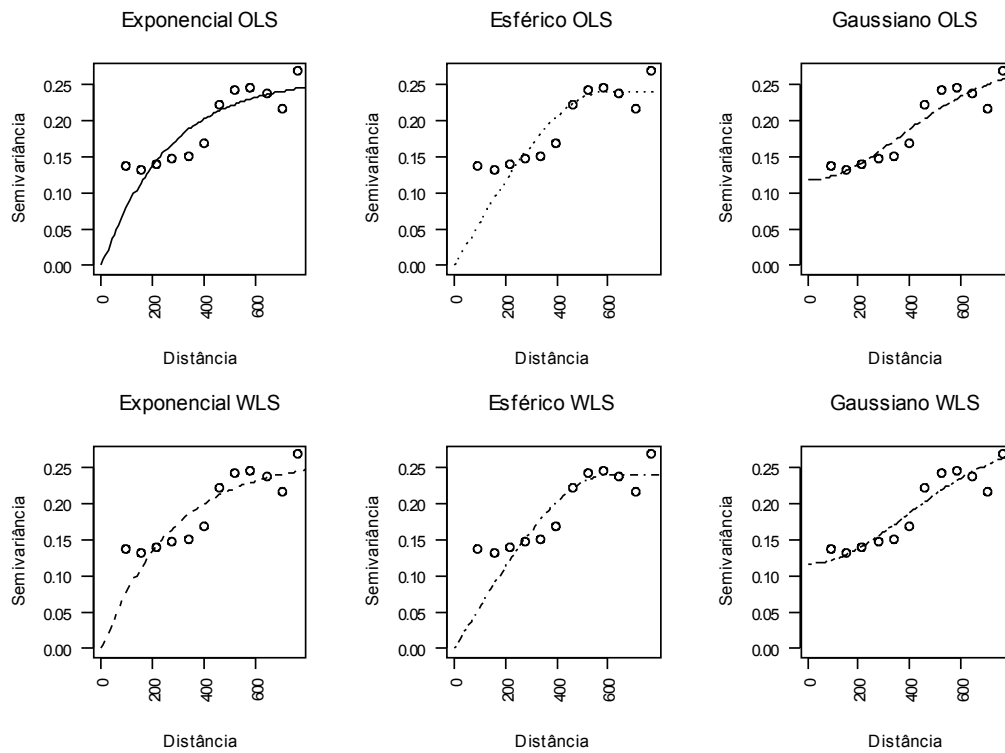


Figura 44 - Semivariogramas para os dados de resistência do solo à penetração na profundidade 20 a 30 cm.

Os resultados da aplicação das técnicas de validação de ajustes de modelos geoestatísticos encontram-se na Tabela 15.

Tabela 15 - Soma dos quadrados dos resíduos, AIC (\hat{A}), Filliben (r^*) e validação cruzada, para a resistência do solo à penetração, nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm

Variáveis estudadas	Modelos	N	Soma quadrados residuais		\hat{A}		r^*	
			OLS	WLS1	OLS	WLS1	OLS	WLS1
RSP 00-10	Exponencial	93	1,06E-02	3,01E+00	-416,8	108,4	0,9885185	0,9885360
	Esférico	93	1,02E-02	2,87E+00	-420,7	103,9	0,9888748	0,9888973
	Gaussiano	93	1,03E-02	2,92E+00	-419,3	105,6	0,9890416	0,9889842
RSP 10-20	Exponencial	90	1,34E-03	3,50E-01	-589,3	-88,4	0,9858902	0,9870525
	Esférico	90	1,20E-03	2,89E-01	-599,5	-105,6	0,9848515	0,9854298
	Gaussiano	90	1,25E-03	2,96E-01	-595,6	-103,6	0,9852789	0,9856522
RSP 20-30	Exponencial	88	4,31E-03	1,17E+00	-473,4	19,8	0,9200281	0,9201249
	Esférico	88	4,26E-03	1,17E+00	-474,3	19,5	0,9193549	0,9198493
	Gaussiano	88	3,56E-03	9,46E-01	-490,2	1,1	0,9095376	0,9095502

VALIDAÇÃO CRUZADA : ANÁLISE DOS RESÍDUOS						
MODELOS		MIN	MEDIANA	MÉDIA	MAX	D PADRÃO
RSP 00-10	Exponencial OLS	-2,21E+00	-3,21E-02	8,33E-04	1,54E+00	6,65E-01
	Exponencial WLS1	-2,20E+00	-3,69E-02	7,42E-04	1,52E+00	6,63E-01
	Esférico OLS	-2,19E+00	-3,05E-02	9,00E-04	1,52E+00	6,62E-01
	Esférico WLS1	-2,18E+00	-3,11E-02	8,93E-04	1,52E+00	6,62E-01
	Gaussiano OLS	-2,20E+00	-2,12E-02	9,98E-04	1,54E+00	6,65E-01
	Gaussiano WLS1	-2,18E+00	-1,33E-02	8,97E-04	1,53E+00	6,65E-01
RSP 10-20	Exponencial OLS	-1,71E+00	-8,04E-02	-2,36E-04	1,17E+00	5,04E-01
	Exponencial WLS1	-1,68E+00	-7,52E-02	-3,63E-04	1,17E+00	5,05E-01
	Esférico OLS	-1,73E+00	-7,83E-02	-1,16E-04	1,19E+00	5,05E-01
	Esférico WLS1	-1,71E+00	-8,17E-02	-1,12E-04	1,19E+00	5,04E-01
	Gaussiano OLS	-1,73E+00	-7,81E-02	-4,15E-06	1,21E+00	5,05E-01
	Gaussiano WLS1	-1,72E+00	-8,13E-02	1,63E-05	1,20E+00	5,04E-01
RSP 20-30	Exponencial OLS	-1,65E+00	-3,95E-02	-7,01E-03	2,98E+00	5,62E-01
	Exponencial WLS1	-1,65E+00	-3,87E-02	-7,13E-03	2,98E+00	5,63E-01
	Esférico OLS	-1,79E+00	-4,53E-02	-9,11E-03	3,01E+00	5,72E-01
	Esférico WLS1	-1,79E+00	-4,29E-02	-8,94E-03	3,00E+00	5,73E-01
	Gaussiano OLS	-1,40E+00	-6,76E-02	-9,57E-04	2,99E+00	5,32E-01
	Gaussiano WLS1	-1,41E+00	-6,71E-02	-9,39E-04	2,99E+00	5,32E-01

Notas: Em negrito a melhor escolha; N: número de pontos amostrados; RSP: resistência do solo à penetração; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados; \hat{A} : critério de Akaike; r^* , teste de Filliben, compara-se com $r = 0,9766$ (RSP 0-10), $0,9760$ (RSP 10-20) e $0,9752$ (RSP 20-30), (Anexo).

No caso dessa variável nas profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm, os valores \hat{A} do critério de Akaike são aqueles correspondentes a $-420,7$, $-599,5$ e $-490,2$, respectivamente, que, conforme a Tabela 15, apontou para os modelos: esférico, com parâmetros estimados pelo método OLS, esférico,

com parâmetros estimados pelo método OLS; e gaussiano, com parâmetros estimados pelo método OLS.

Por meio do critério de Filliben, testaram-se os resíduos ortonormais para examinar se seguiam uma distribuição normal, ao nível de 5% de significância. Para as profundidades 0 a 10 cm, 10 a 20 cm e 20 a 30 cm, com 93, 90 e 88 valores amostrados, respectivamente, e para um nível de significância de 5 %, o valor de r , obtido da tabela apresentada no Anexo, é de 0,9766, 0,9760 e 0,9752, nessa ordem. Comparando-se com os valores r^* , apresentados na Tabela 15, nas colunas r^* , OLS e WLS1, verificou-se que os valores r^* , que eram maiores do que r , indicaram que a hipótese de que os erros seguiam uma distribuição normal não foi rejeitada para um nível de significância de 5 %. Por esse critério, os seguintes modelos de semivariogramas teóricos puderam ser aceitos: nas profundidades de 0 a 10 cm e 10 a 20 cm, todos os modelos, e para a profundidade 20 a 30 cm, nenhum modelo foi aceito.

O critério de validação cruzada sugeriu que o menor valor da média dos erros de estimação indicou o modelo com melhor ajuste. À profundidade 0 a 10 cm, a escolha aponta para o modelo exponencial, com parâmetros estimados pelo método WLS1 e na profundidade 10 a 20 cm, a escolha apontou para o modelo gaussiano, com parâmetros estimados pelo método OLS. Para a profundidade 20 a 30 cm, a escolha apontou para o modelo gaussiano com parâmetros estimados pelo método WLS1.

Além dos critérios apresentados, para fins de comparação, optou-se por calcular o soma de quadrados mínimos relativos aos erros de estimação. Por esse critério, ao se procurar por um modelo, buscou-se aquele que apresentasse o menor valor e, observando-se a Tabela 15, a escolha apontou para os modelos: esférico, com parâmetros estimados pelo método OLS; esférico, com parâmetros estimados pelo método OLS; e gaussiano, com parâmetros estimados pelo método OLS.

Na Tabela 16, é apresentado um resumo com os melhores modelos escolhidos usando-se cada um dos métodos de ajuste.

Tabela 16 - Modelos escolhidos com melhor ajuste

ATRIBUTO	AIC	FILLIBEN	VALIDAÇÃO CRUZADA
RSP 00-10	ESF OLS	TODOS	EXP WLS1
RSP 10-20	ESF OLS	TODOS	GAU OLS
RSP 20-30	GAU OLS	NENHUM	GAU WLS1

Notas: AIC: critério de Akaike; ESF: modelo esférico; EXP: modelo exponencial; GAU: modelo gaussiano; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados.

Para a confecção dos mapas, foi utilizado o melhor modelo indicado pela validação cruzada. A Figura 45 mostra o mapa referente ao atributo resistência do solo à penetração, na profundidade 0 a 10 cm, ajustado por um modelo de semivariograma teórico exponencial, com parâmetros estimados pelo método WLS1.

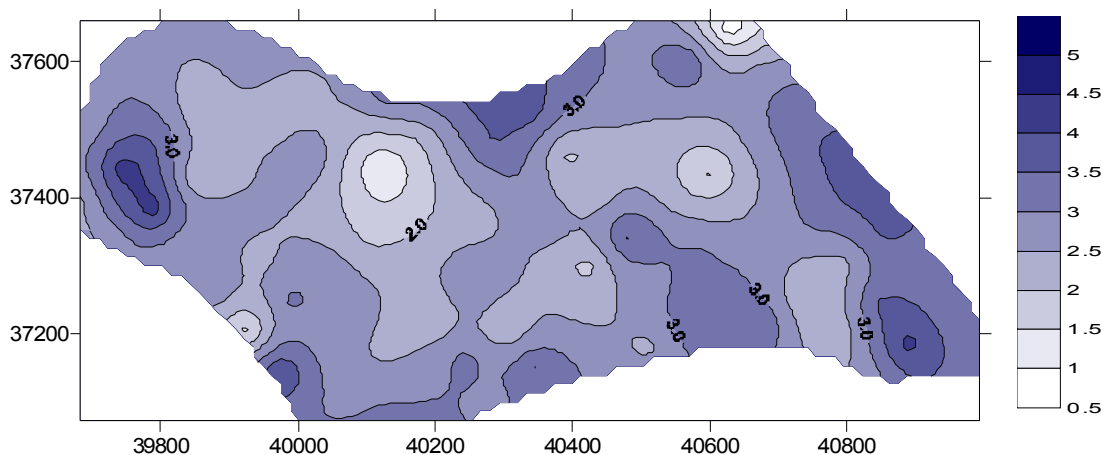


Figura 45 - Mapa para o atributo resistência do solo à penetração na profundidade 0 a 10 cm; unidade do atributo: [Mpa].

A Figura 46 refere-se ao mapa do atributo resistência do solo à penetração, na profundidade 10 a 20 cm, ajustado por um modelo de semivariograma teórico gaussiano, com parâmetros estimados pelo método OLS. O mapa referente ao atributo resistência do solo à penetração, na profundidade 20 a 30 cm, ajustado por um modelo de semivariograma teórico gaussiano, com parâmetros estimados pelo método WLS1, é apresentado na Figura 47.

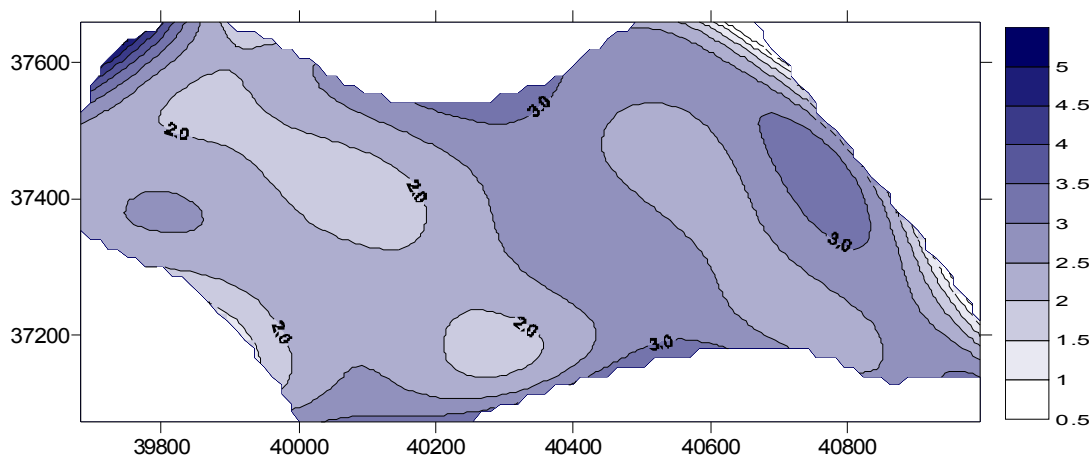


Figura 46 - Mapa para o atributo resistência do solo à penetração na profundidade 10 a 20 cm; unidade do atributo: [Mpa].

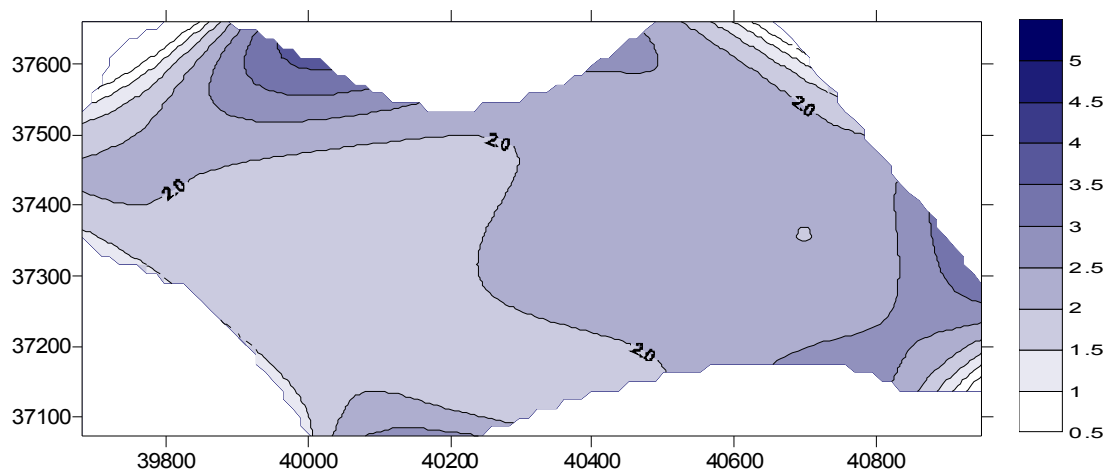


Figura 47 - Mapa para o atributo resistência do solo à penetração na profundidade 20 a 30 cm; unidade do atributo: [Mpa].

Produtividade

O semivariograma experimental para a variável produtividade foi calculado pelo estimador de Matheron, pois os dados apresentaram normalidade. A Figura 48 ilustra o semivariograma experimental obtido.

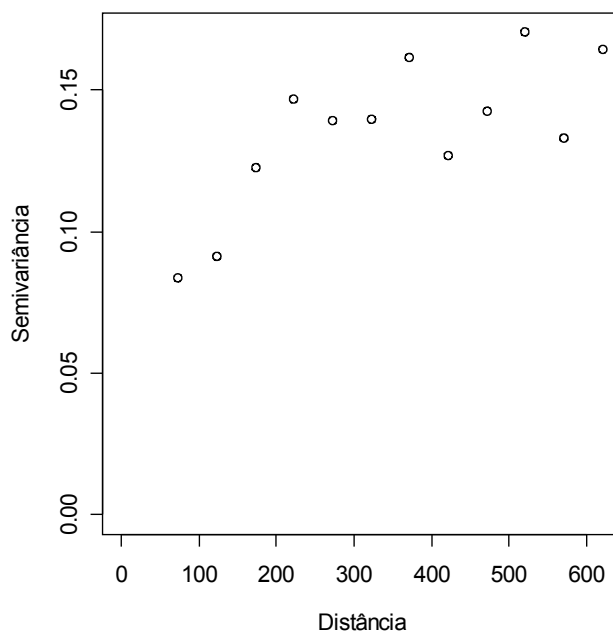


Figura 48 - Semivariograma experimental para a variável produtividade.

A esse semivariograma experimental foram ajustados três modelos teóricos, com o objetivo de se aplicarem os critérios de validação em estudo: exponencial, esférico e gaussiano, considerados adequados aos dados em análise. Para a estimação dos parâmetros utilizaram-se dois métodos de ajuste: o dos mínimos quadrados ordinários (OLS) e o método dos mínimos quadrados ponderados (WLS1).

A Tabela 17 apresenta os parâmetros efeito pepita (C_0), patamar ($C_0 + C_1$) e alcance (a), para os modelos ajustados. A Figura 49 mostra os semivariogramas obtidos por ajuste, sobrepostos ao semivariograma experimental.

Tabela 17 - Parâmetros dos modelos semivariográficos pelos métodos OLS e WLS1 para os dados de produtividade

Variáveis estudadas	Modelos	(C_0)		$(C_0 + C_1)$		(a)	
		OLS	WLS1	OLS	WLS1	OLS	WLS1
Produtividade	Exponencial	0,0096	0,0000	0,1519	0,1532	341	320
	Esférico	0,0394	0,0358	0,1477	0,1490	309	300
	Gaussiano	0,0624	0,0664	0,1485	0,1505	283	294

Notas: OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados.

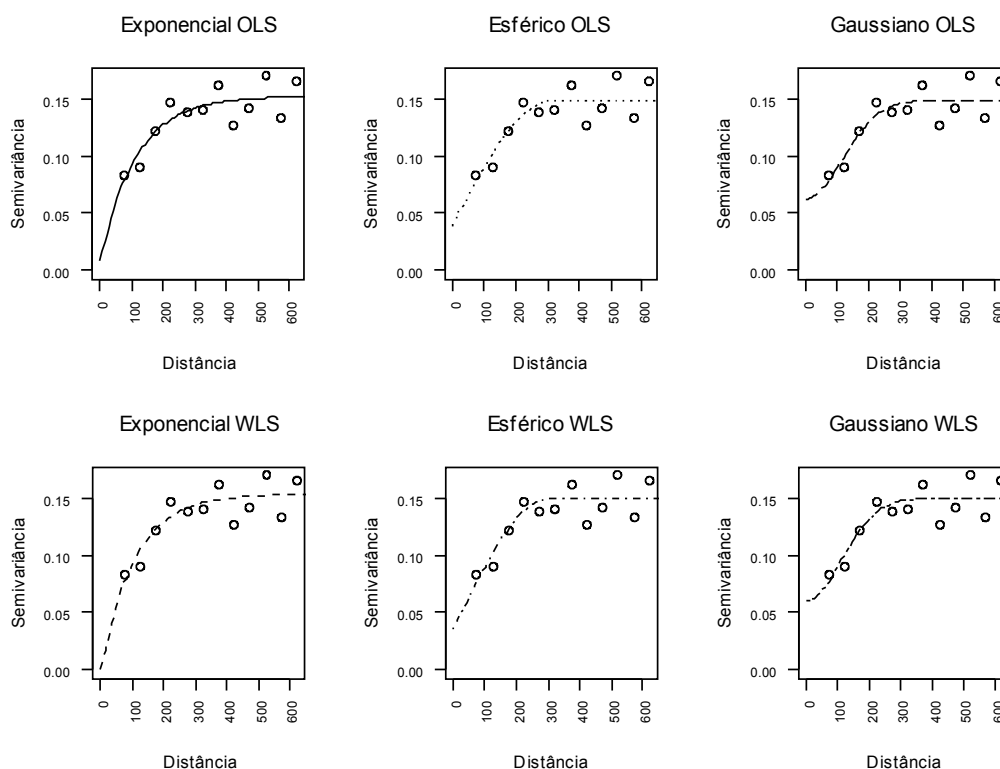


Figura 49 - Semivariogramas para os dados de produtividade.

Para aplicarem-se as técnicas de validação de ajustes de modelos geoestatísticos, necessitou-se analisar os resultados dos diversos métodos e proceder-se à comparação de valores. Os resultados apurados encontram-se na Tabela 18.

Tabela 18 - Soma dos quadrados dos resíduos, AIC (\hat{A}), Filliben (r^*) e validação cruzada para a produtividade

Variáveis estudadas	Modelos	N	Soma quadrados residuais		\hat{A}		r^*	
			OLS	WLS1	OLS	WLS1	OLS	WLS1
Prod	Exponencial	66	2,06E-03	2,84E-01	-402,1	-77,1	0,9866310	0,9878670
	Esférico	66	2,01E-03	2,88E-01	-403,9	-76,2	0,9836391	0,9848132

VALIDAÇÃO CRUZADA : ANÁLISE DOS RESÍDUOS						
MODELOS	MIN	MEDIANA	MÉDIA	MAX	D PADRÃO	
Prod	Exponencial OLS	-1,10E+00	-9,33E-04	-4,66E-03	8,39E-01	3,55E-01
	Exponencial WLS1	-1,11E+00	8,00E-03	-5,01E-03	8,50E-01	3,57E-01
	Esférico OLS	-1,11E+00	-5,32E-02	-4,72E-03	8,34E-01	3,56E-01
	Esférico WLS1	-1,12E+00	-5,10E-02	-5,03E-03	8,42E-01	3,58E-01
	Gaussiano OLS	-1,11E+00	-4,41E-02	-4,86E-03	8,12E-01	3,54E-01
	Gaussiano WLS1	-1,11E+00	-4,11E-02	-4,46E-03	8,04E-01	3,53E-01

Notas: Em negrito a melhor escolha; N: número de pontos amostrados; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados; \hat{A} : critério de Akaike; r^* , teste de Filliben, compara-se com $r = 0,9674$ (Anexo).

Para essa variável o valor de \hat{A} obtido foi de $-405,5$ que, pela Tabela 18, apontou para o modelo gaussiano, com seus parâmetros estimados pelo método OLS.

O critério de Filliben aplicado testou se os resíduos ortonormais seguiam uma distribuição normal a um determinado nível de significância. Para o caso dos dados da produtividade, com 66 valores amostrados, e para um nível de significância de 5 %, o valor de r , obtido da tabela apresentada no Anexo, foi de 0,9674. Comparando-se com os valores r^* , apresentados na Tabela 18, nas colunas Filliben, OLS e WLS1, verificou-se que todos os valores r^* são maiores do que r e constatou-se, então, que a hipótese de que os resíduos seguiam uma distribuição normal não foi rejeitada, a 5 %. Por esse critério, todos os modelos de semivariogramas teóricos dessa variável puderam ser aceitos.

O critério de validação cruzada, aplicado aos modelos em estudo, sugeriu que o menor valor da média dos erros de estimação indicou o modelo com melhor ajuste. Para essa variável a escolha aponta para o modelo gaussiano com parâmetros estimados pelo método WLS1.

A soma de quadrados mínimos, relativa aos erros de estimação, apontou para um modelo gaussiano, com os parâmetros estimados pelo método OLS.

Na Tabela 19, apresenta-se um resumo com os melhores modelos escolhidos por meio dos métodos de ajuste.

Tabela 19 - Modelos escolhidos com melhor ajuste

VARIÁVEL	AIC	FILLIBEN	VALIDAÇÃO CRUZADA
Produtividade	GAU OLS	TODOS	GAU WLS1

Notas: AIC: critério de Akaike; ESF: modelo esférico; EXP: modelo exponencial; GAU: modelo gaussiano; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados.

O melhor modelo indicado pela validação cruzada foi escolhido para representar a variável produtividade na confecção do mapa temático. A Figura 50 mostra o mapa referente à variável produtividade, ajustada por um modelo de semivariograma teórico gaussiano, com parâmetros estimados pelo método WLS1.

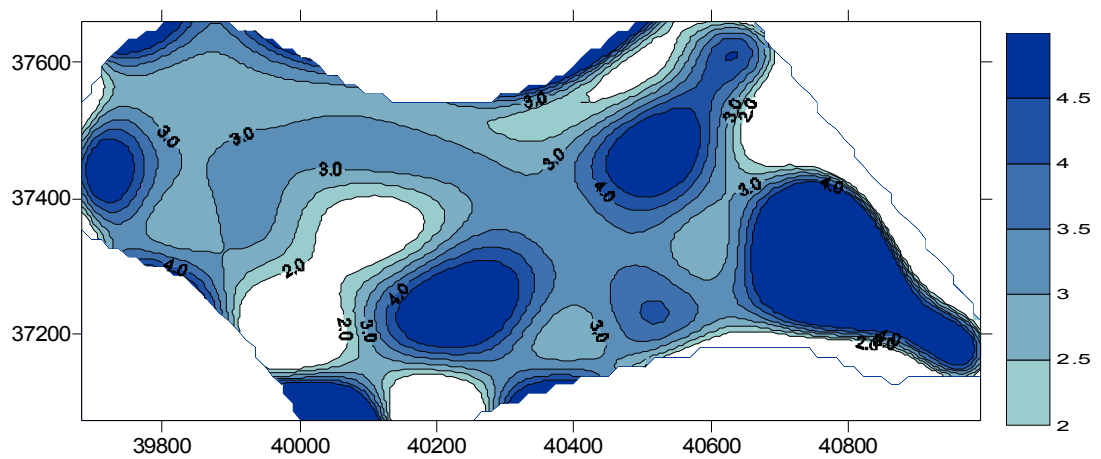


Figura 50 - Mapa para a variável produtividade; unidade do atributo: [Mg ha⁻¹].

5 CONCLUSÕES

Com base nos dados obtidos na pesquisa, é possível concluir que:

- Para o caso da simulação, o critério da validação cruzada alcançou com precisão o modelo esférico de dependência espacial proposto. O critério de Filliben mostrou-se não restritivo, indicando que todos os modelos ajustados são válidos. O critério de Akaike foi o menos preciso para este caso, afastando-se do modelo original, pois indicou um modelo gaussiano como o melhor.
- Construíram-se modelos empíricos de variabilidade espacial dos atributos físicos do solo e da produtividade. Foi verificado, ao se aplicar os critérios de validação cruzada Akaike e Filliben, que eles não convergem para um mesmo modelo. Essa constatação corrobora as distintas naturezas dos critérios. No caso do critério de Akaike, que estabelece uma penalidade para o acréscimo de parâmetros aos modelos, se todos tiverem o mesmo número de parâmetros, buscar o que tem menor valor para o AIC não traz vantagens sobre o cálculo da soma de quadrados dos resíduos. A aplicação do critério de Filliben para as variáveis estudadas não foi eletivo de uma forma precisa, isto é, não apontou para um modelo determinado e, sim, para vários simultaneamente.
- O critério da validação cruzada foi considerado o mais adequado para a escolha do melhor ajuste e os mapas temáticos foram construídos utilizando-se uma estrutura de dependência espacial, escolhida segundo esse critério, e permitiram uma noção visual do

comportamento dos atributos e da variável produtividade na área estudada.

- À semelhança da construção de modelos empíricos de variabilidade espacial aos semivariogramas experimentais, em que o conhecimento do pesquisador, a respeito da variável em estudo, conduz a resultados melhores, também no caso de verificação de qualidade dos ajustes, o conhecimento e o bom senso do pesquisador são de grande valia.

REFERÊNCIAS

AKAIKE, H. Information theory as an extension of the maximum likelihood principle. INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY, 2. 1973, Budapest. **Anais...** Budapest: B. Petrov and F. Csaki (eds.), Akadémia Kiadó, p.267-281, 1973.

ANTUNES, J. M. **Rendimento da soja aumenta 25 Kg/ha ao ano.** In: EMBRAPA Soja. Disponível em: www.cnpso.embrapa.br/noticias. Acesso em: 12 de outubro de 2004.

AZEVEDO, A. C. D.; DALMOLLIN, R. S. D. **Solos e ambiente:** uma introdução. Santa Maria, RS: Palloti, 2004. 100 p.

BALASTREIRE, L. A. Estudo de caso: uma pesquisa brasileira em agricultura de precisão. In: CONGRESSO BRASILEIRO DE ENGENHARIA AGRÍCOLA, 27. 1998, Poços de Caldas. **Anais...** Lavras: Universidade Federal de Lavras, 1998. p. 203-231.

BARNES, R. J. The variogram sill and the sample variance. **Mathematical Geology**, New York, v. 23, n. 4, p.673-8, 1991.

BURROUGH, P. A. **Principles of geographical information systems for land resources assessment.** Oxford: Clarendon Press, 1987. 193 p.

CAMARGO, E. C. G. **Desenvolvimento, implementação e teste de procedimentos geoestatísticos (krigagem) no sistema de processamento de informações georreferenciadas (Spring).** São José dos Campos, SP, 1997. 123 f. Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais.

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 America Journal**, Madison, v. 58, n. 5, p. 1501-1511, 1994.

CARVALHO, J. R. P. de; ASSAD, E. D. Análise espacial de precipitação pluviométrica no estado de São Paulo: comparação de interpoladores. In: CONGRESSO BRASILEIRO DE ENGENHARIA AGRÍCOLA, 32. 2003, Goiânia. **Anais...** Goiânia: SBEA, 2003. CD-ROM.

CARVALHO, J. R. P.; SILVEIRA, P. M.; VIEIRA, S. R. Geoestatística na determinação da variabilidade espacial de características químicas do solo sob diferentes preparos. **Pesquisa Agropecuária Brasileira**, Brasília - DF, v. 37, n. 8, p. 1151-1159, ago. 2002.

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

COMPANHIA NACIONAL DE ABASTECIMENTO - CONAB. **Série histórica de produtividade**. Disponível em: <http://www.conab.gov.br/download/safra/SojaSerieHist.xls>. Acesso em: 12 de janeiro de 2006.

CRESSIE, N. Fitting variogram models by weighted least squares. **Mathematical Geology**, New York, v. 17, n. 4, p. 563-586, 1985.

CRESSIE, N. Spatial prediction and ordinary kriging. **Mathematical Geology**, New York, v. 20, n. 4, p. 405-21, 1988. Erratum. **Mathematical Geology**, New York, v. 21, n. 4, p. 493-4, 1989.

CRESSIE, N. **Statistics for spatial data**. New York: John Wiley, 1993. 900 p.

CRESSIE, N.; HAWKINS, D. M. Robust estimation of the variogram. **Mathematical Geology**. New York, v. 12, n. 2, p. 115-125, 1980.

DAVID, M. The Practice of Kriging. In: GUARASCIO, M.; DAVID, M.; HUIJBREGTS, C. (Eds.), **Advanced geostatistics in the mining industry**: Boston: D. Reidell, 1976. p.31-48.

DAVIS, B. M. Uses and abuses of cross-validation in Geoestatistics. **Mathematical Geology**, New York, v. 19, n. 13, p. 1-8, 1987.

DELFINER, P. Linear estimation of nonstationary spatial phenomena. In: GUARASCIO, M.; DAVID, M.; HUIJBREGTS, C. (Eds.). **Advanced geostatistics in the mining industry**. Boston: D. Reidell, 1976. p. 49-68.

DIGGLE, P. J.; RIBEIRO JUNIOR, P. J. **Model-based geostatistics**. Caxambu: Associação Brasileira de Estatística, 2000. 129 p.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA. Serviço Nacional de Levantamento e Conservação de Solos. **Manual de métodos de análise do solo**. Rio de Janeiro: Embrapa - Centro Nacional de Pesquisas de Solos, 1997. 212 p.

FILLIBEN, J. J. The probability plot correlation coefficient test for normality. **Technometrics**, Alexandria, v. 17, n. 1, p. 111-117, 1975.

FONTANA, D. C.; WEBER, E.; DUCATI, J.; FIGUEIREDO, D. C.; BERLATO, M. A.; BERGAMASCHI, H. Previsão da safra de soja no Brasil: 1999/2000. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 12; REUNIÃO LATINO-AMERICANA DE AGROMETEOROLOGIA, 3., 2001. Fortaleza. **Anais...** Fortaleza: Sociedade Brasileira de Agrometeorologia, 2001. p.585-586.

GOMES, F. P. **A estatística moderna na pesquisa agropecuária**. Piracicaba: Potafós, 2000. 160 p.

GONÇALVES, A. C. A. **A variabilidade espacial de propriedades físicas do solo pra fins de manejo da irrigação**. Piracicaba, 1997. 118 f. Tese (Doutorado em Agronomia) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.

GONÇALVES, A. C. A.; FOLEGATTI, M. V. Variabilidade espacial de propriedade físicas do solo, ao longo de uma transeção em área irrigada por pivô central. In: CONGRESSO BRASILEIRO DE ENGENHARIA AGRÍCOLA, 24. 1995, Viçosa – MG. **Anais...** Viçosa – MG: Sociedade Brasileira de Engenharia Agrícola, 1995. CD-ROM.

GUERRA, P. A. G. **Geoestatística operacional**. Brasília: Ministério das Minas e Energia, Departamento Nacional da Produção Mineral, 1988. 145 p.

ISAAKS, E. H.; SRIVASTAVA, R. M. **An introduction to applied geostatistics**. New York: Oxford University Press, 1989. 561 p.

JOURNAL, A. G.; HUIJBREGTS, C. J. **Mining geostatistics**. London: Academic Press, 1978. 600 p.

KILPP, A. R. Fieldstar: Solução em agricultura de precisão. In: BALASTREIRE, L. A. **O estado-da-arte da agricultura de precisão no Brasil**. Piracicaba, SP: L. A. Balastreire, 2000. p. 39-70.

KLEIN, V. A. **Propriedades físico-hídrico-mecânicas de um latossolo roxo, sob diferentes sistemas de uso e manejo**. Piracicaba, SP, 1999. 150 f. Tese (Doutorado em Agronomia – Solos e Nutrição de Plantas) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.

KRIGE, D. G. **A statistical approach to some basic mine evaluation problems on the Witwatersrand**. Johannesburg: Johannesburg Chemistry Metallurgy Mining Society South African, 1951. p. 119-139.

LEE, S. Validation of geostatistical models using the Filliben test for orthonormal residuals. **Journal of Hydrology**, Amsterdam, v. 158, p. 319-332, 1994.

LI, D.; LAKE, L. W. A moving window semivariance estimator. **Water Resources Research**, Washington DC - EUA, v. 30, n. 5, p. 1479-1490, 1994.

McBRATNEY, A. B.; WEBSTER, R. Choosing functions for semivariograms of soil properties and fitting them to samples estimates. **Journal of Soil Science**. Oxford, v. 37, p. 617-39, 1986.

MATHERON, G. Principles of geostatistics. **Economic Geology**, Lancaster, v. 58, p. 1246-1266, 1963.

MILANI, D. I. C. **Variabilidade espacial da infiltração da água no solo e correlação com parâmetros físicos e produtividade do milho**. Cascavel, PR, 2005. 119 f. Dissertação (Mestrado em Engenharia Agrícola) - Centro de Ciências Exatas e Tecnológicas, Universidade Estadual do Oeste do Paraná.

MOTTA, J. L. G.; FONTANA, D. C.; WEBER, E. Verificação da acurácia da estimativa de área cultivada com soja através de classificação digital em imagens Landsat. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 10., 2001, Foz do Iguaçu. **Anais...** Foz do Iguaçu – PR: INPE, 2001. p. 123-129

OLEA, R. A. **Optimum mapping techniques using regionalized variable theory**. Kansas: Kansas Geological Survey, 1975. 137 p. (Series on Spatial Analysis, 2).

_____. **Measuring spatial dependence with semivariograms**, Kansas: Kansas Geological Survey, 1977. 29 p. (Series on Spatial Analysis 3).

PANNATIER, Y. **Variowin 2.2**. Software for special data analysis in 2D. New York: Springer-Verlag, 1996. 96 p.

PARKER, H. M.; JOURNAL, A. G.; DIXON, W. C. The use of conditional lognormal probability distribution for the estimation of open-pit ore reserves in strata-bound uranium deposits – A case study. In: APCOM, 16., 1979, New York. **Proceedings...** New York: T. J. O’Neil (Ed.), Society of Mining Engineers., 1979. p. 133-148.

QUENOUILLE, M. Notes on bias in estimation. **Biometrika**, Edinburgo, v. 43, p. 353-360, 1956.

R DEVELOPMENT CORE TEAM. 2005. **R: A language and environment for statistical computing**. R Vienna, Austria: Foundation for Statistical Computing, Disponível em: <http://www.R-project.org>. Acesso em: 26 de setembro de 2005. ISBN 3-900051-07-0.

REICHARDT, K.; TIMM, L. C. **Solo, planta e atmosfera: conceitos, processos e aplicações**. Barueri-SP: Editora Manole Ltda, 2004. 478 p.

RIBEIRO JR. P. J. **Métodos geoestatísticos no estudo da variabilidade espacial de parâmetros do solo**. Piracicaba, SP, 1995. 99 f. Dissertação (Mestrado em Agronomia) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.

RIBEIRO JR., P. J.; DIGGLE, P. J. GeoR: A package for geostatistical analysis. **R-News**, Vienna - Austria, v. 1, n. 2. 2001. R: Foundation for Statistical Computing, Disponível em: <http://www.R-project.org>. Acesso em: 26 de setembro de 2005. ISSN 1609-3631.

SAKAMOTO, Y.; ISHIGURO, M.; KITAGAWA, G. **Akaike information criterion statistics**. New York: Springer. 1986. 320 p.

SECCO, D. **Estados de compactação de dois latossolos sob plantio direto e suas implicações no comportamento mecânico e na produtividade de culturas**. Santa Maria, RS, 2003. 108 f. Tese (Doutorado em Agronomia) - Programa de Pós-Graduação em Agronomia, Universidade Estadual de Santa Maria.

SILVA, V. R. D.; REICHERT, J. M.; REINERT, D. J. Variabilidade espacial da resistência do solo à penetração em plantio direto. **Ciência Rural**, Santa Maria, v. 34, n. 2, p. 399 - 406, 2004.

SOBER, E. Instrumentalism, parsimony, and the Akaike framework. **Philosophy of science**, University of Chicago Press, v. 69, p. S112-S123, 2002.

TIEPPO, R. C. **Desenvolvimento de um penetrógrafo eletrônico**. Cascavel, PR, 2004. 85 f. Dissertação (Mestrado em Engenharia Agrícola) - Centro de Ciências Exatas e Tecnológicas, Universidade Estadual do Oeste do Paraná.

TORRES, E.; SARAIVA, O. F. **Camadas de impedimento do solo em sistemas agrícolas com a soja**. Londrina: Embrapa Soja, 1999. 58 p. (Circular Técnica, 23)

TUKEY, J. **Bias and confidence in not quite large samples**, Ann. Math. Stat., v. 29, 1958. 614 p.

VIEIRA, S. R. Variabilidade espacial de argila, silte e atributos químicos em uma parcela experimental de um latossolo roxo de Campinas (SP). **Bragantia**, Campinas, v. 56, n. 1, p. 1-17, 1997.

VIEIRA, S. R. **Geoestatística aplicada à agricultura de precisão**. São Paulo: Instituto Agrônomo de Campinas, 1998. 53 p.

VIEIRA, S. R.; CARVALHO, J. R. P. de. **Estudo de periodicidade temporal de chuvas em bacia hidrográfica dos Rios Turvo/Grande – uma proposta**. Campinas: Embrapa Informática Agropecuária, 2001. 17 p. (Série Documentos, 10).

VIEIRA, S. R.; CARVALHO, J. R. P.; MORAES, J. R. P. de. Uso de *Jack knifing* para validação geoestatística. In: CONGRESSO BRASILEIRO DE AGRICULTURA DE PRECISÃO, CONBAP, 2004. Piracicaba – SP. **Anais...** Piracicaba – SP: Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, 2004.

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

WEBSTER, R. Quantitative spatial analysis of soil in the field. **Advances in Soil Science**. New York, v.3, p.1-70, 1985.

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

APÊNDICES

APÊNDICE A - GRÁFICOS DE DISPERSÃO

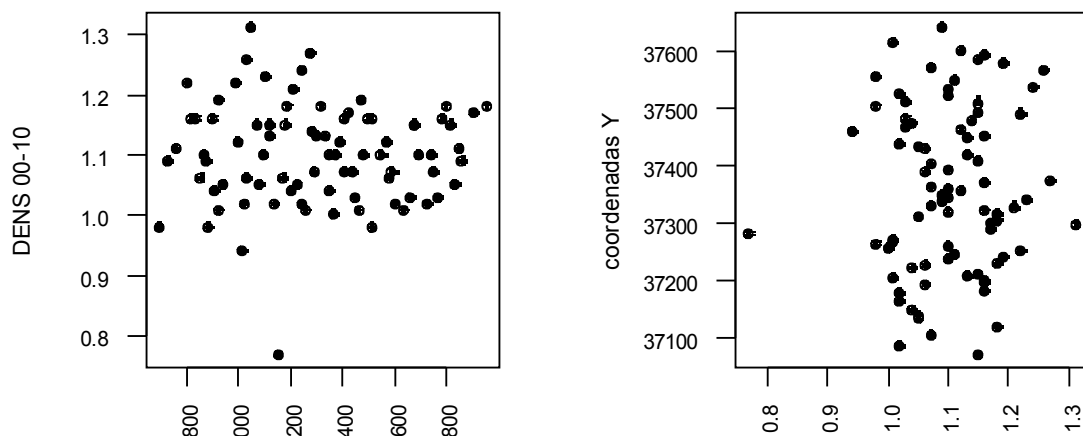


Figura 1A - Gráficos de dispersão para a densidade do solo na profundidade 0 a 10 cm, *versus* linha à esquerda e *versus* coluna à direita.

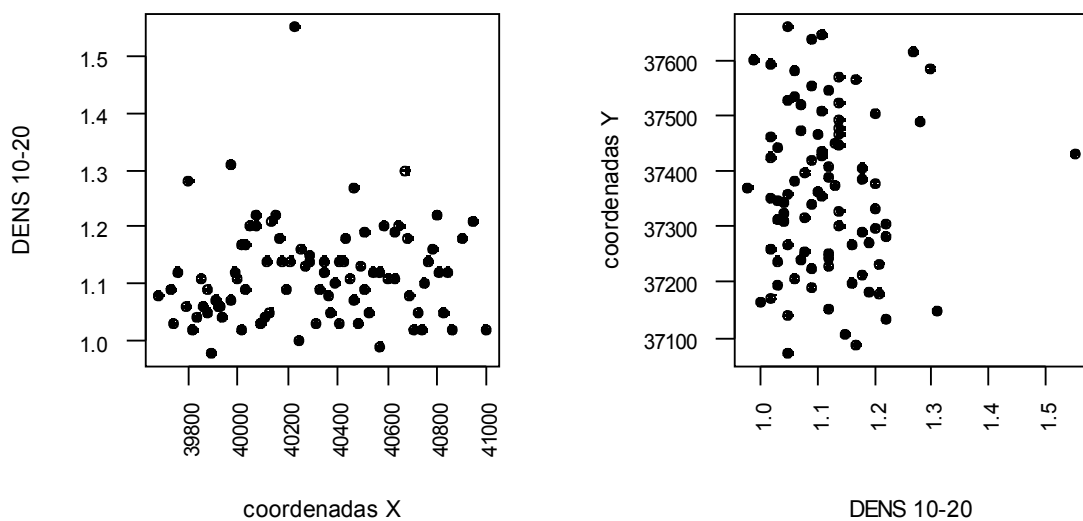


Figura 2A - Gráficos de dispersão para a densidade do solo na profundidade 10 a 20 cm, *versus* linha à esquerda e *versus* coluna à direita.

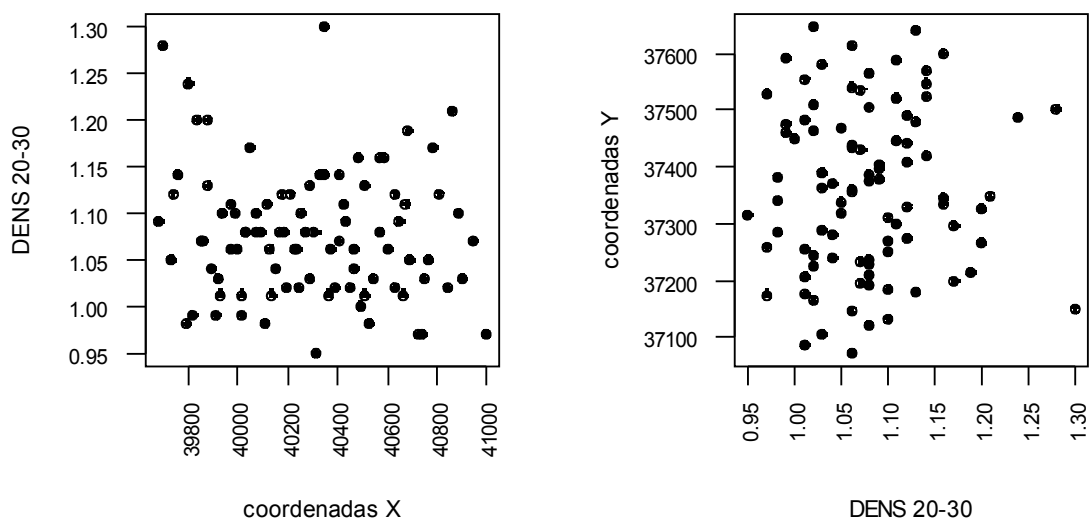


Figura 3A - Gráficos de dispersão para a densidade do solo na profundidade 20 a 30 cm, *versus* linha à esquerda e *versus* coluna à direita.

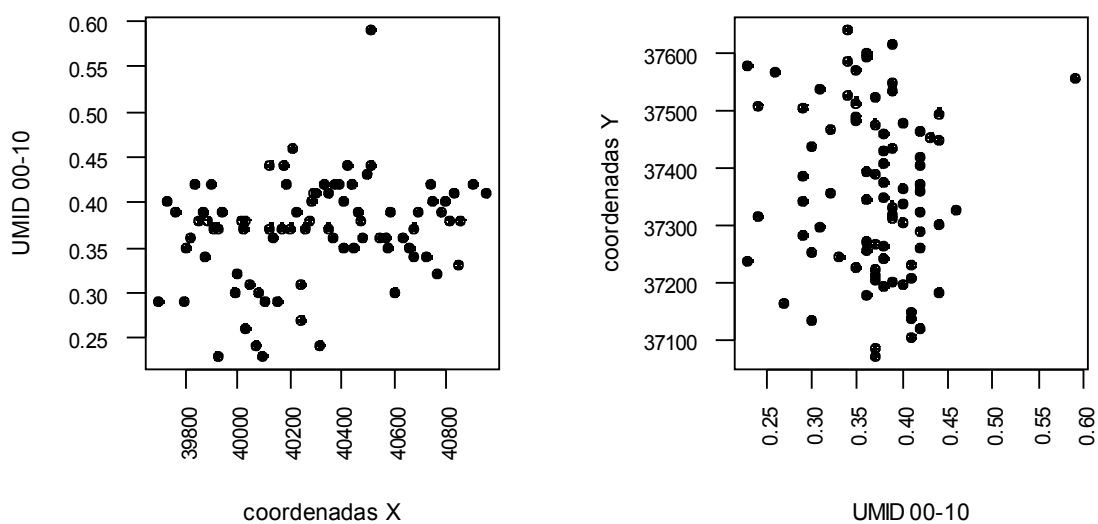


Figura 4A - Gráficos de dispersão para a umidade do solo na profundidade 0 a 10 cm, *versus* linha à esquerda e *versus* coluna à direita.

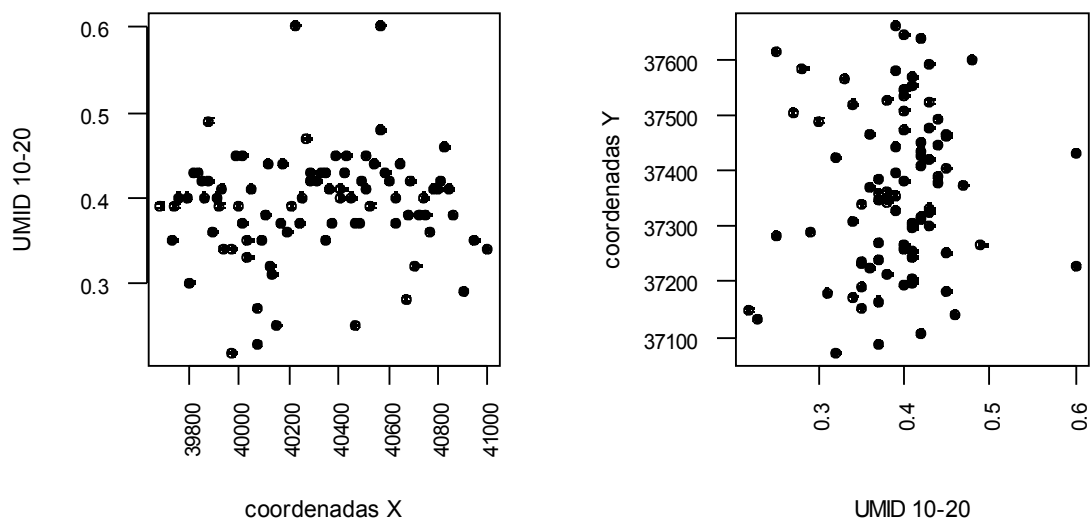


Figura 5A - Gráficos de dispersão para a umidade do solo na profundidade 10 a 20 cm, *versus* linha à esquerda e *versus* coluna à direita.

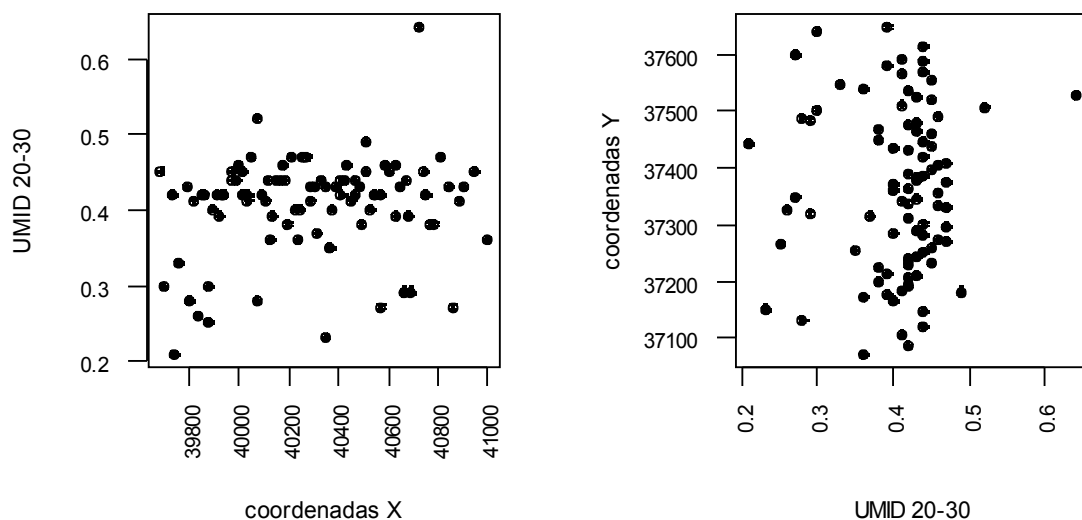


Figura 6A - Gráficos de dispersão para a umidade do solo na profundidade 20 a 30 cm, *versus* linha à esquerda e *versus* coluna à direita.

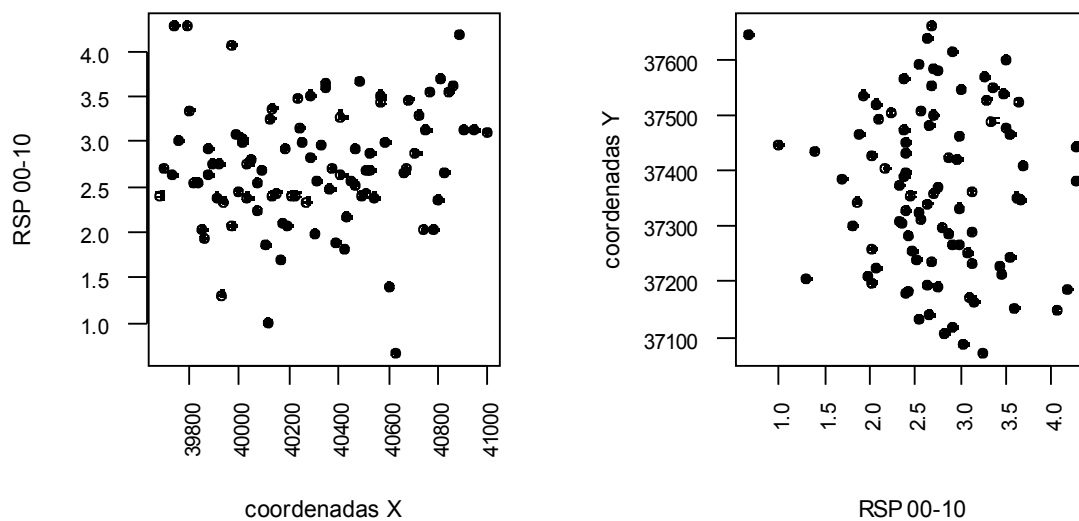


Figura 7A - Gráficos de dispersão para a resistência do solo à penetração na profundidade 0 a 10, cm *versus* linha à esquerda e *versus* coluna à direita.

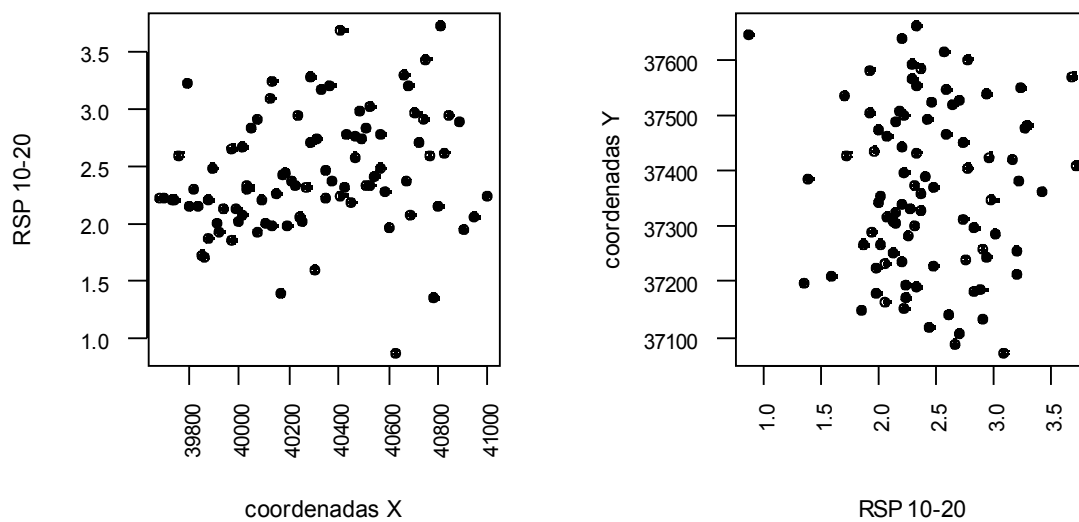


Figura 8A - Gráficos de dispersão para a resistência do solo à penetração na profundidade 10 a 20 cm, *versus* linha à esquerda e *versus* coluna à direita.

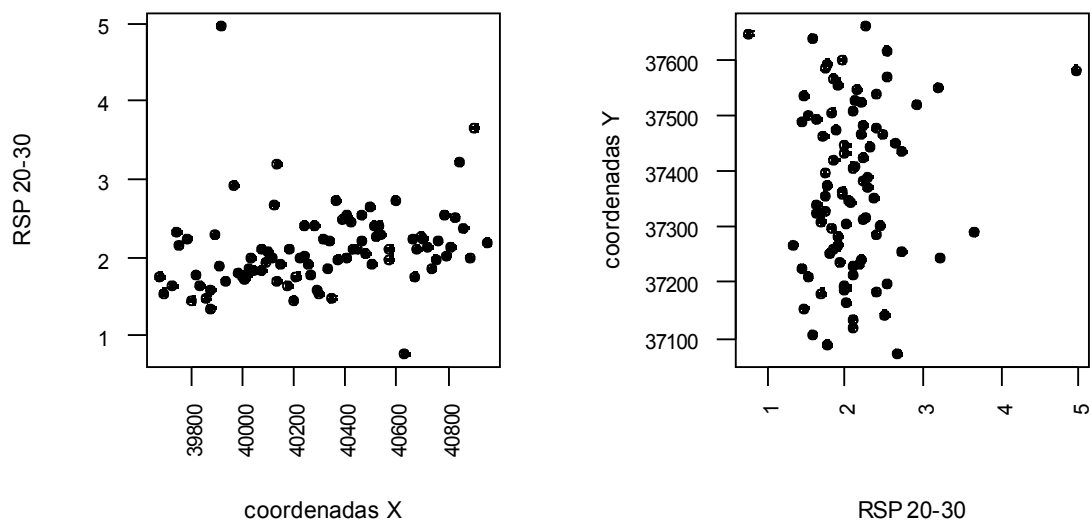


Figura 9A - Gráficos de dispersão para a resistência do solo à penetração na profundidade 20 a 30 cm, *versus* linha à esquerda e *versus* coluna à direita.

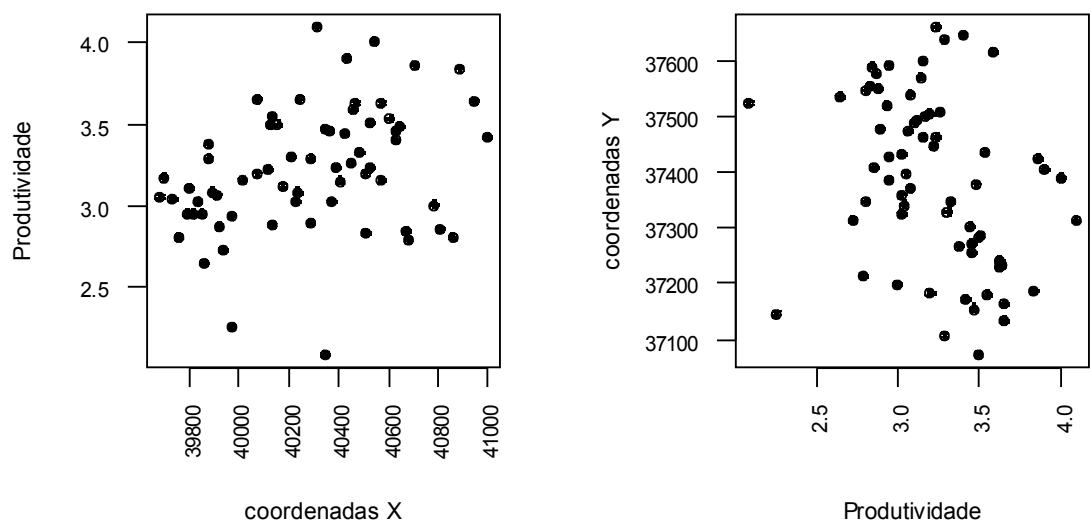


Figura 10A - Gráficos de dispersão para a produtividade da soja *versus* linha à esquerda e *versus* coluna à direita.

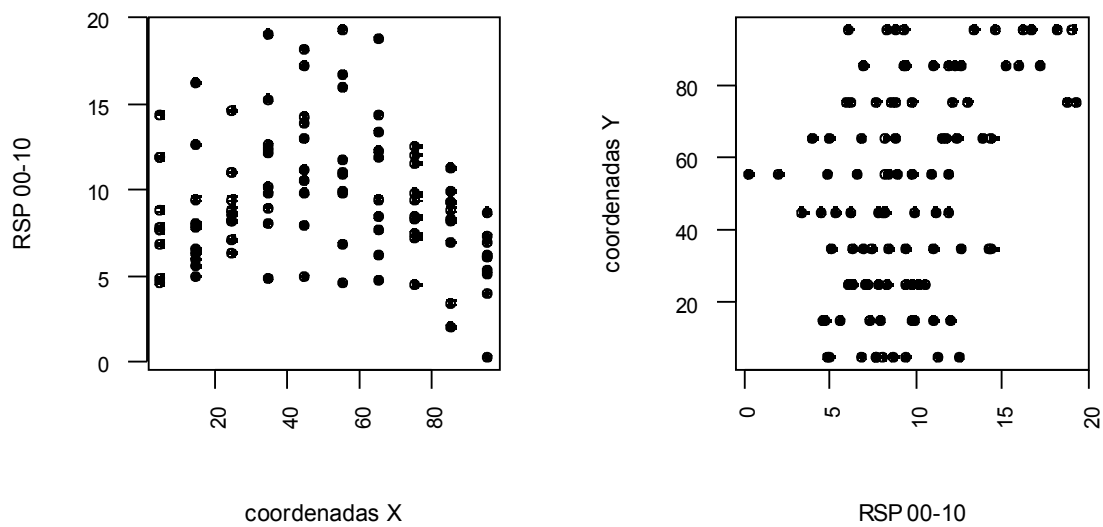


Figura 11A - Gráficos de dispersão para os dados simulados versus linha à esquerda e versus coluna à direita.

APÊNDICE B - SCRIPTS DO R

```

# Análise Exploratória dos dados da DENS 00-10 de
# um experimento realizado em uma área comercial do município de
# Cascavel-PR - Área de 57 ha - 2004/05
# Procedendo a leitura dos dados no formato usado pelo geoR
require(geoR)

arq=read.geodata("M:\DENS 75 00 10.txt", head=F, coords.col=2:3, data.col=4)
bordas=read.table("M:\COORDENADAS DA AREA.txt", head=F)
# Examinando a área de coleta de dados (POST-PLOT)
# O tamanho dos pontos é proporcional ao seu quartil
points(arq,bordas=bordas,pt.div="quartile", xlab="coordenadas XX", ylab="coordenadas
YY")

legend(40800,37000, c("P-1 ", "P-2 ", "P-3 ", "P-4 "), fill =
c("yellow","green","red","blue"))

# Examinando a presença de valores discrepantes.
# Valores discrepantes podem provocar uma assimetria na distribuição
# dos dados.

boxplot(arq$data,main ="DENS 00-10")

# Examinando a forma da distribuição de frequências
# Isto permitirá avaliarmos a simetria e o achatamento da curva
# de distribuição de probabilidade aderente ao histograma
hist(arq$data,ylab="Frequência", xlab="DENS 00-10", main = paste("Histograma de DENS 00-
10"))

# Examinando algumas estatísticas (média, variância, mínimo, ....
summary(arq)
sd(arq$data)
var(arq$data)

# Avaliando a homogeneidade dos dados
# Para CV menores que 30% supomos homogeneidade dos dados
arq.var=var(arq$data)
CV=sqrt(arq.var)/mean(arq$data)*100
CV

# Avaliando a normalidade na distribuição dos dados
# Se p-valor for maior que 5% (0.05) rejeitamos a não-normalidade

```

```

shapiro.test(arq$data)

# Verificação de tendência para a média. Precisamos avaliar se a média
# é constante ou precisaremos eliminar a tendência e trabalharmos com
# o resíduo.

arq.coord=arq$coords

XX=arq.coord[,1]
YY=arq.coord[,2]

par(mfrow=c(1,2))

par(las=2)

par(cex.axis=0.6)

par(mex=1.4)

plot(XX,arq$data,ylab="DENS 00-10",xlab="coordenadas xx")
plot(arq$data,YY,xlab="DENS 00-10",ylab="coordenadas yy")

# fazendo teste de correlação de Pearson

# Aqui avaliamos o comportamento dos dados segundo os eixos ordenados
# Se p-valor for menor do que 5%, dizemos que a correlação é zero
# Se a correlação for nula então supomos a média constante

cor.test(XX,arq$data)

cor.test(arq$data,YY)

# Fazendo uma análise geoestatística exploratória.

# Construindo a nuvem variográfica

# Utilizaremos um cutoff de 70% que corresponde à distancia de 400 m

par(las=2)

par(mex=1.4)

par(mfrow=c(1,3))

par(cex.axis=0.6)

arq.var.cloud=variog(arq,max.dist=400,option="cloud")

plot(arq.var.cloud,ylab="Semivariância", xlab="Distância",cex.main=1, font.main=1,
main="DENS 00-10 Nuvem")

# Construindo o variograma por intervalos

#MATHERON

arq.var.bin=variog(arq,max.dist=400)

plot(arq.var.bin,ylab="Semivariância", xlab="Distância",cex.main=1, font.main=1,
main="DENS 00-10 Matheron")

#CRESSIE

#arq.var.bin=variog(arq,max.dist=400,estimator.type = "modulus")

```

```

#plot(arq.var.bin,ylab="Semivariância", xlab="Distância",cex.main=1, font.main=1,
main="DENS 00-10 Cressie")

#####

par(mfrow=c(1,3))

# Modelo exponencial ajustado por mínimos quadrados ordinários - OLS
plot(arq.var.bin,cex.lab=.9,ylab="Semivariância",xlab="Distância",cex.main=1,font.main=1
,main="Exponencial OLS")
arq.var1.OLS=variofit(arq.var.bin,ini=c(0.0077,47),fix.nugget
=TRUE,nugget=0,cov.model="exp",weight="equal")
arq.var1.OLS
lines(arq.var1.OLS,lty=3)

# Modelo esférico ajustado por mínimos quadrados ordinários - OLS
plot(arq.var.bin,cex.lab=.9,ylab="Semivariância",xlab="Distância",cex.main=1,font.main=1
,main="Esférico OLS")
arq.var3.OLS=variofit(arq.var.bin,ini=c(0.0077,140),fix.nugget=TRUE,nugget=0,cov.model="
sph",weight="equal")
arq.var3.OLS
lines(arq.var3.OLS,lty=3)

# Modelo gaussiano ajustado por mínimos quadrados ordinários - OLS
plot(arq.var.bin,cex.lab=.9,ylab="Semivariância",xlab="Distância",cex.main=1,font.main=1
,main="Gaussiano OLS")
arq.var5.OLS=variofit(arq.var.bin,ini=c(0.0077,80),fix.nugget=TRUE,nugget=0,cov.model="g
au",weight="equal")
arq.var5.OLS
lines(arq.var5.OLS,lty=3)

par(mfrow=c(1,3))

# Modelo exponencial ajustado por mínimos quadrados ponderados - WLS
plot(arq.var.bin,cex.lab=.9,ylab="Semivariância",xlab="Distância",cex.main=1,font.main=1
,main="Exponencial WLS")
arq.var2.WLS=variofit(arq.var.bin,ini=c(0.0057,47),cov.model="exp")
arq.var2.WLS
lines(arq.var2.WLS,lty=3)

# Modelo esférico ajustado por mínimos quadrados ponderados - WLS
plot(arq.var.bin,cex.lab=.9,ylab="Semivariância",xlab="Distância",cex.main=1,font.main=1
,main="Esférico WLS")
arq.var4.WLS=variofit(arq.var.bin,ini=c(0.0057,140),cov.model="sph")
arq.var4.WLS
lines(arq.var4.WLS,lty=3)

# Modelo gaussiano ajustado por mínimos quadrados ponderados - WLS

```

```

plot(arq.var.bin,cex.lab=.9,ylab="Semivariância",xlab="Distância",cex.main=1,font.main=1
,main="Gaussiano WLS")

arq.var6.WLS=variofit(arq.var.bin,ini=c(0.0077,80),fix.nugget=TRUE,nugget=0,
cov.model="gau")

arq.var6.WLS

lines(arq.var6.WLS,lty=3)

summary(arq.var1.OLS)

summary(arq.var2.WLS)

summary(arq.var3.OLS)

summary(arq.var4.WLS)

summary(arq.var5.OLS)

summary(arq.var6.WLS)

#####
# Fazendo a validação cruzada e avaliando a qualidade do
# modelo variográfico escolhido
#####

# Gera um arquivo com a estimação de cada ponto pelo método de
# remover o ponto e estima-lo através do modelo adotado, repetindo
# o processo para todos os pontos

#####

arq.cross=xvalid(arq,model=arq.var1.OLS)

summary(arq.cross)

#####

r=dim(6)

# ERRO ORDENADO

EO=sort(arq.cross$std.error)

n=length(arq.cross$std.error)

m=dim(n)

m[n]=(0.5)^(1/n)

m[2:(n-1)]=(2:(n-1))-0.3175)/(n+0.365)

m[1]=1-m[n]

M=qnorm(m)

M2=M^2

MX=M*EO

SEQ=sum((EO-mean(EO))^2)

r[1]=sum(MX)/sqrt(sum(M2)*SEQ)

```

```
#####
arq.cross=xvalid(arq,model=arq.var2.WLS)
summary(arq.cross)
#####
# ERRO ORDENADO
EO=sort(arq.cross$std.error)
n=length(arq.cross$std.error)
m=dim(n)
m[n]=(0.5)^(1/n)
m[2:(n-1)]=((2:(n-1))-0.3175)/(n+0.365)
m[1]=1-m[n]
M=qnorm(m)
M2=M^2
MX=M*EO
SEQ=sum((EO-mean(EO))^2)
r[2]=sum(MX)/sqrt(sum(M2)*SEQ)
#####
arq.cross=xvalid(arq,model=arq.var3.OLS)
summary(arq.cross)
#####
# ERRO ORDENADO
EO=sort(arq.cross$std.error)
n=length(arq.cross$std.error)
m=dim(n)
m[n]=(0.5)^(1/n)
m[2:(n-1)]=((2:(n-1))-0.3175)/(n+0.365)
m[1]=1-m[n]
M=qnorm(m)
M2=M^2
MX=M*EO
SEQ=sum((EO-mean(EO))^2)
r[3]=sum(MX)/sqrt(sum(M2)*SEQ)
#####
arq.cross=xvalid(arq,model=arq.var4.WLS)
summary(arq.cross)
#####
# ERRO ORDENADO
EO=sort(arq.cross$std.error)
n=length(arq.cross$std.error)
```

```

m=dim(n)
m[n]=(0.5)^(1/n)
m[2:(n-1)]=((2:(n-1))-0.3175)/(n+0.365)
m[1]=1-m[n]
M=qnorm(m)
M2=M^2
MX=M*EO
SEQ=sum((EO-mean(EO))^2)
r[4]=sum(MX)/sqrt(sum(M2)*SEQ)
#####
arq.cross=xvalid(arq,model=arq.var5.OLS)
summary(arq.cross)
#####
# ERRO ORDENADO
EO=sort(arq.cross$std.error)
n=length(arq.cross$std.error)
m=dim(n)
m[n]=(0.5)^(1/n)
m[2:(n-1)]=((2:(n-1))-0.3175)/(n+0.365)
m[1]=1-m[n]
M=qnorm(m)
M2=M^2
MX=M*EO
SEQ=sum((EO-mean(EO))^2)
r[5]=sum(MX)/sqrt(sum(M2)*SEQ)
#####
arq.cross=xvalid(arq,model=arq.var6.WLS)
summary(arq.cross)
#####
# ERRO ORDENADO
EO=sort(arq.cross$std.error)
n=length(arq.cross$std.error)
m=dim(n)
m[n]=(0.5)^(1/n)
m[2:(n-1)]=((2:(n-1))-0.3175)/(n+0.365)
m[1]=1-m[n]
M=qnorm(m)
M2=M^2
MX=M*EO

```



```

SEQ=sum((EO-mean(EO))^2)
r[6]=sum(MX)/sqrt(sum(M2)*SEQ)
#####

# Análise Exploratória dos dados SIMULADOS
#
# Procedendo a leitura dos dados no formato usado pelo geoR
require(geoR)
area=expand.grid((0:9)*10+5,(0:9)*10+5)
plot(area,cex=0.5,pch=19)
set.seed(875) # fixa uma semente para o gerador de nrs aleatórios
arq=grf(200,grid=area,cov.pars=c(15,60),nugget=0,cov.model="sph")
arq$data=arq$data +10
# Examinando a área de coleta de dados (POST-PLOT)
# O tamanho dos pontos é proporcional ao seu quartil
points(arq,area,pch=21,col=c("yellow","green","red","blue"),pt.div="quartile",xlab="coordenadas XX",ylab="coordenadas YY")
#legend(105, 30,c("1° QUARTIL ","2° QUARTIL ","3° QUARTIL ","4° QUARTIL"),fill=c("yellow","green","red","blue"))

#####
#####
par(las=1)
par(mfrow=c(2,3))
#####
#####

# Modelo exponencial ajustado por mínimos quadrados ordinários - OLS
plot(arq.var.bin,cex.lab=.9, cex.axis=.8, font.axis=1, ylab="Semivariância",
xlab="Distância",cex.main=1, font.main=1, main="Exponencial OLS")
arq.var1.OLS=variofit(arq.var.bin,ini=c(16,17),fix.nug=TRUE,nugget=.00,cov.model="exp",weight="equal")
arq.var1.OLS
lines(arq.var1.OLS,lty=1)

# Modelo esférico ajustado por mínimos quadrados ordinários - OLS

```

```

plot(arq.var.bin,cex.lab=.9, cex.axis=.8, font.axis=1, ylab="Semivariância",
xlab="Distância",cex.main=1, font.main=1, main="Esférico OLS")

arq.var3.OLS=variofit(arq.var.bin,ini=c(15,50),cov.model="sph",weight="equal")

arq.var3.OLS

lines(arq.var3.OLS,lty=3)

# Modelo gaussiano ajustado por mínimos quadrados ordinários - OLS
plot(arq.var.bin,cex.lab=.9, cex.axis=.8, font.axis=1, ylab="Semivariância",
xlab="Distância",cex.main=1, font.main=1, main="Gaussiano OLS")

arq.var5.OLS=variofit(arq.var.bin,ini=c(14,29),cov.model="gau",weight="equal")

arq.var5.OLS

lines(arq.var5.OLS,lty=5)

# Modelo exponencial ajustado por mínimos quadrados ponderados - WLS
plot(arq.var.bin,cex.lab=.9, cex.axis=.8, font.axis=1, ylab="Semivariância",
xlab="Distância",cex.main=1, font.main=1, main="Exponencial WLS")

arq.var2.WLS=variofit(arq.var.bin,ini=c(16,17), cov.model="exp")

arq.var2.WLS

lines(arq.var2.WLS,lty=2)

# Modelo esférico ajustado por mínimos quadrados ponderados - WLS
plot(arq.var.bin,cex.lab=.9, cex.axis=.8, font.axis=1, ylab="Semivariância",
xlab="Distância",cex.main=1, font.main=1, main="Esférico WLS")

arq.var4.WLS=variofit(arq.var.bin,ini=c(15,50), cov.model="sph")

arq.var4.WLS

lines(arq.var4.WLS,lty=4)

# Modelo gaussiano ajustado por mínimos quadrados ponderados - WLS
plot(arq.var.bin,cex.lab=.9, cex.axis=.8, font.axis=1, ylab="Semivariância",
xlab="Distância",cex.main=1, font.main=1, main="Gaussiano WLS")

arq.var6.WLS=variofit(arq.var.bin,ini=c(14,29), cov.model="gau")

arq.var6.WLS

lines(arq.var6.WLS,lty=6)

#####

# Fazendo a validação cruzada e avaliando a qualidade do
# modelo variográfico escolhido

#####

```

```
# Gera um arquivo com a estimaco de cada ponto pelo mtodo de  
# remover o ponto e estima-lo atravs do modelo adotado  
  
arq.cross=xvalid(arq,model=arq.var1.OLS)  
summary(arq.cross)  
arq.cross=xvalid(arq,model=arq.var2.WLS)  
summary(arq.cross)  
arq.cross=xvalid(arq,model=arq.var3.OLS)  
summary(arq.cross)  
arq.cross=xvalid(arq,model=arq.var4.WLS)  
summary(arq.cross)  
arq.cross=xvalid(arq,model=arq.var5.OLS)  
summary(arq.cross)  
arq.cross=xvalid(arq,model=arq.var6.WLS)  
summary(arq.cross)
```

APÊNDICE C - DADOS SIMULADOS

X	Y	Z	X	Y	Z
1	1	7,7070	1	6	4,8449
2	1	8,0912	2	6	6,6264
3	1	9,3635	3	6	8,1174
4	1	4,7724	4	6	8,8561
5	1	4,9621	5	6	9,7459
6	1	6,7884	6	6	10,9496
7	1	7,6382	7	6	11,8471
8	1	12,5166	8	6	8,4509
9	1	11,2669	9	6	2,0349
10	1	8,6851	10	6	0,3044
1	2	4,6098	1	7	6,7731
2	2	5,5612	2	7	4,9332
3	2	11,0045	3	7	8,8311
4	2	9,8039	4	7	12,3621
5	2	7,9139	5	7	13,8352
6	2	4,5330	6	7	11,7706
7	2	4,6771	7	7	14,3338
8	2	12,0333	8	7	11,5268
9	2	9,9373	9	7	8,2354
10	2	7,2615	10	7	3,9079
1	3	7,7558	1	8	7,6331
2	3	6,3571	2	8	5,9546
3	3	7,0495	3	8	8,5614
4	3	10,1797	4	8	12,1910
5	3	10,4661	5	8	12,9630
6	3	9,8186	6	8	19,2642
7	3	9,4039	7	8	18,7890
8	3	7,2330	8	8	9,8414
9	3	8,3014	9	8	8,7843
10	3	6,0740	10	8	6,2544
1	4	14,3382	1	9	11,9370
2	4	9,4114	2	9	12,6242
3	4	6,3636	3	9	11,0485
4	4	12,6799	4	9	15,1762
5	4	14,2142	5	9	17,2204
6	4	10,9985	6	9	15,9535
7	4	8,4133	7	9	12,2270
8	4	7,4272	8	9	9,3776
9	4	6,9526	9	9	9,2426
10	4	5,0566	10	9	6,8890
1	5	11,9003	1	10	8,7551
2	5	7,8107	2	10	16,1651
3	5	8,1907	3	10	14,6573
4	5	8,0289	4	10	19,0970
5	5	11,0887	5	10	18,2526
6	5	9,9123	6	10	16,7200
7	5	6,2384	7	10	13,3486
8	5	4,4502	8	10	8,2893
9	5	3,3260	9	10	9,2709
10	5	5,3455	10	10	6,0557

ANEXO

ANEXO - TABELA DO COEFICIENTE DE CORRELAÇÃO R

n	.00 0	.00 5	.01 0	.02 5	.05 0	.75 0	.90 0	.950	.975	.990	.995
3	.86 6	.86 7	.86 9	.87 2	.87 9	.99 1	.99 9	1,00 0	1,00 0	1,00 0	1,00 0
4	.78 4	.81 3	.82 2	.84 5	.86 8	.97 9	.99 2	.996	.998	.999	1,00 0
5	.72 6	.80 3	.82 2	.85 5	.87 9	.97 7	.98 8	.992	.995	.997	.998
6	.68 3	.81 8	.83 5	.86 8	.89 0	.97 7	.98 6	.990	.993	.996	.997
7	.64 8	.82 8	.84 7	.87 6	.89 9	.97 8	.98 6	.990	.992	.995	.996
8	.61 9	.84 1	.85 9	.88 6	.90 5	.97 9	.98 6	.990	.992	.995	.996
9	.59 5	.85 1	.86 8	.89 3	.91 2	.98 0	.98 7	.990	.992	.994	.995
10	.57 4	.86 0	.87 6	.90 0	.91 7	.98 1	.98 7	.990	.992	.994	.995
11	.55 6	.86 8	.88 3	.90 6	.92 2	.98 2	.98 8	.990	.992	.994	.995
12	.53 9	.87 5	.88 9	.91 2	.92 6	.98 2	.98 8	.990	.992	.994	.995
13	.52 5	.88 2	.89 5	.91 7	.93 1	.98 3	.98 8	.991	.993	.994	.995
14	.51 2	.88 8	.90 1	.92 1	.93 4	.98 4	.98 9	.991	.993	.994	.995
15	.50 0	.89 4	.90 7	.92 5	.93 7	.98 4	.98 9	.991	.993	.994	.995
16	.48 9	.89 9	.91 2	.92 8	.94 0	.98 5	.98 9	.991	.993	.994	.995
17	.47 8	.90 3	.91 6	.93 1	.94 2	.98 6	.99 0	.992	.993	.994	.995
18	.46 9	.90 7	.91 9	.93 4	.94 5	.98 6	.99 0	.992	.993	.995	.995
19	.46 0	.90 9	.92 3	.93 7	.94 7	.98 7	.99 0	.992	.993	.995	.995
20	.45 2	.91 2	.92 5	.93 9	.95 0	.98 7	.99 1	.992	.994	.995	.995
21	.44 5	.91 4	.92 8	.94 2	.95 2	.98 7	.99 1	.993	.994	.995	.996
22	.43 7	.91 8	.93 0	.94 4	.95 4	.98 8	.99 1	.993	.994	.995	.996
23	.43 1	.92 2	.93 3	.94 7	.95 5	.98 8	.99 1	.993	.994	.995	.996
24	.42 4	.92 6	.93 6	.94 9	.95 7	.98 8	.99 2	.993	.994	.995	.996
25	.41 8	.92 8	.93 7	.95 9	.95 8	.98 9	.99 2	.993	.994	.995	.996
26	.41 2	.93 0	.93 9	.95 2	.95 9	.98 9	.99 2	.993	.994	.995	.996
27	.40 7	.93 2	.94 1	.95 3	.96 0	.98 9	.99 2	.994	.995	.995	.996
28	.40 2	.93 4	.94 3	.95 5	.96 2	.98 0	.99 2	.994	.995	.995	.996
29	.39 7	.93 7	.94 5	.95 6	.96 2	.98 0	.99 2	.994	.995	.995	.996
30	.39 2	.93 8	.94 7	.95 7	.96 4	.98 0	.99 3	.994	.995	.996	.996
31	.38 8	.93 9	.94 8	.95 8	.96 5	.98 0	.99 3	.994	.995	.996	.996

32	.38	.93	.94	.95	.96	.99	.99				
	3	9	9	9	6	0	3	.994	.995	.996	.996
33	.37	.94	.95	.96	.96	.99	.99				
	9	0	0	0	7	1	3	.994	.995	.996	.996
34	.37	.94	.95	.96	.96	.99	.99				
	5	1	1	0	7	1	3	.994	.995	.996	.996
35	.37	.94	.95	.96	.96	.99	.99				
	1	3	2	1	8	1	3	.995	.995	.996	.997
36	.36	.94	.95	.96	.96	.99	.99				
	7	5	3	2	8	1	4	.995	.996	.996	.997
37	.36	.94	.95	.96	.96	.99	.99				
	4	7	5	2	9	1	4	.995	.996	.996	.997
38	.36	.94	.95	.96	.97	.99	.99				
	0	8	6	4	0	2	4	.995	.996	.996	.997
39	.35	.94	.95	.96	.97	.99	.99				
	7	9	7	5	1	2	4	.995	.996	.996	.997
40	.35	.94	.95	.96	.97	.99	.99				
	4	9	8	6	2	2	4	.995	.996	.996	.997
41	.35	.95	.95	.96	.97	.99	.99				
	1	0	8	7	2	2	4	.995	.996	.996	.997
42	.34	.95	.95	.96	.97	.99	.99				
	8	1	9	7	3	2	4	.995	.996	.997	.997
43	.34	.95	.95	.96	.97	.99	.99				
	5	3	9	7	3	2	4	.995	.996	.997	.997
44	.34	.95	.96	.96	.97	.99	.99				
	2	4	0	8	3	2	4	.995	.996	.997	.997
45	.33	.95	.96	.96	.97	.99	.99				
	9	5	1	9	4	3	4	.995	.996	.997	.997
46	.33	.95	.96	.96	.97	.99	.99				
	6	6	2	9	4	3	5	.995	.996	.997	.997
47	.33	.95	.96	.97	.97	.99	.99				
	4	6	3	0	4	3	5	.995	.996	.997	.997
48	.33	.95	.96	.97	.97	.99	.99				
	1	7	3	0	5	3	5	.996	.996	.997	.997
49	.32	.95	.96	.97	.97	.99	.99				
	9	7	4	1	5	3	5	.996	.996	.997	.997
50	.32	.95	.96	.97	.97	.99	.99				
	6	9	5	2	7	3	5	.996	.996	.997	.997
55	.31	.96	.96	.97	.97	.99	.99				
	5	2	7	4	8	4	5	.996	.997	.997	.997
60	.30	.96	.97	.97	.98	.99	.99				
	5	5	0	6	0	4	5	.996	.997	.997	.988
65	.29	.96	.97	.97	.98	.99	.99				
	6	7	2	7	1	4	6	.996	.997	.997	.988
70	.28	.96	.97	.97	.98	.99	.99				
	8	9	4	8	2	5	6	.997	.997	.998	.988
75	.28	.97	.97	.97	.98	.99	.99				
	1	1	5	9	3	5	6	.997	.997	.998	.988
80	.27	.97	.97	.98	.98	.99	.99				
	4	3	6	0	4	5	6	.997	.997	.998	.988
85	.26	.97	.97	.98	.98	.99	.99				
	8	4	7	1	5	5	7	.997	.997	.998	.988
90	.26	.97	.97	.98	.98	.99	.99				
	3	6	8	2	5	6	7	.997	.998	.998	.988
95	.25	.97	.97	.98	.98	.99	.99				
	7	7	9	3	6	6	7	.997	.998	.998	.988
100	.25	.97	.98	.98	.98	.99	.99				
	2	9	1	4	7	6	7	.998	.998	.998	.988

Fonte: FILLIBEN (1975).