

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

**AUTOCORRELAÇÃO ESPACIAL DIRECIONAL PARA ANÁLISE DA ANISOTROPIA COM
DADOS AGRÍCOLAS**

Dyogo Lesniewski Ribeiro

CASCAVEL – Paraná – Brasil

2017

DYOGO LESNIEWSKI RIBEIRO

**AUTOCORRELAÇÃO ESPACIAL DIRECIONAL PARA ANÁLISE DA ANISOTROPIA COM
DADOS AGRÍCOLAS**

Dissertação apresentada ao programa de Pós-Graduação em Engenharia Agrícola em cumprimento parcial aos requisitos para obtenção ao título de Mestre em Engenharia Agrícola, área de concentração em Engenharia de Sistemas Agroindustriais.

**Orientadora: Dra. Luciana Pagliosa
Carvalho Guedes.**

CASCADEL – Paraná – Brasil

SETEMBRO 2017

Dados Internacionais de Catalogação-na-Publicação (CIP)
(Sistema de Bibliotecas – UNIOESTE – Campus Cascavel)

Ribeiro, Dyogo Lesniewski.

R368a Autocorrelação espacial direcional para análise da anisotropia com dados agrícolas / Dyogo Lesniewski Ribeiro. --- Cascavel (PR), 2017.

58 f.:il.

Orientadora: Dra. Luciana Pagliosa Carvalho Guedes.

Dissertação (Mestrado) – Universidade Estadual do Oeste do Paraná, Campus de Cascavel, 2017.

Programa de Pós-Graduação em Engenharia Agrícola, área de concentração em Engenharia de Sistemas Agroindustriais.
Inclui Bibliografia

1. Agricultura de precisão. 2. Química do solo-Qualidade. 3. Análise espacial (Estatística). 4. Moran, Índice direcional de. I. Guedes, Luciana Pagliosa Carvalho. II. Universidade Estadual do Oeste do Paraná. III. Título.

Rosângela A. A. Silva – CRB 9ª/1810

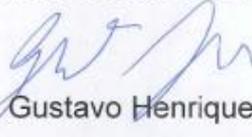
DYOGO LESNIEWSKI RIBEIRO

Autocorrelação Espacial Direcional para Análise da Anisotropia com Dados Agrícolas

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Agrícola em cumprimento parcial aos requisitos para obtenção do título de Mestre em Engenharia Agrícola, área de concentração Sistemas Biológicos e Agroindustriais, linha de pesquisa Geoprocessamento, Estatística Espacial e Agricultura de Precisão, APROVADO(A) pela seguinte banca examinadora:



Orientador(a) – Profa. Dra. Luciana Pagliosa Carvalho Guedes
Universidade Estadual do Oeste do Paraná - Campus de Cascavel (UNIOESTE)



Prof. Dr. Gustavo Henrique Dalposso
Universidade Tecnológica Federal do Paraná (UTFPR)



Prof. Dr. Miguel Angel Uribe Opazo
Universidade Estadual do Oeste do Paraná - Campus de Cascavel (UNIOESTE)

Cascavel, 15 de setembro de 2017

BIOGRAFIA

Dyogo Lesniewski Ribeiro, nascido em Boa Vista da Aparecida/PR, em 30 de dezembro de 1989, graduado em Matemática na Universidade Paranaense - UNIPAR (2014). Possui Especializações em Educação do Campo e Educação de jovens e adultos nas Faculdades integradas do Vale do Ivaí - UNIVALE (2014). Desde 2015 é discente em nível de Mestrado no Programa de Pós-Graduação em Engenharia Agrícola, área de concentração Engenharia de Sistemas Agroindustriais, linha de pesquisa Estatística Espacial na UNIOESTE - Universidade Estadual do Oeste do Paraná / Campus Cascavel. Atualmente exerce a docência em matemática no ensino regular estadual.

Tentar e falhar é, pelo menos, aprender.
Não chegar a tentar é sofrer a inestimável
perda do que poderia ter sido.

(Geraldo Eustáquio)

AGRADECIMENTOS

Primeiramente, agradeço a Deus por tudo e mais essa benção conquistada em minha vida.

Agradeço aos meus pais, que iluminam o caminho de minha vida e também a meus familiares e amigos, a quem espero fazer o mesmo.

Agradeço também a minha orientadora Luciana, grande responsável pelo meu sucesso nesse estudo.

Aos amigos do laboratório de estatística aplicada.

Ao programa de Pós-graduação em Engenharia Agrícola, seu corpo docente, direção e administração.

À CAPES pelo apoio financeiro.

E ao senhor Agassiz Linhares neto, pela parceria na implantação dos experimentos em sua propriedade.

AUTOCORRELAÇÃO ESPACIAL DIRECIONAL PARA ANÁLISE DA ANISOTROPIA COM DADOS AGRÍCOLAS

RESUMO

As técnicas de geoestatística contribuem para o entendimento das características da área em estudo, facilitam as tomadas de decisões em relação ao gerenciamento do sistema de produção agrícola e contribuem para o desenvolvimento sustentável em agricultura de precisão. A anisotropia é uma característica que influencia na precisão dos mapas temáticos que representam a variabilidade espacial do fenômeno estudado. Assim, esse trabalho tem por escopo utilizar o índice de Moran direcional na análise de anisotropia em variáveis georreferenciadas. O índice de Moran direcional foi calculado considerando modelos geoestatístico isotrópicos e anisotrópicos, com o intuito de evidenciar a diferença direcional que existe nos mapas temáticos quando se incorpora ou não a anisotropia no modelo geoestatístico. Para isso, foram utilizados dados simulados a partir de uma configuração amostral irregular, com cem pontos. Os dados foram simulados com uma estrutura de dependência espacial anisotrópica (geométrica) de acordo com um modelo exponencial, com ângulo de maior continuidade espacial igual a 90° (azimute) e variação do fator de anisotropia. O índice de Moran direcional foi calculado para os valores amostrais dos dados simulados, como ferramenta de auxílio na tomada de decisão quanto à existência de anisotropia. Posteriormente, esse processo também foi utilizado para os atributos químicos do solo observados em uma área agrícola com plantação de soja, referente ao ano agrícola de 2014/2015. A autocorrelação espacial direcional se apresentou eficaz para os dados simulados e os atributos químicos do solo, quanto à identificação da anisotropia geométrica e também para evidenciar a diferença direcional que existe nos mapas temáticos, quando se considera (ou não) a existência de anisotropia no modelo geoestatístico.

Palavras-chave: agricultura de precisão; atributos químicos do solo; dependência espacial; geoestatística; índice de Moran direcional.

DIRECTIONAL SPATIAL AUTOCORRELATION FOR ANISOTROPIC ANALYSIS WITH AGRICULTURAL DATA

ABSTRACT: Geostatistical techniques have contributed on acquainting the studied area characteristics. They have made the decisions easier to be taken regarding the management of the agricultural yield system and contributed to sustainable development in precision agriculture. Anisotropy is a characteristic that has influenced the precision of thematic maps that represent spatial variability of the studied phenomenon. Thus, this trial aimed at using Moran directional index in anisotropy analysis in georeferenced variables. Moran directional index was calculated considering isotropic and anisotropic geostatistical models to highlight the directional difference in thematic maps when anisotropy is incorporated or not in the geostatistical model. Thus, simulated data were used considering an irregular sample configuration, with 100 points. Data were simulated with an anisotropic (geometric) spatial dependence structure following an exponential model, with an angle of greater spatial continuity equal to 90° (azimuth) and varying the anisotropy factor. Moran directional index was calculated for sampled values of simulated data, as a tool to assist in decision making regarding the existence of anisotropy. Then, this process was also used for soil chemical attributes, observed in an agricultural area with soybean cropping, referring to the agricultural year of 2014/2015. The directional spatial autocorrelation was effective in identifying geometric anisotropy for simulated data and soil chemical attributes. It also highlighted the directional difference among the thematic maps, when the existence of anisotropy is considered or not in the geostatistical model.

KEYWORDS: precision agriculture; soil chemical attributes; spatial dependence; geostatistics.

SUMÁRIO

LISTA DE TABELAS	VIII
LISTA DE FIGURAS	IX
1 INTRODUÇÃO	10
2 OBJETIVOS	11
2.1 Objetivo geral	11
2.2 Objetivos específicos	11
3 REVISÃO DE LITERATURA	12
3.1 Geoestatística	12
3.2 Hipóteses de estacionaridade e Função Semivariância	14
3.3 Semivariograma experimental	16
3.3.1 Parâmetros do modelo teórico da função semivariância	18
3.4 Isotropia e Anisotropia	19
3.5 Autocorrelação espacial direcional	20
3.6 Modelos Teóricos	22
3.7 Métodos de Ajuste ou estimação de parâmetros	24
3.7.1 Estimação de parâmetros por máxima verossimilhança - MV	24
3.8 Critérios de escolha do melhor modelo	25
3.8.1 Validação cruzada	25
3.8.2 Critério de informação de Akaike e Bayesiano	26
3.9 Krigagem	27
3.9.1 Krigagem ordinária	27
3.10 Comparação de mapas temáticos	28
4 MATERIAL E MÉTODOS	30
4.1 Simulações de processos espaciais gaussianos	30
4.2 Dados agrícolas	31
4.2.1 Caracterização da área e análises geoestatísticas	31
5 RESULTADOS	33
5.1 Estudos das simulações	33
5.2 Estudo prático	40
5.2.1 Análise descritiva e espacial	40
6 CONCLUSÃO	54
7 REFERÊNCIAS	55
8 APÊNDICE	58

LISTA DE TABELAS

Tabela 1 Representação genérica de uma matriz de erros.....	29
Tabela 2 Classificação do índice Kappa.....	29
Tabela 3 Porcentagem de simulações com valores significativos em cada distância de corte, nas simulações com modelo anisotrópico em 90° azimute, para os dados amostrais.	33
Tabela 4 Porcentagem de simulações com valores significativos em cada distância de corte, nas simulações com modelo anisotrópico em 90° azimute, para os valores preditos.	36
Tabela 5 Porcentagem de simulações com valores significativos em cada distância de corte, nas simulações com modelo anisotrópico em 90° azimute, para os dados amostrais.	37
Tabela 6 Porcentagem de simulações com valores significativos em cada distância de corte, nas simulações com modelo anisotrópico em 90° azimute, para os valores preditos.	39
Tabela 7 Análise exploratória descritiva dos atributos químicos solo: (a) Zinco Zn (mg dm^{-3}), (b) Ferro Fe (mg dm^{-3}), (c) Manganês Mn (mg dm^{-3}), (d) Fósforo P (mg dm^{-3}), (e) Carbono C (g dm^{-3}), (f) Cálcio Ca (cmolc dm^{-3}), (g) Alumínio Al (cmolc dm^{-3}) e (h) Potássio K (cmolc dm^{-3}).	41
Tabela 8 Índice de Moran direcional para os dados amostrais Carbono (C).....	44
Tabela 9 Índice de Moran direcional para os dados amostrais Cálcio (Ca) e Potássio (K). ...	45
Tabela 10 Valores estimados dos parâmetros dos modelos geoestatísticos isotrópicos para os atributos químicos do solo Zn, Fe, Al, P e Mn e anisotrópicos para os atributos C, Ca e K, considerando como 90° a direção de anisotropia geométrica.	49
Tabela 11 Medidas de acurácia para comparação dos mapas temáticos considerando modelos estimados anisotrópicos e isotrópicos para os atributos químicos do solo carbono, cálcio e potássio.	52
Tabela 12 Valor estimado do índice de Moran direcional para os valores preditos considerando modelos estimados anisotrópicos e isotrópicos para os atributos químicos do solo carbono, cálcio e potássio.	53

LISTA DE FIGURAS

Figura 1 Relação entre as funções covariância e a semivariância.	15
Figura 2 Gráfico semivariograma experimental omnidirecional.	17
Figura 3 Tolerância angular em relação ao vetor de distância h.	17
Figura 4 Modelo teórico ajustado ao semivariograma experimental.	18
Figura 5 Área de estudo e localização dos pontos amostrais.	31
Figura 6 Boxplot I de Moran direcional calculado com os valores amostrados das simulações, considerando (F_a, α) : (a) (1, 0°); (b) (1, 90°); (c) (2, 0°); (d) (2, 90°); (e) (3, 0°); (f) (3, 90°); (g) (4, 0°) e (h) (4, 90°).	34
Figura 7 Boxplot I de Moran direcional calculado com os valores preditos das simulações, considerando (F_a, α) : (a) (1, 0°); (b) (1, 90°); (c) (2, 0°); (d) (2, 90°); (e) (3, 0°); (f) (3, 90°); (g) (4, 0°) e (h) (4, 90°).	36
Figura 8 Boxplot I de Moran direcional calculado com os valores amostrados das simulações, considerando (F_a, α) : (a) (1, 0°); (b) (1, 90°); (c) (2, 0°); (d) (2, 90°); (e) (3, 0°); (f) (3, 90°); (g) (4, 0°) e (h) (4, 90°).	38
Figura 9 Boxplot I de Moran direcional calculado com os valores preditos das simulações, considerando (F_a, α) : (a) (1, 0°); (b) (1,90°); (c) (2, 0°); (d) (2, 90°); (e) (3, 0°); (f) (3, 90°); (g) (4, 0°) e (h) (4, 90°).	40
Figura 10 Gráfico boxplot dos atributos químicos do solo: (a) Zinco Zn (mg dm^{-3}), (b) Ferro Fe (mg dm^{-3}), (c) Manganês Mn (mg dm^{-3}), (d) Fósforo P (mg dm^{-3}), (e) Carbono C (g dm^{-3}), (f) Cálcio Ca (cmolc dm^{-3}), (g) Alumínio Al (cmolc dm^{-3}) e (h) Potássio K (cmolc dm^{-3}).	42
Figura 11 Gráficos post-plot dos atributos químicos do solo: (a) Zinco Zn, (b) Ferro Fe, (c) Manganês, (d) Fósforo P, (e) Carbono C, (f) Cálcio Ca, (g) Alumínio Al e (h) Potássio K. Círculo vermelho indica os pontos discrepantes identificados pelo gráfico box plot.	43
Figura 12 Gráfico de envelopes das variáveis: (a) Zinco Zn (mg dm^{-3}), (b) Ferro Fe (mg dm^{-3}), (c) Manganês Mn (mg dm^{-3}), (d) Fósforo P (mg dm^{-3}), (e) Carbono C (g dm^{-3}), (f) Cálcio Ca (cmolc dm^{-3}), (g) Alumínio Al (cmolc dm^{-3}) e (h) Potássio K (cmolc dm^{-3}).	46
Figura 13 Semivariograma direcionais dos seguintes atributos químicos do solo: (a) e (b) Zinco, (c) e (d) Ferro, (e) e (f) Alumínio e (g) e (h) Fósforo.	47
Figura 14 Semivariograma direcionais dos seguintes atributos químicos do solo: (a) e (b) Manganês, (c) e (d) Carbono, (e) e (f) Cálcio e (g) e (h) Potássio.	48
Figura 15 Mapa temático das variáveis: (a) Zinco (mg dm^{-3}), (b) Ferro (mg dm^{-3}), (c) Alumínio (cmolc dm^{-3}), (d) Fósforo (mg dm^{-3}), (e) Manganês (mg dm^{-3}), considerando o modelo isotrópico estimado.	50
Figura 16 Mapa temático dos seguintes atributos químicos do solo: considerando o modelo geoestatístico estimado isotrópico (a) Carbono, (c) Cálcio, (e) Potássio, e considerando um modelo geoestatístico estimado anisotrópico (b) Carbono, (d) Cálcio, (f) Potássio.	51

1 INTRODUÇÃO

A agricultura de precisão é fundamentada na coleta e no processamento de dados espacialmente referenciados, cujo objetivo central é a obtenção de informação que conduza a maximizar a rentabilidade a partir do manejo localizado das culturas. Aperfeiçoar a rentabilidade está vinculado a dois pontos cruciais e complementares na agricultura de precisão, quais sejam: o aumento da produtividade associada à melhor qualidade da cultura e à redução do custeio associado a uma aplicação localizada de insumos agrícolas. Tais manejos reduzem os impactos ambientais causados pela atividade agrícola.

Uma das maneiras para o processamento de dados espacialmente referenciados é a partir de técnicas de geoestatística. A geoestatística é um conjunto de ferramentas estatísticas aplicadas a amostras espacialmente referenciadas, quando se considera que cada amostra coletada carrega consigo informações sobre sua vizinhança. Assim, é possível descrever a continuidade espacial da variável de interesse em toda área de estudo para a coleta de apenas algumas amostras.

Na aplicação das técnicas de geoestatística, é possível construir semivariograma para determinada direção (semivariogramas direcionais), que permite uma análise preliminar quanto à existência e ao tipo de anisotropia no fenômeno estudado (GUEDES et al., 2008). Os semivariogramas direcionais podem apresentar comportamento semelhante em todas as direções. Esse comportamento é dito isotrópico e é representado por um único semivariograma que contém informações de todas as direções (semivariograma omnidirecional). Posteriormente, seguem-se as etapas de: ajuste do modelo teórico, que descreve a semivariância em função de distância bem como realiza-se a predição de variável georreferenciada em locais não amostrados. Assim, é possível a construção do mapa temático da continuidade espacial do fenômeno em estudo.

Identificada a presença de anisotropia no fenômeno estudado, faz-se necessária a incorporação dessa na estrutura do modelo geoestatístico. Logo, é possível a representação de todo o conjunto de dados por uma única função semivariância válida para todas as direções. Desse modo, são obtidas estimativas com maior precisão em locais não amostrados, pela técnica de krigagem. Por conseguinte, melhora-se a acurácia na construção de mapas de variabilidade espacial, os quais representam o comportamento dos atributos químicos do solo na área em estudo.

Um obstáculo comumente encontrado por pesquisadores é a definição para qual conjunto de parâmetros anisotrópicos é aceitável assumir a suposição de anisotropia. Assim, a autocorrelação espacial direcional descreve a influência exercida pela anisotropia geométrica na acurácia da estimação espacial de atributos químicos do solo, os quais

expressam as diferenças direcionais quando se incorpora a anisotropia no modelo geoestatístico.

2 OBJETIVOS

2.1 Objetivo geral

Este trabalho tem por objetivo utilizar medidas de autocorrelação espacial direcional na análise da anisotropia.

2.2 Objetivos específicos

- i. Identificar a presença de anisotropia em variáveis georreferenciadas simuladas e de atributos químicos do solo;
- ii. Calcular e utilizar o índice de Moran direcional para os valores amostrais de dados simulados e dos atributos químicos do solo, como ferramenta de auxílio na tomada de decisão quanto à existência de anisotropia;
- iii. Realizar a modelagem anisotrópica e ajustar um modelo teórico para a função semivariância, para dados simulados e atributos químicos do solo;
- iv. Calcular o índice de Moran direcional para os valores preditos, considerando-se os dados simulados e atributos químicos do solo;
- v. Identificar a diferença direcional que existe nos mapas temáticos quando se incorpora ou não a anisotropia no modelo geoestatístico usando medidas de autocorrelação espacial direcional.

3 REVISÃO DE LITERATURA

3.1 Geoestatística

O estudo da variabilidade espacial do solo é uma das preocupações desde o início do século XX. Trabalhos como os de Smith (1910) e Montgomery (1913) entre outros baseavam-se em estatística clássica que considera os parâmetros, média e desvio padrão, para a descrição dos fenômenos estudados. Embora a variação espacial do solo tenha sido reconhecida, a falta de ferramentas apropriadas quanto ao estudo da variabilidade espacial impossibilitava a sua quantificação e caracterização (BISWAS et al., 2014).

Na década de 50, Daniel Gerhardus Krige (engenheiro de minas) e Herbert Simon Sichel (estatístico) trabalharam com dados de concentração de ouro e concluíram que não havia sentido na variância sem se considerar a distância entre as amostras. A partir de então, surge o conceito de geoestatística, o qual considera a localização geográfica e a dependência espacial (REZENDE et al., 2012).

Quando comparada à estatística clássica, a geoestatística tem como diferencial auxiliar justamente nos casos em que se considera a dependência espacial entre as amostras. A geoestatística é definida como um conjunto de ferramentas estatísticas aplicadas a dados georreferenciados, porém, cada observação é descrita de acordo com sua posição e não apenas pelo seu valor. Consequentemente, observações geograficamente mais próximas tendem a apresentar valores mais semelhantes (VIEIRA, 2000).

O termo geoestatística foi utilizado nas publicações realizadas pelo matemático francês George Matheron que, inspirado pelo trabalho de Krige, desenvolveu no início dos anos 60 a “Teoria das Variáveis Regionalizadas”, a primeira a considerar a dependência espacial entre as amostras. No início, a utilização das técnicas de geoestatística era restrita a pesquisadores com a compreensão da análise numérica para programar os métodos, entretanto, nos últimos 20 anos, o avanço computacional favoreceu o uso e a disseminação das técnicas de geoestatística (OLIVER; WEBSTER, 2014).

A geoestatística possui conceitos baseados em funções e variáveis aleatórias. Um valor medido $z(s)$ é considerado uma realização do processo estocástico $Z(s)$, sendo s uma posição fixa. Pressupõe-se que cada dado $z(s)$ é modelado como uma variável aleatória $Z(s)$, denominada variável regionalizada. A soma dos três seguintes componentes expressa a variação espacial de uma variável aleatória:

- Um componente aleatório – em que a irregularidade não permite prever as variações de um ponto a outro;

- Um componente estruturado – que reflete as características estruturais do fenômeno regionalizado;
 - Um ruído – aleatório ou erro residual (DEL MONEGO; RIBEIRO JUNIOR, 2014).
- Logo, se s_i representa a i -ésima posição em primeira (IR), segunda (IR^2) ou terceira (IR^3) dimensão entre as amostradas em n localizações ($i = 1, \dots, n$), então o valor de Z , em s_i , pode ser descrito pelo modelo espacial linear gaussiano (Equação 1) (CRESSIE, 2015).

$$Z(s_i) = \mu(s_i) + \mathcal{E}'(s_i) + \mathcal{E}''(s_i), \quad (1)$$

em que:

$\mu(s_i)$ é uma função determinística que descreve o componente estrutural de Z em s_i ;

$\mathcal{E}'(s_i)$ é um termo estocástico correlacionado que varia localmente;

$\mathcal{E}''(s_i)$ é um ruído aleatório não correlacionado, com distribuição normal de probabilidade, média zero e variância σ^2 .

A dependência espacial define que cada ponto no espaço não apresenta um único valor e sim única distribuição de probabilidade de ocorrência de valores. Assim, o entendimento dessa distribuição de probabilidade pode descrever melhor o padrão espacial do fenômeno em estudo. Tal conceito diferencia-se do tradicional, em que cada observação é considerada como resultado independente de uma variável (BORSSOI, 2007). A variável regionalizada possui três características:

- Localização geográfica – no ponto s_i , com ($i = 1, \dots, n$), $Z(s_i)$ é uma variável aleatória que está associada ao tamanho e à forma da amostra, ou seja, a variável aleatória leva em consideração a geometria da distribuição espacial dos dados amostrais (CRESSIE, 2015);
- Estrutura espacial – é expressa pela correlação espacial entre as variáveis aleatórias, sendo que, as amostras próximas no tempo e/ou espaço sejam mais semelhantes entre si do que as que estiverem distantes (CRESSIE, 2015);
- Isotropia – quando a variabilidade espacial de uma área apresenta comportamento semelhante para todas as direções. Quando o fenômeno estudado apresenta variabilidade espacial mais intensa para alguma direção, esse é dito anisotrópico. Assim, é necessário realizar a modelagem da anisotropia para que um único semivariograma chamado de omnidirecional represente o fenômeno em estudo (GUEDES et al., 2008, 2013).

As técnicas de geoestatística possuem aplicações em diversas áreas, por exemplo: Grzegozewski et al. (2017) trabalharam com a correlação espacial da produtividade de soja; Yamamoto et al. (2016) trabalharam com a uniformidade de distribuição de água de irrigação por aspersores compensantes e não compensantes em diferentes alturas e Sobjak et al. (2016) usaram a geoestatística para verificar a qualidade das zonas de manejo. Portanto, conhecer a variabilidade espacial das variáveis de interesse possibilita, por

exemplo, a aplicação localizada de insumos agrícolas, proporciona aumento racional da produtividade de culturas, a redução nos custos de produção e minimização dos impactos negativos no meio ambiente (KESTRING et al., 2015).

3.2 Hipóteses de estacionaridade e Função Semivariância

Em geoestatística, geralmente, há n variáveis aleatórias com apenas um ensaio, mas não existe repetição de uma mesma variável. Logo, para que seja possível a realização de inferências estatísticas, é preciso que o processo estocástico Z atenda a algumas hipóteses de estacionaridade. Portanto, um processo estocástico é definido por $Z = \{Z(\mathbf{s}_i), \mathbf{s}_i \in D\}$, em que: o valor regionalizado do atributo Z , obtido na localização \mathbf{s}_i do espaço em estudo R subconjunto de \mathbb{R}^p , $p \geq 1$ (DE BASTIANI, 2012).

Desse modo, a análise geoestatística demanda no máximo a estacionaridade de segunda ordem, isto é, que os dois primeiros momentos da função aleatória sejam constantes em relação à origem. Logo, quando um fenômeno é considerado estacionário de ordem k , ele também é estacionário para ordens inferiores a k (NOGUEIRA et al., 2013). É importante ressaltar que processos não estacionários podem apresentar trechos estacionários.

Segundo Cressie (2015), a hipótese de estacionariedade de segunda ordem é satisfeita se:

- A esperança matemática de $Z(\mathbf{s})$ existe e não depende da posição \mathbf{s} . Ou seja:

$$E[Z(\mathbf{s})] = \mu = \text{constante}, \quad \forall \mathbf{s} \in D \subset \mathbb{R}^p, \quad (2)$$

- Para cada par $\{Z(\mathbf{s}), Z(\mathbf{s} + \mathbf{h})\}$, a covariância existe e depende somente de \mathbf{h} .

$$\text{Cov}[Z(\mathbf{s}), Z(\mathbf{s} + \mathbf{h})] = E[Z(\mathbf{s} + \mathbf{h}), Z(\mathbf{s})] - \mu^2 = C(\mathbf{h}). \quad (3)$$

em que $h = \|\mathbf{h}\|$ é a distância euclidiana entre duas localizações desconhecidas e \mathbf{h} é o vetor distância entre essas duas localizações (x_i, y_i) e $(x_i + h, y_i + h)$

Desse modo, define-se o covariograma $C(h)$, que consiste em um gráfico dos valores da covariância (eixo das ordenadas), em função das distâncias (eixo das abcissas) (Figura 1).

Uma estacionaridade mais fraca que a estacionaridade de segunda ordem foi definida e denominada hipótese intrínseca devido à necessidade de tornar os modelos de

inferências mais abrangentes. Na literatura, também é conhecida como estacionaridade fraca ou dos incrementos, expressa por (CRESSIE, 2015):

- A esperança matemática existe e não depende da posição s :

$$[Z(\mathbf{s})] = \mu, \quad \forall \mathbf{s} \in D \subset \mathbb{R}^P, \text{ e} \quad (4)$$

- Para todo \mathbf{h} , a variância da diferença $[Z(\mathbf{s}) - Z(\mathbf{s} + \mathbf{h})]$ existe e não depende de s :

$$\frac{1}{2} \text{Var} [Z(\mathbf{s}) - Z(\mathbf{s} + \mathbf{h})] = \frac{1}{2} E([Z(\mathbf{s}) - Z(\mathbf{s} + \mathbf{h})]^2) = \gamma(\mathbf{h}), \quad (5)$$

em que $\gamma(\mathbf{h})$ representa a função semivariância. O prefixo “semi” existe pelo fato de que $\gamma(\mathbf{h})$ representa a metade da variância (Equação 5). Desta forma, o semivariograma consiste em um gráfico dos valores da semivariância em função das distâncias h (Figura 1).

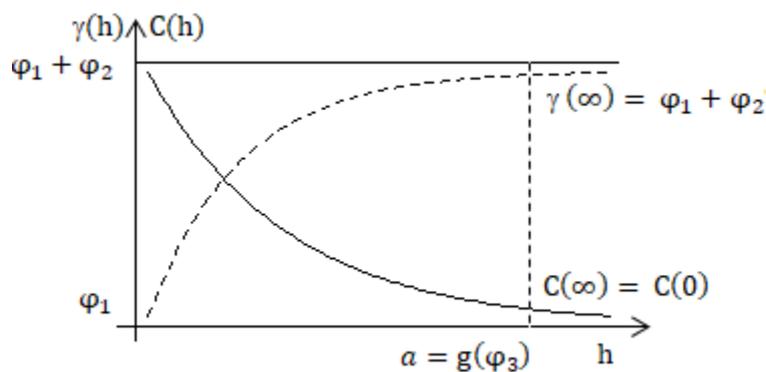


Figura 1 Relação entre as funções covariância e a semivariância.

A estacionaridade de segunda ordem é restritiva quando o fenômeno estudado possui dispersão ilimitada, portanto, não é possível associar uma variância finita a esse tipo de fenômeno. Uma vez que essa estacionaridade não é atendida, o correlograma não pode ser utilizado, pois é necessária a utilização do semivariograma para as análises futuras. Caso ocorra estacionaridade de segunda ordem, o correlograma e o semivariograma são ditos ferramentas correspondentes, para que possam estabelecer uma relação (Figura 1) entre a função semivariância $\gamma(\mathbf{h})$ e a covariância $C(\mathbf{h})$, expressa por (ALMEIDA et al., 2013):

$$C(\mathbf{h}) = C(0) - \gamma(\mathbf{h}), \text{ onde: } C(0) = \text{var}(Z(\mathbf{s})) \quad (6)$$

O processo estocástico de segunda ordem permite o uso do coeficiente de correlação, definido como: a razão entre a covariância dos valores da variável Z nas posições \mathbf{s} e $\mathbf{s} + \mathbf{h}$ e a variância dessa variável em função da distância h . Esta função é adimensional e está limitada entre os valores -1 e 1, além de possibilitar comparações entre

variáveis e também inferências sobre o grau de associação espacial (GRZEGOZEWSKI et., 2013).

$$\rho(h) = \frac{C(h)}{C(0)} = \frac{C(0) - \gamma(h)}{C(0)} = 1 - \frac{\gamma(h)}{C(0)}. \quad (7)$$

3.3 Semivariograma experimental

A semivariância é uma medida de dissimilaridade entre as amostras e leva em consideração a distância que as separa. Sua representação gráfica (Figura 2) é denominada semivariograma, portanto, é um gráfico de dispersão dos valores da semivariância em função das distâncias h (ISAAKS e SRIVASTAVA, 1989).

Em geoestatística, considerando-se dados amostrais georreferenciados, os valores da semivariância são calculados utilizando-se estimadores. O estimador de Matheron (1963) (Equação 8) é um dos mais utilizados na literatura para o conjunto de dados de um processo estocástico gaussiano (CRESSIE, 2015).

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(\mathbf{s}_i) - Z(\mathbf{s}_i + \mathbf{h})]^2, \quad (8)$$

em que: $\gamma(h)$ é o valor da semivariância, obtida pelos valores amostrados; $N(h)$ é o número de pares de valores medidos separados por uma distância h ; $Z(\mathbf{s}_i)$ e $Z(\mathbf{s}_i + \mathbf{h})$ são os valores da variável nas posições \mathbf{s}_i e $\mathbf{s}_i + \mathbf{h}$, de tal modo que esses pontos estejam separados por uma distância h .

O estimador da função semivariância de Matheron (1963) permite estruturar duas versões do semivariograma: o experimental direcional e o experimental omnidirecional. O semivariograma experimental direcional permite analisar o comportamento espacial da variável georreferenciada em diferentes direções. Já o semivariograma experimental omnidirecional é definido com informações de todas as possíveis direções utilizadas para os cálculos das semivariâncias (SOARES, 2014).

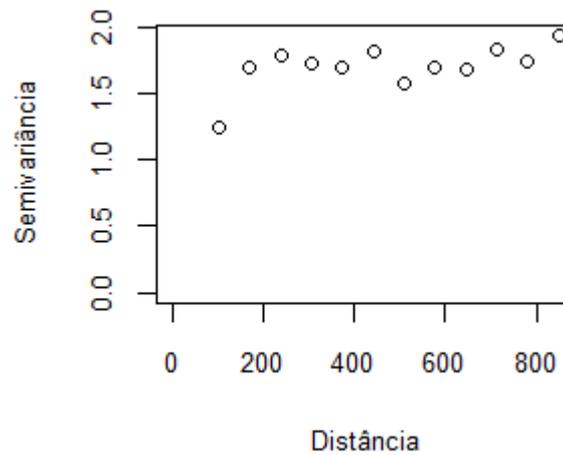


Figura 2 Gráfico semivariograma experimental omnidirecional.

Na construção do semivariograma, visando tornar o resultado consistente, pode ser estabelecido um número mínimo de 30 pares de pontos georreferenciados, espaçados pela mesma distância, como suficiente para o cálculo da função semivariância (ANDRIOTTI, 2004).

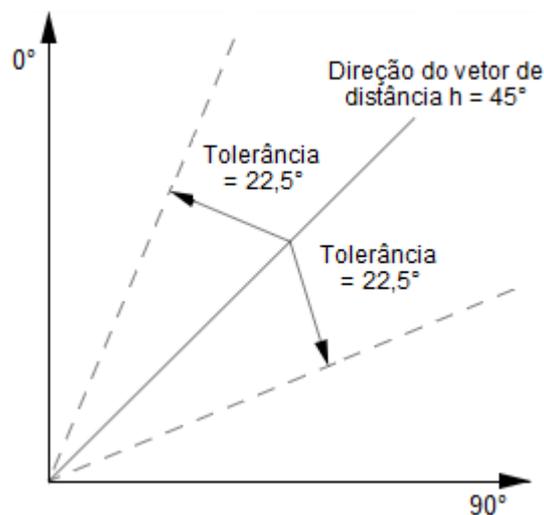


Figura 3 Tolerância angular em relação ao vetor de distância h .

Esse resultado é facilmente obtido em amostras regulares, uma vez que todos os pares de pontos amostrais estão separados pela mesma distância h . Porém, em grades irregulares, pode existir um número muito pequeno de pares de pontos amostrais espaçados exatamente pela mesma distância h . Assim, MCBratney e Webster (1986) recomendam que seja admitida uma tolerância Δh na distância e uma tolerância $\Delta\theta$ na direção, quanto ao ângulo, assim, os autores propõem uma tolerância de 40° .

3.3.1 Parâmetros do modelo teórico da função semivariância

A semivariância é uma função crescente da distância h , logo, quanto mais próximos estiverem dois pontos amostrados, menor será o valor da semivariância entre eles.

A Figura 3 representa o modelo teórico do semivariograma e seus respectivos parâmetros (GRZEGOZEWSKI et al., 2013), sendo eles:

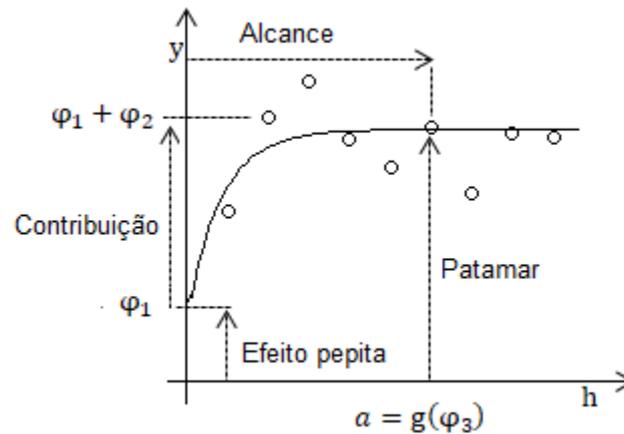


Figura 4 Modelo teórico ajustado ao semivariograma experimental.

- Alcance ($a = g(\varphi_3)$) – distância máxima em que as amostras se apresentam correlacionadas espacialmente. Elas são consideradas independentes nas distâncias superiores ao alcance. O alcance é a função do parâmetro φ_3 , em que g representa uma função (ROSSONI; LIMA, 2012).
- Efeito pepita (φ_1) - é uma descontinuidade na origem. Teoricamente $\gamma(0) = 0$, porém, na prática, isso não ocorre, à medida que h tende a 0 (zero), $\gamma(h)$ se aproxima de um valor positivo denominado efeito pepita, isso ocorre devido aos erros de medição ou da variabilidade em pequena escala não captada pela amostragem (NOGUEIRA et al., 2013).
- Contribuição (φ_2) – é a amplitude da função semivariância, ou seja, a diferença entre o patamar e o efeito pepita.
- Patamar ($C = \varphi_1 + \varphi_2$) – é o valor que a semivariância $\gamma(h)$ se torna constante, deste ponto em diante considera-se que não existe mais dependência espacial entre as amostras (ROSSONI; LIMA, 2012).

3.4 Isotropia e Anisotropia

Considerando-se os modelos estimados para a função semivariância em direções ortogonais, uma variável regionalizada é dita isotrópica quando a continuidade espacial possui comportamentos semelhantes para todas as direções analisadas. Logo, $\gamma(h)$ depende somente do módulo do vetor \mathbf{h} , em que $h = \|\mathbf{h}\|$; e os valores estimados do alcance, efeito pepita e patamar são semelhantes em todas as direções quando se ajusta um modelo para a semivariância experimental direcional. No entanto, se a estrutura de dependência espacial difere em relação à direção, então o processo é dito anisotrópico (GUEDES et al., 2013).

Três tipos de anisotropia são citados na literatura: a anisotropia geométrica, a zonal e a combinada (GUEDES et al., 2013).

A anisotropia é dita geométrica quando uma direção se apresenta com maior continuidade espacial, ou seja, maior valor do alcance (a) no semivariograma em determinada direção (ROSSONI; LIMA, 2012).

Igualmente, a anisotropia é dita zonal quando os semivariogramas construídos sobre distintas direções apresentam diferentes características de variabilidade, isto é, diferentes patamares ($\varphi_1 + \varphi_2$) e com alcance e efeito pepita constantes (GUEDES et al., 2008). Por fim, nos fenômenos da natureza estudados, normalmente detecta-se uma combinação entre a anisotropia geométrica e anisotropia zonal, e essa combinação é denominada anisotropia combinada (GUEDES et al., 2008).

Semivariogramas experimentais direcionais são construídos para se iniciar a análise de anisotropia, geralmente para as direções 0° , 45° , 90° e 135° (CÂMARA; MEDEIROS, 1998) considerando o sistema azimute. Estima-se o modelo para a função semivariância para cada direção e, em seguida, para cada par de direções ortogonais, são comparados os valores estimados do alcance, efeito pepita e patamar (GUEDES et al., 2013).

Na literatura, existem várias fórmulas de transformações propostas para correção da anisotropia geométrica. A transformação de coordenadas (Equação 9), proposta por Diggle e Ribeiro Junior (2007), rotaciona a coordenada do ponto no sentido horário ao se usar o ângulo α ($\frac{\pi}{2} \leq \alpha \leq \frac{\pi}{2}$), em que α é o ângulo de maior continuidade espacial. E, posteriormente, aplica-se a transformação de dilatação na coordenada usando-se a matriz \mathbf{F}_1 .

$$\mathbf{h}' = (x, y) \times \mathbf{R}_{-\alpha} \times \mathbf{F}_1 \quad (9)$$

em que, $\mathbf{R}_{-\alpha} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$ é a matriz de rotação; $\mathbf{F}_1 = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{F_a} \end{bmatrix}$ a matriz de contração;
 $F_a = \frac{\text{alcançe}_{\max}}{\text{alcançe}_{\min}}$ é o fator de anisotropia, com $F_a \geq 1$.

As diferenças entre mapas temáticos quanto à incorporação ou não da anisotropia geométrica foram apresentadas na literatura nos estudos de Guedes et al. (2013), estudo no qual os autores elucidam, a partir de observações dos mapas temáticos, que tais diferenças se intensificam quando o fator de anisotropia apresentado é maior que 2.

3.5 Autocorrelação espacial direcional

O índice de Moran é utilizado para indicar a possibilidade de autocorrelação espacial. Tal índice objetiva indicar se os dados se distribuem de forma aleatória ou não no espaço (DALPOSSO et al., 2012). No entanto, para a obtenção do cálculo da autocorrelação espacial direcional, é necessário realizar uma adaptação na matriz de peso, para que o Índice global de Moran seja capaz de descrever a autocorrelação espacial considerando-se a distância e a direção.

Rosenberg (2000) apresentou tal adaptação para o I de Moran ao se associar a matriz de peso \mathbf{W} a um ângulo de interesse e a um fator de ponderação, considerando-se simultaneamente a distância e a direção, nas quais as amostras são espaçadas, de modo que, para cada par de pontos i e j , o peso w_{ij} é 1 se a distância entre eles for menor ou igual a distância de corte (d) e 0 se a distância entre os pontos for maior que a distância de corte.

O índice de Moran direcional calculado para uma distância de corte (d) e uma direção de interesse (θ) é definido pela seguinte expressão:

$$I(d) = \frac{\left[\frac{1}{W(d)} \right] \sum_{i=1}^n \sum_{j=1, j \neq i}^n \mathbf{W}_{ij}(d) (m_i - \bar{m})(m_j - \bar{m})}{\frac{1}{n} \sqrt{\sum_{i=1}^n (m_i - \bar{m})^2}} \quad (10)$$

em que:

- n é o número de pontos de interesse para o cálculo do índice de Moran;
- m_i e m_j são os valores do atributo nos pontos i e j ;
- \bar{m} é média do atributo dos pontos usados para o cálculo do índice de Moran;
- $W(d)$ é o somatório dos elementos W_{ij} da matriz de pesos na distância d e
- \mathbf{W}_{ij} é a matriz de pesos na distância d , de dimensão $n \times n$, em que cada elemento w'_{ij} é definido por:

$$w'_{ij} = w_{ij} \cos^2(\alpha_{ij} - \theta) \quad (11)$$

em que: w_{ij} é elemento de uma matriz binária, e $w_{ij} = \begin{cases} 0, & \text{se } d_{ij} > d \\ 1, & \text{se } d_{ij} \leq d \end{cases}$, d_{ij} é a distância euclidiana entre os pares de pontos i e j , e d é a distância de corte; α_{ij} é a direção angular entre os pontos i e j , calculados de forma $\alpha_{ij} = \arctan \left[\frac{y_j - y_i}{x_j - x_i} \right] \times \frac{180}{\pi}$, sendo que $\alpha_{ij} < 0 + \pi$; em que (x_i, y_i) e (x_j, y_j) são as coordenadas dos pontos i e j , e θ é a direção de interesse para o cálculo do índice de Moran direcional.

Segundo Rosenberg (2000), conforme a direção angular (α_{ij}) se afasta da direção de interesse θ , os pares de pontos i e j passam a serem punidos baseando-se na baixa associação com a direção testada, sem afetar pares de pontos que possuem α_{ij} e θ próximos.

Os resultados obtidos pelo índice de Moran variam de -1 a +1 e podem ser interpretados da seguinte forma: quando o I de Moran for positivo ao apresentar valores elevados, deve possuir em seu entorno valores elevados (autocorrelação direta). No mesmo sentido, quando I de Moran for positivo ao apresentar valores baixos deve possuir em seu entorno valores baixos, de modo a se caracterizarem grupos. Entretanto, quando I de Moran apresentar resultados negativos, espera-se que valores elevados sejam cercados por valores baixos (autocorrelação inversa) e valores baixos sejam cercados por valores elevados (ANSELIN, 1995).

O índice de Moran direcional pode ser aplicado como técnica para identificar a presença de anisotropia geométrica em uma variável georreferenciada com dependência espacial. Desse modo, quando comparados os valores do índice de Moran direcional calculados para as direções ortogonais, espera-se a obtenção de valores mais elevados para direção de maior continuidade espacial, e ainda, conforme aumenta-se a distância de corte espera-se menor variação dos valores deste índice para a direção de maior continuidade espacial.

No mesmo sentido, Falsetti e Sokal (1993) introduziram um método denominado procedimento de rolamento também calculado pela equação 10, porém, com diferenças na obtenção da matriz de peso, a qual foi obtida pelo produto $\mathbf{W}_{ij} = \mathbf{D}_{n \times n} \times \mathbf{L}_{n \times n}$ em que: $\mathbf{D}_{n \times n}$ é

a matriz de distância binária de modo que $d_{ij} = \begin{cases} 0, & \text{se } d_{ij} > d \\ 1, & \text{se } d_{ij} \leq d \end{cases}$ sendo d_{ij} elemento da matriz

$\mathbf{D}_{n \times n}$ gerada com base nas distâncias de corte d ; $\mathbf{L}_{n \times n}$ é a matriz de direção angular α_{ij} entre as localidades i e j , associada a um ângulo de interesse θ e uma tolerância angular β , de

modo que, $l_{ij} = \begin{cases} 0, & \text{se } \alpha_{ij} > \theta \pm \beta \\ 1, & \text{se } \alpha_{ij} \leq \theta \pm \beta \end{cases}$ sendo l_{ij} elemento da matriz $\mathbf{L}_{n \times n}$.

3.6 Modelos Teóricos

Existem dois tipos de modelos para a função semivariância, os modelos com patamar e os modelos sem patamar, assim como o modelo potência. Neste trabalho, somente os modelos com patamar serão apresentados. Os modelos esférico, exponencial, gaussiano e da família Matérn possuem as próprias características e na literatura foram descritos considerando que a variável georreferenciada é isotrópica (ISAACS e SRIVASTAVA, 1989; CRESSIE, 2015).

O modelo esférico apresenta as seguintes características: o crescimento rápido na origem; atinge o patamar a $\frac{2}{3}$ do alcance e é válido para IR , IR^2 e IR^3 (CRESSIE, 2015).

As equações 12, 13 e 14 representam respectivamente as funções de semivariância ($\gamma(h)$), covariância ($C(h)$) e correlação ($\rho(h)$) para um modelo esférico com os valores de efeito pepita igual a φ_1 , patamar igual a $\varphi_1 + \varphi_2$ e alcance $a = g(\varphi_3)$.

$$\gamma(h) = \begin{cases} 0, & h = 0 \\ \varphi_1 + \varphi_2 \left[\frac{3}{2} \left(\frac{h}{\varphi_3} \right) - \frac{1}{2} \left(\frac{h}{\varphi_3} \right)^3 \right], & 0 < h \leq \varphi_3, \\ \varphi_1 + \varphi_2, & h > \varphi_3 \end{cases} \quad (12)$$

$$C(h) = \begin{cases} \varphi_1 + \varphi_2, & h = 0 \\ \varphi_2 \left[1 - \frac{3}{2} \left(\frac{h}{\varphi_3} \right) + \frac{1}{2} \left(\frac{h}{\varphi_3} \right)^3 \right], & 0 < h \leq \varphi_3, \\ 0, & h > \varphi_3 \end{cases} \quad (13)$$

$$\rho(h) = \begin{cases} 1, & h = 0 \\ 1 - \frac{3}{2} \left(\frac{h}{\varphi_3} \right) + \frac{1}{2} \left(\frac{h}{\varphi_3} \right)^3, & 0 < h \leq \varphi_3. \end{cases} \quad (14)$$

O modelo exponencial apresenta como características: o comportamento aproximadamente linear na origem; atinge o patamar assintoticamente com alcance prático definido como a distância na qual o valor do modelo é de 95% de φ_2 , cujo alcance prático é dado por $a = 3\varphi_3$ (CRESSIE, 2015).

As equações 15, 16 e 17 representam respectivamente as funções de semivariância ($\gamma(h)$), covariância ($C(h)$) e correlação ($\rho(h)$) para um modelo exponencial com os valores de efeito pepita igual a φ_1 , patamar igual a $\varphi_1 + \varphi_2$.

$$\gamma(h) = \begin{cases} 0, & h = 0 \\ \varphi_1 + \varphi_2 \left[1 - \exp\left(-\frac{h}{\varphi_3}\right) \right], & 0 < h \leq \varphi_3, \\ \varphi_1 + \varphi_2, & h > \varphi_3 \end{cases} \quad (15)$$

$$C(h) = \begin{cases} \varphi_1 + \varphi_2, & h = 0 \\ \varphi_2 \left[\exp - \left(\frac{h}{\varphi_3}\right) \right], & 0 < h \leq \varphi_3, \\ 0, & h > \varphi_3 \end{cases} \quad (16)$$

$$\rho(h) = \begin{cases} 1, & h = 0 \\ \exp\left(\frac{h}{\varphi_3}\right), & 0 < h \leq \varphi_3. \end{cases} \quad (17)$$

O modelo gaussiano apresenta as seguintes características: o comportamento parabólico na origem; atingindo o patamar apenas assintoticamente e tem alcance prático dado por $a = \sqrt{3}\varphi_3$. O modelo gaussiano é utilizado para modelar um fenômeno extremamente contínuo e seus parâmetros φ_1 e φ_2 são obtidos da mesma maneira que no modelo esférico. As equações 18, 19 e 20 representam respectivamente as funções de semivariância ($\gamma(h)$), covariância ($C(h)$) e correlação ($\rho(h)$) para um modelo gaussiano (URIBE-OPAZO et al., 2012).

$$\gamma(h) = \begin{cases} 0, & h = 0 \\ \varphi_1 + \varphi_2 \left\{ 1 - \exp\left[-\left(\frac{h}{\varphi_3}\right)^2\right] \right\}, & 0 < h \leq \varphi_3, \\ \varphi_1 + \varphi_2, & h > \varphi_3 \end{cases} \quad (18)$$

$$C(h) = \begin{cases} \varphi_1 + \varphi_2, & h = 0 \\ \varphi_2 \left\{ \exp\left[-\left(\frac{h}{\varphi_3}\right)^2\right] \right\}, & 0 < h \leq \varphi_3, \\ 0, & h > \varphi_3 \end{cases} \quad (19)$$

$$\rho(h) = \begin{cases} 1, & h = 0 \\ \exp\left[-\left(\frac{h}{\varphi_3}\right)^2\right], & 0 < h \leq \varphi_3. \end{cases} \quad (20)$$

As equações 21, 22 e 23 representam respectivamente as funções de semivariância ($\gamma(h)$), covariância ($C(h)$) e correlação ($\rho(h)$) para um modelo Matérn de ordem k (CRESSIE, 2015).

$$\gamma(h) = \begin{cases} 0, & h = 0 \\ \varphi_1 + \varphi_2 \left[\left(1 - 2^{k-1}\Gamma(k)\right)^{-1} \left(\frac{h}{\varphi_3}\right)^k k_k\left(\frac{h}{\varphi_3}\right) \right], & h > 0, \end{cases} \quad (21)$$

$$C(h) = \begin{cases} 0, & h = 0 \\ \varphi_2 \left[\left(2^{k-1}\Gamma(k)\right)^{-1} \left(\frac{h}{\varphi_3}\right)^k k_k\left(\frac{h}{\varphi_3}\right) \right], & h > 0, \end{cases} \quad (22)$$

$$\rho(h) = \left\{ \left(2^{k-1}\Gamma(k)\right)^{-1} \left(\frac{h}{\varphi_3}\right)^k k_k\left(\frac{h}{\varphi_3}\right) \right\} h > 0. \quad (23)$$

em que: φ_1 , φ_2 , φ_3 e k , são parâmetros da função semivariância; K_k é a função de Bessel de terceiro tipo de ordem k , sendo $k_k^\mu = \frac{1}{2} \int_0^\infty x^{k-1} e^{-\frac{1}{2} \left(x + \frac{1}{x}\right)} dx$. A família Matérn é válida para $\varphi_3, k > 0$ e corresponde à função correlação exponencial quando $k = 0,5$; e à função correlação gaussiana quando $k \rightarrow \infty$ (DIGGLE e RIBEIRO JÚNIOR, 2007).

3.7 Métodos de Ajuste ou estimação de parâmetros

É necessário que se estimem os parâmetros efeito pepita, contribuição e um valor que é função de alcance, a fim de se estruturar a dependência espacial para ajustar o modelo teórico à função semivariância. A estimação dos parâmetros pode ser realizada por vários métodos, entre eles o método de Máxima Verossimilhança (MV).

3.7.1 Estimação de parâmetros por máxima verossimilhança - MV

Um campo aleatório estacionário permite que se utilizem estimadores de máxima verossimilhança para estimar os parâmetros desconhecidos $\theta = (\beta^T, \phi^T)^T \in \Theta$ (espaço paramétrico). De modo que a melhor estimativa para o vetor θ de parâmetros será aquela que maximiza o logaritmo da função verossimilhança (Equação 24), ou seja, $l(\hat{\theta}) = \max l(\theta)$, em que, θ pertence a um espaço paramétrico e $\hat{\theta}$ é o estimador de máxima verossimilhança (CRESSIE, 2015).

$$l(\theta) = -\left(\frac{n}{2}\right) \log(2\pi) - \frac{1}{2} \log|\Sigma| - \frac{1}{2} (\mathbf{Z} - \mathbf{X}\beta)^T \Sigma^{-1} (\mathbf{Z} - \mathbf{X}\beta) \quad (24)$$

em que : $\theta = (\beta^T, \phi^T)^T$, $\beta = (\beta_1, \dots, \beta_n)^T$ e $\phi = (\varphi_1, \varphi_2 \text{ e } \varphi_3)^T$, Σ é a matriz de covariância e Z é o vetor dos valores da variável.

3.8 Critérios de escolha do melhor modelo

3.8.1 Validação cruzada

A validação cruzada consiste em avaliar os erros de estimativas (Equação 25) comparando-se os valores estimados e medidos. De tal forma que uma amostra $Z(\mathbf{s}_i)$ é temporariamente retirada da amostra de dados. Então é feita uma estimação de $Z(\mathbf{s}_i)$, denotada por $\hat{Z}(\mathbf{s}_{(i)})$ usando-se um interpolador e as informações dos demais elementos amostrais.

A estimação é feita e em seguida o valor amostral é reintroduzido nos dados e esse processo é repetido para todos os dados amostrais para que os resultados apresentados para média dos erros de estimação sejam nulos, a variância constante e a distribuição normal de probabilidade (ISAAKS e SRIVASTAVA, 1989).

$$\varepsilon(\mathbf{s}_i) = Z(\mathbf{s}_i) - \hat{Z}(\mathbf{s}_{(i)}). \quad (25)$$

em que $\hat{Z}(\mathbf{s}_{(i)})$ é o valor estimado por krigagem sem a i -ésima observação $Z(\mathbf{s}_i)$.

Nesse processo, além da eficácia do ajuste do modelo teórico, a função semivariância é avaliada quanto à estacionaridade e à presença de dados atípicos. A comparação é realizada pelo erro médio (EM) (Equação 26), erro médio reduzido (\overline{ER}) (Equação 27), desvio padrão dos erros médio (S_{EM}) (Equação 28), desvio padrão dos erros reduzidos (S_{ER}) (Equação 29) e desvio do erro absoluto (EA) (Equação 30) (FARACO et al., 2008).

$$EM = \frac{1}{n} \sum_{i=1}^n (Z(\mathbf{s}_i) - \hat{Z}(\mathbf{s}_{(i)})), \quad (26)$$

$$\overline{ER} = \frac{1}{n} \sum_{i=1}^n \left(\frac{(Z(\mathbf{s}_i) - \hat{Z}(\mathbf{s}_{(i)}))}{\sigma(\mathbf{s}_{(i)})} \right), \quad (27)$$

$$S_{EM} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Z(\mathbf{s}_i) - \hat{Z}(\mathbf{s}_i))^2}, \quad (28)$$

$$S_{ER} = \sqrt{\frac{1}{n} \sum_{i=1}^n \frac{(Z(\mathbf{s}_i) - \hat{Z}(\mathbf{s}_i))^2}{\sigma(\hat{Z}(\mathbf{s}_i))}}, \quad (29)$$

$$EA = \frac{1}{n} \sum_{i=1}^n |Z(\mathbf{s}_i) - \hat{Z}(\mathbf{s}_i)|. \quad (30)$$

em que: n é o número de observações amostrais; $Z(\mathbf{s}_i)$ é o valor observado no ponto \mathbf{s}_i ; $\hat{Z}(\mathbf{s}_i)$ é o valor estimado por krigagem no ponto \mathbf{s}_i , sem considerar a observação $Z(\mathbf{s}_i)$ na estimação do modelo para a função semivariância e no processo da krigagem; e $\sigma(\hat{Z}(\mathbf{s}_i))$ é o desvio padrão da krigagem no ponto \mathbf{s}_i , sem se considerar a observação $Z(\mathbf{s}_i)$.

É estritamente necessário que EM e \overline{ER} possuam valores mais próximos de zero para que os resultados não sejam tendenciosos, S_{EM} tenha menor valor e S_{ER} tenha valor mais próximo de 1 (um), este é o critério de escolha do melhor modelo ajustado. O erro absoluto mede a magnitude dos erros a partir dos valores medidos e preditos por krigagem, logo, é considerado o melhor modelo aquele que tiver o menor valor de EA (FARACO et al., 2008).

3.8.2 Critério de informação de Akaike e Bayesiano

O critério de informação de Akaike indica que a escolha entre modelos teóricos para função semivariância deve ser realizada ao se considerar a simplicidade do modelo, impondo uma penalidade para a complexidade, e cujo modelo escolhido for aquele que apresentar menor valor de AIC:

$$AIC = -2l(\boldsymbol{\theta}) + 2r \quad (31)$$

em que: $l(\boldsymbol{\theta})$ é o valor do logaritmo da função verossimilhança; r é o número de parâmetros do modelo estimado.

O critério de Informação Bayesiano (BIC) penaliza a função verossimilhança para que um modelo mais parcimonioso seja selecionado. Assim, o modelo que apresentar a menor estimativa de BIC é considerado o modelo de melhor ajuste (DEL MONEGO; RIBEIRO JUNIOR, 2014).

$$\text{BIC} = \frac{(-2l(\boldsymbol{\theta}))}{n} + \frac{(k \log n)}{n} \quad (32)$$

em que: k é o número de restrições; n é o tamanho da amostra; $l(\boldsymbol{\theta})$ é o valor do logaritmo da função verossimilhança obtido na otimização por ML e $\boldsymbol{\theta}$ é o vetor de parâmetros.

3.9 Krigagem

Em geoestatística, para se descrever a análise da variabilidade espacial do fenômeno estudado, são utilizados mapas de variabilidade espacial construídos pelo uso de interpoladores. A aplicação dessa técnica permite que sejam estimados valores em locais não foram amostrados, para possibilitar a criação do mapa de variabilidade espacial em toda área de estudo (SILVEIRA et al., 2014). A krigagem engloba um conjunto de métodos de estimação, a saber: krigagem simples, krigagem ordinária, krigagem universal, Co-krigagem, krigagem disjuntiva, etc. Este trabalho limita-se à apresentação da krigagem ordinária (OLIVER; WEBSTER, 2014).

A diferença entre a krigagem e outros métodos de interpolação é a maneira como os pesos λ_i são atribuídos às diferentes amostras. O procedimento é semelhante ao de interpolação por média móvel ponderada, exceto que, na krigagem, os pesos são determinados a partir de uma análise espacial, por isso, é necessário que a soma dos pesos λ_i seja igual a 1. Além disso, a krigagem fornece em média uma estimativa não tendenciosa e com variância mínima (ROSSONI; LIMA, 2012).

3.9.1 Krigagem ordinária

A krigagem ordinária é fundamentada na ideia de regressão linear, em que, a predição de um valor desconhecido da variável georreferenciada Z em uma localização não amostrada \mathbf{s}_0 ($\hat{Z}(\mathbf{s}_0)$) é obtida pela combinação linear de n valores conhecidos $Z(\mathbf{s}_i)$, com $i = 1, 2, \dots, n$ adicionada a um parâmetro λ_0 (CÂMARA et al., 1998). Ou seja:

$$\hat{Z}(\mathbf{s}_0) = \sum_{i=1}^n \lambda_i Z(\mathbf{s}_i) \quad (33)$$

em que: $\hat{Z}(s_0)$ é o valor predito da variável georreferenciada Z na posição não amostrada s_0 ; n é o número de valores amostrados; λ_i são os pesos associados a cada valor de $Z(s_i)$ ($i = 1, 2, \dots, n$); $Z(s_i)$ são os valores conhecidos da variável georreferenciada nos locais amostrados ($i = 1, 2, \dots, n$). Sendo satisfeita se e somente se (DE BASTIANI, 2012):

$$\sum_{i=1}^n \lambda_i = 1. \quad (34)$$

Minimizando-se a variância do erro ($\text{Var}[\hat{Z}(s_0) - Z(s_0)]$) de maneira que a soma dos pesos seja igual a 1 (Equação 34), os pesos λ_i podem ser obtidos pela resolução do sistema de equação lineares e expressos por:

$$\begin{cases} \sum_{j=1}^n \lambda_j \gamma(s_i, s_j) - \alpha = \gamma(s_i, s_0) & \text{para } i = 1, \dots, n \\ \sum_{j=1}^n \lambda_j = 1 \end{cases} \quad (35)$$

em que: $\gamma(s_i, s_j)$ e $\gamma(s_i, s_0)$ são, respectivamente, o valor de semivariância entre os pontos s_i e s_j e entre os pontos s_i e s_0 ($i = 1, \dots, n$); α é o multiplicador de lagrange necessário para a minimização da variância do erro.

3.10 Comparação de mapas temáticos

Dois mapas (um anisotrópico e um isotrópico) de variabilidade espacial podem ser construídos para uma variável georreferenciada, pelo interpolador de krigagem.

A avaliação da acurácia entre esses mapas temáticos é obtida a partir dos elementos da matriz de contingência das classes, conhecida como matriz de erros (Tabela 1). Assim, assume-se um mapa referência em que os pixels definidos como uma área mínima que pode ser georreferenciada, (considerados nesta a informação a respeito de uma variável regionalizada), são divididos em classes iguais e quantificados nas colunas, enquanto os pixels de um mapa modelo são quantificados nas linhas (YAMAMOTO e LANDIM, 2013). Desta matriz, podem-se derivar várias medidas de precisão da classificação como a Exatidão Global (Equação 36) e os índices de concordância Kappa e Tau (Equações 37 e 38) (DALPOSSO et al., 2012).

Tabela 1 Representação genérica de uma matriz de erros.

Mapa Modelo	Classe	Mapa de Referência				Total por Linha $n_{i\cdot}$
		C_1	C_2	...	C_m	
	C_1	n_{11}	n_{12}	...	n_{1m}	$n_{1\cdot}$
	C_2	n_{21}	n_{22}	...	n_{2m}	$n_{2\cdot}$
	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
	C_m	n_{m1}	n_{m2}	...	n_{mm}	$n_{m\cdot}$
	Total por Coluna $n_{\cdot j}$	$n_{\cdot 1}$	$n_{\cdot 2}$...	$n_{\cdot m}$	n

$$EG = \sum_{i=1}^m \frac{n_{ii}}{n}. \quad (36)$$

$$K = \frac{n \sum_{i=1}^m n_{ii} - \sum_{i=1}^m (n_{i\cdot} \cdot n_{\cdot i})}{n^2 - \sum_{i=1}^m (n_{i\cdot} \cdot n_{\cdot i})}. \quad (37)$$

$$T = \frac{\frac{\sum_{i=1}^m n_{ii}}{n} - p_i}{1 - p_i}. \quad (38)$$

em que: m é o número de classes; $n_{i\cdot}$ e $n_{\cdot j}$ são os subtotais da linha i e coluna j ; n_{ij} é o número de pixels classificados na classe i do mapa modelo e na classe j do mapa de referência (quando as probabilidades, a priori, forem iguais para as classes, tem-se $p_i = \frac{1}{m}$) (DE BASTIANI et al., 2012).

A exatidão global EG (Equação 36) representa o percentual de acerto da classificação entre o mapa referência e o mapa modelo, considerando-se como valor aceitável EG acima de 0,85 (ANDERSON et al., 1976).

A análise do índice de concordância Kappa e Tau (Equação 37 e 38) é uma técnica utilizada para se mensurar a exatidão das classificações temáticas. Esse índice fornece uma medida de concordância entre os valores do mapa modelo e os valores do mapa referência utilizando todos os elementos da matriz de erro. O índice de concordância Kappa varia entre 0 e 1, assim, é possível transformá-lo em percentual ao multiplicá-lo por 100, e classificado conforme Tabela 2 (KRIPPENDORFF, 1980):

Tabela 2 Classificação do índice Kappa

Índice Kappa	Desempenho de classificação
maior que 0,80	concordância excelente
entre 0,67 e 0,80	concordância regular a boa
menor que 0,67	concordância ruim

4 MATERIAL E MÉTODOS

O presente trabalho foi dividido em duas etapas: um estudo de simulações e um estudo prático com dados reais provenientes de determinada área agrícola.

4.1 Simulações de processos espaciais gaussianos

Um conjunto de dados foi simulado e obtido a partir do experimento de Monte Carlo para a concretização do estudo, ao se considerar que a variável georreferenciada possui estrutura de dependência espacial anisotrópica (CRESSIE, 2015).

O processo da simulação consiste em definir as populações de interesse definindo os parâmetros para distribuição de probabilidade da variável georreferenciada, modelo da função semivariância, efeito pepita, contribuição, patamar, alcance, média, fator de anisotropia e ângulo de maior continuidade. Para cada simulação, foi considerada uma configuração amostral irregular totalizando 100 pontos com distâncias de 0 a 1 nos eixos O_x e O_y . Foram realizadas 100 simulações com parâmetros pré-definidos do modelo geoestatístico exponencial: Média (20), Efeito Pepita (0,0), Patamar (1,0) Alcance Prático (0,6). Além disso, em todas as configurações amostrais, considerou-se o ângulo de maior continuidade espacial (90° azimuth), fator de anisotropia (1 (isotropia); 2; 3 e 4).

O ajuste do modelo teórico à função semivariância foi realizado pela metodologia de estimação por máxima verossimilhança em cada simulação (Equação 24), o qual ocorreu em duas situações: considerando-se um modelo isotrópico da função semivariância e outro anisotrópico. Para cada situação, a predição de valores da variável georreferenciada em locais não amostrados foi realizada de acordo com a técnica de krigagem ordinária (Equação 33).

Além disso, para cada simulação, foi calculado o I de Moran direcional (Equação 10) para as direções de 0° e 90° (azimute) de acordo com as metodologias propostas por Rosenberg (2000) e Falsetti e Sokal (1993) para os valores amostrais. O I de Moran direcional variou a distância de corte (0,15; 0,30; 0,45; 0,60 e 0,75) e foi calculado na amostra com o intuito de verificar a existência de anisotropia geométrica, tal como foi pré-definido nos conjuntos de dados simulados, na direção de 90° (azimute). Foi verificado se os valores do índice de Moran direcional são elevados e significativos (teste de Pseudo significância), para maiores distâncias de cortes, quando comparadas as direções de maior (90°) e menor continuidade (0°) espacial.

As metodologias para o cálculo do I de Moran direcional foram utilizadas para os valores preditos em duas situações: ao se considerar um modelo isotrópico da função semivariância e outro anisotrópico, a fim de que fossem verificadas as diferenças direcionais que ocorrem quando a anisotropia é incorporada ao modelo geoestatístico. Para que fossem comparados os resultados do I de Moran direcional nas duas situações, primeiro foi comparada a porcentagem de simulações, nas quais foram obtidos valores significativos, depois foi realizada uma análise visual ao serem comparados os boxplots obtidos com os resultados simulados.

4.2 Dados agrícolas

4.2.1 Caracterização da área e análises geoestatísticas

Os dados agrícolas utilizados nesse trabalho foram coletados em uma área comercial de 167,35 ha, localizada no município de Cascavel, região Oeste do Paraná. A área possui localização geográfica aproximada de 24,95° Sul de latitude, 53,57° Oeste de longitude e altitude média de 650m em relação ao nível do mar.

O solo da região é classificado como latossolo vermelho distroférico de textura argilosa e apresenta solos profundos, com boa capacidade de retenção de água, aeração e permeabilidade (IAPAR, 2016). O clima da região é classificado como mesotérmico superúmido, tipo climático Cfa (koeppen) e temperatura anual média de 21 °C.

Os pontos amostrais foram georreferenciados com o auxílio do receptor (GPS) GEOEXPLORE 3, utilizando-se o sistema espacial de coordenadas UTM. No total, foram amostrados 102 pontos, considerando-se a amostragem *lattice plus close pairs* que consiste em uma grade regular, com distância mínima entre os pontos de 141 metros. E, de forma aleatória, foram escolhidos dezenove (19) locais para adição de pontos com distâncias menores (75 e 50m entre os pontos), como apresentados na Figura 4.

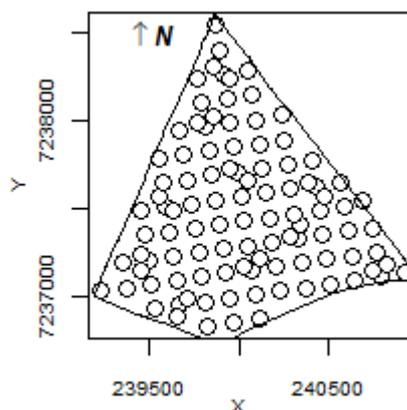


Figura 5 Área de estudo e localização dos pontos amostrais.

O experimento foi conduzido por pesquisadores do grupo de pesquisa do Laboratório de Estatística Espacial (LEE) e Laboratório de Estatística Aplicada (LEA) da Universidade Estadual do Oeste do Paraná – UNIOESTE, campus de Cascavel.

Os dados amostrais correspondem ao ano safra 2014/2015 e foram obtidos por análise química de rotina a partir de amostras coletadas em cada ponto demarcado. Foram coletadas cinco subamostras de solo, de 0,0 a 0,2 m de profundidade nas proximidades dos pontos, totalizando aproximadamente 500 g, armazenadas em sacos plásticos para compor a amostra representativa da parcela. As análises foram realizadas no laboratório de análise do solo da COODETEC (Cooperativa Central de Pesquisa Agrícola) para definir o teor de cada variável nas amostras coletadas.

A área foi cultivada com soja e os atributos químicos estudados foram: zinco (Zn) (mg dm^{-3}), ferro (Fe) (mg dm^{-3}), manganês (Mn) (mg dm^{-3}), fósforo (P) (mg dm^{-3}), carbono (C) (g dm^{-3}), cálcio (Ca) (cmolc dm^{-3}), Alumínio (Al) (cmolc dm^{-3}) e Potássio (K) (cmolc dm^{-3}). E, para os atributos químicos do solo, foram calculadas as medidas de posição, medidas de dispersão e as medidas de forma, os quais possibilitaram ao pesquisador uma visualização geral do comportamento da variável em estudo.

Cada variável em estudo foi submetida à análise geoestatística para se verificar a dependência espacial com o objetivo de encontrar variáveis anisotrópicas. E, para a identificação da anisotropia, analisou-se simultaneamente a tendência direcional a partir do gráfico postplot e do comportamento dos semivariogramas direcionais construídos com o estimador de Matheron (1963).

Posteriormente, foi realizado o ajuste dos modelos teóricos Esférico, Exponencial, Gaussiano e Família Matérn com $k=1$ pelo Método de Máxima Verossimilhança. O melhor modelo foi escolhido utilizando-se três critérios: validação cruzada (Equações 26, 27, 28, 29 e 30), informação de Akaike (Equação 31) e informação Bayesiana (Equação 32). A krigagem ordinária foi utilizada para interpolação dos dados (Equação 33) e, para cada variável anisotrópica, dois mapas foram construídos para a estimação do modelo da função semivariância: um modelo isotrópico e outro anisotrópico.

Por fim, foram calculadas as medidas de precisão Exatidão Global (Equação 36) e os índices de concordância Kappa e Tau (Equações 37 e 38).

Foi também calculado o I de Moran direcional (Equação 10) para os dados agrícolas amostrais e para os valores preditos pelos dois modelos citados acima. Para isso, foram utilizadas as metodologias proposta por Rosenberg (2000) e Falsetti e Sokal (1993).

Todas as simulações e análises dos dados foram feitas no software R (R DEVELOPMENT CORE TEAM, 2015) – pacote geoR (RIBEIRO JUNIOR e DIGLLE, 2001).

5 RESULTADOS

5.1 Estudos das simulações

Os resultados obtidos para as simulações, de acordo com o cálculo do índice de Moran direcional para os dados amostrais simulados com o método proposto por Rosenberg (2000) (Equação 10) calculado para as direções de (0° e 90°), são apresentados na Tabela 3 e Figura 6.

Para todos os fatores de anisotropia, a porcentagem de valores significativos do índice de Moran direcional nas simulações diminui conforme aumenta a distância de corte, na direção de 0° . O mesmo resultado não é observado na direção 90° , ou seja, mesmo para a distância 0,75 existiu uma autocorrelação espacial significativa, na direção 90° , na maioria das simulações (Tabela 3). Porém, quando se comparam as direções de interesse (0° e 90°), considerando-se um mesmo fator de anisotropia, o I de Moran direcional sofre uma taxa de decaimento mais suavizada na direção definida com maior continuidade espacial (90°).

Tabela 3 Porcentagem de simulações com valores significativos em cada distância de corte, nas simulações com modelo anisotrópico em 90° azimute, para os dados amostrais.

Fator de Anisotropia simulado (F_a)	Direção Azimute	% de valores significativos				
		Distância de Corte (d)				
		0,15	0,30	0,45	0,60	0,75
$F_a = 1$	0°	99	88	73	50	36
$F_a = 1$	90°	98	90	71	45	26
$F_a = 2$	0°	99	89	63	40	26
$F_a = 2$	90°	100	99	91	80	59
$F_a = 3$	0°	98	83	54	33	15
$F_a = 3$	90°	100	100	98	91	75
$F_a = 4$	0°	98	78	44	26	12
$F_a = 4$	90°	100	100	98	93	86

Nota: Significativo a 5%, por meio do teste de Pseudo significância.

Cada gráfico boxplot (Figura 6) representa os valores calculados do índice de Moran direcional proposto por Rosenberg (2000) para os dados amostrais para as direções 0° e 90° e distâncias de corte 0,15; 0,30; 0,45; 0,60 e 0,75 em 100 simulações agrupadas para cada fator de anisotropia.

Observa-se que, nas duas direções, conforme se aumentou a distância de corte, houve diminuição dos valores do I de Moran direcional em todos os fatores de anisotropia

considerados nas simulações. Isso pode ser explicado a partir do conceito básico de geoestatística de que dados mais próximos se parecem mais que dados mais distantes (VIEIRA, 2000). Tal conceito também explica a influência da distância de corte na magnitude do índice, qual seja, quanto menor a distância de corte maior é o resultado calculado para o índice de Moran, ou seja, maior é a autocorrelação espacial da variável georreferenciada. Nota-se ainda que o I de Moran direcional tende a se estabilizar próximo de zero conforme a distância de corte aumenta.

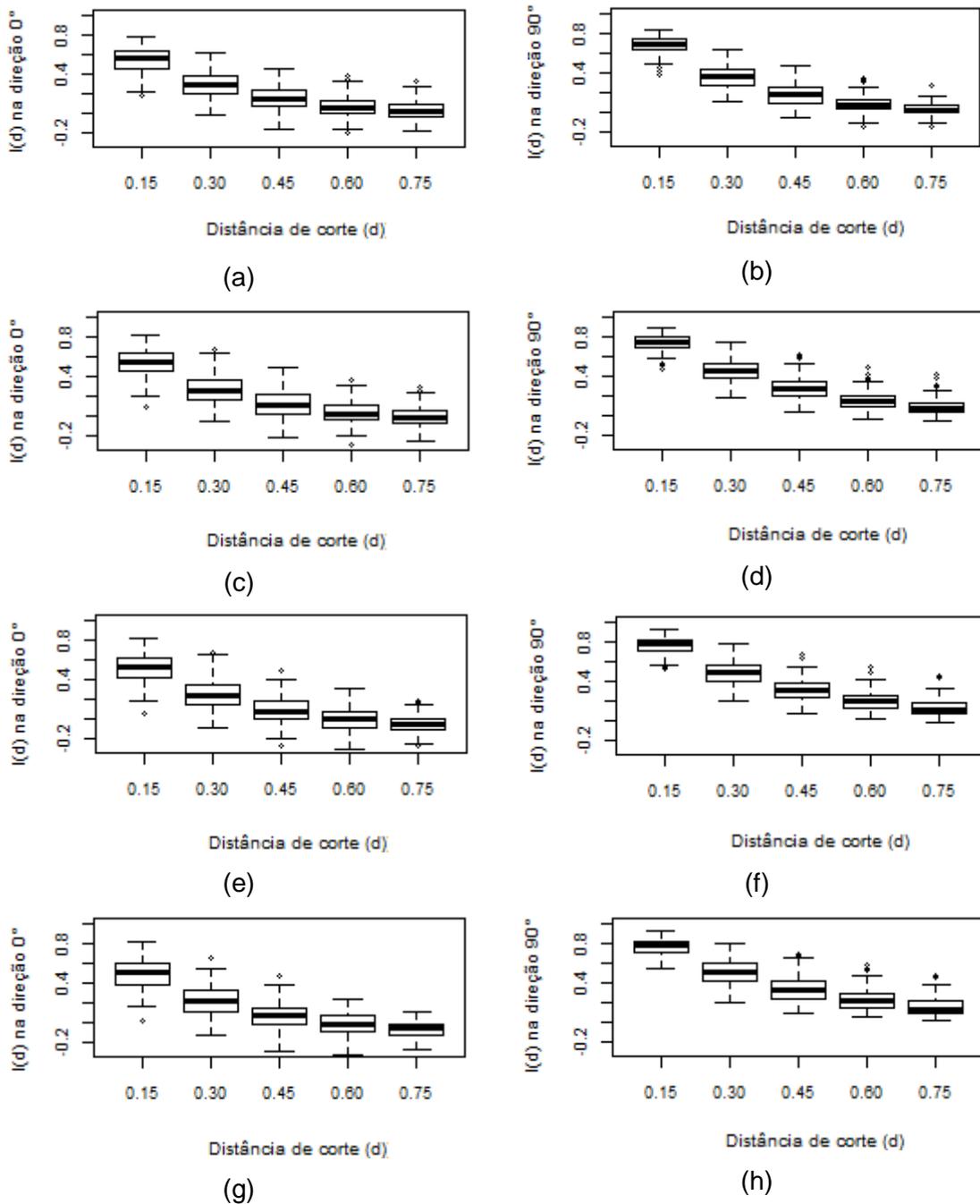


Figura 6 Boxplot I de Moran direcional calculado com os valores amostrados das simulações, considerando (F_α, α) : (a) $(1, 0^\circ)$; (b) $(1, 90^\circ)$; (c) $(2, 0^\circ)$; (d) $(2, 90^\circ)$; (e) $(3, 0^\circ)$; (f) $(3, 90^\circ)$; (g) $(4, 0^\circ)$ e (h) $(4, 90^\circ)$.

A partir dos valores do I de Moran direcional, tem-se que a direção de 90° apresentou os maiores valores calculados para todos os fatores de anisotropia simulados (Figuras 6d, 6f e 6h), mesmo com o aumento da distância de corte, assim, é possível que essa é a direção de maior continuidade espacial.

Além disso, quando os resultados obtidos são comparados entre as direções, para as simulações com estrutura de dependência espacial isotrópica ($F_a=1$) (Figura 6a e 6b), o comportamento foi semelhante nas duas direções, em relação aos valores obtidos do índice de Moran direcional para cada distância de corte, e em relação à taxa de decaimento dos valores do I de Moran direcional, conforme a distância de corte aumenta.

Quando se varia o fator de anisotropia (2, 3 e 4) (Figuras 6c a 6h), nota-se que: quanto maior o fator de anisotropia, menor foi a declinação sofrida na magnitude do I de Moran direcional, ou seja, menor foi a taxa de decaimento dos valores do I de Moran direcional quando se aumentou a distância de corte na direção de 90°.

A taxa de decaimento é menor na direção de maior continuidade espacial (90°) quando se aumenta F_a , mas em relação à direção de menor continuidade espacial (0°) para qualquer F_a , a taxa de decaimento do I de Moran direcional, quando se aumenta a distância de corte, é maior quando comparada com a direção de 90°.

Segundo Guedes et al (2013), quando a anisotropia geométrica é incorporada no modelo geoestatístico, nota-se a formação de sub-regiões nos mapas temáticos, as quais se estendem mais na direção de maior continuidade espacial. Assim, o Índice de Moran direcional foi calculado para os valores preditos, com intuito de expressar a diferença direcional quando a anisotropia geométrica foi incorporada no modelo geoestatístico.

Os resultados obtidos com a aplicação do Índice de Moran direcional proposto por Rosenberg para os valores preditos (Tabela 4 e Figura 7) foram semelhantes aos resultados obtidos para os dados amostrais (Tabela 3 e Figura 6). Dessa forma, para ambas as direções, o número de valores significativos apresentou taxa de decaimento ainda mais suavizada com o aumento da distância de corte, quando comparado aos resultados obtidos para os valores amostrais.

O I de Moran direcional proposto por Rosenberg (2000) apresenta valores observados mais elevados para a direção de maior continuidade espacial, tanto para o cálculo realizado com os valores amostrais simulados, quanto com os valores preditos. Ademais, todas as simulações realizadas com a aplicação do método apresentaram o comportamento similar.

Quando comparadas as simulações anisotrópica versus isotrópica, os resultados apresentam amplas diferenças. Todavia, para as análises isotrópicas, o valor calculado do I de Moran direcional se estabilizou próximo a zero para menores distâncias de corte, analisadas com os valores amostrais simulados e preditos.

Tabela 4 Porcentagem de simulações com valores significativos em cada distância de corte, nas simulações com modelo anisotrópico em 90° azimute, para os valores preditos.

Fator de Anisotropia simulado (F_a)	Direção Azimute	Número de valores significativos				
		Distância de Corte (d)				
		0,15	0,30	0,45	0,60	0,75
$F_a = 1$	0°	100	100	95	82	60
$F_a = 1$	90°	100	100	92	83	56
$F_a = 2$	0°	100	99	82	57	37
$F_a = 2$	90°	100	100	98	96	84
$F_a = 3$	0°	100	95	70	42	25
$F_a = 3$	90°	100	100	100	99	95
$F_a = 4$	0°	100	90	59	39	19
$F_a = 4$	90°	100	100	100	100	98

Nota: Significativo a 5%, por meio do teste de Pseudo significância.

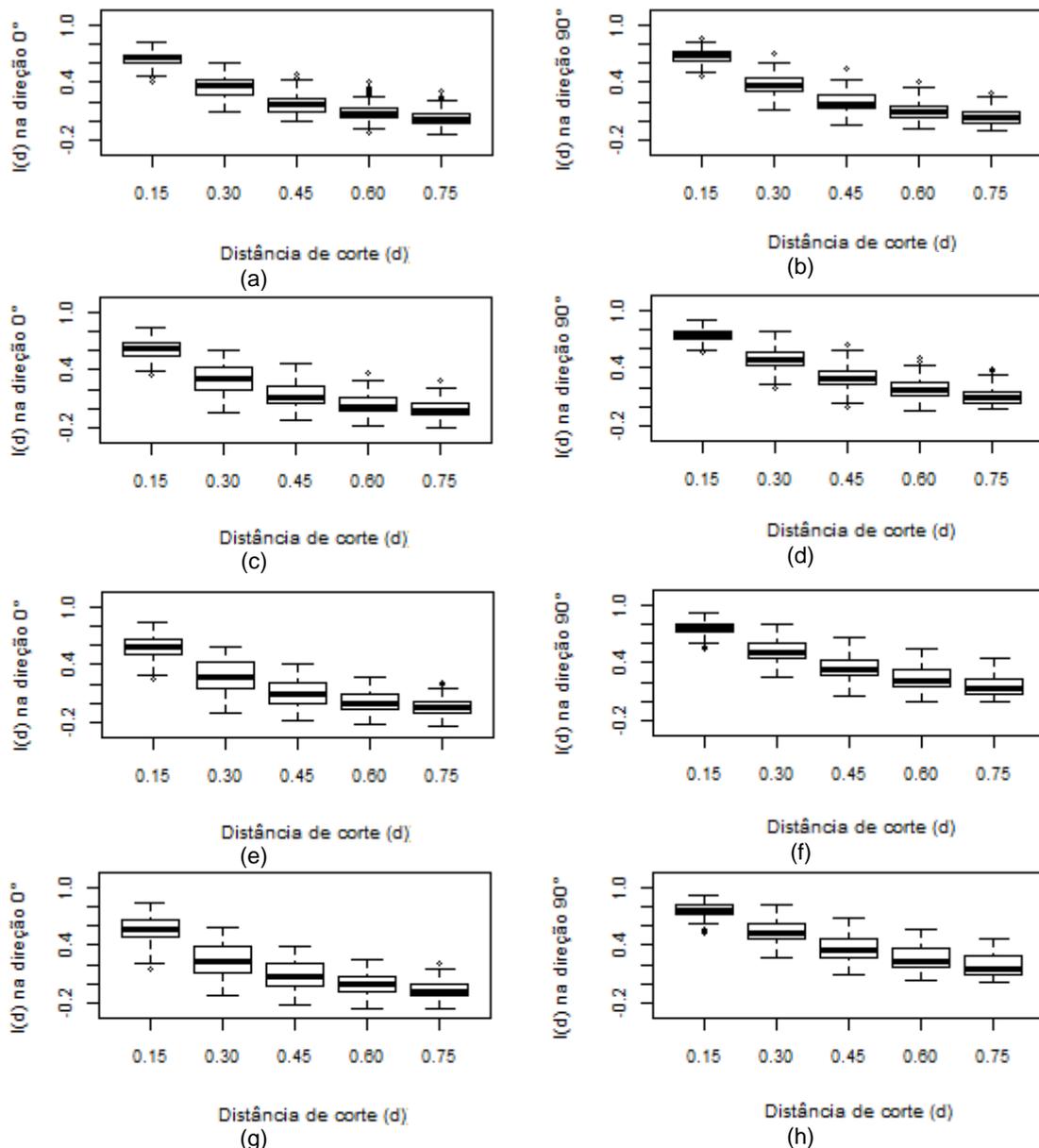


Figura 7 Boxplot I de Moran direcional calculado com os valores preditos das simulações, considerando (F_a , α): (a) (1, 0°); (b) (1, 90°); (c) (2, 0°); (d) (2, 90°); (e) (3, 0°); (f) (3, 90°); (g) (4, 0°) e (h) (4, 90°).

Os valores da autocorrelação espacial direcional calculados para os dados amostrais simulados, usando a técnica apresentada por Falsetti e Sokal (1993), são apresentados na Tabela 5 e Figura 7. A porcentagem de simulações com valores significativos do I de Moran direcional diminui com o aumento da distância de corte para a direção de 0° em todos os fatores de anisotropia, porém, para as simulações realizadas nos dados amostrais considerando a direção de 90°, a porcentagem de simulações com valores significativos aumentou na distância de corte 0,30 em todos os fatores de anisotropia. E, para maiores distâncias de cortes, o comportamento foi semelhante ao apresentado na direção de 0°. Pode-se observar que quando as porcentagens de valores significativos da Tabela 3 foram comparadas aos dados de Rosenberg (2000), as porcentagens registradas na Tabela 5 apresentaram respostas mais significativas.

A técnica de Falsetti e Sokal (1993) analisa os pares de pontos encontrados para uma mesma direção e dentro de uma tolerância angular, pois existe um aumento considerável nos números de pares analisados para distâncias de corte maiores, os quais possibilitam a obtenção de maiores números de valores observados significativos para o I de Moran direcional com o aumento da distância de corte, como mostra a Tabela 5.

Tabela 5 Porcentagem de simulações com valores significativos em cada distância de corte, nas simulações com modelo anisotrópico em 90° azimute, para os dados amostrais.

Fator de Anisotropia simulado (F_a)	Direção Azimute	Número de valores significativos				
		Distância de Corte (d)				
		0,15	0,30	0,45	0,60	0,75
$F_a = 1$	0°	62	59	46	35	37
$F_a = 1$	90°	30	34	23	21	14
$F_a = 2$	0°	50	42	30	24	19
$F_a = 2$	90°	61	75	63	51	44
$F_a = 3$	0°	44	36	20	13	9
$F_a = 3$	90°	77	95	85	69	62
$F_a = 4$	0°	43	30	13	10	4
$F_a = 4$	90°	86	98	94	85	76

Nota: Significativo a 5%, por meio do teste de Pseudo significância.

O comportamento do valor observado para I de Moran direcional calculado para os dados amostrais apresenta grande diferença quando comparado às direções de menor e maior continuidade espacial (0° e 90°, respectivamente Figura 8). Nota-se, ainda, que conforme se aumenta distância de corte, os valores do I de Moran direcional tendem a se estabilizar próximo a zero para ambas as direções. Logo, há indícios de maiores porcentagens de valores não significativos para maiores distâncias de cortes.

Para as análises que desconsideram os parâmetros de anisotropia (Figuras 8a e 8b), o comportamento é semelhante quanto às análises realizadas para os fatores de anisotropia 2, 3 e 4, ou seja, todas caminham em direção a zero. Contudo, quando são

adotados parâmetros de anisotropia, a diferença ocorre quanto à taxa de decaimento do I de Moran direcional, sendo que a direção de 0° se apresentou maior, quando comparada com a direção de 90° , principalmente para os fatores de anisotropia 2, 3 e 4. Os maiores valores calculados para o I de Moran direcional simulado foram obtidos na direção de maior continuidade espacial (90°) e, todas as simulações realizadas com a aplicação do método proposto por Falsetti e Sokal (1993) apresentaram resultados semelhantes.

Quando comparadas as simulações anisotrópica versus isotrópica, os resultados apresentam diferenças relevantes. Para as análises isotrópicas, o valor calculado do I de Moran direcional se estabilizou próximo a zero para menores distâncias de corte, indicando menor autocorrelação espacial direcional na variável georreferenciada. E a técnica de Falsetti e Sokal (1993) apresentou menor taxa de decaimento na direção de 90° , para todos os fatores de anisotropia quando comparada com Rosenberg (2000).

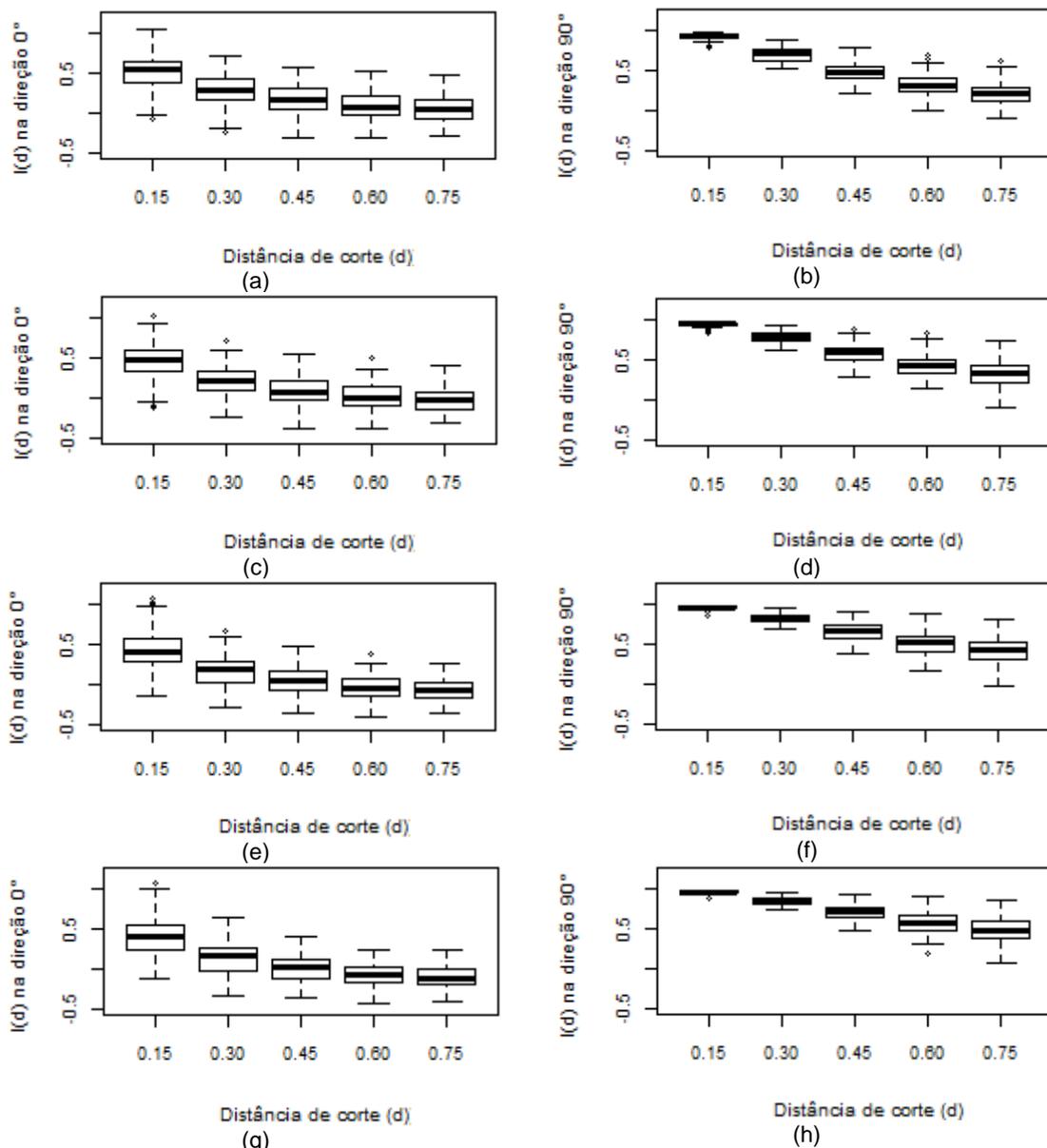


Figura 8 Boxplot I de Moran direcional calculado com os valores amostrados das simulações, considerando (F_a , α): (a) (1, 0°); (b) (1, 90°); (c) (2, 0°); (d) (2, 90°); (e) (3, 0°); (f) (3, 90°); (g) (4, 0°) e (h) (4, 90°).

Os resultados obtidos com a aplicação do Índice de Moran direcional proposto por Falsetti e Sokal para os valores preditos (Tabela 6 e Figura 9) foram semelhantes aos resultados obtidos para os dados amostrais (Tabela 5 e Figura 8). Assim, o número de valores significativos apresentou menor taxa de decaimento dos valores I de Moran direcional com o aumento da distância de corte, quando comparados aos resultados obtidos para os valores amostrais. Nota-se que o índice consegue descrever com clareza, de acordo com uma medida, as diferenças apresentadas quanto à existência de direções com regiões de maior ou menor continuidade espacial, conforme foi identificado visualmente em mapas temáticos por Guedes et al., (2013).

Na Tabela 6, é importante ressaltar que na direção de 90°, conforme o fator de anisotropia aumentou a porcentagem de valores significativos também aumentou, entretanto, para o fator de anisotropia 4, praticamente todas as distâncias de cortes obtiveram 100% dos valores do I de Moran direcional significativos, com exceção da distância de corte 0,75, que obteve 99 % das simulações com valores significativos.

Tabela 6 Porcentagem de simulações com valores significativos em cada distância de corte, nas simulações com modelo anisotrópico em 90° azimute, para os valores preditos.

Fator de Anisotropia simulado (F_a)	Direção Azimute	Número de valores significativos				
		Distância de Corte (d)				
		0,15	0,30	0,45	0,60	0,75
$F_a = 1$	0°	100	99	88	69	51
$F_a = 1$	90°	100	98	88	68	51
$F_a = 2$	0°	100	99	74	43	35
$F_a = 2$	90°	100	100	99	94	86
$F_a = 3$	0°	100	98	59	32	25
$F_a = 3$	90°	100	100	100	100	96
$F_a = 4$	0°	100	97	54	36	21
$F_a = 4$	90°	100	100	100	100	99

Nota: Significativo a 5%, por meio do teste de Pseudo significância.

Para a direção de maior continuidade espacial, a taxa de decaimento sofrida pelos valores do I de Moran direcional se apresentou menor conforme se aumentou o fator de anisotropia, indicando maior autocorrelação espacial entre os valores.

A técnica de Falsetti e Sokal (1993) apresentou menor taxa de decaimento conforme se aumentou a distância de corte, quanto às direções de 0 e 90° quando comparada com Rosenberg (2000).

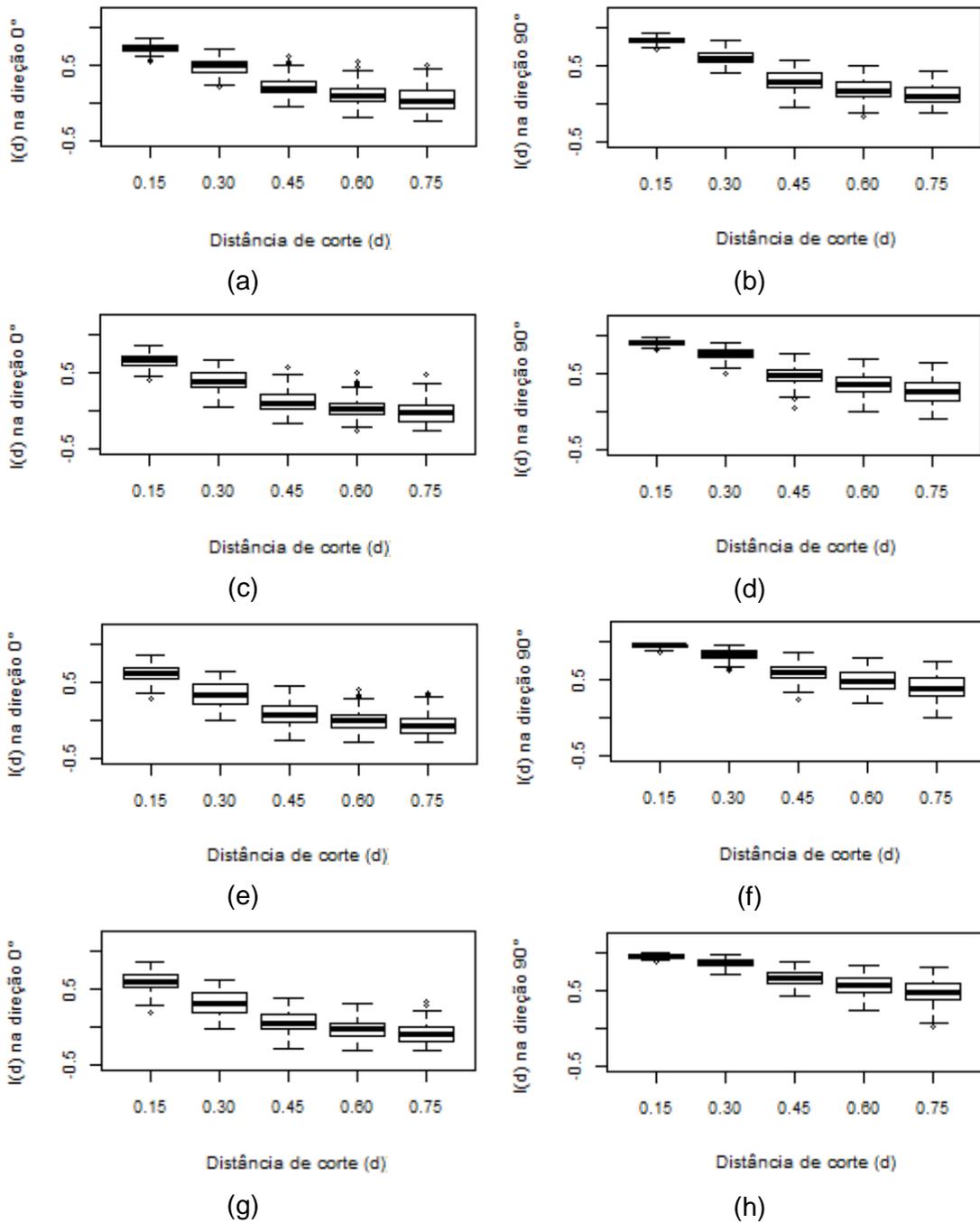


Figura 9 Boxplot I de Moran direcional calculado com os valores preditos das simulações, considerando (F_{α}, α) : (a) $(1, 0^{\circ})$; (b) $(1, 90^{\circ})$; (c) $(2, 0^{\circ})$; (d) $(2, 90^{\circ})$; (e) $(3, 0^{\circ})$; (f) $(3, 90^{\circ})$; (g) $(4, 0^{\circ})$ e (h) $(4, 90^{\circ})$.

5. 2 Estudo prático

5.2.1 Análise descritiva e espacial

Na Tabela 7, são apresentados as estatísticas descritivas e os valores do coeficiente linear de Pearson associado às coordenadas x e y, respectivamente, para os atributos

químicos Zinco (Zn), Ferro (Fe), Manganês (Mn), Fósforo (P), Carbono (C), Cálcio (Ca), Alumínio (Al) e Potássio (K).

Todos os atributos químicos do solo obtiveram valores próximos entre a média e mediana. Todos os valores foram baixos (próximo a zero) para os valores do coeficiente de correlação linear de Pearson indicando a ausência de associação linear destes atributos com os eixos x e y. Quanto ao coeficiente de variação, segundo Pimentel Gomes (2000), as variáveis ferro (Fe) e carbono (C) apresentaram dispersão média ($10\% \leq CV \leq 20\%$). Enquanto as variáveis manganês (Mn) e o cálcio (Ca) apresentaram alta dispersão ($20\% \leq CV \leq 30\%$). As variáveis zinco (Zn), Fósforo (P), alumínio (Al) e potássio (K) apresentaram variação muito alta ($CV > 30\%$) em relação às suas respectivas médias.

Tabela 7 Análise exploratória descritiva dos atributos químicos solo: (a) Zinco Zn (mg dm^{-3}), (b) Ferro Fe (mg dm^{-3}), (c) Manganês Mn (mg dm^{-3}), (d) Fósforo P (mg dm^{-3}), (e) Carbono C (g dm^{-3}), (f) Cálcio Ca (cmolc dm^{-3}), (g) Alumínio Al (cmolc dm^{-3}) e (h) Potássio K (cmolc dm^{-3}).

	Zn	Fe	Mn	P	C	Ca	Al	K
Média	2,82	57,76	76,54	16,29	29,42	5,39	0,28	0,30
Valor mínimo	0,50	29,09	38,79	3,40	22,40	2,25	0,00	0,10
1º Quartil	1,75	49,96	62,29	11,52	27,48	4,47	0,07	0,20
Mediana	2,36	57,40	71,40	16,90	29,33	5,32	0,16	0,26
3º Quartil	3,37	63,12	88,40	23,82	31,32	6,11	0,37	0,38
Valor máximo	11,04	103,20	136,80	60,00	45,22	8,76	2,01	0,67
Variância	3,01	121,89	440,97	123,31	13,90	6,11	0,13	0,08
CV (%)	61,61	19,11	27,43	57,57	12,67	25,18	126,28	45,04
r(x)	0,31	-0,10	0,08	-0,09	0,23	0,22	-0,15	0,09
r(y)	0,04	0,15	-0,02	0,01	-0,11	0,03	0,18	-0,26

Nota: r(x) e r(y) coeficiente de correlação linear de Pearson dos atributos em relação respectivamente ao eixo x e ao eixo y. CV coeficiente de variação.

A análise dos pontos discrepantes foi realizada por gráfico boxplot (Figura 10). O gráfico evidencia que todos os atributos químicos do solo apresentam pontos discrepantes. Quanto à distribuição dos dados, os atributos Zinco (Figura 10 a), Manganês (Figura 10 c) e Alumínio (Figura 10 g) apresentaram índice de distribuição assimétrica à direita e as variáveis Ferro (Figura 10 b), Fósforo (Figura 10 d), Carbono (Figura 10 e), Cálcio (Figura 10 f) e Potássio (Figura 10 h) apresentaram índices de distribuição simétrica.

Foram identificados os seguintes pontos discrepantes: Zn (pontos 100; 70; 47; 24 e 32, com valores respectivos de 11,04; 9,13; 8,03; 7,64 e 6,26 mg dm^{-3}). Três deles estão localizados na região oeste, um ao norte e um na região central; Fe (pontos 96; 78 e 34, com valores respectivos de 103,22; 29,09 e 89,64 mg dm^{-3}), nas regiões norte, leste e central; Mn (pontos 97; 7; 5 e 4, com valores respectivos de 128,60; 135,97; 136,76 e 128,36 mg dm^{-3}), e três pontos vizinhos estão localizados na região sul e um ponto localizado ao norte; P (pontos 86; 71; 70 e 64, com valores respectivos de 60; 58,60; 52,50 e 50,30 mg dm^{-3}), localizados nas regiões leste, central e oeste; C (pontos 83 e 73, com

valores respectivos de 45,22 e 38,50 g dm⁻³), localizados na região central; Ca (ponto 83, com valor respectivo de 8,76 cmolc dm⁻³), localizados na região central; Al (pontos 101; 95; 81; 80; 71; 52; 49; 34; 21 e 13, com valores respectivos de 2,01; 1,24; 1,21; 1,05; 0,85; 0,97; 1,17; 0,84; 1,02 e 1,46 cmolc dm⁻³), com seis pontos vizinhos ao leste, dois pontos ao norte, um ponto na região central e um ponto na região oeste; K (ponto 8, com valores respectivos de 0,67 cmolc dm⁻³), localizados na região sul.

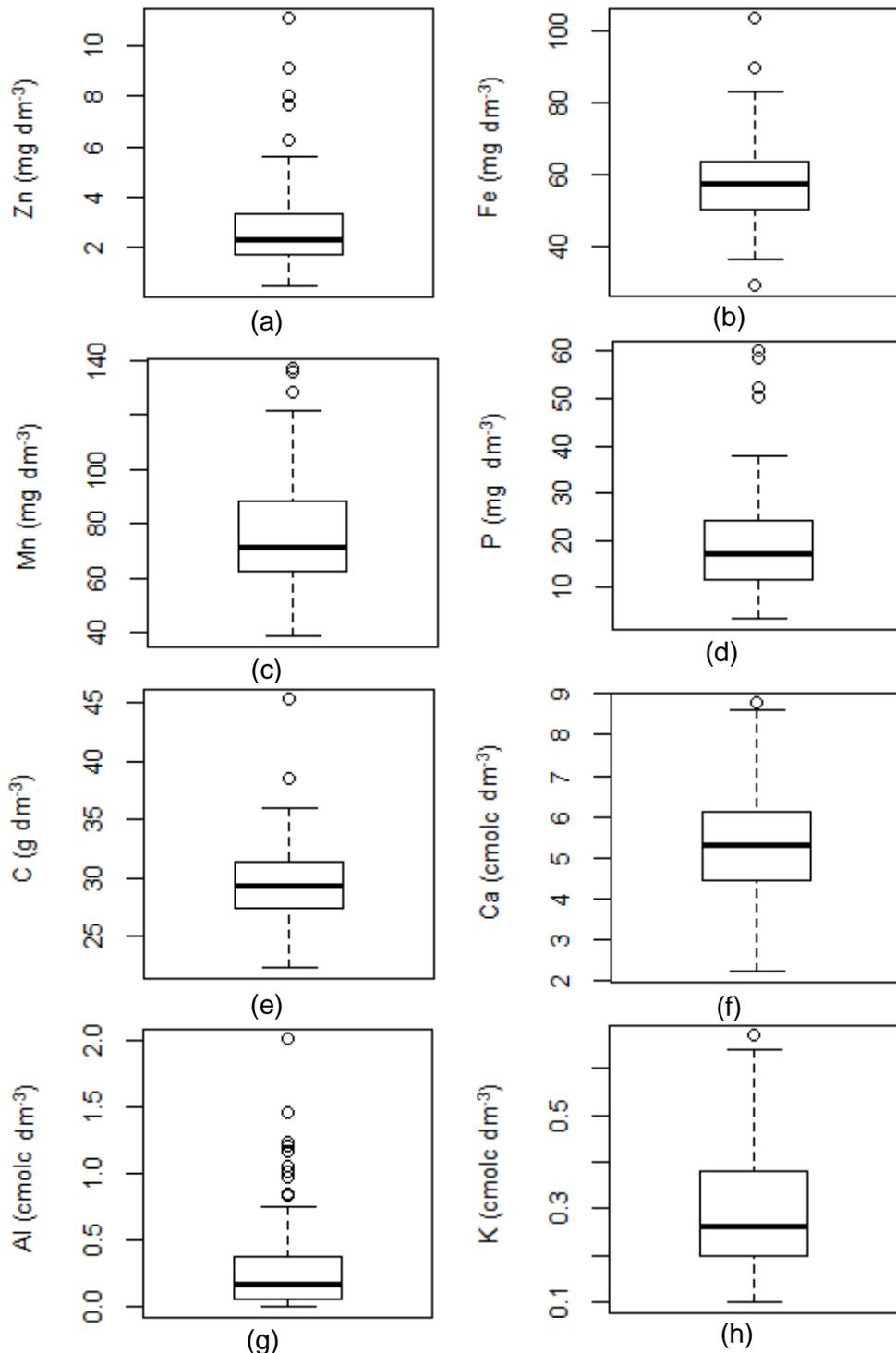


Figura 10 Gráfico boxplot dos atributos químicos do solo: (a) Zinco Zn (mg dm⁻³), (b) Ferro Fe (mg dm⁻³), (c) Manganês Mn (mg dm⁻³), (d) Fósforo P (mg dm⁻³), (e) Carbono C (g dm⁻³), (f) Cálcio Ca (cmolc dm⁻³), (g) Alumínio Al (cmolc dm⁻³) e (h) Potássio K (cmolc dm⁻³).

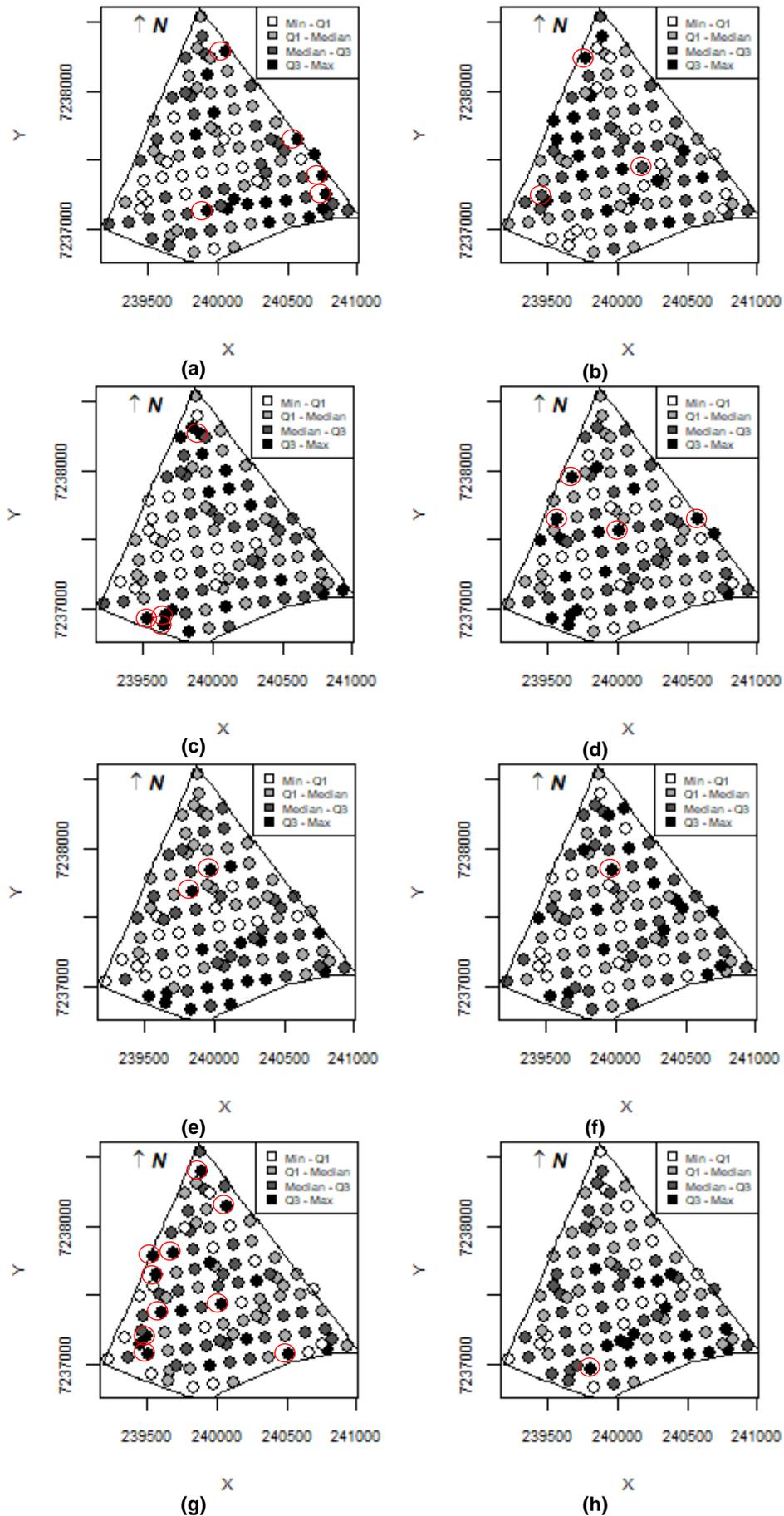


Figura 11 Gráficos post-plot dos atributos químicos do solo: (a) Zinco Zn, (b) Ferro Fe, (c) Manganês, (d) Fósforo P, (e) Carbono C, (f) Cálcio Ca, (g) Alumínio Al e (h) Potássio K. Círculo vermelho indica os pontos discrepantes identificados pelo gráfico box plot.

Iniciou-se a análise de tendência direcional quanto à continuidade espacial com objetivo de encontrar quais variáveis possuem estrutura de dependência espacial anisotrópica, a partir da construção do gráfico espacial da área em estudo (Figura 11), denominado de Postplot. Observa-se que não foi identificado qualquer agrupamento de valores semelhantes, os quais se estenderam especificamente em uma determinada direção. Tais resultados indicam que os atributos químicos do solo (Zn, Fe, Mn, P e Al) não apresentam tendência direcional (Figuras 11 a; b; c; d e g).

Nos atributos C, Ca e K (Figuras 11 d; f e h, respectivamente) ocorre um oscilante agrupamento de valores semelhantes na direção de 90°, o que indica a presença de tendência direcional quanto à continuidade espacial nesses atributos. Para identificar tal tendência direcional, foram calculados para os dados amostrais destes atributos, medidas de autocorrelação espacial direcional, variando a distância de corte e o ângulo de interesse.

Os resultados do índice I(d), calculados para os atributos C, Ca e K (Tabelas 8 e 9) apresentaram maiores valores para direção de 90°. Logo, confirmam-se os resultados apresentados nos gráficos postplot. Ainda, na metodologia de Falsetti e Sokal (1993), os valores continuaram mais elevados quando comparados com os dados de Rosenberg.

Tabela 8 Índice de Moran direcional para os dados amostrais Carbono (C).

Variáveis	Ângulo de interesse	Distância de Corte	Rosenberg (2000)		Falsetti e Sokal (1993)		
			I de Moran direcional	P valor	I de Moran direcional	P valor	Nº de Pares
Carbono (C)	0°	150	0,053	0,245	0,012	0,444	71
	90°	150	0,522	0,002*	0,606	0,166	116
	45°	150	0,340	0,017*	-0,072	0,603	11
	135°	150	0,328	0,643	0,276	0,097	16
	0°	300	0,090	0,037*	0,024	0,329	147
	90°	300	0,306	0,004*	0,489	0,177	176
	45°	300	0,256	0,001*	0,177	0,050	94
	135°	300	0,155	0,041*	-0,127	0,862	93
	0°	450	0,007	0,279	-0,038	0,682	223
	90°	450	0,171	0,003*	0,469	0,064	229
	45°	450	0,163	0,001*	0,210	0,002*	171
	135°	450	0,017	0,648	-0,112	0,923	202
	0°	600	-0,019	0,643	-0,056	0,789	326
	90°	600	0,110	0,015*	0,409	0,052	279
	45°	600	0,108	0,003*	0,167	0,011*	256
	135°	600	-0,017	0,954	-0,148	0,992	296
	0°	750	-0,030	0,901	-0,053	0,790	408
	90°	750	0,066	0,067	0,383	0,033*	324
	45°	750	0,076	0,002*	0,095	0,049*	327
	135°	750	-0,046	0,998	-0,157	0,997	400

Nota: asterisco (*) indica valores significativos a 5% de probabilidade pelo teste de Pseudo significância do I de Moran direcional.

Tabela 9 Índice de Moran direcional para os dados amostrais Cálcio (Ca) e Potássio (K).

Variáveis	Ângulo de interesse	Distância de Corte	Rosenberg (2000)		Falsetti e Sokal (1993)		Nº de Pares
			I de Moran direcional	P valor	I de Moran direcional	P valor	
Cálcio (Ca)	0°	150	0,008	0,427	0,028	0,364	71
	90°	150	0,452	0,028*	0,596	0,189	116
	45°	150	0,273	0,133	0,125	0,316	11
	135°	150	0,293	0,179	0,435	0,036	16
	0°	300	-0,010	0,494	-0,009	0,506	147
	90°	300	0,236	0,035*	0,504	0,070	176
	45°	300	0,137	0,081	0,095	0,169	94
	135°	300	0,108	0,191	-0,080	0,740	93
	0°	450	-0,018	0,595	0,014	0,369	223
	90°	450	0,126	0,044*	0,426	0,087	229
	45°	450	0,054	0,186	0,039	0,233	171
	135°	450	0,040	0,322	-0,044	0,668	202
	0°	600	0,009	0,179	0,010	0,337	326
	90°	600	0,102	0,019*	0,350	0,134	279
	45°	600	0,027	0,281	0,010	0,380	256
	135°	600	0,065	0,03*	-0,039	0,699	296
	0°	750	0,002	0,223	0,011	0,307	408
	90°	750	0,063	0,043*	0,253	0,395	324
	45°	750	0,008	0,538	-0,041	0,671	327
	135°	750	0,037	0,051	-0,044	0,740	400
Potássio (K)	0°	150	0,009	0,389	0,052	0,284	71
	90°	150	0,394	0,196	0,587	0,214	116
	45°	150	0,282	0,089	0,019	0,467	11
	135°	150	0,244	0,200	-0,130	0,650	16
	0°	300	0,044	0,158	0,033	0,303	147
	90°	300	0,296	0,003*	0,555	0,016*	176
	45°	300	0,205	0,006*	0,106	0,154	94
	135°	300	0,166	0,024*	0,116	0,137	93
	0°	450	0,003	0,278	0,017	0,352	223
	90°	450	0,155	0,01*	0,494	0,013*	229
	45°	450	0,109	0,012*	0,065	0,182	171
	135°	450	0,057	0,158	0,048	0,226	202
	0°	600	0,008	0,200	0,017	0,315	326
	90°	600	0,126	0,005*	0,395	0,023*	279
	45°	600	0,090	0,011*	0,033	0,284	256
	135°	600	0,045	0,089	-0,004	0,439	296
	0°	750	0,015	0,074	0,007	0,356	408
	90°	750	0,101	0,006*	0,354	0,033*	324
	45°	750	0,076	0,006*	0,020	0,334	327
	135°	750	0,038	0,073	0,034	0,191	400

Nota: asterisco (*) indica valores significativos a 5% de probabilidade pelo teste de Pseudo significância do I de Moran direcional.

A identificação de forma exploratória da existência da continuidade espacial de cada atributo químico do solo foi obtida pela técnica dos envelopes, proposta por Diggle e Ribeiro Junior (2007). Os semivariogramas de envelopes (Figura 12) foram construídos apenas para o atributo Potássio (Figura 12 h) e o semivariograma não apresentou pontos fora ou na linha do envelope, o que sugere indício de não existência de dependência espacial.

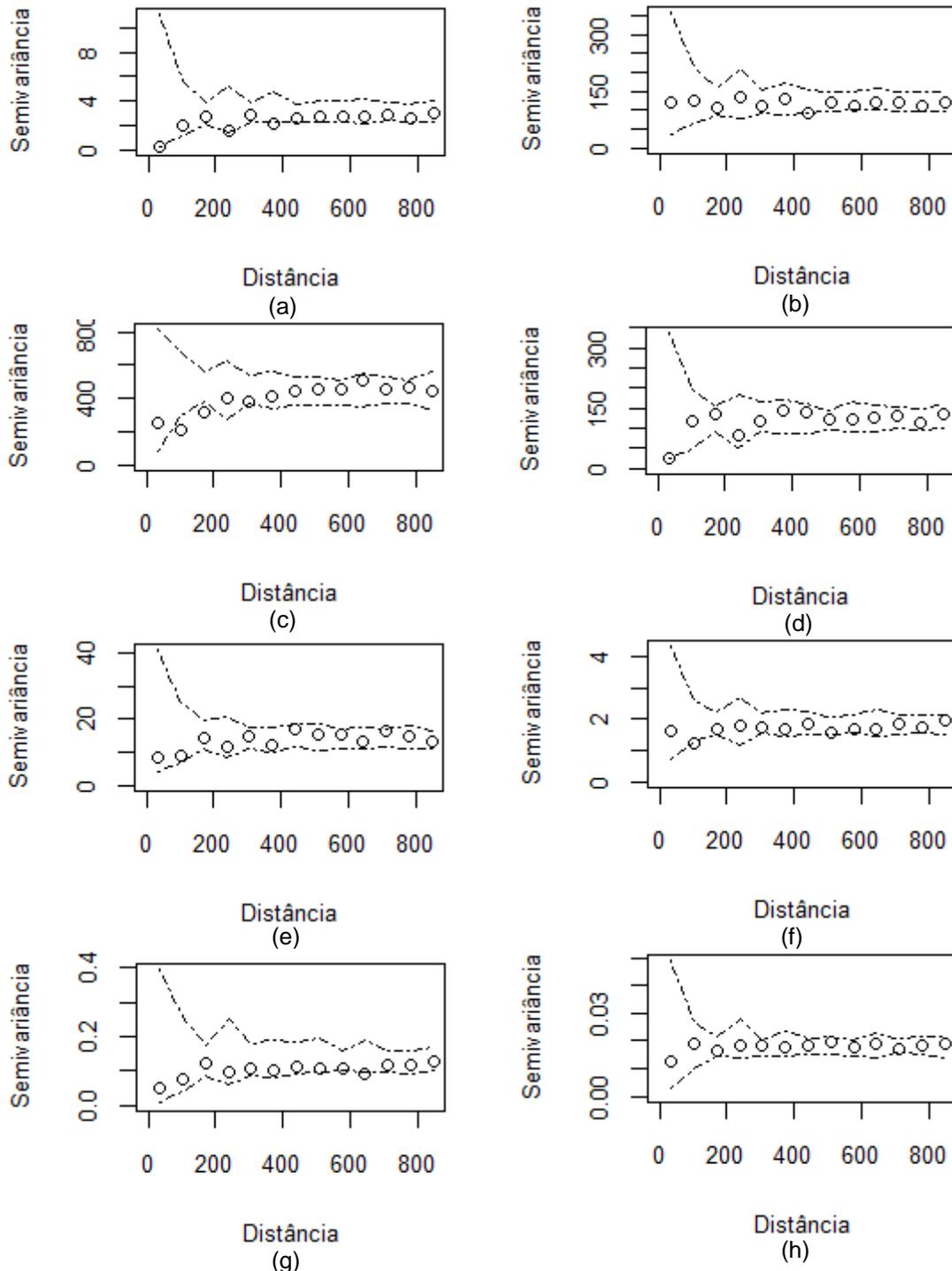


Figura 12 Gráfico de envelopes das variáveis: (a) Zinco Zn (mg dm^{-3}), (b) Ferro Fe (mg dm^{-3}), (c) Manganês Mn (mg dm^{-3}), (d) Fósforo P (mg dm^{-3}), (e) Carbono C (g dm^{-3}), (f) Cálcio Ca (cmolc dm^{-3}), (g) Alumínio Al (cmolc dm^{-3}) e (h) Potássio K (cmolc dm^{-3}).

Os semivariogramas direcionais foram construídos dois a dois, com base nas direções ortogonais (0° e 90° , 45° e 135°), para facilitar a análise visual daqueles. Assim, observou-se nos atributos químicos do solo Zinco (Figuras 13 a, e b), Ferro (Figuras 13 c, e d), Alumínio (Figuras 13 e, e f) e Fósforo (Figuras 13 g, e h) comportamento semelhante nas direções ortogonais, nas quatro direções analisadas, indicando que a estrutura de dependência espacial destas variáveis é isotrópica.

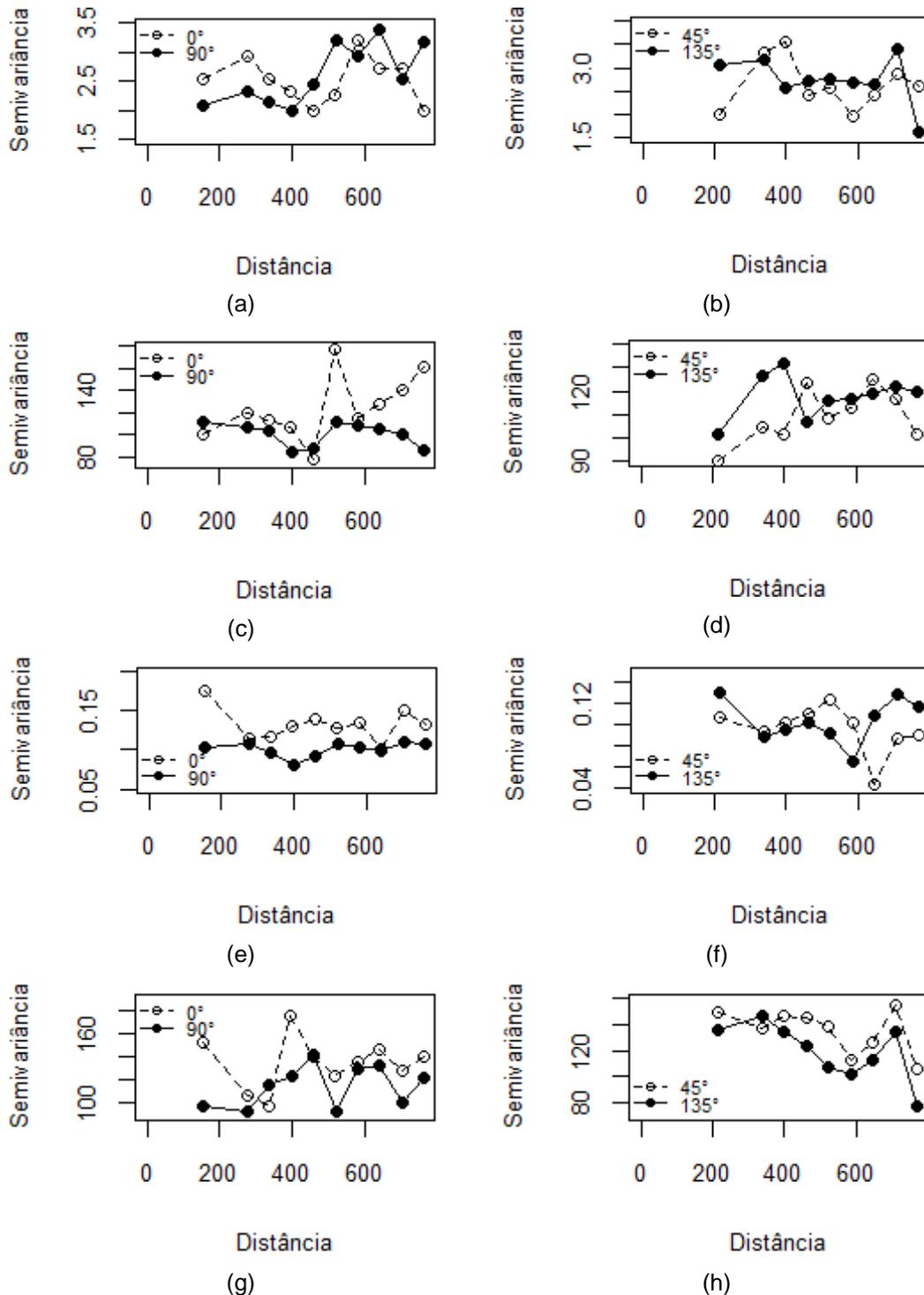


Figura 13 Semivariogramas direcionais dos seguintes atributos químicos do solo: (a) e (b) Zinco, (c) e (d) Ferro, (e) e (f) Alumínio e (g) e (h) Fósforo.

A análise visual dos semivariogramas direcionais foi associada às análises dos parâmetros para os seguintes atributos químicos do solo: Manganês (Figuras 14 a, e b), Carbono (Figuras 14 c, e d), Cálcio (Figuras 14 e, e f) e Potássio (Figuras 14 g, e h) (Tabela 10).

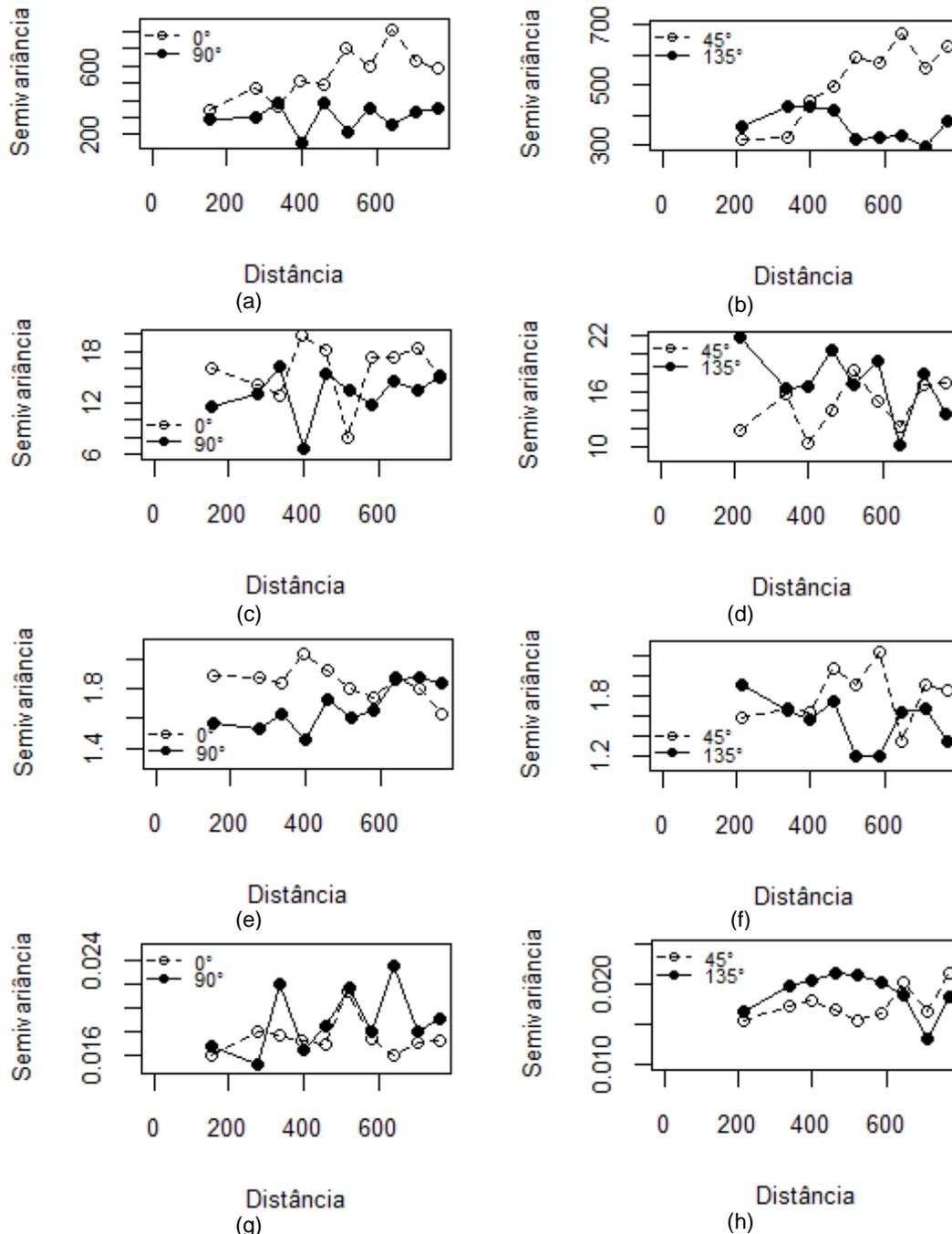


Figura 14 Semivariogramas direcionais dos seguintes atributos químicos do solo: (a) e (b) Manganês, (c) e (d) Carbono, (e) e (f) Cálcio e (g) e (h) Potássio.

Ainda na Figura 14, para o atributo manganês (Figura 14 a e b), foram detectadas diferenças nos alcances analisados, conseqüentemente foram obtidos índices de anisotropia combinada. Porém, o valor estimado foi obtido para o fator de anisotropia quando se estima

o modelo com anisotropia geométrica na direção de 90°. Toma-se essa direção como a de maior continuidade espacial, tal qual foi indicada pelo postplot e pelas medidas de autocorrelação espacial direcional para esse atributo, F_a foi considerado baixo ($F_a = 1,431$), desconsiderando-se assim a presença de anisotropia (Tabela 10). Os atributos Carbono, Cálcio e Potássio confirmaram os indícios de maior continuidade espacial na direção de 90°, (Figuras 14 c a h) (Tabela 10). O mesmo foi verificado nos gráficos postplot (Figura 11) e também pelos valores no I de Moran direcional (Tabelas 8 e 9).

O melhor modelo ajustado foi definido por critérios de validação cruzada, Akaike (AIC) e Bayesiano (BIC) (Apêndice 1), cujo modelo exponencial foi o que melhor se ajustou para os atributos zinco e alumínio. Por outro lado, o modelo esférico foi o que melhor se ajustou para ferro, fósforo e manganês e o modelo gaussiano foi considerado como o melhor para carbono, cálcio e potássio.

Os parâmetros obtidos dos semivariogramas direcionais para estes atributos apresentaram diferenças entre os alcances, ambos para direção ortogonal (0° e 90°), caracterizando-se anisotropia geométrica com ângulo de maior continuidade espacial na direção de 90° e fatores de anisotropia iguais a $F_a = 6,08$ para Carbono, $F_a = 4,55$ para Cálcio e $F_a = 3,754$ para potássio Tabela 10.

A Tabela 10 também apresenta os valores obtidos do melhor modelo ajustado para os parâmetros: média (μ), efeito pepita (φ_1), contribuição (φ_2), Patamar ($\varphi_1 + \varphi_2$), alcance prático (a) e efeito pepita relativo ($EPR = \varphi_1 / (\varphi_1 + \varphi_2) \times 100$), que medem o grau de dependência do atributo na área em estudo.

Tabela 10 Valores estimados dos parâmetros dos modelos geoestatísticos isotrópicos para os atributos químicos do solo Zn, Fe, Al, P e Mn e anisotrópicos para os atributos C, Ca e K, considerando como 90° a direção de anisotropia geométrica.

Var.	Mod.	Média (μ)	($\hat{\varphi}_1$)	($\hat{\varphi}_2$)	$\hat{\varphi}_1 + \hat{\varphi}_2$	F_a	\hat{a}	EPR
Zn	Exp	2,944	0,000	3,127	3,127	1,000	270,542	0,00%
Fe	Sph	57,981	97,348	22,992	120,340	1,000	335,245	79,06%
Al	Exp	0,283	0,126	0,000	0,126	1,000	255,904	100,0%
P	Sph	19,300	111,610	10,520	122,130	1,000	199,244	91,39%
Mn	Sph	76,480	109,720	299,810	409,530	1,000	292,433	29,79%
C	gaus	29,443	9,027	4,497	13,524	6,078	149,575	66,75%
Ca	gaus	5,388	1,401	0,390	1,791	4,550	175,962	78,24%
K	gaus	0,292	0,015	0,003	0,018	3,754	371,890	81,67%

Nota: Efeito pepita ($\hat{\varphi}_1$), contribuição ($\hat{\varphi}_2$), Patamar $\hat{\varphi}_1 + \hat{\varphi}_2$, fator de anisotropia (F_a), alcance prático (\hat{a}) e efeito pepita relativo ($EPR = \varphi_1 / (\varphi_1 + \varphi_2) \times 100$).

Definidos os melhores modelos teóricos ajustados em todas as variáveis, realizou-se a interpolação a partir da técnica de krigagem ordinária. Após, foram construídos os mapas temáticos (Figura 14 e 15), os quais representam o comportamento dos atributos do solo na área em estudo. Vale ressaltar que as regiões mais escuras nestes mapas indicam teores

elevados e as regiões mais claras indicam teores mais baixos dos atributos químicos no solo estudado. Na Figura 14, os mapas temáticos dos atributos Zn, Fe, Al, P e Mn foram construídos considerando no ajuste um modelo com estrutura de dependência espacial isotrópica, portanto, observa-se que há um comportamento similar para continuidade espacial das sub-regiões, que apresentam formato circular.

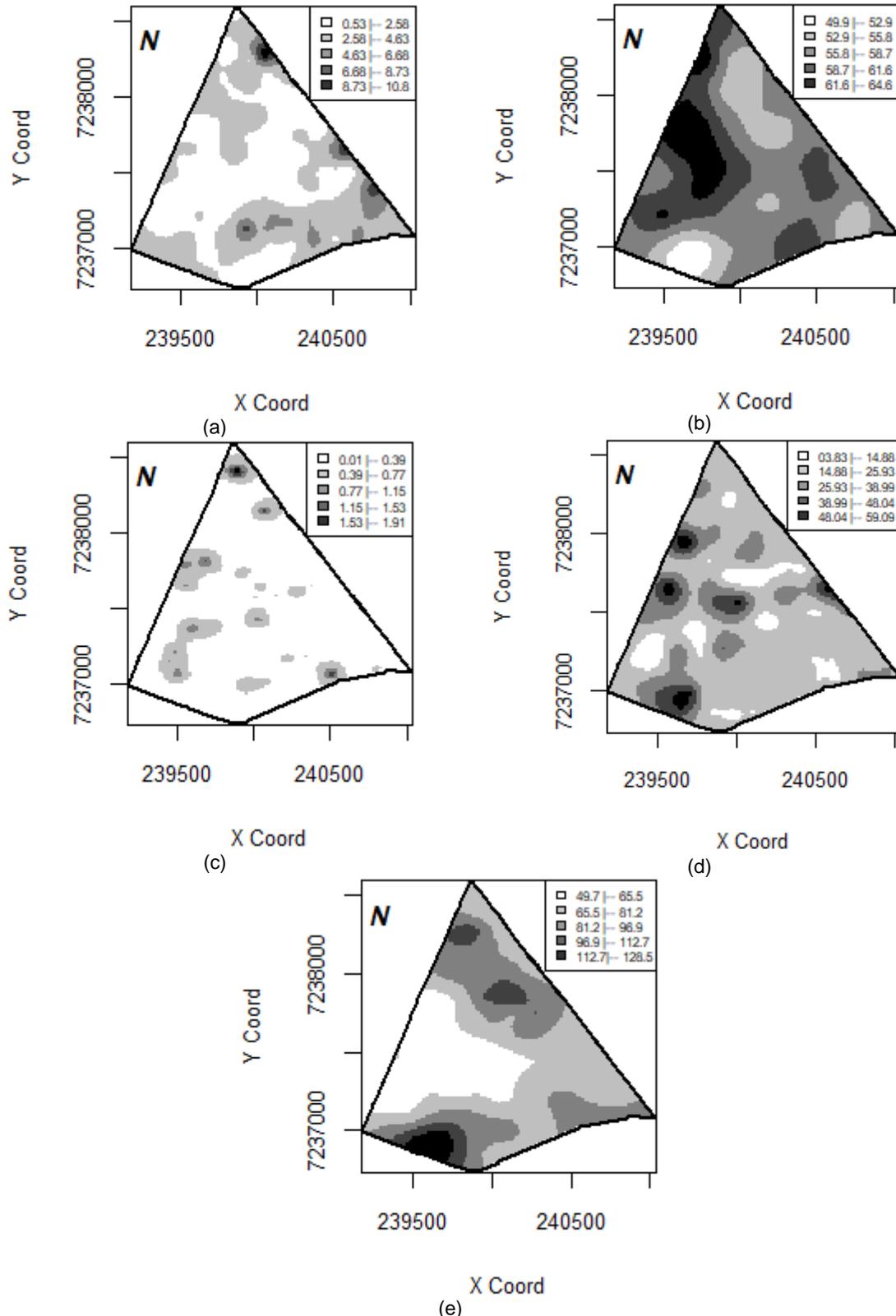


Figura 15 Mapa temático das variáveis: (a) Zinco (mg dm^{-3}), (b) Ferro (mg dm^{-3}), (c) Alumínio (cmolc dm^{-3}), (d) Fósforo (mg dm^{-3}), (e) Manganês (mg dm^{-3}), considerando o modelo isotrópico estimado.

Os comportamentos das sub-regiões apresentaram maior continuidade espacial para a direção 90° nos mapas temáticos construídos para os atributos com estrutura de dependência espacial anisotrópica geométrica (Figuras 15 b, d e f), em formato elíptico para as sub-regiões. O comportamento das sub-regiões em variáveis isotrópica e anisotrópica já foi descrito na literatura por (GUEDES et al., 2013).

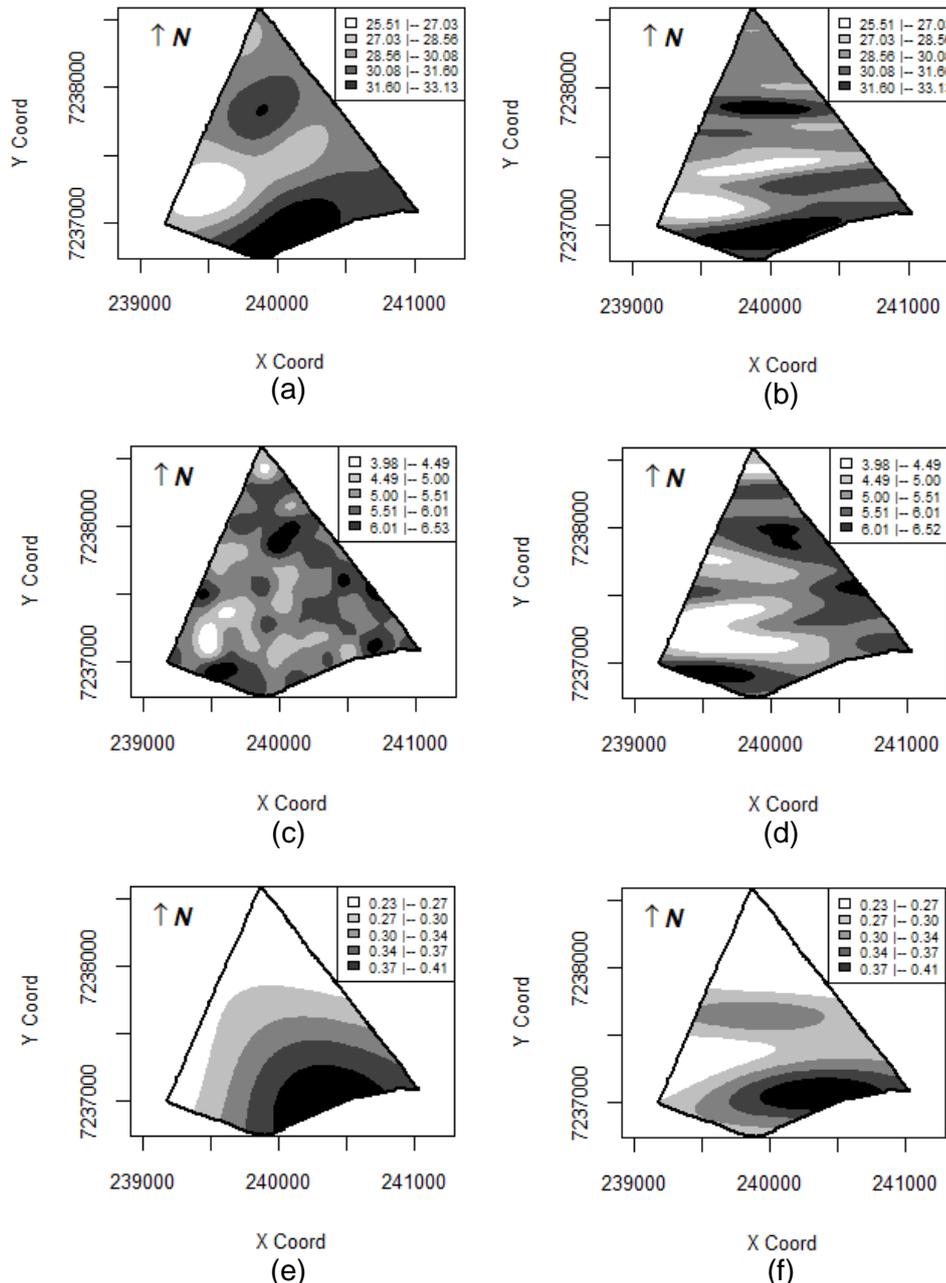


Figura 16 Mapa temático dos seguintes atributos químicos do solo: considerando-se o modelo geoestatístico estimado isotrópico (a) Carbono, (c) Cálcio, (e) Potássio, e considerando um modelo geoestatístico estimado anisotrópico (b) Carbono, (d) Cálcio, (f) Potássio.

Mapas temáticos considerando-se os modelos isotrópicos para os atributos carbono, cálcio e potássio (Figuras 16 a; c; e, respectivamente) foram construídos a fim de se ressaltar a importância da incorporação da anisotropia geométrica em variáveis

georreferenciadas. E, para se considerar a variável isotrópica, os parâmetros F_a e α foram retirados da função de máxima verossimilhança. Observou-se, também, que as sub-regiões passaram a apresentar forma circular, indicando diferenças visuais relevantes quanto à incorporação ou não da anisotropia geométrica.

A média Exatidão Global e os índices de concordância Kappa e Tau foram calculados para quantificar as diferenças apresentadas entre os mapas temáticos considerando anisotropia geométrica (Figuras 16 b, d e f), e os mapas temáticos considerando a isotropia (Figuras 16 a; c; e), os quais foram construídos para os atributos químicos carbono, cálcio e potássio (Tabela 11). Esses, por sua vez, indicaram baixa precisão quanto à similaridade entre os mapas anisotrópicos e isotrópicos construídos para carbono (C), cálcio (Ca) e Potássio (K), com valores da Exatidão Global menores que 0,85 (ANDERSON et al., 1976). Os valores obtidos para os índices de concordância Kappa e Tau indicaram baixa exatidão quanto à comparação dos mapas. Assim, pode-se verificar que existem diferenças na classificação dos mapas temáticos quanto à incorporação ou não da anisotropia geométrica.

Tabela 11 Medidas de acurácia para comparação dos mapas temáticos considerando modelos estimados anisotrópicos e isotrópicos para os atributos químicos do solo carbono, cálcio e potássio.

Variáveis	Exatidão Global	Kappa	Tau
Carbono	0.51	0.31	0.38
Cálcio	0.58	0.38	0.47
Potássio	0.48	0.32	0.36

As diferenças apresentadas na estimação espacial dos atributos químicos do solo em estudo quanto à incorporação ou não da anisotropia geométrica foram analisadas visualmente nos mapas temáticos (Figura 16) e comprovadas com os baixos resultados obtidos nas aplicações das medidas de acurácia. Assim, utilizou-se o I de Moran para verificar diferenças na estimação desses atributos químicos do solo para uma direção de interesse.

A Tabela 12 apresenta os valores do I de Moran direcional, calculados para os valores preditos nos mapas temáticos, considerando-se ou não a anisotropia dos atributos carbono, cálcio e potássio.

Nota-se para todos os atributos químicos do solo analisados que praticamente todos os valores do I de Moran direcional obtidos foram significativos a 5% de probabilidade. O I de Moran direcional apresentou valores mais elevados na direção de maior continuidade espacial (90°). Porém, os valores do I de Moran direcional entre as direções ortogonais 0° e 90° foram mais próximos quando calculados para valores preditos considerando-se o modelo isotrópico, os quais descrevem a forma circular que as sub-regiões se apresentam.

Em contrapartida, os valores do I de Moran direcional entre as direções ortogonais 0° e 90° apresentam diferenças relevantes com maiores valores para a direção 90° para os valores preditos considerando-se um modelo anisotrópico. Isso indica que nos mapas temáticos ocorre maior continuidade espacial nessa direção. Dessa forma, esta medida, de acordo com as duas metodologias, conseguiu quantificar a diferença existente nos mapas temáticos quando se considera ou não a anisotropia no modelo geoestatístico.

Tabela 12 Valor estimado do índice de Moran direcional para os valores preditos considerando modelos estimados anisotrópicos e isotrópicos para os atributos químicos do solo carbono, cálcio e potássio.

Variável	d	α	Rosenberg (2000)		Falsetti e Sokal (1993)		N° de Pares	
			(I) Isotrópico	(I) Anisotrópico	(I) Isotrópico	(I) Anisotrópico		
Carbono	150m	0°	0,896*	0,538*	0,910*	0,598*	4944	
		90°	0,942*	0,838*	0,955*	0,973*	3155	
	300m	0°	0,657*	0,257*	0,606*	0,249*	15238	
		90°	0,781*	0,591*	0,860*	0,915*	8293,5	
	450m	0°	0,346*	0,091*	0,321*	0,073*	27277	
		90°	0,537*	0,404*	0,745*	0,849*	14286	
	600m	0°	0,121*	-0,042	0,092*	-0,070	44592	
		90°	0,327*	0,267*	0,593*	0,741*	22868	
	750m	0°	0,008	-0,096*	0,060*	-0,088	59604	
		90°	0,168*	0,158*	0,484*	0,651*	30287	
	Cálcio	150m	0°	0,531*	0,607*	0,617*	0,636*	4944
			90°	0,629*	0,864*	0,781*	0,969*	3155
300m		0°	0,168*	0,248*	0,150*	0,203*	15238	
		90°	0,320*	0,605*	0,456*	0,846*	8293,5	
450m		0°	0,061*	0,093*	0,055*	0,072*	27277	
		90°	0,192*	0,404*	0,306*	0,695*	14286	
600m		0°	0,040*	0,042*	0,043*	0,041*	44592	
		90°	0,136*	0,277*	0,183*	0,485*	22868	
750m		0°	0,029*	0,023*	0,032*	0,013	59604	
		90°	0,089*	0,180*	0,103*	0,329*	30287	
Potássio		150m	0°	0,963*	0,838*	0,970*	0,847*	4944
			90°	0,972*	0,945*	1,000*	1,000*	3155
	300m	0°	0,869*	0,538*	0,852*	0,445*	15238	
		90°	0,899*	0,799*	0,636*	0,967*	8293,5	
	450m	0°	0,715*	0,275*	0,701*	0,226*	27277	
		90°	0,784*	0,621*	0,869*	0,920*	14286	
	600m	0°	0,542*	0,186*	0,482*	0,176*	44592	
		90°	0,661*	0,498*	0,774*	0,850*	22868	
	750m	0°	0,359*	0,147*	0,300*	0,136*	59604	
		90°	0,526*	0,390*	0,695*	0,789*	30287	

Nota (d) distância de corte; (α) ângulo de interesse; (I) índice de Moran direcional; asterisco (*) indica os resultados significativos a 5% de probabilidade (p-valor < 0,05) pelo teste de Pseudo sindicância do índice de Moran direcional.

6 CONCLUSÃO

Conclui-se que, quanto aos resultados das simulações, o índice de Moran direcional proposto por Rosenberg (2000) identificou a anisotropia geométrica pré-definida para todas as situações analisadas com valores elevados para a direção de maior continuidade espacial.

O índice de Moran direcional proposto por Rosenberg (2000), calculado para direções ortogonais de variáveis isotrópicas, apresentou valores próximos em todas as situações analisadas e descreveu o comportamento circular quanto à continuidade espacial, apresentado por variáveis isotrópicas. Nesse diapasão, a técnica pode ser utilizada para a identificação da anisotropia geométrica.

Quanto à análise para o (I) índice de Moran direcional proposto por Falsetti e Sokal (1993), este apresentou resultados semelhantes aos de Rosenberg, entretanto, a magnitude dos valores obtidos para o índice de Moran direcional apresenta menores diferenças conforme a distância de corte aumenta.

A variação do fator de anisotropia, quando se aumenta a distância de corte para ambas as metodologias estudadas, indicou que quanto maior o fator de anisotropia menor a taxa de decaimento sofrida nos valores do I de Moran direcional.

Para o estudo prático, a modelagem geoestatística pela construção de semivariograma direcionais identificou a presença de anisotropia geométrica para as variáveis: carbono, cálcio e potássio.

Em todos os atributos foram obtidos valores mais elevados e significativos para direção de maior continuidade espacial com a aplicação das técnicas propostas pelo índice de Moran direcional para os dados amostrais, dos atributos químicos do solo com presença de anisotropia geométrica. Conclui-se, portanto, que o índice é eficaz para auxiliar na identificação da anisotropia geométrica.

O Índice de Moran direcional, calculado para os valores preditos para expressar a diferença direcional que existe nos mapas temáticos, quando se incorpora ou não à anisotropia no modelo geoestatístico se apresentou eficiente e descreveu o comportamento elíptico encontrado nas sub-regiões dos mapas temáticos.

Por fim, a partir do estudo apresentado, conclui-se que as metodologias estudadas se mostraram uma ferramenta eficaz para expressar a diferença direcional que existe nos mapas temáticos e, também, para a identificação ou não de anisotropia geométrica.

7 REFERÊNCIAS

- ANSELIN, L. Local Indicators of Spatial Association-LISA. **Geographical Analysis** vol. 27, p. 93-115, 1995.
- ALMEIDA, M. F. F.; SANTOS, G. R.; PINHEIRO, W. R. F.; VALENTE, D. S. M. Análise geoestatística dos níveis de suficiência de Fósforo, Potássio, Cálcio, Magnésio e a acidez/alcalinidade na presença da altitude, que interferem no cultivo de café: um estudo feito por meio de mapas de Krigagem Indicativa. **Matemática e Estatística em Foco**, v. 1, p. 1-3, 2013.
- ANDERSON J. F.; HARDY, E. E.; ROACH J. T.; WITMER, R. R. A land use and land cover classification system for use with remote sensor data. U.S. **Geological Survey Professional paper**, n. 964. USGS, Washington, D.C, 1976. 28p.
- ANDRIOTTI, J. L. S. **Fundamentos de Estatística e Geoestatística**. Rio Grande do Sul: Unisinos, 2004. 165 p.
- BISWAS A; CRESSWELL H. P; ROSSEL R. A. V; SI B. C. Curvelet transform to study scale-dependent anisotropic soil spatial variation. **Geoderma**. Canada, v. 213, p. 589-599, 2014.
- BORSSOI J. A. **Técnicas de diagnóstico em modelos espaciais lineares gaussianos**. 2007. 168 f. Dissertação, Programa de Pós-Graduação em Engenharia Agrícola, Universidade Estadual do Oeste do Paraná, Cascavel. 2007.
- CÂMARA, G.; MEDEIROS, J. S. Princípios básicos em geoprocessamento. In: ASSAD, E. D.; SANO, E. E. **Sistemas de informações geográficas: aplicações na agricultura**. 2. ed. Brasília: Embrapa – SPI, 1998. p. 3-11.
- CRESSIE, N. A. C. **Statistics for spatial data**. Reviewed Edition. Wiley, New York, 2015.
- DALPOSSO, G. H., M. A. URIBE-OPAZO, E. MERCANTE, J. A. BORSSOI, J. A. JOHANN. 2012. **Comparison measures of maps generated by geostatistical methods**. Engenharia Agrícola (Impresso), v. 32, p. 174-183.
- DE BASTIANI, F. **Influência local em modelos espaciais lineares com distribuição da família de contornos elípticos**. 2012, 156 f. Dissertação, Programa de Pós-Graduação em Mestrado em estatística, Universidade Federal de Pernambuco, 2012.
- DE BASTIANI, F., M. A. URIBE-OPAZO, G. H. DALPOSSO. Comparison of maps of spatial variability of soil resistance to penetration constructed with and without covariables using a spatial linear model. **Engenharia Agrícola** (Impresso), v. 32, p. 394-404, 2012.
- DEL MONEGO M.; RIBEIRO JUNIOR, P. J. Comparing the performance of geostatistical models with additional information from covariates for sewage plume characterization. **Springer**. Berlin, v. 22, p. 5850-5863, 2014.
- DIGGLE, P. J.; RIBEIRO JUNIOR P. J.. **Model-based Geostatistics**. 230 p. Springer, New York, EUA, 2007.
- FALSETTI A. B.; SOKAL R. R. Genetic structure of human populations in the British isles. **Annals of Human Biology**, v. 20, p. 215-229, 1993.

FARACO, M. A.; URIBE-OPAZO, M. A. SILVA, E. A.; JOHANN, J. A.; BORSSOI, J. A. Seleção de modelos de variabilidade espacial para elaboração de mapas temáticos de atributos físicos do solo e produtividade da soja. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 32, n. 2, p. 463-476, 2008.

GRZEGOZEWSKI, D. M.; URIBE-OPAZO, M. A.; JOHANN, J. A.; GUEDES, L. P. C. spatial correlation of soybean productivity, enhanced vegetation index (evi) and agrometeorological variables. **Engenharia Agrícola**, v. 37, p. 541-555, 2017.

GRZEGOZEWSKI, D. M.; URIBE-OPAZO, MIGUEL A; DE BASTIANI, F.; GALEA, M. Local influence when fitting Gaussian spatial linear models: an agriculture application. **Ciencia e Investigación Agraria** (En línea. Edición en Español), v. 40, p. 537-545, 2013.

GUEDES, L. P. C.; URIBE-OPAZO, M. A.; JOHANN, J. A.; SOUZA, E. G. Anisotropia no estudo da variabilidade espacial de algumas variáveis químicas do solo. **Revista Brasileira de Ciência do Solo**, v. 7, p. 2217-2226, 2008.

GUEDES, L. P. C.; URIBE-OPAZO, M. A. RIBEIRO JUNIOR, P. J. Influence of incorporating geometric anisotropy on the construction of thematic maps of simulated data and chemical attributes of soil. **Chilean Journal of Agricultural Research**. v. 73, n. 4, 2013.

IAPAR – Instituto Agrônomo do Paraná (2016) <http://www.iapar.br/modules/conteudo/conteudo.php?conteudo=597>> Acesso em 25 de janeiro de 2016.

ISAAKS, E. H.; SRIVASTAVA, R. M. **An Introduction to Applied Geostatistics**. New York: Oxford, Oxford University Press, 1989.

KESTRING, F. B. F.; GUEDES, L. P. G.; DE BASTINI, F.; URIBE-OPAZO, M. A. Comparação de mapas temáticos de diferentes grades amostrais para a produtividade da soja. **Revista Engenharia Agrícola**, Jaboticabal, v. 35, n. 4, p.733-743, 2015.

KRIPPENDORFF, K. 1980. **Content Analysis: an introduction to its methodology**. 189 p. Beverly Hills: Sage Publications.

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

McBRATNEY, A. B.; WEBSTER, R. **Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates**. Journal of Soil Science. Oxford. v. 37, p. 617-639.

MORAN, P. A. P. Notes on continuous stochastic phenomena. **Biometrika**, v. 37, p. 17-23, 1950.

NOGUEIRA C. H; LIMA R. R; OLIVEIRA M. S. Aprimoramento da análise de variância: a influência da proximidade espacial. **Rev. Bras. Biom**, São Paulo v. 31 n. 3, p. 408-422, 2013.

OLIVER M. A.; WEBSTER R. A tutorial guide to geoestatistics: Computing and modelling variograms and kriging. **Catena**. v. 113, p. 56-69, 2014.

PIMENTEL GOMES, F. **Curso de estatística experimental**. 13ª ed. São Paulo: Nobel, 2000, 479p.

REZENDE, F. C.; RIBEIRO, V. B.; ÁVILA, L. F.; FARIA, M. A.; SILVA, E. L. Variabilidade espacial do pH em área com cafeeiro fertirrigado e sistema tradicional. **Coffee Science**, Lavras, v. 7, n. 3, p. 198-207, 2012.

ROSENBERG, M. S. The bearing correlogram: a new method of analyzing directional spatial autocorrelation. **Geographical Analysis**, Columbus, v. 32, n. 3, p. 267-278, Mar. 2000.

ROSSONI, D. F.; LIMA, R. R. A importância da correção da anisotropia em análises geoestatísticas. **Revista da Estatística da Universidade Federal de Ouro Preto**, v. 2, p. 217-221, 2012.

SILVEIRA, T. A.; PORTUGAL, J. L.; SA, L. A. C. M.; VITAL, S. R. O. Análise Estatística Espacial Aplicada a Construção de Superfícies Batimétricas. **Geociências** (São Paulo. Online), v. 33, p. 596-615, 2014.

SOARES, A. **Coleção ensino da ciência e da tecnologia**: Geoestatística para as ciências da terra e do ambiente. Lisboa. Técnico Lisboa, 3ª Edição. 2014.

SOBJAK, R.; DE SOUZA, E. G.; BAZZI, C. L.; URIBE-OPAZO, M. A.; BETZEK, N. M. Redundant variables and the quality of management zones. **Engenharia Agrícola (Online)**, v. 36, p. 78-93, 2016.

URIBE-OPAZO, M. A.; BORSSOI, J. A.; GALEA, M. Influence diagnostics in gaussian spatial linear models. **Journal of Applied Statistics**. v. 39, n. 3, p. 615-630, 2012.

VICENTE, J. Estudo comparativo de métodos geoestatísticos aplicados em agricultura de precisão. 2004. 163 f. Dissertação, Programa de Pós-Graduação em Ciências Cartográficas. Universidade Estadual Paulista, Presidente Prudente.

VIEIRA, S. R. **Geoestatística em estudos de variabilidade espacial do solo**. In: Tópicos em ciência do solo, Viçosa, Sociedade Brasileira de Ciência do Solo, 2000. v. 1, p. 3-4.

YAMAMOTO, J. K.; URIBE-OPAZO, M. A.; JOHANN, J. A.; VILAS BOAS, M. A. Uniformidade de distribuição de água de irrigação por aspersores compensantes e não compensantes em diferentes alturas. **Irriga**, v. 21, p. 631-647, 2016.

8 APÊNDICE

Apêndice 1 Medidas para a escolha do melhor modelo ajustado por critérios de validação cruzada, AIC e BIC.

Variáveis	Modelos	EM	EMR	SEM	SEMR	EA	AIC	BIC
Zn	Sph	0,0304	0,0100	1,6840	1,0520	115,0171	399,2	412,3
	Exp	0,0299	0,0098	1,6687	1,0291	114,0908	401,0	414,1
	Gaus	0,0289	0,0088	1,7072	1,0888	115,5681	399,3	412,4
	Matérn	0,0343	0,0114	1,6909	1,0635	114,6526	400,1	413,2
Fe	Sph	0,0290	0,0013	10,8355	1,0068	801,4655	786,1	799,2
	Exp	0,0298	0,0014	10,8768	1,0064	806,0089	786,5	799,7
	Gaus	0,0274	0,0013	10,8456	1,0069	802,4368	786,2	799,3
	Matérn	0,0290	0,0013	10,8685	1,0065	804,8065	786,4	799,6
P	Sph	0,0002	0,0000	11,2042	1,0098	852,8917	789,5	802,6
	Exp	0,0000	0,0000	11,2148	1,0099	851,8921	789,6	802,7
	Gaus	0,0000	0,0000	11,2147	1,0099	851,8862	789,6	802,7
	Matérn	0,0000	0,0000	11,2147	1,0099	851,8907	789,6	802,7
Al	Sph	0,0021	0,0030	0,3614	1,0303	25,5220	86,3	99,4
	Exp	0,0011	0,0016	0,3601	1,0198	25,1593	87,1	100,2
	Gaus	0,0018	0,0027	0,3615	1,0297	25,4554	86,5	99,6
	Matérn	0,0015	0,0021	0,3612	1,0264	25,3863	86,9	100,0
Mn	Sph	0,0290	0,0013	10,8686	1,0065	804,8065	891,3	904,4
	Exp	-0,2430	-0,0076	16,7741	1,0279	1333,7890	892,7	905,8
	Gaus	-0,1621	-0,0049	17,0045	1,0308	1358,1100	891,4	904,5
	Matérn	-0,2105	-0,0065	16,8000	1,0295	1336,3450	891,3	904,4
C	Sph	-0,0064	-0,0009	3,3933	1,0234	267,3245	556,8	569,9
	Exp	-0,0088	-0,0013	3,3756	1,0157	267,5155	555,6	568,7
	Gaus	-0,0113	-0,0016	3,3716	1,0066	265,8829	555,1	568,2
	Matérn	-0,0093	-0,0013	3,3731	1,0143	267,106	555,4	568,5
Ca	Sph	0,0017	0,0007	1,3313	1,0182	108,179	357,3	370,4
	Exp	-0,0007	-0,0004	1,3251	1,0166	107,531	357,0	370,1
	Gaus	-0,0020	-0,0009	1,3186	1,0184	106,758	356,5	369,7
	Matérn	-0,0008	-0,0004	1,3233	1,0179	107,407	356,8	369,9
K	Sph	-0,0005	-0,0020	0,1271	1,0021	10,3356	-118,1	-105,0
	Exp	-0,0005	-0,0019	0,1280	1,0060	10,3963	-117,2	-104,1
	Gaus	-0,0005	-0,0021	0,1266	1,0013	10,2790	-118,3	-105,2
	Matérn	-0,0005	-0,0020	0,1275	1,0050	10,3459	-117,5	-104,4