

DANILO PEREIRA BARBOSA

**MODELAGEM DE SEMIVARIOGRAMA CONSIDERANDO ANISOTROPIA
E DADOS DISCREPANTES NO ESTABELECIMENTO DE ZONAS DE
MANEJO**

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

VIÇOSA
MINAS GERAIS – BRASIL
2018

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

T

B238m
2018
Barbosa, Danilo Pereira, 1984-
Modelagem de semivariograma considerando anisotropia e
dados discrepantes no estabelecimento de zonas de manejo /
Danilo Pereira Barbosa. – Viçosa, MG, 2018.
xv, 63f. : il. (algumas color.) ; 29 cm.

Inclui apêndices.

Orientador: Nerilson Terra Santos.

Tese (doutorado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. Sistemas de informação geográfica. 2. Geologia -
Métodos estatísticos. 3. Anisotropia. I. Universidade Federal de
Viçosa. Departamento de Estatística. Doutorado em Estatística
Aplicada e Biometria. II. Título.

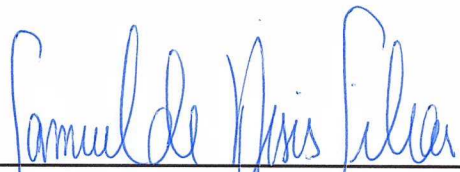
CDD 22 ed. 526

DANILO PEREIRA BARBOSA

**MODELAGEM DE SEMIVARIOGRAMA CONSIDERANDO ANISOTROPIA
E DADOS DISCREPANTES NO ESTABELECIMENTO DE ZONAS DE
MANEJO**

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


APROVADA: 07 de março de 2018.




Samuel de Assis Silva




Domingos Sárvio M. Valente



José Ivo Ribeiro Júnior



Antônio Policarpo Souza Carneiro



Nerilson Terra Santos
(Orientador)

“... a inteligência é humana quando o conhecimento é aprendido, é divina, quando não depende de aprendizado.”

(MOKITI OKADA)

A Deus,

A Meishu Sama

A minha mãe Marli e

minha avó Benvinda com amor

Dedico

AGRADECIMENTOS

A Deus e Meishu Sama pela proteção e orientação durante todo o percurso.

A minha mãe Marli Pereira, que sempre com seu amor incondicional esteve ao meu lado, dando forças e auxiliando com sabedoria.

A minha avó Benvinda Augusta Ferreira de Jesus, que sempre esteve preocupada comigo. Buscava sempre entender o que é doutorado e me acompanhava com dedicação e carinho. E que infelizmente não conseguiu esperar a conclusão deste.

Aos meus ancestrais e antepassados, pois reconheço que: "... sou a soma dos meus antepassados. Por isso, sou representante de todos eles e reconheço que eles se manifestam através dos meus pensamentos e sentimentos..." (IMMB).

Ao Instituto Federal de Educação Ciência e Tecnologia Goiano – Rio Verde, por todo o apoio na liberação para capacitação docente, e demais colegas de trabalho.

Ao meu orientador Nerilson Terra Santos, pelo qual tenho grande estima e respeito. E que, por seu intermédio, aprendi grandes ensinamentos tanto acadêmicos quanto profissional.

Aos meus coorientadores Eduardo Leonel Bottega e Wellington Donizete Guimarães pela prontidão e solicitude que sempre me atenderam.

Aos professores Domingos Sárvio Magalhães Valente, João Marcos Louzada e Gerson Rodrigues dos Santos pela colaboração de forma significativa diante conhecimentos acadêmico-teóricos transmitidos. Aos professores José Ivo Ribeiro Júnior, Antônio Policarpo Souza Carneiro, Domingos Sárvio Magalhães Valente e Samuel de Assis Silva pela participação e colaborações na banca.

A Lidiane Maria Ferraz e demais amigos que sempre estiveram ao meu lado.

Ao secretário do programa de Pós-Graduação em Estatística Aplicada e Biometria, Junior José Pires e a Maria Anita Silva de Paiva, secretária do respectivo departamento. Pelos quais tenho imensa gratidão por sempre ser atendido com educação e presteza.

A todos que de forma direta ou indireta contribuíram na realização deste trabalho.

BIOGRAFIA

Danilo Pereira Barbosa, filho de Marli Pereira e Getúlio Oliveira Barbosa, nasceu em Rio Verde, Goiás, no dia 28 de agosto de 1984.

Em agosto de 2002 iniciou o curso de Ciências – Habilitação em Matemática na Universidade de Rio Verde – UniRV, Goiás, o qual concluiu em julho de 2006.

Em setembro de 2006, iniciou o Mestrado em Estatística Aplicada e Biometria, na Universidade Federal de Viçosa, Minas Gerais, defendendo a dissertação em Fevereiro de 2009.

Em 2014, no primeiro semestre, ingressou no doutorado em Estatística Aplicada e Biometria na Universidade Federal de Viçosa, Minas Gerais, defendendo a tese em Março de 2018.

SUMÁRIO

LISTA DE FIGURAS	ix
LISTA DE TABELAS.....	xi
RESUMO.....	xii
ABSTRACT	xiv
INTRODUÇÃO GERAL	1
O problema	2
Hipóteses	3
Objetivos	4
<i>Objetivo geral:</i>	4
<i>Objetivos específicos:</i>	4
Sistematização do trabalho.....	4
REFERÊNCIAS	6
ARTIGO I – ANISOTROPIA GEOMÉTRICA NA DELIMITAÇÃO DE ZONAS DE MANEJO.....	9
Resumo	9
Abstract	10
Introdução.....	11
Material e Métodos	13

Resultados e Discussão.....	19
Conclusão	28
Referências	29
ARTIGO II – KRIGAGEM ROBUSTA À <i>OUTLIERS</i> , NO DELINEAMENTO DE ZONAS HOMOGÊNEAS.	34
Resumo	34
Abstract	35
Introdução.....	36
Material e Métodos	38
Resultados e Discussão.....	43
Conclusões.....	53
Referências	54
CONCLUSÕES GERAIS	59
APÊNDICE.....	60
ARTIGO I	60
1 - Verificação da continuidade espacial	60
2 – Obtenção dos modelos de semivariogramas omnidirecionais sem correção da anisotropia geométrica.	62
3 – Obtenções dos modelos de semivariogramas direcionais e omnidirecionais após transformações lineares (correção da anisotropia geométrica).	62
4 – Obtenção dos mapas de zonas de manejo para a produtividade e CEa, com e sem correção da anisotropia geométrica, com respectivos números de classes.	63

LISTA DE FIGURAS

- Figura 1 - Malha amostral, regularmente espaçada em 50 metros. Coordenadas UTM, zona 21 Sul. 13
- Figura 2 - Índice de Performance Fuzzy (FPI) e Entropia da Partição Modificada (MPE), relativo ao número de classes na delimitação de zonas de manejo para as variáveis produtividade e condutividade elétrica aparente (CEa), com e sem correção da anisotropia..... 24
- Figura 3 - Índice kappa, em função do número de classes, comparando mapas com e sem correção da anisotropia, para as variáveis produtividade e condutividade elétrica aparente (CEa)..... 26
- Figura 4 - Mapas de zonas de manejo para as variáveis produtividade e CEa obtidos sem correção da anisotropia geométrica. 28
- Figura 5 – Histograma e box-plot da distribuição da condutividade elétrica aparente (CEa)..... 44
- Figura 6 – Malha amostral contendo as posições (marcadas com X) dos respectivos *outliers* identificados para a condutividade elétrica aparente do solo, cujas dimensões das circunferências são proporcionais aos respectivos valores..... 45
- Figura 7 - Semivariograma experimental para o estimador de semivariância de Matheron (MM estimator), Genton (Qn estimator), Cressie Hawkins (Ch estimator) e Dowd (MAD estimator) aos dados amostrais da CEa. 46
- Figura 8 – Modelo cúbico ajustado às semivariâncias obtidas pelo estimador de Cressie Hawkins para a variável condutividade elétrica aparente (CEa)..... 47

Figura 9 – Modelo cúbico ajustado às semivariâncias obtidas pelo estimador de Matheron para a variável condutividade elétrica aparente (CEa).....	48
Figura 10 - Mapa interpolado considerando a análise robusta à <i>outliers</i> , definindo a estrutura da variabilidade espacial da condutividade elétrica aparente (CEa).....	48
Figura 11 - Mapa interpolado considerando a análise não robusta à <i>outliers</i> , definindo a estrutura da variabilidade espacial da condutividade elétrica aparente (CEa).....	49
Figura 12 - Mapas com 2, 3, 4 e 5 zonas de manejo dentro da área em estudo, gerados com uso do algoritmo <i>fuzzy k-means</i> para a variável CEa, considerando a análise robusta à <i>outliers</i> (ARob) e não robusta à <i>outliers</i> (ANRob).....	50
Figura 13 - Mapas com duas zonas de manejo para a condutividade elétrica aparente (CEa) obtidos com a análise robusta à <i>outliers</i> (ARob) e com a análise não robusta à <i>outliers</i> (ANRob).....	52
Figura 14 - Envelopes definidos com valores máximos e mínimos de semivariogramas empíricos resultantes da permutação de coordenadas dos valores amostrais de Produtividade e CEa.....	61
Figura 15 - Ajuste das semivariâncias ao modelo Gaussiano sem correção da anisotropia, para as variáveis produtividade e CEa.	62
Figura 16 – Semivariogramas direcionais e omnidirecionais do modelo gaussiano, após correção da anisotropia geométrica para as variáveis produtividade e CEa.....	62
Figura 17 - Mapas de zonas de manejo, para as variáveis produtividade e CEa, com 2 a 5 classes, com estruturas isotrópica e anisotrópica.....	63

LISTA DE TABELAS

Tabela 1 - Estatísticas descritivas associadas às variáveis produtividade e condutividade elétrica aparente.....	19
Tabela 2 - Especificações dos modelos omnidirecionais ajustados sem a correção da anisotropia.....	20
Tabela 3 - Parâmetros dos semivariogramas direcionais relativos aos ângulos de maior e menor continuidade espacial, após transformações lineares nas coordenadas georreferenciadas das variáveis.....	22
Tabela 4 - Parâmetros dos semivariogramas omnidirecionais após transformações lineares nas coordenadas georreferenciadas dos pontos amostrados onde foram mensurados as variáveis produtividade e CEa.....	23
Tabela 5 – Estatísticas descritivas da condutividade elétrica aparente (CEa) amostrada na profundidade 0 a 20cm.....	43
Tabela 6 - Médias em cada zona de manejo (ZM), para a condutividade elétrica aparente, nas análises robusta à outliers (ARob) e não robusta à outliers (ANRob), com respectivos limites superiores e inferiores dos intervalos de confiança.....	52

RESUMO

BARBOSA, Danilo Pereira, D.Sc., Universidade Federal de Viçosa, março de 2018. **Modelagem de semivariograma considerando anisotropia e dados discrepantes no estabelecimento de zonas de manejo.** Orientador: Nerilson Terra Santos. Coorientadores: Eduardo Leonel Bottega, Wellington Donizete Guimarães e Gerson Rodrigues dos Santos.

Com o estabelecimento da agricultura de precisão, a heterogeneidade do solo tornou-se um parâmetro expressivo quanto ao seu manuseio. Frente a este cenário, destaca-se a utilização massiva das zonas de manejo (ZM). As ZM são sub-regiões do campo com necessidades específicas quanto as variáveis analisadas, permitindo o controle da heterogeneidade do solo, maximização produtiva e sustentabilidade agrícola. Entretanto, sua aplicabilidade esta condicionada ao mapeamento do padrão de variabilidade espacial dos atributos físico-químicos presentes no solo. Este mapeamento tem sido resultante da utilização contínua de métodos geoestatísticos, dos quais apresentam pressuposições inexploradas em suas aplicações, conduzindo assim, o objetivo desta pesquisa. E consequentemente norteou os específicos objetivos: a) avaliar alterações em mapas de ZM devido à correção da anisotropia e b) avaliar variações em mapas de ZM quanto à utilização de metodologia robusta à *outliers*. Para tanto, 160 pontos amostrais regularmente espaçados, relativos à condutividade elétrica aparente do solo (CEa), e produtividade de soja foram utilizados. Quanto à verificação de alterações em mapas de ZM devido à correção da anisotropia, os mesmos foram interpolados sem e com correção da anisotropia geométrica para cada variável. Na sequencia foram então utilizados para o delineamento das ZM por meio do método *fuzzy k-means*. As ZM para cada variável, com e sem correção da anisotropia geométrica, foram avaliadas quanto as suas semelhanças pelo índice kappa. Para a avaliação de variações em mapas de ZM quanto à ocorrência de *outliers* utilizaram-se dois tipos de análises, robusta a presença de *outliers* (ARob) e não robusta à *outliers* (ANRob). Na ARob utilizaram-se estimadores robustos de

semivariâncias e o plug-in de krigagem de deriva externa para a geração de mapas de variabilidade espacial da CEa. Para a ANRob utilizou-se o estimador de semivariâncias de Matheron e a krigagem ordinária. Posteriormente os mapas obtidos foram submetidos ao delineamento de zonas de manejo pelo classificador *fuzzy k-means*. E de maneira conclusiva, os mapas obtidos em ambas as análises (ARob e ANRob) foram confrontados quanto à significância do nível de concordância entre suas classes pelo índice Kappa. Os resultados obtidos na verificação de alterações em mapas de ZM devido à correção da anisotropia foram: a) utilizou-se o modelo gaussiano na constituição dos mapas de variabilidade espacial para a CEa e para a produtividade, tanto para os dados corrigidos à anisotropia quanto aos não corrigidos; b) conforme os índices FPI e MPE, definiram-se duas classes para o delineamento de ZM para os dados corrigidos à anisotropia, quanto aos não corrigidos; c) a comparação entre os mapas (corrigido e não corrigido à anisotropia) pelo índice Kappa apresentou concordância significativa entre classes de ZM a 5% de probabilidade. Concluindo assim que, no caso em estudo, a correção da anisotropia geométrica não apresentou alterações significativas nos mapas de ZM. Os resultados obtidos na avaliação de variações em mapas de ZM quanto à ocorrência de *outliers* foram: a) na ARob selecionou-se o estimador de semivariâncias de Cressie Hawkins dentre os demais estimadores robustos avaliados. Na predição do mapa de estrutura de variabilidade espacial da CEa utilizou-se o plug-in de krigagem de deriva externa. Os índices FPI, MPE, Fukuyama Sugento e Xie beni definiram duas classes de ZM. b) na ANRob utilizou-se o estimador de semivariâncias de Matheron e a krigagem ordinária na composição do mapa de variabilidade espacial da CEa. Os índices avaliados definiram duas classes de ZM. c) os mapas obtidos em ambas as análises (ARob e ANRob) apresentaram concordância significativa entre classes de ZM pelo índice Kappa a 1% de probabilidade. Com isso, de maneira conclusiva, para o caso em estudo, o uso da ARob não apresentou variações significativas no estabelecimento das ZM.

ABSTRACT

BARBOSA, Danilo Pereira, D.Sc., Universidade Federal de Viçosa, March, 2018. **Semivariogram modeling considering anisotropy and discrepant data in management zones establishment.** Adviser: Nerilson Terra Santos. Co-advisers: Eduardo Leonel Bottega, Wellington Donizete Guimarães and Gerson Rodrigues dos Santos.

With the precision agriculture establishment, soil heterogeneity became an expressive parameter for its handling. In front this scenario, the massive use management zones is highlighted. The ZM are field sub-regions with specific needs as the variables analyzed. Therefore, ZM allows control of soil heterogeneity, productive maximization and agricultural sustainability. Therefore, ZM allows control of soil heterogeneity, productive maximization and agricultural sustainability. However, its applicability is conditioned to spatial variability pattern mapping of the physico-chemical attributes in soil present. This mapping, has been result of the continuous use of geostatistical methods of which present untapped assumptions of these methods have been unexplored in their applications, thus conducting the purpose of this research. And consequently guided the specific objectives: a) evaluate alterations in ZM maps due to anisotropy correction and b) evaluate variations in ZM maps regarding use of a outliers robust methodology. Therefore, 160 regularly spaced sampling points, relative to soil apparent electrical conductivity (CEa), and soybean productivity, in different periods, were used. As for the verification of alterations in ZM maps due to the anisotropy correction, they were interpolated without and with correction of the geometric anisotropy for each variable. In sequence was then used for ZM delineation by fuzzy K-means method. The ZM for each attribute, with and without geometric anisotropy correction, were evaluated as similarities by Kappa Index. For the evaluation of variations in ZM maps for the occurrence of outliers, two types of analysis were used, robust to outliers (ARob) and non-robust to outliers (ANRob). In ARob we used robust semivariance estimators and the external drift kriging plug-in to generate

spatial variability maps of CEa. For the ANRob, the Matheron semivariance estimator and the ordinary kriging were used. Afterwards the maps obtained were submitted to the management zones delineation by the fuzzy k-means classifier. And conclusively, the maps obtained in both analyzes (ARob and ANRob) were statistically compared, using Kappa index, in relation their zones composition. The results obtained in the verification of alterations in ZM maps due to anisotropy correction were: a) the Gaussian model was used in the constitution of spatial variability maps for CEa and productivity, both for anisotropy-corrected and uncorrected data; b) the FPI and MPE indices defined two classes for ZM delineation for corrected data for the anisotropy and the uncorrected; c) the comparison between the maps (corrected and uncorrected to the anisotropy) by the Kappa index showed significant concurrence between ZM classes at 5% probability. In conclusion, in this case, the geometric anisotropy correction did not present significant concurrence in the ZM maps. The results obtained in the evaluation of variations in ZM maps regarding the occurrence of outliers were the following: a) in ARob was selected the semivariances estimator Cressie Hawkins among the other robust estimators evaluated. In the prediction of the spatial variability structure map of the CEa, the external drift kriging plug-in was used. The FPI, MPE, Fukuyama Sugento and Xie beni indices defined two classes of ZM. b) in the ANRob was used the Matheron's semivariance estimator and the ordinary kriging in the spatial variability map of the CEa. The evaluated indices defined two classes of ZM. c) the maps obtained in both analyzes (ARob and ANRob) showed significant agreement between ZM classes by Kappa index at 1% probability. Thus, for the case under study, the use of ARob did not show significant differences in the establishment of the ZM maps.

INTRODUÇÃO GERAL

Na agricultura convencional, toda a área ocupada por uma determinada cultura é tratada de maneira uniforme. Esta abordagem despreza as necessidades específicas que podem ocorrer em regiões peculiares do terreno devido à variabilidade existente no campo.

O apontamento mais antigo de que produtores já se davam conta desta variabilidade existente no solo é datado em 1929, em um boletim do campo experimental de Illinois (BERNARDI et al., 2014). No entanto, foi em meados de 1970, que se iniciou a conscientização da necessidade de registros dessa variabilidade (ROBERT, 1999). E no período de 1980 a 1990, surge à implementação de dispositivos inteligentes para coleta de dados (BRASE, 2005).

E posteriormente, a agricultura de precisão veio auxiliar na logística de manuseio do campo, uma vez que considera e agrega a variabilidade existente no solo para sua gestão (HAGHVERDI et al., 2015). Tal variabilidade espacial pode ser caracterizada pela condutividade elétrica aparente do solo (CORWIN; LESCH, 2003). Isso porque, a condutividade elétrica aparente do solo (CEa), apresenta confiabilidade na estimativa indireta da variabilidade existente da unidade produtora, baixo custo (CORWIN; LESCH, 2003) e viabiliza a delimitação de zonas de manejo (SHANER *et al.*, 2008).

As zonas de manejo podem ser definidas como sub-regiões do campo com potenciais produtivos distintos (HAGHVERDI et al., 2015). Sua utilização permite a aplicação de insumos às taxas variáveis, e conseqüentemente favorece a conservação

do solo (DALCHIAVON *et al.*, 2012), redução de custos e de contaminação ambiental (BOTTEGA, 2014; CÓRDOBA *et al.*, 2013).

Entretanto, o emprego expressivo de tal metodologia, favorável à agricultura de precisão, deve-se à concepção adequada de mapas que expressem a estrutura de variabilidade espacial do solo. A modelagem desta estrutura de variabilidade espacial está diretamente relacionada com uso de metodologias em geoestatística.

O problema

Metodologias geoestatísticas têm sido amplamente utilizadas na elaboração de mapas de variabilidade espacial, que conseqüentemente, estes mapas viabilizam o delineamento das zonas de manejo.

Em diversos estudos relacionados à delimitação destas zonas de manejo, a gama de interpoladores geoestatísticos é variada, dentre eles, destacam-se a krigagem ordinária (ALVES *et al.*, 2013; CHANG *et al.*, 2014; FU; WANG; JIANG., 2010; MORAL; TERRÓN; REBOLLO, 2011; SAFANELLI; BOESING; BOTTEGA, 2015; TRIPATHI *et al.*, 2015), a cokrigagem (MORARI; CASTRIGNANO; PAGLIARIN, 2009) e a regressão krigagem (MORAL; TERRÓN; DA SILVA, 2010).

Entretanto, considerando estas e outras diversas pesquisas, nota-se que apresentam ausência de pressuposições estatísticas que interferem na acurácia dos mapas de estruturas de variabilidade espacial. Dentre estas pressuposições, é fundamental, caso haja necessidade, a correção da anisotropia. Em especial a anisotropia geométrica por sua maior frequência. E assim, com a correção da

anisotropia geométrica obter maior exatidão na composição destes mapas temáticos (GUEDES et al., 2013).

Além da correção da anisotropia, é relevante a verificação da distribuição dos dados quanto à simetria. Isso porque a variável em estudo, em especial a condutividade elétrica aparente, pode assumir forma simétrica (CÓRDOBA et al., 2013) ou assimétrica (SCUDIERO et al., 2013; YAO et al., 2014; BOTTEGA, 2014; SHAHBAZI; ALIASGHARZAD; EBRAHIMZAD, 2013) dependendo das condições climáticas.

A assimetria altera expressivamente o ajuste do modelo teórico do semivariograma (FU et al., 2016). Além do que, compromete as estatísticas da validação cruzada da krigagem (KERRY e OLIVER, 2007) e a precisão do mapa de variabilidade espacial. E conseqüentemente implica no delineamento das zonas de manejo prejudicando desta forma a gestão no campo.

Na presença de assimetria, é importante verificar a ocorrência de *outliers*. Isso porque, sua exclusão tende a normalizar os dados evitando a necessidade de transformação dos mesmos (FU et al., 2016).

Entretanto, a exclusão incorreta de *outliers* influencia a inferência estatística (MARONA; MARTIN; YORAI, 2006), comprometendo assim as estimativas dos parâmetros do semivariograma e conseqüentemente a precisão do preditor linear geoestatístico.

Hipóteses

Em face da problemática abordada diante as implicações relacionadas quanto à anisotropia e outliers, a presente tese alicerçou-se nas seguintes hipóteses:

- ✓ Corrigir a anisotropia presente na variabilidade espacial dos dados modifica significativamente o estabelecimento das zonas de manejo?
- ✓ Utilizar metodologias robustas a *outliers* altera significativamente o estabelecimento das zonas de manejo?

E com base nessas premissas, estabeleceram-se os seguintes objetivos:

Objetivos

Objetivo geral: verificar possíveis alterações significativas no estabelecimento das zonas de manejo, considerando aspectos relevantes dos dados georreferenciados.

Objetivos específicos: Visando a compreensão do objetivo geral mencionado, estabeleceram-se os seguintes objetivos específicos:

- ✓ Avaliar alterações em mapas de zonas de manejo devido à correção da anisotropia;
- ✓ Avaliar modificações em mapas de zonas de manejo com uso de metodologia robusta à presença de outliers para a modelagem da dependência espacial.

Sistematização do trabalho

Em favor de uma percepção aprazível na leitura, a presente pesquisa revelou-se inicialmente por uma introdução geral, na qual foram apresentadas as indagações que a emergiu, e a exposição dos objetivos específicos supracitados. Que posteriormente, o mesmo compôs-se por dois artigos, mencionados aqui por Artigo I e Artigo II.

No Artigo I são citadas questões relacionadas à correção da anisotropia. Em específica a anisotropia geométrica, cujo objetivo é verificar se a correção ou não desta anisotropia altera a composição do mapa de zonas de manejo.

Além desta avaliação, o Artigo II versa sobre dados com distribuições assimétricas. Neste capítulo expõem-se as implicações ao corrigir a assimetria. Dentre elas, destaca-se a exclusão de *outliers*. Uma vez que a exclusão incorreta destes valores atípicos afeta a inferência estatística, propôs-se a utilização de uma metodologia robusta à *outliers*, fundamentada na teoria da geoestatística baseada em modelos (GBM) e na teoria de aproximações robustas.

E, de maneira conclusiva, ao final deste trabalho, expõem-se as conclusões gerais diante das hipóteses mencionadas. E com isso, esclarecer possíveis implicações quanto à delimitação de zonas de manejo.

REFERÊNCIAS

- ALVES, S.M.F.; ALCÂNTARA, G.R.; REIS, E.F.; QUEIROZ, D.M.; VALENTE, D.S.M. Definição de zonas de manejo a partir de mapas de condutividade elétrica e matéria orgânica. *Bioscience Journal*, v. 29, n. 1, p. 104-114, 2013.
- BERNARDI, A. C. C. et al. Agricultura de precisão: resultados de um novo olhar. Brasília, DF: Embrapa, 2014.
- BRASE, T. Precision Agriculture, Delmar Publishers Inc, 288 p, 2005.
- CHANG, Dong, et al. Delineation of management zones using an active canopy sensor for a tobacco field. *Computers and Electronics in Agriculture*, v. 109, p. 172-178, 2014.
- BOTTEGA, E. L. Utilização de zonas de manejo para a produção de soja no cerrado brasileiro. Tese (Doutorado em Engenharia Agrícola) – Universidade Federal de Viçosa. Viçosa, 2014.
- CÓRDOBA, M., et al. Subfield management class delineation using cluster analysis from spatial principal components of soil variables. *Computers and electronics in agriculture*, 2013, 97: 6-14.
- CORWIN, D.L.; LESCH, S.M. Application of Soil Electrical Conductivity to Precision Agriculture: Theory, Principles, and Guidelines. *Agronomy Journal*, v.95, n. 3, p.471-471. 2003.
- FU, Q.; WANG, Z.; QIUXIANG, J. Delineating soil nutrient management zones based on fuzzy clustering optimized by PSO. *Mathematical and Computer Modelling*, v.51, p. 1299-1305, 2010.
- FU, Weijun et al. Outlier identification of soil phosphorus and its implication for spatial structure modeling. *Precision agriculture*, v. 17, n. 2, p. 121-135, 2016.

GUEDES, Carvalho, et al. Influence of incorporating geometric anisotropy on the construction of thematic maps of simulated data and chemical attributes of soil. *Chilean journal of agricultural research*, 2013, 73.4: 414-423.

HAGHVERDI, Amir et al. Perspectives on delineating management zones for variable rate irrigation. *Computers and Electronics in Agriculture*, v. 117, p. 154-167, 2015.

KERRY, R.; OLIVER, M. A. Comparing sampling needs for variograms of soil properties computed by the method of moments and residual maximum likelihood. *Geoderma*, v. 140, n. 4, p. 383-396, 2007.

MARONNA, R. A. R. D.; MARTIN, R. Douglas; YOHAI, Victor. *Robust statistics*. John Wiley & Sons, Chichester. ISBN, 2006.

MORAL, F. J.; TERRÓN, J. M.; DA SILVA, JR Marques. Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. *Soil and Tillage Research*, 2010, 106.2: 335-343.

MORAL, F. J.; TERRÓN, J. M.; REBOLLO, F. J. Site-specific management zones based on the Rasch model and geostatistical techniques. *Computers and Electronics in Agriculture*, 2011, 75.2: 223-230.

MORARI, F., CASTRIGNANO, A., PAGLIARIN, C. Application of multivariate geostatistics in delineating management zones within a gravelly vineyard using geoelectrical sensors. *Computers and Electronics in Agriculture* 68, 97 – 107, 2009.

ROBERT, P. C., Precision agriculture: Research needs and status in the USA. In J. V. Stafford (Ed.), *Precision Agriculture '99* (pp. 19–33). Sheffield, UK: Sheffield Academic Press, 1999.

SAFANELLI, J. L., BOESING, B. F. B., & BOTTEGA, E. L. Estabelecimento de zonas de manejo a partir da resposta espectral do solo relacionada ao teor de matéria orgânica. Anais: XVII Simpósio Brasileiro de Sensoriamento Remoto. INPE, 2015.

SCUDIERO, Elia et al. Delineation of site-specific management units in a saline region at the Venice Lagoon margin, Italy, using soil reflectance and apparent electrical conductivity. *Computers and electronics in agriculture*, v. 99, p. 54-64, 2013.

SHAHBAZI, F. et al. Geostatistical analysis for predicting soil biological maps under different scenarios of land use. *European journal of soil biology*, v. 55, p. 20-27, 2013.

SHANER, D. L.; FARAHANI, H. J.; BUCHLEITER, G. W. Predicting and Mapping Herbicide–Soil Partition Coefficients for EPTC, Metribuzin, and Metolachlor on Three Colorado Fields. *Weed Science*. v. 56, p. 133– 139, 2008.

TRIPATHI, RAHUL, et al. Delineation of soil management zones for a rice cultivated area in eastern India using fuzzy clustering. *Catena*, 2015, 133: 128-136.

YAO, Rong-Jiang et al. Determination of site-specific management zones using soil physico-chemical properties and crop yields in coastal reclaimed farmland. *Geoderma*, v. 232, p. 381-393, 2014.

ARTIGO I – ANISOTROPIA GEOMÉTRICA NA DELIMITAÇÃO DE ZONAS DE MANEJO.

Resumo - O manejo adequado do solo, além de permitir a redução de contaminantes e a maximização da produtividade agrícola, está diretamente relacionado ao conhecimento da variabilidade espacial dos atributos do solo. Esta variabilidade espacial pode expressar-se de forma isotrópica e anisotrópica, sendo a última, negligenciada nas pesquisas relacionadas ao delineamento de zonas de manejo. Nesse contexto, o presente trabalho objetivou avaliar alterações em mapas de zonas de manejo devido à correção da anisotropia. A metodologia foi aplicada sobre um banco de dados de produtividade de soja e condutividade elétrica aparente (CEa) de uma propriedade rural em Ponta Porã – MS. Por meio deste banco de dados georreferenciados interpolaram-se mapas com uso da krigagem ordinária. Para cada combinação, variável (produtividade e CEa) e número de classes, foram produzidos dois mapas de zonas de manejo, um sem e outro com a correção da anisotropia, em que os mesmos foram comparados por meio do índice kappa. Tal índice foi avaliado quanto à sua significância pelo teste Z, que conseqüentemente permitiu aferir acerca da concordância entre classes nos diferentes mapas. O número de zonas de manejo também foi avaliado quanto ao Índice de Performance Fuzzy (FPI) e a Entropia da

Partição Modificada (MPE). Apesar de estes índices terem apresentados valores mais próximos do ideal para as zonas de manejo geradas com a correção da anisotropia, o índice kappa não detectou diferença significativa em relação aos mapas obtidos sem a correção.

Palavras-chave: Geoestatística. Fator de anisotropia. Variabilidade espacial. Condutividade elétrica aparente. Soja.

Abstract - The proper handling soil allows the reduction of contaminants, maximize agricultural productivity, and is directly related the knowledge spatial variability soil attributes. This spatial variability can express isotropic and anisotropic form. Being the last, neglected in research related management zones delineation. This context, the present study aimed evaluate the correction effect geometric anisotropy on management zone delineation. The methodology was applied under database of soybean productivity and apparent electrical conductivity (CEa) of a rural property in Ponta Porã – MS. By means this georeferenced database, maps was interpolated with ordinary kriging. For each combination, attribute (productivity and CEa) and number of classes, were produced two maps management zones, one without and one with anisotropy correction, the same were compared through the kappa index. The index was evaluated by Z-test, which consequently allowed gauge the similarity between maps. The management zones number was also evaluated by Fuzziness Performance Index (FPI) and the Modified Partition Entropy (MPE). Although those index have presented ideal values closer to maps with the correction anisotropy, the kappa index no detected any significant difference in relation maps obtained without fix.

Key words: Geostatistics. Anisotropy factor. Spatial variability. Apparent electrical conductivity. Soybeans.

Introdução

A geoestatística permite identificar áreas com baixa produtividade bem como identificar fatores que a influenciam. O conhecimento do padrão da variabilidade espacial desses fatores permite o tratamento às taxas variáveis nas sub-regiões do campo.

Estas sub-regiões são denominadas de zonas de manejo. A zona de manejo pode ser definida como uma sub-região do campo que necessita receber de maneira uniforme a dosagem de insumos, pois apresenta os mesmos níveis de fatores limitantes à produtividade das culturas (BOTTEGA, 2014). Sua utilização permite a conservação do solo (DALCHIAVON et al., 2012), redução de custos de produção e dos impactos ambientais (BOTTEGA, 2014).

Dentre as questões em torno da geração de zonas de manejo é fundamental identificar os principais parâmetros em sua implantação (FRIDGEN et al., 2004). Nesse sentido, destaca-se a utilização da condutividade elétrica aparente do solo (CEa) cujo padrão de variabilidade espacial relaciona-se ao padrão de variação espacial das culturas e das propriedades físico-químicas do solo (CORWIN et al., 2003).

Morari, Castrignano e Pagliarin (2009) utilizaram valores de CEa e outras propriedades físicas do solo para a caracterização da variabilidade espacial com uso da krigagem. Os resultados desse estudo permitiram isolar fontes de variações que

atuavam em diferentes escalas espaciais. Estas fontes de variações possibilitaram delimitar zonas de manejo por meio do algoritmo *fuzzy k-means*.

Molin e Castro (2008) e Bottega (2014) também utilizaram dados de CEa e propriedades do solo na caracterização da variabilidade espacial por meio da krigagem ordinária. Os mapas interpolados foram usados na delimitação de zonas de manejo aplicando o método *fuzzy K-means*.

Em face dos estudos mencionados, é notório que a variabilidade dos atributos do solo, em especial a CEa, pode auxiliar a delimitação de zonas de manejo ao usar métodos de classificação. A variabilidade espacial destes atributos tem sido modelada e mapeada por meio de preditores lineares geoestatísticos. O uso destes mapas permite expor as causas da variabilidade da produção (QUEIROZ *et al.*, 2000).

Na delimitação de zonas de manejo os mapas de variabilidade têm sido gerados com uso de diversos interpoladores geoestatísticos como a krigagem ordinária (ALVES *et al.*, 2013; CHANG *et al.*, 2014; FU; WANG; JIANG., 2010; MORAL; TERRÓN; REBOLLO, 2011; SAFANELLI; BOESING; BOTTEGA, 2015; TRIPATHI *et al.*, 2015), a cokrigagem (MORARI; CASTRIGNANO; PAGLIARIN, 2009), a regressão krigagem (MORAL; TERRÓN; DA SILVA, 2010), a programação linear (CID-GARCIA *et al.*, 2013) entre outros.

No entanto, estas pesquisas não avaliam a presença da anisotropia geométrica. A anisotropia geométrica é ocasionada quando o semivariograma apresenta distintos padrões de dependência espacial em diferentes direções. A correção da anisotropia geométrica permite uma maior precisão na construção de

mapas temáticos que descrevem a variabilidade espacial dos atributos do solo (GUEDES et al., 2013).

Nesse contexto, o presente trabalho tem como objetivo avaliar alterações no estabelecimento das zonas de manejo devido à correção da anisotropia.

Material e Métodos

As metodologias neste trabalho foram aplicadas em uma base de dados coletados por Bottega (2014) em uma propriedade rural de Ponta Porã – MS. Essa base de dados contém mensurações, em 160 pontos georreferenciados, das variáveis CEa e produtividade de soja na safra 2011/2012 (Figura 1). As medições da CEa foram obtidas na camada de 0 a 20cm.

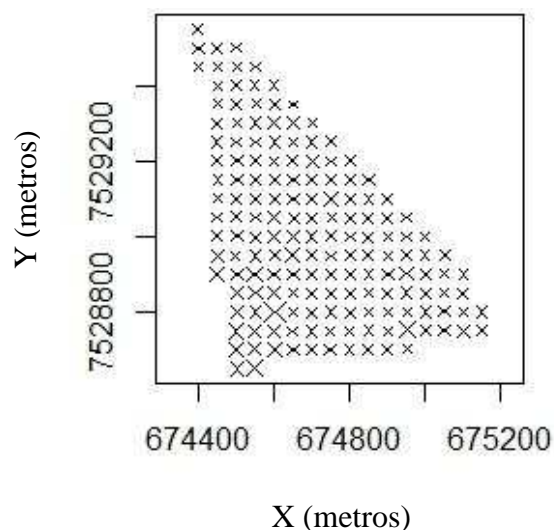


Figura 1 - Malha amostral, regularmente espaçada em 50 metros. Coordenadas UTM, zona 21 Sul.

O georreferenciamento dos pontos foi realizado utilizando um receptor GPS da marca Garmin, modelo GPSMAP 62. Os valores de produtividade foram obtidos manualmente em 160 pontos amostrais georreferenciados, compostos de três linhas de um metro linear, com área de 1,35 m² sendo os valores expressos em kg ha⁻¹.

Para cada uma das variáveis mensuradas foi ajustado um modelo de dependência espacial com e sem a correção da anisotropia. Tal modelo foi ajustado ao semivariograma empírico obtido a partir do cálculo da semivariância $\gamma(h)$, conforme a equação:

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

para diferentes distâncias de separação (MANTO, 2005). Em que h é distância entre pares de observações e $N(h)$ é o número de pares de observações $Z(x_i)$ e $Z(x_i + h)$ observados nas posições x_i e $x_i + h$ e que estão separados por uma distância h .

Para a modelagem da estrutura desconsiderando a correção da anisotropia, foram aplicados os ajustes de mínimos quadrados ponderados (WLS) aos modelos gaussianos

$$y(h) = c_0 + c \left\{ 1 - \exp\left(-\frac{h^2}{a^2}\right) \right\} \text{ para } h > 0, r = \sqrt{3a},$$

exponencial

$$y(h) = c_0 + c \left\{ 1 - \exp\left(-\frac{h}{a}\right) \right\} \text{ para } h > 0, r = 3a$$

e esférico

$$y(h) = c_0 + c \left\{ \frac{3h}{2r} - \frac{1}{2} \left(\frac{h}{r}\right)^3 \right\} \text{ para } 0 < h \leq r.$$

Em que: c_0 define o efeito pepita, c corresponde a contribuição e r define o alcance.

Os modelos ajustados receberam como estimativas iniciais os parâmetros encontrados por Bottega (2014). Cada ajuste foi submetido ao preditor linear geoestatístico da krigagem ordinária, desconsiderando o efeito de tendência.

E como critério de seleção, utilizou-se o modelo que apresentou menor média dos erros de validação cruzada da krigagem (Mean Error: ME):

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

em que $\hat{Z}(x_i)$ é o valor da estimativa no ponto não observado x_i .

Segundo Oliver e Webster (2014), mesmo com uma má escolha do modelo, a média dos erros igual a zero garante que a krigagem seja não viesada. Além da ME, para garantir a escolha correta do modelo, seguiu-se com a sugestão do mesmo autor supracitado e utilizou-se o erro quadrático médio ponderado (Mean Squared Deviation Ratio: MSDR) com valor ideal esperado, o mais próximo de 1.

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

Sendo $\hat{Z}(x_i)$ a estimativa de krigagem da variável amostrada no ponto x_i e $\hat{\sigma}_K^2(x_i)$ a estimativa da variância de krigagem da variável amostrada no ponto x_i .

O modelo selecionado foi então utilizado na mensuração do índice de dependência espacial (IDE) proposto por Zimback (2001):

$$IDE = \left(\frac{c}{c + c_0} \right) 100$$

Tal índice foi definido como a porcentagem da contribuição em relação ao patamar. Assim, a dependência espacial foi classificada como: forte para IDE > 75%, moderada para IDE entre 25% e 75% e fraca para IDE < 25%.

Todos os passos mencionados também foram empregados na modelagem cuja correção da anisotropia foi considerada. Neste caso, o valor do semivariograma não dependeu somente do comprimento do vetor h , mas também da sua direção.

Assim, para cada ajuste referenciado acima, seguiu-se os passos descritos em Manto (2005). Primeiramente identificaram-se os eixos de anisotropia (eixos de maior e menor alcance) por meio de semivariogramas experimentais em diferentes direções (0° , 45° , 90° e 135° com tolerância de $22,5^\circ$).

Em seguida, foi feita uma translação dos eixos da matriz de dados, coincidindo com os eixos da anisotropia por meio da matriz de rotação R:

$$R = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

onde α é o ângulo na direção norte com o maior eixo da elipse. Feito isso, os semivariogramas direcionais são combinados em um único semivariograma, em que o alcance é normalizado para tamanho 1 com uso da matriz T:

$$T = \begin{pmatrix} 1/a_{max} & 0 \\ 0 & 1/a_{min} \end{pmatrix}.$$

em que a_{max} e a_{min} são os eixos da elipse de anisotropia com maior e menor alcance, respectivamente. E então, as matrizes de rotação R e translação de eixos T foram combinadas de forma a obter um único semivariograma isotrópico, dado por $\gamma(x_i - x_j) = \gamma(\|TR(x_i - x_j)\|)$. Diggle e Ribeiro Junior (2007) também utilizaram o mesmo processo usando a transformação linear (rotação e translação) por meio da matriz A:

$$A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha)/F_a & \cos(\alpha)/F_a \end{pmatrix}.$$

E dessa forma, a correção da anisotropia geométrica foi definida pelos parâmetros α (direção de maior continuidade espacial) e fator de anisotropia $F_a =$

a_2/a_1 , sendo a_2 e a_1 respectivamente os valores de maior e menor alcance dos semivariogramas direcionais.

Todos os mecanismos metodológicos utilizados na modelagem da estrutura de dependência espacial foram realizados com uso do Software R (R CORE TEAM, 2016) usando o pacote geoR (RIBEIRO JR.; DIGGLE, 2016).

Após a modelagem da dependência espacial, os mapas interpolados para as variáveis CEa e produtividade da soja foram processados pelo programa computacional KrigMe (VALENTE *et al.*, 2012) o qual utiliza o método de classificação *fuzzy k-means* para delimitar as zonas de manejo.

O algoritmo de classificação *fuzzy k-means* determina o número de zonas de manejo por meio do Índice de Performance Fuzzy (FPI – “Fuzziness Performance Index”) e o Índice de Entropia da Partição Modificada (MPE – “Modified Partition Entropy”), que relacionam a desorganização dos dados quanto à definição das zonas de manejo. Quando estes dois índices apresentam valores mínimos, tem-se o número de classes ótimo estabelecido (SONG *et al.*, 2009).

Para cada combinação, variável e número de classes, foram produzidos dois mapas de zonas de manejo, um sem e outro com a correção da anisotropia, em que os mesmos foram comparados por meio do índice kappa:

$$\hat{K} = \left(n \sum_{i=1}^c x_{ii} - \sum_{i=1}^c x_{i\cdot} x_{\cdot i} \right) / \left(n^2 - \sum_{i=1}^c x_{i\cdot} x_{\cdot i} \right)$$

Onde:

\hat{K} : estimativa do coeficiente kappa; n : número de amostras; c : total de classes; x_{ii} : diagonal da matriz de erros; $x_{i\cdot}$: total da linha i ; $x_{\cdot i}$: total da coluna i .

Tal índice foi avaliado quanto a sua significância pelo teste Z:

$$Z = (\widehat{K}_1 - \widehat{K}_2) / \sqrt{\text{var}(\widehat{K}_1) + \text{var}(\widehat{K}_2)}$$

Sendo Z : valor calculado da estatística Z ; \widehat{K}_1 e \widehat{K}_2 : estimativa dos respectivos Kappa comparados; $\text{var}(\widehat{K}_1)$ e $\text{var}(\widehat{K}_2)$: as correspondentes variâncias das estimativas dos respectivos Kappa comparados. Sendo as variâncias obtidas da seguinte forma:

$$\text{var}(\widehat{K}) = \frac{1}{n} \left[\frac{\theta_1(1 - \theta_1)}{(1 - \theta_2)^2} + \frac{2(1 - \theta_1)(2\theta_1\theta_2 - \theta_3)}{(1 - \theta_2)^3} + \frac{(1 - \theta_1)^2(\theta_4 - 4\theta_2^2)}{(1 - \theta_2)^4} \right]$$

De modo que:

$$\theta_1 = \frac{1}{n} \sum_{i=1}^c x_{ii};$$

$$\theta_2 = \frac{1}{n^2} \sum_{i=1}^c x_{i\circ} x_{\circ i}$$

$$\theta_3 = \frac{1}{n^2} \sum_{i=1}^c x_{ii} (x_{i\circ} + x_{\circ i})$$

$$\theta_4 = \frac{1}{n^3} \sum_{i=1}^c \sum_{j=1}^c x_{ij} (x_{j\circ} + x_{\circ j})^2$$

Sendo: x_{ij} o valor na matriz de erro correspondente na linha i e coluna j ; $x_{j\circ}$ o total correspondente na linha j e $x_{\circ j}$ o total correspondente na coluna j .

Tal índice permitiu aferir acerca da concordância entre classes nos mapas. Dessa forma, mapas cujo coeficiente kappa é igual à zero, foram totalmente discordantes, e quanto mais próximos de um, maior a concordância. Assim, estabeleceram-se limites de semelhanças entre os mapas por meio do coeficiente

kappa conforme propõem Landis e Coch (1977), de modo que, o nível de concordância entre os mapas foi classificado como ruim para valores de kappa entre 0,00 e 0,19, razoável para valores de kappa entre 0,20 a 0,39, bom entre 0,40 a 0,59, muito bom 0,60 a 0,79 e excelente para valores de kappa maiores ou iguais a 0,80.

Resultados e Discussão

A amplitude total apontada para a CEa (Tabela 1) reforça os resultados expostos em Machado et al. (2006) que variaram de 1,9 a 13,7 em torno da média 5,2 mS m⁻¹. Por outro lado, Peralta et al. (2015) encontraram maiores valores médios de CEa, variando de 18,4 a 27,1, em 5 diferentes campos experimentais. Entretanto, na coleta de informações da CEa, Peralta et al. (2015) utilizaram profundidades de 30 e 90 cm.

Tabela 1 - Estatísticas descritivas associadas às variáveis produtividade e condutividade elétrica aparente.

Var ¹	N ²	Mín ³	Max ⁹	Med ⁴	Média	DP ⁵	CV ⁶	CA ⁷	CK ⁸
Prod ¹⁰	160	1225	3.244	2.345	2.302	298	13	-0,35	0,85
CEa ¹¹	160	2,74	19,31	5,94	6,19	2,13	34,44	2,24	9,29

¹Variáveis; ²Número de observações; ³Mínimo; ⁴Mediana; ⁵Desvio-padrão; ⁶Coefficiente de variação (%); ⁷Coefficiente de Assimetria; ⁸Coefficiente de curtose; ⁹Máximo; ¹⁰Prod - Produtividade (kg ha⁻¹); ¹¹CEa - Condutividade elétrica aparente do solo (mS m⁻¹).

Apesar da diferença de escala na mensuração das variáveis, a produtividade apresentou menor variação percentual dos dados em torno da média, que a apresentada para a CEa. Valor semelhante de coeficiente de variação (35,1), para esta mesma variável, também foi verificado por Aggelopoulou et al. (2013). Esse autor também obteve uma distribuição dos valores de CEa com assimetria positiva (Tabela 1). Já a produtividade apresentou distribuição com assimetria negativa. Além disso, as distribuições dos valores de produtividade e de CEa apresentaram diferentes graus de achatamento (curtose).

Na análise variográfica sem correção da anisotropia: na Tabela 2 são apresentados os parâmetros dos modelos do semivariograma omnidirecional selecionado com base nas estatísticas da validação cruzada para as variáveis produtividade e CEa.

Tabela 2 - Especificações dos modelos omnidirecionais ajustados sem a correção da anisotropia.

Variáveis	Parâmetros					
	Modelo	Efeito Pepita	Patamar	Alcance P. ³	ME ⁴	MSDR ⁴
Produtividade ¹	Gaussiano	18.948,7	108.527	374,92	0,48	1,3
CEa ²	Gaussiano	3,11	5,81	836,24	0,00	1,0

¹(kg ha⁻¹); ²Condutividade elétrica aparente do solo (mS m⁻¹); ³Alcance Prático; ⁴Média dos erros da validação cruzada da krigagem; ⁵Erro quadrático médio ponderado da validação cruzada da krigagem.

Para ambas as variáveis foi selecionado o modelo gaussiano ajustado pelo método dos mínimos quadrados ponderados. O critério de seleção dos modelos foi à

média dos erros da validação cruzada da krigagem. Quanto menor a média dos erros da validação cruzada da krigagem, maior a acurácia (SILVA JUNIOR et al., 2012). Este mesmo critério também foi utilizado por Zhang et al. (2016), como uma das estatísticas na seleção de modelos. Adicionalmente, verificou-se o erro quadrático médio ponderado da validação cruzada da krigagem, cujo valor deve ser o mais próximo de um.

Os modelos de semivariograma selecionados para as variáveis produtividade e CEa explicaram respectivamente que, 85% e 65% da variação total dos dados são devidos à dependência espacial, conforme índice de dependência espacial (IDE) proposto por Zimback (2001).

Bottega (2014) selecionou o modelo esférico para a variável produtividade. Apesar da diferença do modelo teórico selecionado, as estimativas dos parâmetros do modelo ajustado por Bottega (2014) foram semelhantes às expostas na Tabela 2, exceto para o parâmetro efeito pepita (100).

Por outro lado, Bottega (2014) também selecionou modelo gaussiano para a variável CEa. As estimativas encontradas para os parâmetros foram próximas às apresentadas na Tabela 2, com exceção do parâmetro alcance (539).

Análise variográfica com correção para a anisotropia: A obtenção de semivariogramas empíricos direcionais constatou a existência da anisotropia. Foi constatado para a variável produtividade, ângulo de anisotropia em 0° e fator de anisotropia 1,5; enquanto que na variável CEa constatou-se ângulo de anisotropia em 90° e fator de anisotropia 2,0.

Os parâmetros da anisotropia, de cada variável, foram aplicados nas transformações lineares de coordenadas espaciais de suas respectivas amostras

georreferenciadas, conforme indicado em Manto (2005) e Diggle e Ribeiro Junior (2007).

Após as transformações lineares nas coordenadas espaciais, novos semivariogramas direcionais foram construídos nas direções de maior e menor continuidade espacial e respectivo fator de anisotropia. Os parâmetros dos semivariogramas direcionais do modelo gaussiano obtidos após transformações das coordenadas espaciais são apresentados na Tabela 3.

Tabela 3 - Parâmetros dos semivariogramas direcionais relativos aos ângulos de maior e menor continuidade espacial, após transformações lineares nas coordenadas georreferenciadas das variáveis.

Variável	Modelo	Ângulo	Efeito Pepita	Patamar	Alcance Prático
Prod ¹ (kg ha ⁻¹)	Gaussiano	0°	20.436	105.025	110
	Gaussiano	90°	19.986	92.442	90
CEa (mS m ⁻¹)	Gaussiano	90°	3,11	5,84	685
	Gaussiano	0°	2,62	3,73	597

¹Produtividade

Ao comparar as estimativas dos parâmetros dos semivariogramas ajustados nos ângulos 0° e 90° para a variável produtividade, observou-se valor semelhante para a estimativa do parâmetro efeito pepita e que o eixo de maior continuidade espacial foi apenas 22% maior que o eixo de menor continuidade espacial. Além disso, o patamar apresentou estimativa 15% maior no eixo de maior dependência espacial.

Comportamento semelhante também foi observado para a variável CEa, em que o aumento percentual na estimativa do efeito pepita, não ultrapassa 20%, ao

contrapor as estimativas do eixo de menor continuidade espacial para o de maior continuidade espacial. Nas estimativas do alcance prático, foi observado apenas um incremento percentual inferior a 0,15%. Entretanto para o patamar, notou-se um aumento relativo próximo a 50%.

Esses resultados indicam que as transformações lineares foram capazes de corrigir o efeito de anisotropia, a tal ponto que, ao ajustar semivariograma nos dois eixos de anisotropia (0° e 90°) os modelos ajustados apresentaram estimativas semelhantes para os parâmetros (pepita, patamar e alcance).

Portanto, com a finalidade de corrigir a anisotropia geométrica, as coordenadas georreferenciadas dos pontos amostrados foram submetidas às transformações lineares e um modelo de semivariograma omnidirecional foi obtido para cada uma das variáveis (Tabela 4).

Tabela 4 - Parâmetros dos semivariogramas omnidirecionais após transformações lineares nas coordenadas georreferenciadas dos pontos amostrados onde foram mensuradas a variável produtividade e CEa.

Variável	Modelo	ME ³	MSDR ⁴	Efeito Pepita	Patamar	Alcance Prático
Prod ¹	Gaussiano	0,09	0,80	20.000	92.250	138
CEa ²	Gaussiano	0,00	1,1	2,97	4,6	635

¹Produtividade (kg ha⁻¹); ²Condutividade elétrica aparente (mS m⁻¹); ³Média dos erros da validação cruzada da krigagem; ⁴Erro quadrático médio ponderado da validação cruzada da krigagem.

Os IDEs (Zimback, 2001) dos semivariogramas omnidirecionais foram 82% e 60% para produtividade e CEa respectivamente. Como pode ser observado, após correção da anisotropia geométrica, houve redução no IDE para ambas variáveis. Em contra partida, para a variável CEa a média dos erros da validação cruzada foi

idêntica a encontrada na análise variográfica que desconsiderou a anisotropia. Já a variável produtividade permaneceu com a mesma classificação de IDE.

Os parâmetros destes modelos de semivariograma omnidirecionais foram então utilizados no preditor linear geoestatístico de krigagem ordinária. A utilização deste preditor permitiu expressar os mapas de variabilidade espacial das variáveis.

Definição das zonas de manejo: tantos os mapas desenvolvidos na análise sem a correção da anisotropia, quanto aos desenvolvidos na análise com correção da anisotropia, foram submetidos ao algoritmo *fuzzy k-means* para o delineamento das zonas de manejo com diferentes números de classes.

Em cada combinação, foram avaliados os índices FPI e MPE para ambas as variáveis, produtividade e CEa (Figura 2).

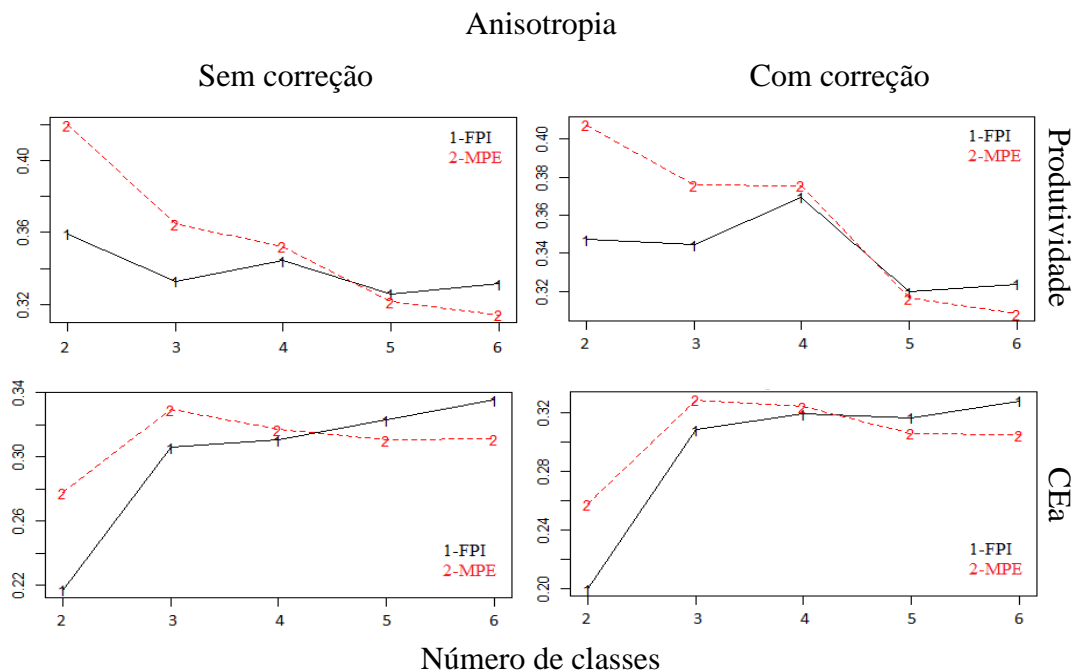


Figura 2 - Índice de Performance Fuzzy (FPI) e Entropia da Partição Modificada (MPE), relativo ao número de classes na delimitação de zonas de manejo para as variáveis produtividade e condutividade elétrica aparente (CEa), com e sem correção da anisotropia.

Assim como Peralta et al. (2015) e Chang et al. (2014), os índices foram avaliados em até seis classes. Como pode ser notado na Figura 2, para a variável produtividade, à medida que aumentou o número de classes, ambos os índices com e sem correção da anisotropia aproximaram-se de zero. Já a variável CEa apresentou comportamento oposto.

Ao comparar os índices com e sem correção da anisotropia os menores valores ocorreram com a correção da anisotropia. Que por sua vez, indicaram melhor organização dos dados (SONG et al., 2009) em relação à outra análise.

Os índices obtidos foram próximos aos encontrados por Bottega (2014) para a produtividade (FPI = 0,3276 e MPE = 0,3215) e CEa (FPI = 0,2071 e MPE = 0,2668). Além da similaridade obtida por Bottega (2014), Valente et al. (2012) também encontraram valores semelhantes de FPI e MPE para a CEa (0,22 e 0,27). Peralta et al. (2015) encontraram valores mais baixos de FPI e MPE para a CEa (0,03 a 0,07), todavia as camadas utilizadas pelos autores foram de 0 a 30cm e 0 a 90cm. Os autores Morari, Castrignano e Pagliarin (2009) por meio de mapas interpolados com a krigagem encontraram valores similares de FPI e MPE aos obtidos por Peralta et al. (2015), variando entre 0,06 e 0,04, entretanto, as variáveis avaliadas foram areia, argila e cascalho.

Para cada número de classes verificado quanto aos índices mencionados (FPI, MPE), produziram-se mapas com suas respectivas zonas de manejo. Os mapas produzidos sem correção da anisotropia foram comparados aos corrigidos para ambas variáveis (produtividade e CEa) para cada número de classes. A comparação deu-se via índice kappa e respectivo teste Z (Figura 3).

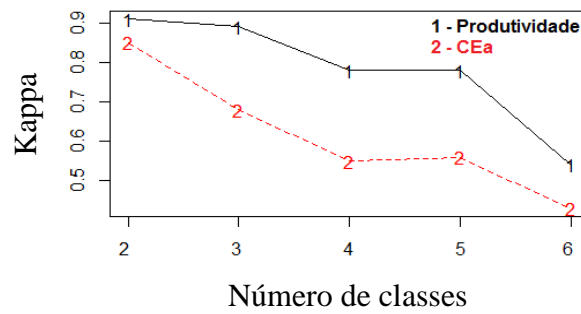


Figura 3 - Índice kappa, em função do número de classes, comparando mapas com e sem correção da anisotropia, para as variáveis produtividade e condutividade elétrica aparente (CEa).

À medida que o número de classes aumentou (Figura 3), os valores de kappa tendem a reduzir, entretanto, independentemente do número de classes, todos os valores de kappa foram significativos a 5% de probabilidade pelo teste Z. Dessa forma, o nível de concordância entre os mapas comparados na variável produtividade apresentaram classificação excelente ao utilizar duas ou três classes, conforme propõe Landis e Koch (1977), entretanto, para quatro ou cinco classes foi classificado como muito bom.

Para a variável CEA, apenas com a utilização de duas classes, o nível de semelhança entre os mapas pode ser classificado como excelente, enquanto que, com três classes houve uma classificação muito boa e com quatro ou cinco classes, classificado como bom.

Os valores de kappa apresentados na Figura 3 foram semelhantes aos obtidos por Guedes et al. (2013) ao comparar mapas de krigagem com estrutura isotrópica e anisotrópica para dados simulados de atributos químicos do solo. Os mesmos autores, ao considerar valores reais de um estudo de caso, obtiveram valores de

kappa entre 0,7 a 1 para uma amostra de 100 observações com fator de anisotropia de 1 a 5.

Em síntese, por mais que os valores de FPI e MPE sejam menores para os mapas corrigidos à anisotropia em relação aos que não apresentam a correção, o índice kappa mostrou que os mapas de zonas de manejo obtidos sem e com correção, apresentaram padrões estatisticamente concordantes. Nesse caso, por mais que a correção da anisotropia aumente a precisão na construção de mapas temáticos (BOISVERT et al., 2009; CHORTI; HRISTOPULOS, 2008; FACAS et al., 2010; GUEDES et al., 2013; ZIMMERMAN, 1993), o delineamento das zonas de manejo não foi afetado.

Resultados semelhantes foram encontrados por Guedes et al. (2013) ao comparar estruturas simuladas de propriedades físico-químicas do solo, de dependência espacial, utilizando o índice kappa. Esse autor observou que, mesmo com valores altos de anisotropia (3,4 e 4,5), há similaridade entre os mapas gerados. Esta mesma similaridade igualmente foi observada por meio da exatidão global e índice *Tau*.

Dessa forma, considerando os menores valores de FPI e MPE, e a significância do kappa, recomenda-se respectivamente, o uso de 2 classes na geração das zonas de manejo para a CEa, e 5 classes para a produtividade. Em ambas as variáveis independem se há ou não correção da anisotropia geométrica. Dessa forma, a Figura 4 exibe de maneira ilustrativa os mapas de zonas de manejo obtidos sem a correção da anisotropia, para as respectivas variáveis.

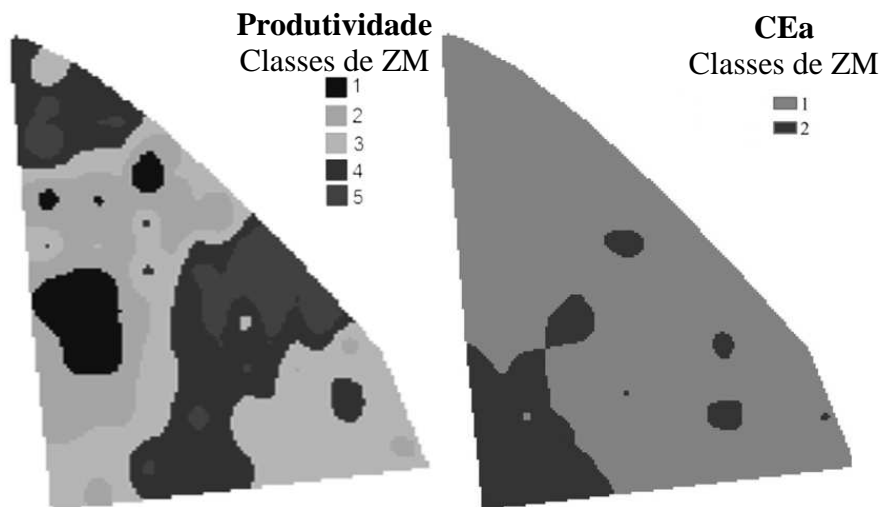


Figura 4 - Mapas de zonas de manejo para a variável produtividade e CEa obtidos sem correção da anisotropia geométrica.

Conclusão

Apesar dos índices FPI e MPE terem apresentados valores mais próximos do ideal com a correção da anisotropia geométrica, o índice kappa não detectou diferença significativa em relação aos que foram obtidos sem a correção.

Assim, ignorar ou transformar a variável anisotrópica geométrica em isotrópica, não afetou significativamente o agrupamento dos dados em classes por meio do algoritmo *fuzzy k-means*, que por sua vez, delimitou as zonas de manejo neste estudo de caso.

Entretanto, conclusões adversas podem ocorrer ao analisar variáveis com distintos padrões de dependência espacial e diferentes valores para o fator de anisotropia.

Referências

- AGGELOPOOULOU, Katerina et al. Delineation of management zones in an apple orchard in Greece using a multivariate approach. *Computers and electronics in agriculture*, v. 90, p. 119-130, 2013.
- ALVES, S.M.F.; ALCÂNTARA, G.R.; REIS, E.F.; QUEIROZ, D.M.; VALENTE, D.S.M. definição de zonas de manejo a partir de mapas de condutividade elétrica e matéria orgânica. *Bioscience Journal*, v. 29, n. 1, p. 104-114, 2013.
- BOISVERT, J.B.; MANCHUK, J.G.; DEUTSCH C.V. 2009. Kriging in the presence of locally varying anisotropy using non-euclidian distances. *Mathematical Geosciences* 41:585-601.
- BOTTEGA, E. L. Utilização de zonas de manejo para a produção de soja no cerrado brasileiro. Tese (Doutorado em Engenharia Agrícola) – Universidade Federal de Viçosa. Viçosa, 2014.
- CHANG, Dong, et al. Delineation of management zones using an active canopy sensor for a tobacco field. *Computers and Electronics in Agriculture*, v. 109, p. 172-178, 2014.
- CHORTI, A.; HRISTOPULOS, D.T. 2008. Nonparametric identification of anisotropic (elliptic) correlations in spatially distributed data sets. *IEEE Transactions on signal processing* 56:4738-4750.

CID-GARCIA, Nestor M., et al. Rectangular shape management zone delineation using integer linear programming. *Computers and electronics in agriculture*, v. 93, p. 1-9, 2013.

CORWIN, D.L., LESCH, S.M., SHOUSE, P.J., SOPPE, R., AYARS, J.E. Identifying soil properties that influence cotton yield using soil sampling directed by apparent soil electrical conductivity. *Agron. J.* 95 (2), 352– 364, 2003.

DALCHIAVON, F.C.; CARVALHO, M.P.; ANDREOTTI, M.; MONTANARI, R. Variabilidade espacial de atributos da fertilidade de um Latossolo Vermelho Distroférico sob Sistema Plantio Direto. *Revista Ciência Agronômica*, v. 43, n. 3, p. 453-461, 2012.

DIGGLE, P.J., e P.J. RIBEIRO JR. *Model-based Geostatistics*. 230 p. Springer, New York, USA, 2007.

FACAS, N.W.; M.A. MOONEY; M. ASCE; R. FURRER. 2010. Anisotropy in the spatial distribution of roller-measured soil stiffness. *International Journal of Geomechanics* 10:129-135.

FRIDGEN, J.J., KITCHEN, N.R., SUDDUTH, K.A., DRUMMOND, S.T., WIEBOLD, W.J., FRAISSE, C. W., 2004. Management zone analyst (MZA). *Agron. J.* 96, 100–108. <http://dx.doi.org/10.2134/agronj2004.1000>.

GUEDES, Carvalho, et al. Influence of incorporating geometric anisotropy on the construction of thematic maps of simulated data and chemical attributes of soil. *Chilean journal of agricultural research*, 2013, 73.4: 414-423.

FU, Q.; WANG, Z.; QIUXIANG, J. Delineating soil nutrient management zones based on fuzzy clustering optimized by PSO. *Mathematical and Computer Modelling*, v.51, p. 1299-1305, 2010.

LANDIS, J.R.; KOCH, G.G. The measurement of observer agreement for categorical data. *Biometrics*, v.33, 1977.

MACHADO, P.L.O.A.; BERNARDI, A.C.C.; VALENCIA, L.I.O.; MOLIN, J.P.; GIMENEZ, L.M.; SILVA, C.A.; ANDRADE, A. G.; MEDARI, B.E.; MEIRELLES, M.S.P. Mapeamento da condutividade elétrica e relação com a argila de Latossolo sob plantio direto. *Pesquisa Agropecuária Brasileira*, Brasília, v.41, n.6, p.1023-1031, 2006.

MANTO, H. Modelling of geometric anisotropic spatial variation. *Mathematical Modelling and Analysis*, 2005, 361-366.

MOLIN, J. P.; CASTRO, C. N. Establishing management zones using soil electrical conductivity and other soil properties by the fuzzy clustering technique. *Scientia Agricola*, v.65, n.6, p.567-573, 2008.

MORAL, F. J.; TERRÓN, J. M.; DA SILVA, JR Marques. Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. *Soil and Tillage Research*, 2010, 106.2: 335-343.

MORAL, F. J.; TERRÓN, J. M.; REBOLLO, F. J. Site-specific management zones based on the Rasch model and geostatistical techniques. *Computers and Electronics in Agriculture*, 2011, 75.2: 223-230.

MORARI, F., CASTRIGNANO, A., PAGLIARIN, C. Application of multivariate geostatistics in delineating management zones within a gravelly vineyard using geoelectrical sensors. *Computers and Electronics in Agriculture* 68, 97 – 107, 2009.

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

PERALTA, NAHUEL RAÚL, et al. Delineation of management zones to improve nitrogen management of wheat. *Computers and Electronics in Agriculture*, 2015, 110: 103-113.

QUEIROZ, D.M., G.P. MANTOVANI, E.C. Agricultura de precisão na produção de grãos. In BORÉM, A.B.; GIÚDICE, M.P.; QUEIROZ, D.M.; MANTOVANI, E.C.; FERREIRA, L.R.; VALLE, F.X.R.; GOMIDE, R.T. Agricultura de Precisão. Viçosa: UFV, p 1-42, 2000.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

RIBEIRO JUNIOR, P. J., DIGGLE, P. J.(2016). geoR: Analysis of Geostatistical Data. R package version 1.7-5.2. <https://CRAN.R-project.org/package=geoR>.

SAFANELLI, J. L., BOESING, B. F. B., & BOTTEGA, E. L. Estabelecimento de zonas de manejo a partir da resposta espectral do solo relacionada ao teor de matéria orgânica. Anais: XVII Simpósio Brasileiro de Sensoriamento Remoto. INPE, 2015.

SILVA JUNIOR, João Fernandes da et al. Simulação geoestatística na caracterização espacial de óxidos de ferro em diferentes pedoformas. *Revista Brasileira de Ciência do Solo*, p. 1690-1703, 2012.

SONG, X.; WANG, J.; HUANG, W.; LIU, L.; YAN, G.; PU, R. The delineation of agricultural management zones with high resolution remotely sensed data. *Precision Agriculture*, v.10, p.471-487, 2009.

TRIPATHI, RAHUL, et al. Delineation of soil management zones for a rice cultivated area in eastern India using fuzzy clustering. *Catena*, 2015, 133: 128-136.

- VALENTE, D.S.M.; QUEIROZ, D.M.; PINTO, F.A.C.; SANTOS, N.T.; SANTOS, F.L. The relationship between apparent soil electrical conductivity and soil properties. *Revista Ciência Agronômica*, v. 43, n. 4, p. 683-690, 2012.
- ZHANG, Xiaohu et al. An improved method of delineating rectangular management zones using a semivariogram-based technique. *Computers and Electronics in Agriculture*, v. 121, p. 74-83, 2016.
- ZIMBACK, C.R.L. Análise espacial de atributos químicos de solos para fins de mapeamento da fertilidade do solo. 2001. 114 f. Tese (Livre-Docência) – Faculdade de Ciências Agronômicas, Universidade Estadual Paulista, Botucatu, 2001.
- ZIMMERMAN, D. 1993. Another look at anisotropy in geostatistics. *Mathematical Geology* 25:453-470.

ARTIGO II – KRIGAGEM ROBUSTA À *OUTLIERS*, NO DELINEAMENTO DE ZONAS HOMOGÊNEAS.

Resumo – Diversas pesquisas utilizam medidas de condutividade elétrica aparente do solo (CEa) como indicador da variabilidade espacial existente em um campo de produção. Com base nestas medidas, zonas de manejo (ZM) são delineadas para o aperfeiçoamento da gestão agrícola. Entretanto, estas amostras têm apresentado presença de *outliers*. Todavia, a presença ou incorreta detecção e exclusão de *outliers* altera o formato do semivariograma, exibindo estimativas viesadas para os seus parâmetros. Dessa forma, objetivou-se nesta pesquisa, avaliar variações em mapas de zonas de manejo quanto à utilização de metodologia robusta à *outliers*. Para tanto, utilizou-se de duas análises comparativas, robusta à *outliers* e não robusta a *outliers*. Na abordagem robusta utilizaram-se estimadores robustos de semivariâncias de Cressie Hawkins, Genton's e MAD Dowd e o plug-in de krigagem de deriva externa. Na abordagem não robusta à presença de *outliers* utilizou-se o estimador de semivariâncias de Matheron e a krigagem ordinária. Ambas as metodologias conduziram ao delineamento de duas zonas de manejo. Estas zonas apresentaram médias estatisticamente diferentes entre si, dentro de cada análise. Entretanto, os mapas delineados em cada análise apresentaram-se significativamente concordantes

entre si pelo índice Kappa a 1% de probabilidade. Com isso, para estes pontos amostrados na área em estudo, concluiu-se que, a utilização da metodologia robusta à *outliers*, não alterou significativamente o mapa de ZM.

Palavras-chave: Geoestatística robusta. Zonas de Manejo. Condutividade elétrica aparente. Variabilidade espacial. *Fuzzy k-means*.

Abstract – Several researches use soil apparent electrical conductivity measures (CEa) as indicator of the spatial variability existing in a field of production. And with this, management zones (ZM) are designed to improve agricultural management. However, these samples have *outliers*. In this case, the presence or incorrect detection and deletion of *outliers* change the variogram format and report unbiased estimates for its parameters. This way, was aimed this research, evaluate variations in management zones maps regarding the use of an outliers robust methodology. For this reason, two analyses were used, robust to the outliers and not robust to outliers. In robust approach was used robust estimator of semivariance of Cressie Hawkins, Genton's and MAD Dowd and plug-in of kriging with external drift. And no-robust approach, was used semivariances estimator of Matheron and ordinary kriging. Both methodologies delineated two management zones. These zones presented statistically different means between them, within each analysis. However, the maps delineated in each analysis presented themselves significantly similar by Kappa index at 1% probability. With this, for these points sampled in the area under study, it's concluded that, the use of an outliers robust methodology did not significantly modify the ZM map.

Keyword: Robust geostatistics. Management zones. Apparent soil electrical conductivity. Spatial variability. *Fuzzy k-means*.

Introdução

Com o advento da agricultura de precisão, a heterogeneidade do solo passou a ser mapeada, tornando-se um fator consideravelmente relevante na gestão agrícola. Com isso, pode-se definir a agricultura de precisão como um procedimento sistemático, que visa inspecionar e incorporar o conhecimento da variabilidade espacial do solo para o manejo do campo (HAGHVERDI et al., 2015). Sendo esta variabilidade causada tanto por fatores climáticos, topográficos e biológicos (CÓRDOBA et al., 2013).

A gestão desta variabilidade espacial do campo decorre via aplicação de zonas de manejo (ZM). As ZM são sub-regiões do campo que apresentam necessidades similares quanto aos atributos físico-químico do solo. A definição destas sub-regiões permite identificar particularidades próprias quanto à necessidade de insumos, e com isso, reduz-se consideravelmente a aplicação dos mesmos, favorecendo a sustentabilidade agrícola (CÓRDOBA et al., 2013; BOTTEGA et al., 2017).

Entretanto, para o delineamento das ZM, deve-se conhecer a variabilidade espacial no campo. O conhecimento desta variabilidade espacial, por vez, é então observado via modelagem da estrutura espacial dos fatores que implicam no manejo do campo. Dentre os diversos fatores a serem observados, destaca-se a utilização da condutividade elétrica aparente do solo (CEa) em diferentes pesquisas (CÓRDOBA et al., 2016; PERALTA et al., 2015; HAGHVERDI et al., 2015; CÓRDOBA et al., 2013; SCUDIERO et al., 2013; PERALTA et al., 2013). Sua ampla utilização deve-

se a facilidade amostral, baixo custo e forte interferência no crescimento e rendimento das plantas (PERALTA et al., 2013).

Dependendo das condições climáticas e entre outros fatores, as distribuições dos valores observados da CEa podem exibir forma simétrica (CÓRDOBA et al., 2013) ou assimétrica (SCUDIERO et al., 2013; YAO et al., 2014; BOTTEGA, 2014; SHAHBAZI; ALIASGHARZAD; EBRAHIMZAD, 2014). A assimetria tem efeito significativo no ajuste do modelo teórico do semivariograma (FU et al., 2016), nas estatísticas da validação cruzada da krigagem (KERRY e OLIVER, 2007) e na precisão do mapa de variabilidade espacial. Portanto, a assimetria compromete o delineamento das ZM prejudicando desta forma a sua gestão no campo.

Nesta perspectiva, distribuições assimétricas de dados amostrais devem ser analisadas com cautela devido a irrefutável ocorrência de *outliers*. *Outliers* são valores atípicos quando comparados aos demais (FU et al., 2016; CÓRDOBA et al., 2016). Diversos métodos como Q-Q plot, box plot, LISA e validação cruzada podem ser utilizados para sua detecção (FU et al., 2016).

Uma vez que estes valores atípicos são identificados, os mesmos são excluídos da análise (RAMOS et al., 2017; CÓRDOBA et al., 2016; FU et al., 2016; PICCINI; MARCHETTI; FRANCAVIGLIA, 2014). A exclusão de *outliers* na análise favorece a normalidade dos dados e evita a necessidade de transformação dos mesmos (FU et al., 2016). E dessa forma, viabiliza uma correta interpolação no mapa de variabilidade espacial e conseqüentemente uma correta tomada de decisão com base na análise (TAYLOR; MCBRATNEY; WHELAN, 2007).

Entretanto, a exclusão incorreta de *outlier* prejudica a inferência estatística (MARONA; MARTIN; YORAI, 2006), e conseqüentemente as estimativas dos

parâmetros do semivariograma serão comprometidas, logo, afetará na decisão correta do manejo no campo. Nesta perspectiva, a presente pesquisa objetivou avaliar variações no estabelecimento de zonas de manejo com uso de metodologia robusta à presença de outliers para a modelagem da dependência espacial.

Material e Métodos

Considerando as diversas características favoráveis no uso da CEa para o delineamento das zonas de manejo, apresentadas por Peralta et al. (2013), utilizou-se a mesma variável nesta pesquisa. Os dados georreferenciados relativos à CEa, cedidos por Bottega (2014), foram coletados em uma propriedade rural no município de Ponta Porã no estado do Mato Grosso do Sul. Esta propriedade situa-se nas coordenadas geográficas 22° 32' 09'' na latitude Sul e 55° 43' 33'' na longitude Oeste. As amostras coletadas em fevereiro de 2012 totalizaram em 160 pontos regularmente espaçados em 50 metros. Na constituição destas amostras, utilizou-se um receptor GPS da marca Garmin, modelo GPSMAP 62.

Inicialmente as amostras georreferenciadas foram submetidas à transformação de coordenadas geográficas WGS84 em coordenadas UTM zona 21 Sul. Posteriormente seguiu-se para a análise descritiva das amostras, com foco na ocorrência de assimetria nos dados e detecção de *outliers*, uma vez que estes modificam o formato do semivariograma (OLIVER; WEBSTER, 2014).

Na detecção de *outliers*, inicialmente, utilizou-se o critério apresentado em Córdoba et al. (2016). Este critério consistiu em quantificar e identificar como

outliers, valores não pertencentes ao intervalo compreendido entre a média (\bar{x}) \pm três vezes o valor do desvio padrão (sd).

Identificado à presença de *outliers*, os dados foram submetidos em duas análises: robusta à presença de *outliers* (ARob.) e não robusta à presença de *outliers* (ANRob). Na primeira análise, considerou-se a utilização da geoestatística robusta, por meio do pacote *georob* (PAPRITZ, 2017) no Software R (TEAM, 2016). Esta metodologia, diferente das demais, baseia-se em fundamentos teóricos de abordagem robusta (KÜNCH et al. 2013; KÜNCH et al. 2011) e na geoestatística baseada em modelos (DIGGLE; RIBEIRO, 2007).

Dessa forma, as observações da CEa $y^T = (y(s_1), y(s_2), \dots, y(s_n))$ amostradas nas posições s_i foram utilizadas para obtenção dos semivariogramas experimentais para os seguintes estimadores de semivariâncias: o estimador clássico de Matheron que utiliza o método dos momentos (ISAAKS and SRIVASTAVA, 1989), o estimador robusto de Cressie e Hawkins (CRESSIE; HAWKINS, 1980), que contrário ao estimador clássico de Matheron, não é sensível a *outliers*; o estimador robusto de Genton (GENTON, 1998) e o estimador robusto MAD (DOWD, 1984). Objetivou-se com a utilização destes estimadores de semivariância, verificar qual deles modela satisfatoriamente o padrão de dependência espacial do processo estacionário com menor sensibilidade aos *outliers*.

O estimador de semivariância identificado forneceu as estimativas iniciais dos parâmetros do semivariograma. Estas estimativas iniciais foram usadas no modelo proposto por Künch et al. (2011), $Y(s) = x(s)^T \beta + Z(s) + \varepsilon(s)$ para predizer $Z(s)$ nas posições s_i não observadas. Neste modelo, tem-se $x(s)^T \beta$ definido como tendência, $Z(s)$ processo gaussiano estacionário de média zero e covariância R e

$\varepsilon(s)$ os erros independentes com parâmetro de escala quadrado σ^2 (efeito pepita) (KÜNCH et al. 2011). Sendo a matriz de covariância R em função da distância h e $\theta^T = (\sigma_0^2, \alpha)$, de modo que σ_0^2 é definido como patamar e α como alcance (KÜNCH et al. 2011).

O diferencial neste método é a utilização da máxima verossimilhança restrita gaussiana robusta (RREML) proposto por (KÜNCH et al. 2011). Este método maximiza a log verossimilhança restrita gaussiana, em adicional aos parâmetros apresentados em Diggle e Ribeiro (2007) $(\sigma^2, \sigma_0^2, \alpha)$, a variável latente z .

Com a entrada dos valores iniciais do semivariograma experimental neste modelo, os modelos de semivariogramas que apresentaram convergência nesta metodologia robusta, foram então avaliados. Dentre os modelos avaliados, selecionou-se o que apresentou menor valor de AIC.

Com isso, o modelo com menor valor de AIC forneceu as estimativas dos parâmetros do semivariograma, que foram utilizados no plug-in de krigagem de deriva externa (KÜNCH et al. 2011):

$$\hat{Z}(s_0) = x(s_0)^T \hat{\beta} + \gamma_{\hat{\theta}}^T(s_0) \Gamma_{\hat{\theta}}^{-1} \hat{Z}$$

Em que: $\Gamma_{\hat{\theta}}$ é a estimativa da matriz de covariância de Z e $\gamma_{\hat{\theta}}$ a estimativa do vetor de covariâncias entre Z e $Z(s_0)$. Obtida a krigagem, estatísticas descritivas da validação da krigagem foram obtidas: erro médio de predição (ME) e erro quadrático médio ponderado (MSDR) (OLIVER; WEBSTER, 2014).

Posterior à análise supracitada, seguiu-se para a ANRob, nesta etapa, utilizou-se o estimador de semivariâncias de Matheron na obtenção do semivariograma, e subsequentemente avaliou-se o ajuste de modelos por máxima verossimilhança.

A utilização do método de ajuste de modelos via máxima verossimilhança (MMV) foi empregada na teoria de geoestatística baseada em modelos (GBM) (DIGGLE; RIBEIRO, 2007). Nesta abordagem, todos os pontos amostrados são utilizados na estimativa dos parâmetros do modelo do variograma pelo método do plug-in. E após obtenção dos parâmetros, os mesmos foram utilizados no preditor linear de krigagem ordinária (DIGGLE; RIBEIRO, 2007).

Os mapas obtidos na ARob e ANRob após a modelagem da estrutura espacial da CEa foram submetidos ao algoritmo *fuzzy k-means* com uso do pacote “e1071” (DIMITRIADOU et al., 2015). Este algoritmo classificou os valores interpolados em classes, de modo que, minimizou a soma dos quadrados das distâncias no domínio do cluster definido (SCUDIÉRO et al., 2013). Esta metodologia define um elemento difuso, em que um ponto observado possa pertencer a diferentes classes, dependendo de um expoente de ponderação pré-definido (SCUDIÉRO et al., 2013).

Nesta fase da pesquisa utilizou-se o expoente de ponderação pré-definido com valor 1,35 convencionalmente deliberado por Odeh, Chittleborough e McBratney (1992). Este valor foi então utilizado na constituição dos clusters. Para tanto, foram obtidos mapas com 2 a 5 clusters em ambas as análises. E em seguida o número ótimo de clusters foi definido com uso do Índice de Performance Fuzzy (FPI – “Fuzziness Performance Index”), Índice de Entropia da Partição Modificada (MPE – “Modified Partition Entropy”), Fukuyama-Sugeno (FS) e xie beni (Xb) (RAMOS et al., (2017), CÓRDOBA et al., (2016), YAO et al., (2014)). Os menores valores obtidos para estes índices foram utilizados na definição do número de classes. Os valores mínimos apresentados para estes índices apresentam melhor divisão entre cluster e similaridades dos valores internos aos clusters definidos (RAMOS et al., (2017); SONG et al., (2009)).

Identificado o número de cluster, o respectivo mapa obtido em cada análise passou por um filtro. Com o uso deste filtro espacial, pretendeu-se promover a continuidade das zonas, reduzindo a fragmentação das classes (LARK et al., 2008), conforme o protocolo apresentado em Córdoba et al. (2016). Nesta fase, utilizou-se o filtro mediano espacial com dimensões 9x9 pixels. Assim, um pixel é então definido como a mediana de seus vizinhos (GONZALES; WOODS, 2008).

Após a suavização das classes com a aplicação do filtro mediano, procedeu-se para a validação das ZM. Nesta etapa, objetivou-se verificar diferenças significativas entre as ZM estabelecidas dentro de cada análise. Com este propósito, procurou-se verificar diferenças significativas entre médias das ZM em cada análise. Para tanto, selecionou-se aleatoriamente amostras de cada ZM e Modelos Lineares Mistos (MLM) foram então definidos, considerando as ZM como efeito fixo e os erros aleatórios espacialmente correlacionados.

Esta mesma configuração de análise apresentada em Córdoba et al. (2016) foi definida com os seguintes MLM: modelo exponencial e esférico, com e sem efeito pepita, e o modelo com erros independentes. Estes mesmos modelos também foram aplicados na validação das ZM estabelecidas nesta pesquisa.

Cada um dos modelos estabelecidos foi avaliado quanto ao Critério de Informação de Akaike (AIC). De modo que o menor valor de AIC foi utilizado para escolha do modelo. Escolhido o modelo, as médias das ZM formadas foram apresentadas com seus respectivos intervalos de confiança. Neste processo de validação das ZM foram utilizados respectivamente os pacotes “nlme” e “lsmeans” para definição dos MLM e teste de médias.

Após a verificação das diferenças entre ZM dentro de cada análise (ARob e ANRob), os respectivos mapas obtidos foram comparados quanto ao nível de concordância das classes pelo índice Kappa. E assim, permitir verificar possíveis modificações nos mapas de ZM quanto à presença de *outliers*.

Resultados e Discussão

Com uso da função *spTransform* do pacote “rgdal” (BIVAND et al., 2016) as coordenadas geográficas da CEa foram transformadas de WGS84 para coordenadas UTM zona 21 Sul. E posterior à transformação, estatísticas descritivas da CEa foram obtidas (Tabela 5).

Tabela 5 – Estatísticas descritivas da condutividade elétrica aparente (CEa) amostrada na profundidade 0 a 20cm.

Var ¹	N ²	Mín ³	Média	Med ⁴	Max ⁵	Sd ⁶	CV(%) ⁷	Cs ⁸	Ck ⁹
CEa ¹⁰	160	2,74	6,19	5,94	19,31	2,13	34,44	2,24	9,29

¹Variável; ²Número de amostras; ³Mínimo; ⁴Mediana; ⁵Máximo; ⁶Desvio padrão; ⁷Coeficiente de variação; ⁸Coeficiente de assimetria; ⁹Curtose; ¹⁰Condutividade elétrica aparente do solo no ano de 2012 (mS m⁻¹).

As 160 observações da CEa apresentaram amplitude total de 16,57 mS m⁻¹ com uma variação percentual de 34% em torno do valor médio 6,19 mS m⁻¹ e ocorrência de assimetria positiva nos dados (Tabela 5).

Peralta et al. (2013) encontraram valores semelhantes para a CEa, quanto ao CV, variando de 17,61% a 44,49%. Já a CEa média apresentou valores maiores

variando de 12,79 a 27,42. Estas diferenças observadas são explicadas pela profundidade da coleta de amostras da CEa (0 a 90cm) em Peralta et al. (2013). Além da diferença de profundidade na coleta das amostras, alto teor de areia no solo e baixo teor de argila, proporcionam maior acúmulo de água e conseqüentemente valores mais baixos para a CEa (PERALTA et al. 2013).

Scudiero et al. (2013) encontraram valor próximo para a média da CEa (1,14 dS m⁻¹), assim como também observaram assimetria positiva (1,05). Além de Scudiero et al. (2013), diversos estudos também apresentaram assimetria na distribuição da CEa (YAO et al., 2014; BOTTEGA et al., 2017; SHAHBAZI; ALIASGHARZAD; EBRAHIMZAD, 2014; AGGELOPOULOU et al. 2013).

A distribuição assimétrica positiva (Tabela 5) apresentou valores atípicos em relação ao conjunto geral dos dados (Figura 5).

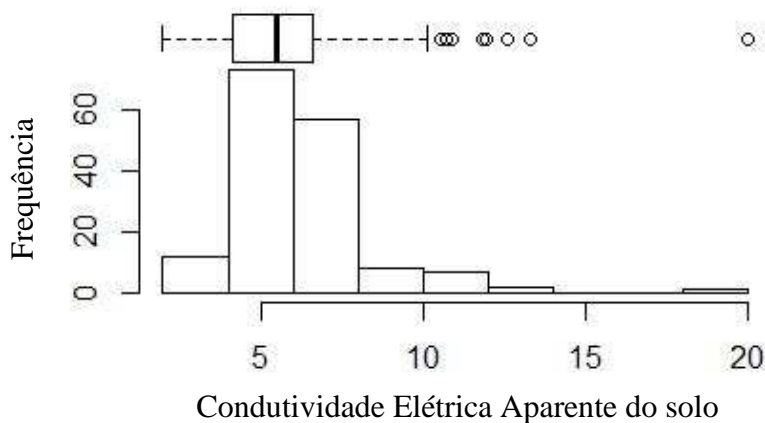


Figura 5 – Histograma e box-plot da distribuição da condutividade elétrica aparente (CEa).

Estes valores atípicos foram então avaliados quanto à possibilidade de possíveis *outliers*. Os possíveis *outliers* podem ser identificados ao compararem com os demais valores em questão, como por exemplo, por meio do box plot (FU et al. 2015; OLIVER; WEBSTER, 2014). Com este propósito, verificou-se a existência de três valores localizados fora do intervalo $[\bar{x} \pm sd: 6,19 \pm 2,13]$ no box plot (Figura 5) com suas respectivas posições (Figura 6).

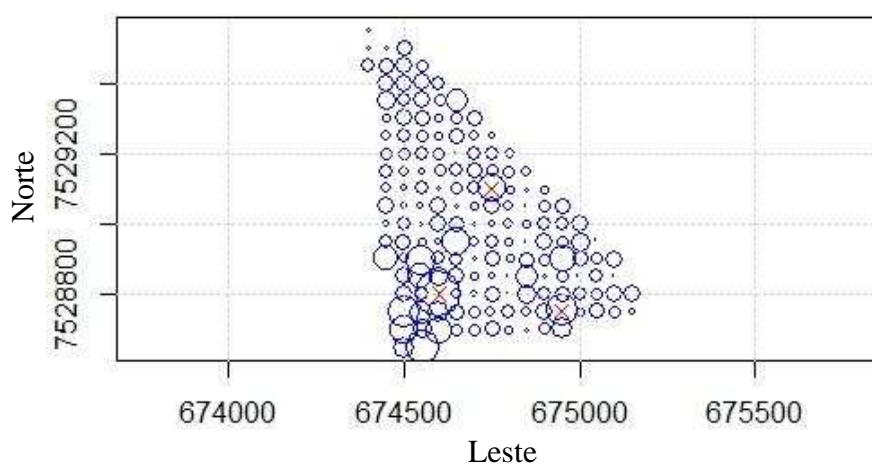


Figura 6 – Malha amostral contendo as posições (marcadas com X) dos respectivos *outliers* identificados para a condutividade elétrica aparente do solo, cujas dimensões das circunferências são proporcionais aos respectivos valores.

Devido à camada utilizada na amostragem da CEa e particulares propriedades físico químicas do solo, Scudiero et al. (2013) e Córdoba et al. (2016) encontraram limiares distintos quanto à detecção de *outliers*, $[\bar{x} \pm sd: 1,14 \pm 0,72]$ e $[\bar{x} \pm sd: 3,8 \pm 7]$ respectivamente.

Identificada a ocorrência de *outliers* na amostra, avaliou-se a influência do mesmo, quanto à forma do semivariograma. Assim, os estimadores de semivariância

de Matheron, Genton, Cressies e Hawkins e MAD Dowd foram comparados segundo o comportamento típico esperado de um semivariograma experimental, considerando os dados da CEa. Estes estimadores de semivariâncias, exceto Matheron, devem ser utilizados em situações relacionadas à pertinência de *outliers* à população (OLIVER; WEBSTER, 2014).

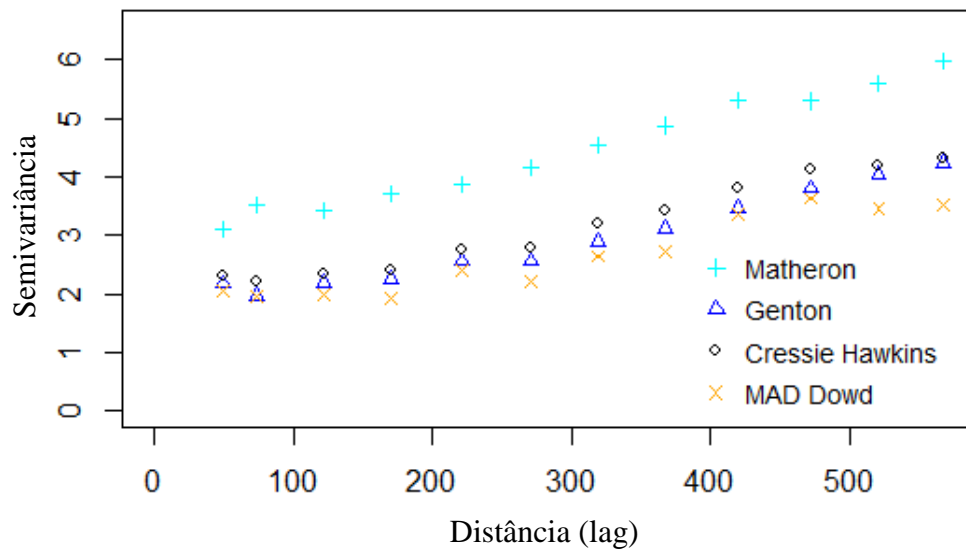


Figura 7 - Semivariograma experimental para o estimador de semivariância de Matheron, Genton, Cressie Hawkins e MAD Dowd aos dados amostrais da CEa.

Aparentemente os *outliers* afetaram o formato do semivariograma obtido pelo estimador de semivariâncias de Matheron, não indicando visivelmente, possível estabilidade para o patamar (Figura 7). Em contrapartida, os demais estimadores de semivariância (Genton, Cressie e Hawkins e Dowd), robustos a presença de *outliers*, apresentam comportamentos semelhantes na forma e estabilidade do patamar (Figura 7). Dentre os estimadores robustos de semivariâncias, utilizou-se o estimador Cressie

Hawkins, visto que, apresentou menor variância de krigagem, erro médio de predição e erro quadrático médio.

Dessa forma, seguiu-se com utilização das duas análises (ARob e ANRob). Na ANRob, fundamentado no semivariograma de Cressie e Hawkins, ajustou-se o modelo cúbico por apresentar menor AIC (577), ao invés dos demais modelos avaliados: gaussiano, gneiting e penta. O ajuste do modelo cúbico (Figura 8) apresentou efeito pepita 1,96; patamar 3,9 e alcance 630.

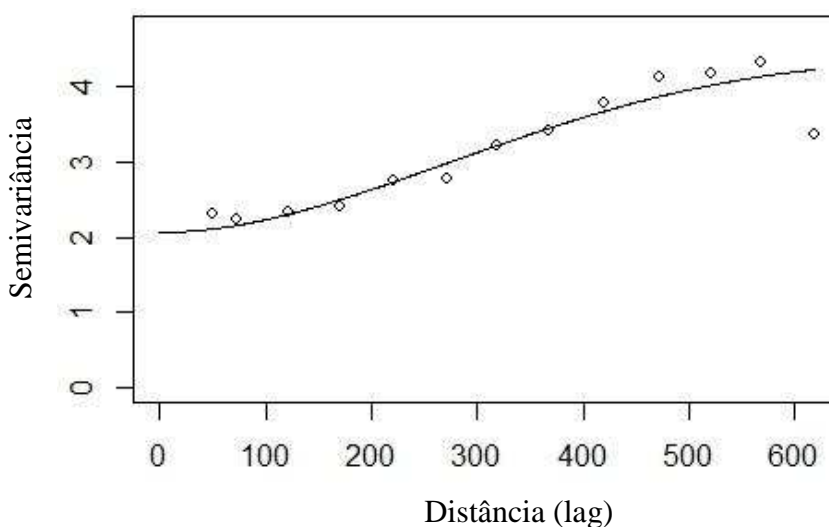


Figura 8 – Modelo cúbico ajustado às semivariâncias obtidas pelo estimador de Cressie Hawkins para a variável condutividade elétrica aparente (CEa).

Seguindo com a ANRob, ajustou-se também o modelo cúbico, do qual apresentou AIC igual a 715, efeito pepita 3,15; patamar 5,54 e alcance prático 648.

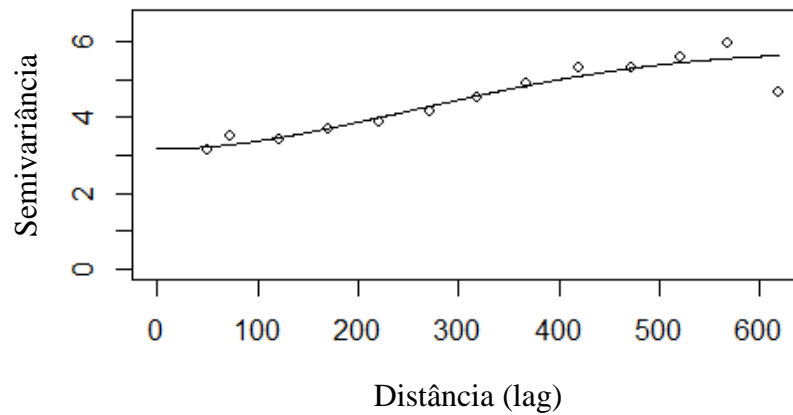


Figura 9 – Modelo cúbico ajustado às semivariâncias obtidas pelo estimador de Matheron para a variável condutividade elétrica aparente (CEa).

Os valores obtidos para as estimativas dos parâmetros do semivariograma (Figura 8) na ARob foram introduzidos como informações de entrada no plug-in de krigagem de deriva externa, com uso do pacote “*georob*” (PAPRITZ, 2017). Com a aplicação desta função foi possível obter o mapa de krigagem da CEa (Figura 10), assim como também estatísticas descritivas da validação cruzada da krigagem (ME = 0,1 e MSDR = 1,66).

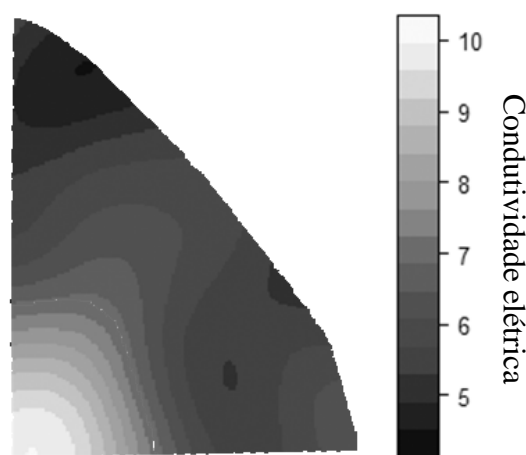


Figura 10 - Mapa interpolado considerando a análise robusta à *outliers*, definindo a estrutura da variabilidade espacial da condutividade elétrica aparente (CEa).

As estimativas obtidas para os parâmetros do modelo cúbico (Figura 9) na ANRob foram introduzidas no preditor linear de krigagem ordinária fundamentado na GBM, por meio do pacote *geoR* (RIBEIRO JR.; DIGGLE, 2016). Que seguidamente, apresentou $ME = 0,5$ e $MSDR = 1,82$.

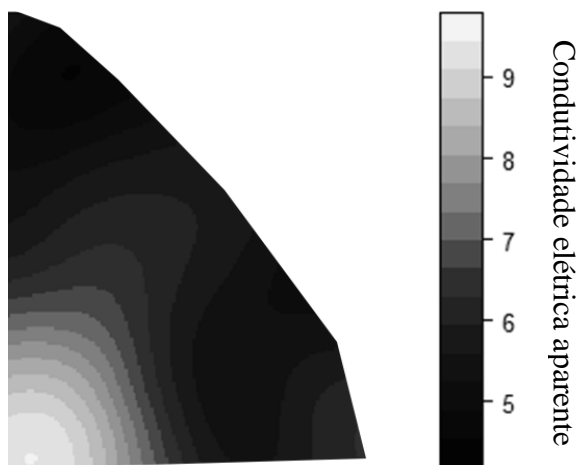


Figura 11 - Mapa interpolado considerando a análise não robusta à *outliers*, definindo a estrutura da variabilidade espacial da condutividade elétrica aparente (CEa).

Os valores sumarizados para a validação cruzada da krigagem legitimam ambas as análises, uma vez que, a média dos erros aproxima de zero e o erro quadrático médio ponderado próximo de um, indicam que a krigagem seja não viesada (OLIVER; WEBSTER, 2014).

Os mapas contendo a estrutura de variabilidade espacial da CEa para ambas as análises foram utilizados como informação complementar no algoritmo *fuzzy k-means*. Com a utilização deste algoritmo, objetiva-se agrupar os valores interpolados em zonas de manejo (RODRIGUES JUNIOR et al. 2011).

O número de zonas de manejo a serem avaliados varia de pesquisa pra pesquisa, no geral se utilizam de 2 a 5 ou 6 zonas de manejo (RAMOS et al. 2017; CÓRDOBA et al. 2016; HAGHVERDI et al. 2015; YAO et al. 2014). Dessa forma, avaliou-se de 2 a 5 zonas de manejo com a função “*cmeans*” e expoente de ponderação 1,35 (ODEH; CHITTLEBOROUGH; MCBRATNEY, 1992) (Figura 12).

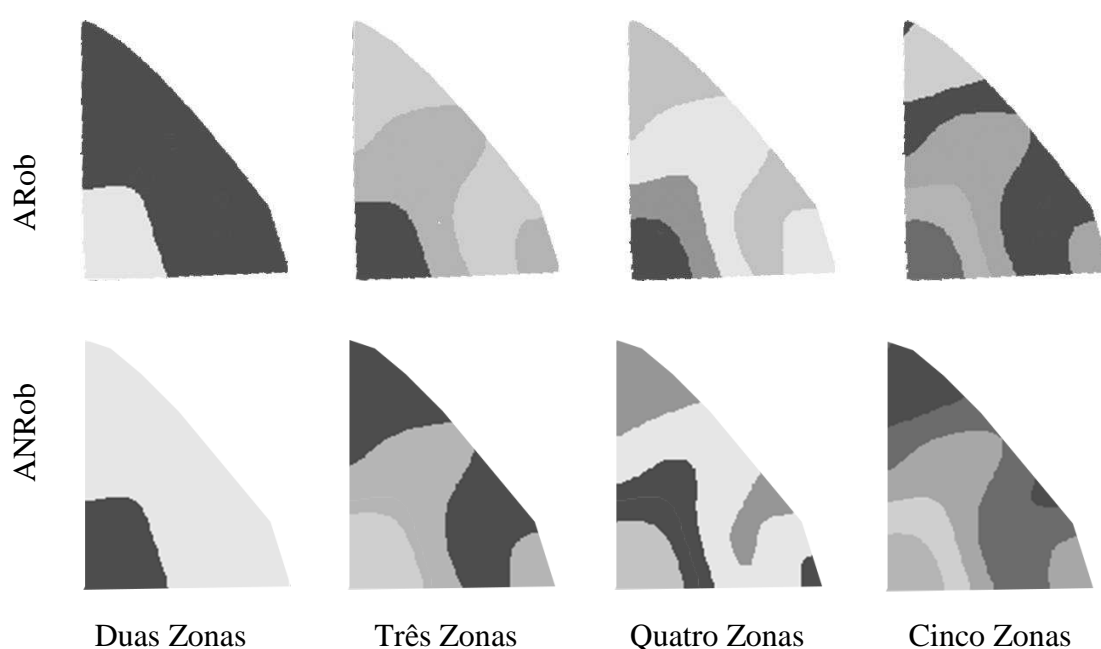


Figura 12 - Mapas com 2, 3, 4 e 5 zonas de manejo dentro da área em estudo, gerados com uso do algoritmo *fuzzy k-means* para a variável CEa, considerando a análise robusta à *outliers* (ARob) e não robusta à *outliers* (ANRob).

O número ótimo de ZM a ser definido é identificado quando os índices FPI, MPE, FS e Xb (RAMOS et al., (2017), CÓRDOBA et al., (2016), YAO et al., (2014)) apresentam valor mínimo (RAMOS et al., (2017); SONG et al., (2009)).

Dessa forma, com uso destes índices, os valores mínimos indicaram a utilização de duas ZM para ambas as análises.

A definição de duas ZM indicadas com uso destes índices é também idêntica ao número de ZM obtidas por Córdoba et al. (2013), Ramos et al. (2017) e YAO et al. (2014). Identificado o número ótimo de ZM, as mesmas devem ser validadas. Com isso, pretende-se verificar o seu potencial quanto a diferenças significativas entre as mesmas. Estas diferenças significativas entre as ZM foram avaliadas em relação a seus valores médios apresentados.

A partir de mil amostras aleatórias de cada ZM estabelecidas, ajustou-se MLM's, e os mesmos foram avaliados quanto ao AIC. Nestes modelos, as ZM foram consideradas de efeito fixo e os erros espacialmente correlacionados com as seguintes estruturas de correlação espacial: exponencial e esférico com e sem efeito pepita e o modelo com erros independentes.

O modelo de erros independentes apresentou menor valor de AIC em ambas as análises (ARob: AIC = 3072; ANRob: AIC = 3098) e, portanto, foi selecionado. Este modelo, por sua vez, foi capaz de detectar diferenças significativas entre as médias nas ZM estabelecidas para análise (Tabela 6).

Tabela 6 - Médias em cada zona de manejo (ZM), para a condutividade elétrica aparente, na análise robusta à outliers (ARob) e não robusta à outliers (ANRob), com respectivos limites superiores e inferiores dos intervalos de confiança.

	ARob		ANRob	
	ZM 1	ZM 2	ZM 1	ZM 2
Média	6,06 ^a	4,88 ^b	6,10 ^a	4,92 ^b
Limite superior	6,13	5,09	6,15	5,07
Limite inferior	5,98	4,67	6,00	4,73

Valores seguidos de mesma letra não diferiram pelo teste t a 1% de probabilidade.

Com a seleção do modelo de erros independentes, foi possível verificar a não sobreposição do intervalo de confiança para as médias obtidas em cada uma das ZM estabelecidas (Tabela 6), ressaltando a heterogeneidade entre elas. E finalmente, após validação verossímil das ZM e identificação do número ótimo das mesmas, pode-se assegurar a utilização do mapa com as duas ZM para ambas as análises (Figura 13).

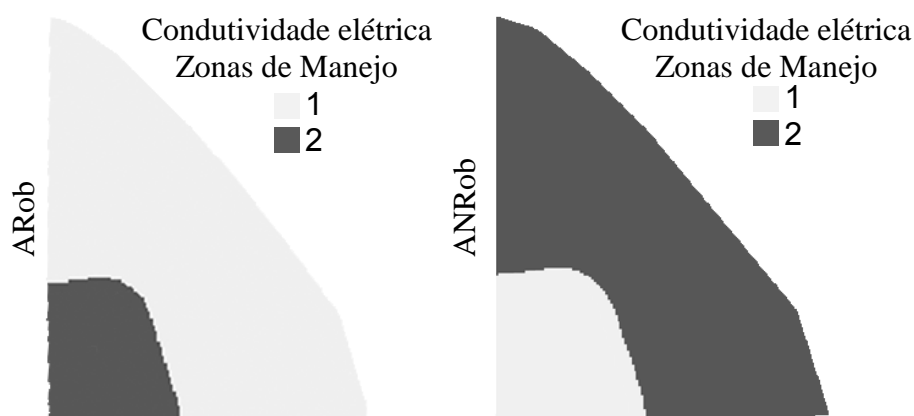


Figura 13 - Mapas com duas zonas de manejo para a condutividade elétrica aparente (CEa) obtidos com a análise robusta à outliers (ARob) e com a análise não robusta à outliers (ANRob).

Os mapas obtidos na Figura 13 apresentam certa continuidade nos limites das ZM obtidas. Isso ocorreu devido à utilização do filtro mediano de suavização (CÓRDOBA et al. 2016), composto por uma janela de 9x9 pixels. Dessa forma, cada pixel constituiu-se da mediana de seus vizinhos internos a janela utilizada.

Os mapas (Figura 13) obtidos foram então avaliados quanto ao nível de concordância entre classes pelo índice Kappa, do qual apresentou semelhança significativa pelo teste Z a 1% de probabilidade.

Conclusões

A análise robusta à presença de *outliers* apresentou-se adequada ao conjunto de dados avaliados, uma vez que expôs de maneira definida o formato do semivariograma. Com isso, a aplicação do método robusto do log da verossimilhança restrita gaussiana implementado no plug-in de krigagem de deriva externa proporcionou ZM bem definidas e de fácil gerenciamento no campo, preconizando sua aplicabilidade.

Entretanto, a utilização da metodologia não robusta à presença de *outliers* apresentou resultados semelhantes às obtidas com a análise robusta, possibilitando o estabelecimento de duas ZM estatisticamente diferentes entre si.

Dessa forma, para o caso em estudo, concluiu-se que, o uso de metodologia robusta a *outliers* resultou em ZM estatisticamente similares às obtidas pela metodologia não robusta.

Referências

- AGGELOPOOULOU, Katerina et al. Delineation of management zones in an apple orchard in Greece using a multivariate approach. *Computers and electronics in agriculture*, v. 90, p. 119-130, 2013.
- BIVAND, R. et al. Bindings for the geospatial Data Abstraction Library. R package version 1.1–10. 2016.
- BOTTEGA, E. L. Utilização de zonas de manejo para a produção de soja no cerrado brasileiro. Tese (Doutorado em Engenharia Agrícola) – Universidade Federal de Viçosa. Viçosa, 2014.
- BOTTEGA, E.L.; QUEIROZ, D.M.; PINTO, F.A.C.; SOUZA, C.M.A.; VALENTE, D.S.M. Precision agriculture applied to soybean: Part I - Delineation of management zones. *Australian Journal of Crop Science*, v. 11, n. 05, p. 573-579, 2017.
- CÓRDOBA, Mariano A. et al. Protocol for multivariate homogeneous zone delineation in precision agriculture. *biosystems engineering*, v. 143, p. 95-107, 2016.
- CÓRDOBA, M. et al. Subfield management class delineation using cluster analysis from spatial principal components of soil variables. *Computers and electronics in agriculture*, v. 97, p. 6-14, 2013.
- CRESSIE, Noel; HAWKINS, Douglas M. Robust estimation of the variogram: I. *Journal of the International Association for Mathematical Geology*, v. 12, n. 2, p. 115-125, 1980.
- DIGGLE, P. J.; RIBEIRO, P. J.; GEOSTATISTICS, Model-based. Springer Series in Statistics. Springer, 2007.

DIMITRIADOU et al. e1071: Misc Functions of the Department of Statistics (e1071), TU Wien. R package version 1.6-7. 2015
<http://CRAN.Rproject.org/package=e1071>.

DOWD, P. A. The variogram and kriging: robust and resistant estimators. In: Geostatistics for natural resources characterization. Springer Netherlands, 1984. p. 91-106.

FU, Weijun et al. Outlier identification of soil phosphorus and its implication for spatial structure modeling. Precision agriculture, v. 17, n. 2, p. 121-135, 2016.

GENTON, Marc G. Highly robust variogram estimation. Mathematical Geology, v. 30, n. 2, p. 213-221, 1998.

GONZALEZ, R. C.; WOODS, R. E. Digital Image Processing, Upper Saddle River, New Jersey 2008.

HAGHVERDI, Amir et al. Perspectives on delineating management zones for variable rate irrigation. Computers and Electronics in Agriculture, v. 117, p. 154-167, 2015.

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

KÜNSCH, Hansruedi et al. Robust estimation of the external drift and the variogram of spatial data. 2013.

KÜNSCH, Hans R. et al. Robust estimation of the external drift and the variogram of spatial data. In: Proceedings of the ISI 58th World Statistics Congress of the International Statistical Institute, Dublin, <http://dx.doi.org/10.3929/ethz-a-009900710>, available at: <http://e-collection.library.ethz.ch/eserv/eth.2011>.

LARK, R. M. Forming spatially coherent regions by classification of multi-variate data: an example from the analysis of maps of crop yield. *International Journal of Geographical Information Science*, v. 12, n. 1, p. 83-98, 1998.

MARONNA, R. A. R. D.; MARTIN, R. Douglas; YOHAI, Victor. *Robust statistics*. John Wiley & Sons, Chichester. ISBN, 2006.

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

ODEH, I. O. A.; CHITTLEBOROUGH, D. J.; MCBRATNEY, A. B. Soil pattern recognition with fuzzy-c-means: application to classification and soil-landform interrelationships. *Soil Science Society of America Journal*, v. 56, n. 2, p. 505-516, 1992.

PAPRITZ, A. (2017). *georob: Robust Geostatistical Analysis of Spatial Data*. R package version 0.3-4. <https://CRAN.R-project.org/package=georob>

PERALTA, Nahuel Raúl et al. Delineation of management zones to improve nitrogen management of wheat. *Computers and Electronics in Agriculture*, v. 110, p. 103-113, 2015.

PERALTA, Nahuel Raúl et al. Delineation of management zones with measurements of soil apparent electrical conductivity in the southeastern pampas. *Canadian Journal of Soil Science*, v. 93, n. 2, p. 205-218, 2013.

PICCINI, Chiara; MARCHETTI, Alessandro; FRANCAVIGLIA, Rosa. Estimation of soil organic matter by geostatistical methods: Use of auxiliary information in agricultural and environmental assessment. *Ecological Indicators*, v. 36, p. 301-314, 2014.

RAMOS, Fabricio Tomaz et al. DEFINING MANAGEMENT ZONES BASED ON SOIL ATTRIBUTES AND SOYBEAN PRODUCTIVITY. *Revista Caatinga*, v. 30, n. 2, p. 427-436, 2017.

RODRIGUES JUNIOR, F. A. et al. Geração de zonas de manejo para cafeicultura empregando-se sensor SPAD e análise foliar. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 2011.

SCUDIERO, Elia et al. Delineation of site-specific management units in a saline region at the Venice Lagoon margin, Italy, using soil reflectance and apparent electrical conductivity. *Computers and electronics in agriculture*, v. 99, p. 54-64, 2013.

SHAHBAZI, F. et al. Geostatistical analysis for predicting soil biological maps under different scenarios of land use. *European journal of soil biology*, v. 55, p. 20-27, 2013.

SONG, Xiaoyu et al. The delineation of agricultural management zones with high resolution remotely sensed data. *Precision agriculture*, v. 10, n. 6, p. 471-487, 2009.

TEAM, R. Core. R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. 2016.

TAYLOR, J. A.; MCBRATNEY, A. B.; WHELAN, B. M. Establishing management classes for broadacre agricultural production. *Agronomy Journal*, v. 99, n. 5, p. 1366-1376, 2007.

YAO, Rong-Jiang et al. Determination of site-specific management zones using soil physico-chemical properties and crop yields in coastal reclaimed farmland. *Geoderma*, v. 232, p. 381-393, 2014.

CONCLUSÕES GERAIS

Diante das indagações anunciadas no Capítulo 1, e os subsequentes objetivos mencionados, o estudo de caso apresentado nesse trabalho permite concluir que:

- ✓ A correção da anisotropia geométrica (Artigo I) não alterou de forma significativa o mapa de ZM, facilitando assim a modelagem da dependência espacial e conseqüentemente a prática da gestão do campo.
- ✓ O uso de uma metodologia robusta à *outliers* (Artigo II) estabeleceu zonas de manejo estatisticamente similares às estabelecidas por uma metodologia não robusta.

APÊNDICE

ARTIGO I

1 - Verificação da continuidade espacial

Após a análise descritiva dos atributos avaliados (Tabela 1), utilizou-se a técnica exploratória proposta por Diggle e Ribeiro Junior (2007) que permite a identificação da continuidade espacial.

Tal metodologia consiste em um teste simples de simulação de Monte Carlo para identificação de correlação espacial sob a hipótese de nulidade de ausência de correlação espacial.

Neste teste permutam-se as coordenadas dos valores amostrados, gerando diferentes conjuntos de dados com respectivos semivariogramas. Os valores máximo e mínimo destes semivariogramas são utilizados na composição de um envelope. A área delimitada com este envelope define uma região de independência espacial. E sob este envelope plota-se o semivariograma dos dados originais. Basta que, pelo menos uma semivariância do semivariograma plotado esteja fora do envelope para

rejeitar a hipótese de nulidade, e assim concluir que os dados apresentam correlação espacial.

Neste contexto, o resultado da aplicação desta metodologia pode ser verificado na Figura 14 onde é apresentado o semivariograma omnidirecional dos atributos produtividade e CEa. Nota-se em cada atributo que pelo menos uma semivariância se encontra fora da área de independência espacial, indicando continuidade nos dados amostrais e existência da dependência espacial nestes atributos.

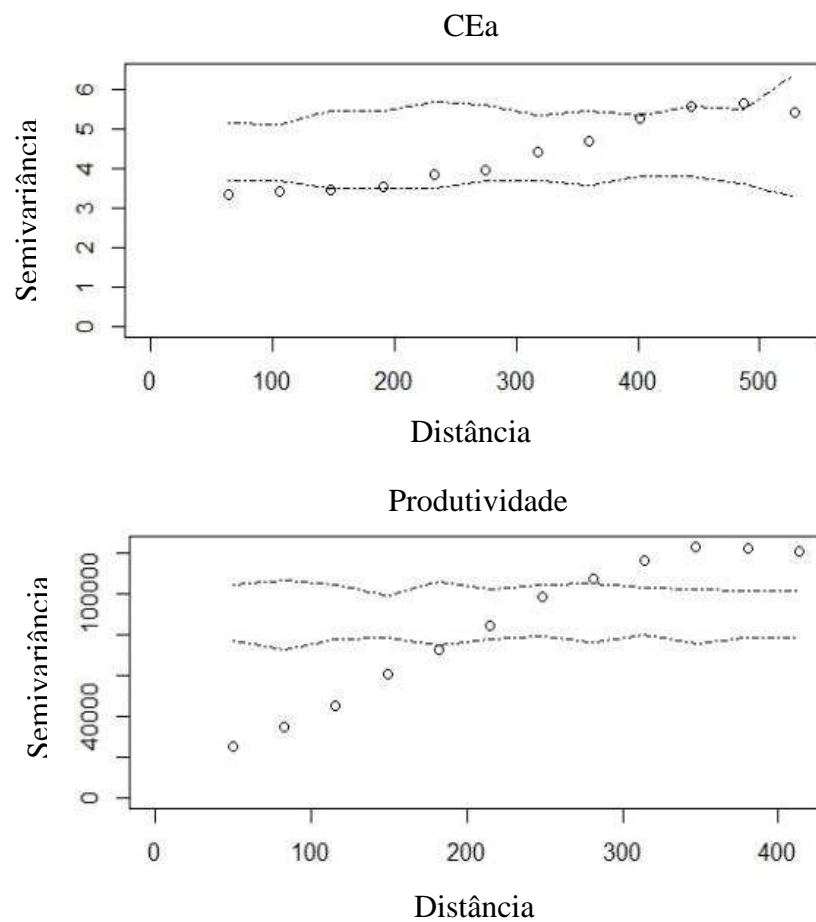


Figura 14 - Envelopes definidos com valores máximos e mínimos de semivariogramas empíricos resultantes da permutação de coordenadas dos valores amostrais de Produtividade e CEa.

2 – Obtenção dos modelos de semivariogramas omnidirecionais sem correção da anisotropia geométrica.

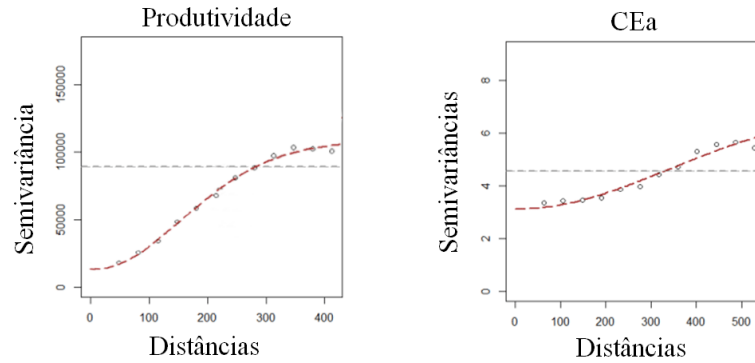


Figura 15 - Ajuste das semivariâncias ao modelo Gaussiano sem correção da anisotropia, para a variável produtividade e CEa.

3 – Obtenções dos modelos de semivariogramas direcionais e omnidirecionais após transformações lineares (correção da anisotropia geométrica).

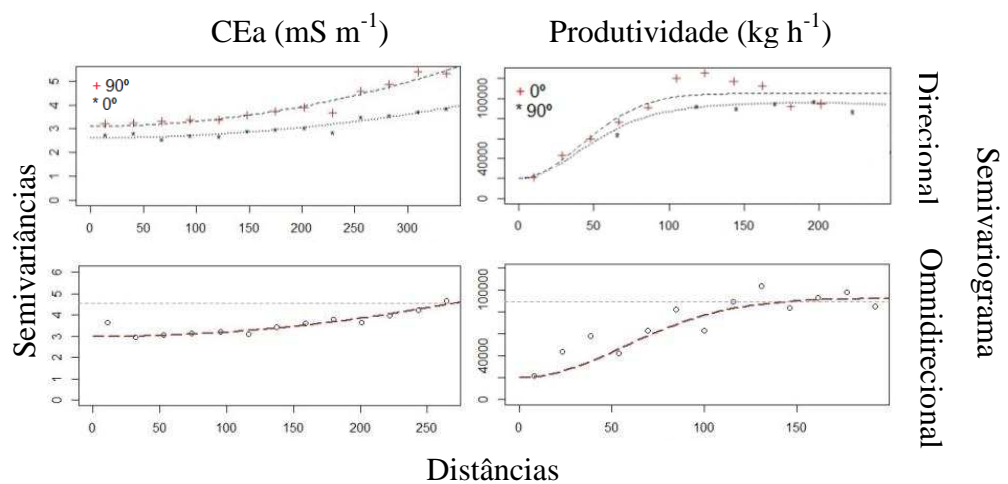


Figura 16 – Semivariogramas direcionais e omnidirecionais do modelo gaussiano, após correção da anisotropia geométrica para a variável produtividade e CEa.

4 – Obtenção dos mapas de zonas de manejo para a produtividade e CEa, com e sem correção da anisotropia geométrica, com respectivos números de classes.

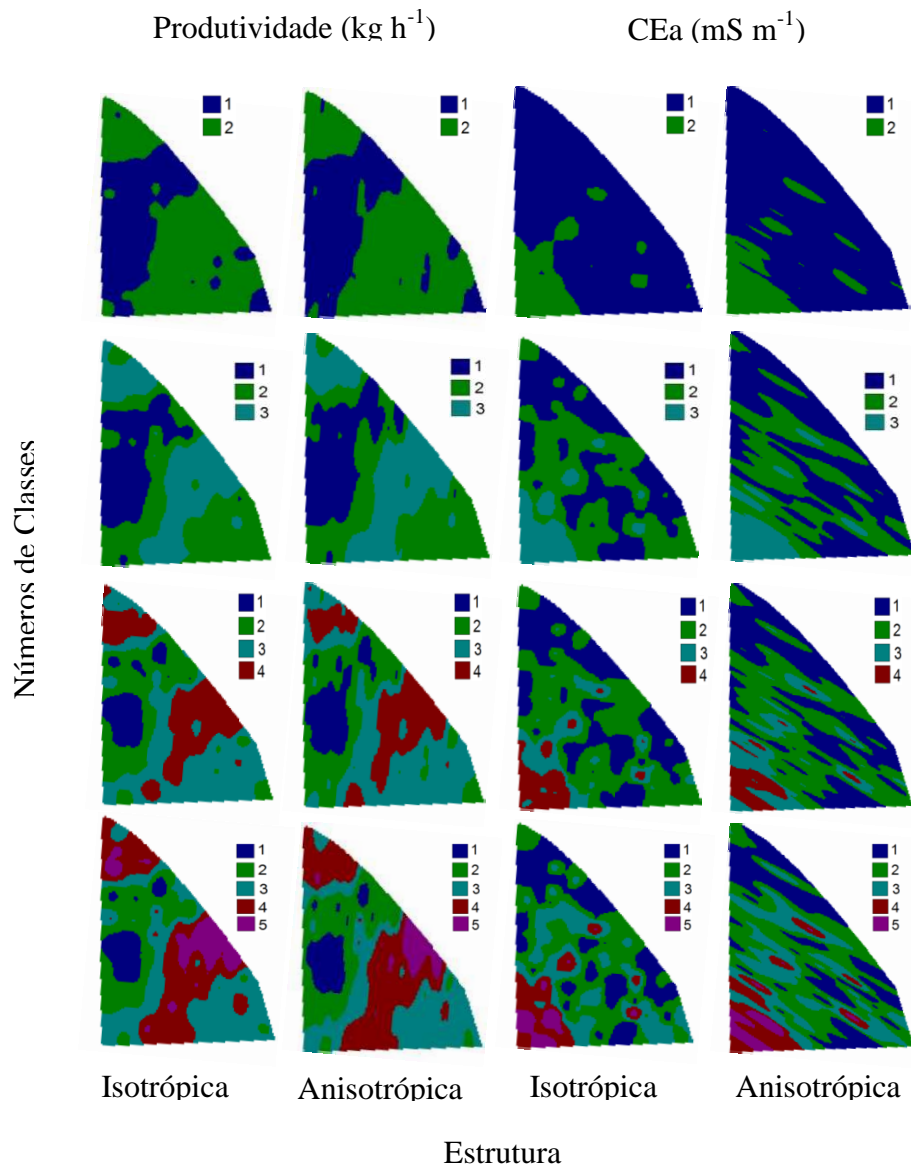


Figura 17 - Mapas de zonas de manejo, para a variável produtividade e CEa, com 2 a 5 classes, com estruturas isotrópica e anisotrópica.