

ALGORITMOS GENÉTICOS À DETERMINAÇÃO DA MELHOR CONFIGURAÇÃO E O MENOR TAMANHO AMOSTRAL NA ANÁLISE DA VARIABILIDADE ESPACIAL DE ATRIBUTOS QUÍMICOS DO SOLO

TAMARA C. MALTAURO¹, LUCIANA P. C. GUEDES², MIGUEL A. URIBE-OPAZO³

¹Matemática, Discente do Programa de Pós - Graduação em Engenharia Agrícola –PGEAGRI, da Universidade Estadual do Oeste do Paraná – UNIOESTE, Cascavel – PR, Fone: (45) 3220-7320, tamara_ma02@hotmail.com.

²Matemática, Prof^a. Dr^a. Adjunta, Pós Graduação em Engenharia Agrícola – PGEAGRI/ UNIOESTE, Cascavel-PR.

³Estatístico, Prof. Dr. Associado, Pós Graduação em Engenharia Agrícola – PGEAGRI/ UNIOESTE, Cascavel-PR.

Apresentado no
XLVI Congresso Brasileiro de Engenharia Agrícola - CONBEA 2017
30 de julho a 03 de agosto de 2017 - Maceió - AL, Brasil

RESUMO: O objetivo deste trabalho foi redimensionar uma configuração amostral, em uma área comercial composta por 102 pontos, reduzindo seu tamanho amostral. Foi utilizado o processo de otimização chamado Algoritmo Genético (AG) e as informações dos atributos químicos que apresentam dependência espacial. A otimização foi realizada pela maximização da medida de acurácia Exatidão Global, que compara os valores preditos obtidos pela configurações inicial e otimizada. O AG foi eficiente quanto à redução do tamanho amostral (53,92% á 68,63% da malha original). No entanto, todos os atributos químicos com tamanho amostral reduzido apresentaram diferenças no modelo geoestatístico estimado e na predição espacial, quando comparados com o modelo obtido usando a configuração original.

PALAVRAS-CHAVE: Agricultura de precisão, geoestatística, otimização.

GENETIC ALGORITHMS FOR THE DETERMINATION OF THE SAMPLE DESIGN SIZE AND SHAPE IN THE ANALYSIS OF THE SPATIAL VARIABILITY OF CHEMICAL PROPERTIES OF THE SOIL

ABSTRACT: The objective of this work was to resize a sample configuration, in a comercial área composed by 102 points, reducing its sample size, using the optimization process Genetic Algorithm (GA) and the information of the chemical attributes with spatial dependence. The GA was made by the maximization of the Overall accuracy measure, which compares the predicted values obtained by the initial and optimized configurations. GA was efficient in reducing sample size (53.92% to 68.63% of the original design). However, all the chemical attributes with reduced sample size presented differences in the estimated geostatistical model and the spatial prediction, when compared with the original design.

KEYWORDS: Geostatistic, optimization, precision agriculture.

INTRODUÇÃO: O estudo da variabilidade espacial de variáveis georreferenciadas contínuas é feito por meio das técnicas de Geoestatística (CÂMARA e MEDEIROS, 1998). Para isso, é necessário determinar um esquema de amostragem com o menor tamanho amostral possível,

que minimize os custos operacionais e maximize a qualidade dos resultados da análise. A escolha de uma configuração eficiente para a predição espacial, consiste num problema de otimização, que a partir de uma malha amostral inicial, escolhe-se uma configuração amostral que minimize as perdas quanto à acurácia dos resultados da predição espacial (GUEDES *et al.*, 2016). Um desses processos de otimização é chamado de Algoritmo Genético (AG), que é uma técnica computacional inspirada na evolução natural (LINDEN, 2012). O objetivo deste trabalho foi determinar uma redução do tamanho amostral, bem como escolher a melhor configuração dos pontos amostrais, a partir de uma malha amostral previamente pesquisada em uma área comercial de cultivo de soja, onde se tem uma aplicação localizada de insumos, utilizando o AG e as informações de atributos químicos do solo que apresentam dependência espacial.

MATERIAL E MÉTODOS: A área comercial em estudo possui 167,35 ha, é composta por 102 pontos e está localizada na Fazenda Agassiz em Cascavel, PR. Estudou-se os atributos químicos do solo teor de Cálcio (Ca) [cmol dm^{-3}], teor de Carbono (C) [g dm^{-3}], teor de Cobre (Cu) [cmol dm^{-3}], $\text{H}+\text{Al}^3$ [cmol dm^{-3}], teor de Manganês (Mn) [cmol dm^{-3}] e pH, do ano safra 2010/2011. Para cada atributo, realizaram-se as análises exploratória e de geoestatística, na qual estimaram-se modelos à função semivariância pelo método de máxima verossimilhança, escolheu-se o melhor modelo pelo método da validação cruzada, e pela krigagem confeccionaram-se os mapas temáticos (CRESSIE, 2015). Em seguida, para cada variável aplicou-se o AG. O tamanho inicial da amostra reduzida do algoritmo foi igual a 20 pontos. Nesta configuração aplicou-se o AG, com 2000 cruzamentos e um ponto de corte (LINDEN, 2012). O AG foi realizado sucessivamente, aumentando-se o tamanho amostral da configuração amostral otimizada, em 5% do total da amostral inicial, até que o algoritmo atingiu o valor mínimo da Exatidão global igual a 0,85. Ao fim do processo de otimização, se obteve uma configuração amostral de tamanho reduzido, na qual, realizaram-se novamente as análises exploratória e geoestatística. As rotinas computacionais do AG, as análises estatísticas e geoestatísticas foram realizadas no software R (R DEVELOPMENT CORE TEAM, 2006) utilizando o pacote geoR (RIBEIRO JR. e DIGGLE, 2001).

RESULTADOS E DISCUSSÕES: Pelos valores obtidos pelo desvio padrão e o coeficiente de variação da configuração original, observou-se que apenas o atributo teor de Mn apresentou alta dispersão dos dados (Tabela 1). Na estimação do modelo geoestatístico, todos os atributos químicos apresentaram moderada dependência espacial ($25\% < EPR \leq 75\%$, CAMBARDELLA *et al.*, 1994). E observou-se um raio de dependência espacial variando de 223,25m (para Ca) a 875,47m (para pH) (Tabela 1). Por meio do mapa temático dos atributos (Figuras 1-a á 1-f) observou-se que os maiores valores de Ca e Mn concentraram-se na região norte da área, Cu possuiu concentração de seus maiores valores na região sul, e C, $\text{H}+\text{Al}^3$ e pH apresentaram concentrações mais espalhadas pela área. Pela utilização do AG, as configurações amostrais foram reduzidas para todos os atributos pesquisados, e tiveram um tamanho amostral entre 55 e 70 pontos, que corresponde respectivamente em 53,92% á 68,63% da malha inicial. Além disso, para todos os atributos químicos, não se identificou um padrão quanto à disposição dos pontos escolhidos, ocorrendo um espalhamento dos pontos em toda a área (Figuras 1-g á 1-l). Pelos valores obtidos com a configuração amostral de tamanho reduzido observou-se pelo desvio padrão e o coeficiente de variação que apenas a variável Mn ainda apresenta alta dispersão dos dados (Tabela 1). Quanto à análise de dependência espacial, observa-se uma escolha diferente do melhor modelo na maioria das variáveis, quando foi utilizado a configuração inicial ou a reduzida. Além disso, tem-se que os atributos

químicos C, Cu, H+Al³ e Mn apresentaram moderada dependência espacial tanto para as variáveis originais quanto para as amostras de tamanho reduzido (Tabela 1). No entanto, as variáveis Ca e pH apresentaram resultados distintos, quando analisados com a configuração reduzida, apresentando fraca ($EPR > 75\%$; CAMBARDELLA *et al.*, 1994) e forte ($EPR \leq 25\%$; CAMBARDELLA *et al.*, 1994) dependência espacial respectivamente (Tabela 1). Percebe-se comparando os alcances estimados dos atributos originais com as amostras de tamanho reduzido, que o alcance para os atributos Ca, Cu, H+Al³ e pH diminuíram, com destaque na redução do alcance do atributo Ca, enquanto os alcances dos atributos C e Mn aumentaram (Tabela 1). Ao construir os mapas usando na krigagem ordinária a configuração otimizada reduzida (Figuras 1-m á 1-r), percebe-se que todos os atributos apresentaram diferenças visuais, quando comparados com os mesmos mapas construídos com a configuração inicial (Figuras 1-a á 1-f), fato também evidenciado pelos valores de exatidão global (EG) abaixo de 0,85 que indica que os mapas não são similares (ANDERSON *et al.*, 1976), valores variando de 0,49 à 0,79. Isto pode ter ocorrido devido as seguintes diferenças que ocorreram entre as análises geoestatísticas realizadas com a configuração amostral inicial e a otimizada: escolha do modelo, valores estimados dos parâmetros (efeito pepita, patamar e alcance), e diferenças em relação a distância máxima das amostras, como também pelo fato dos mapas terem sido construídos com um número de classes diferentes a utilizado no processo de otimização.

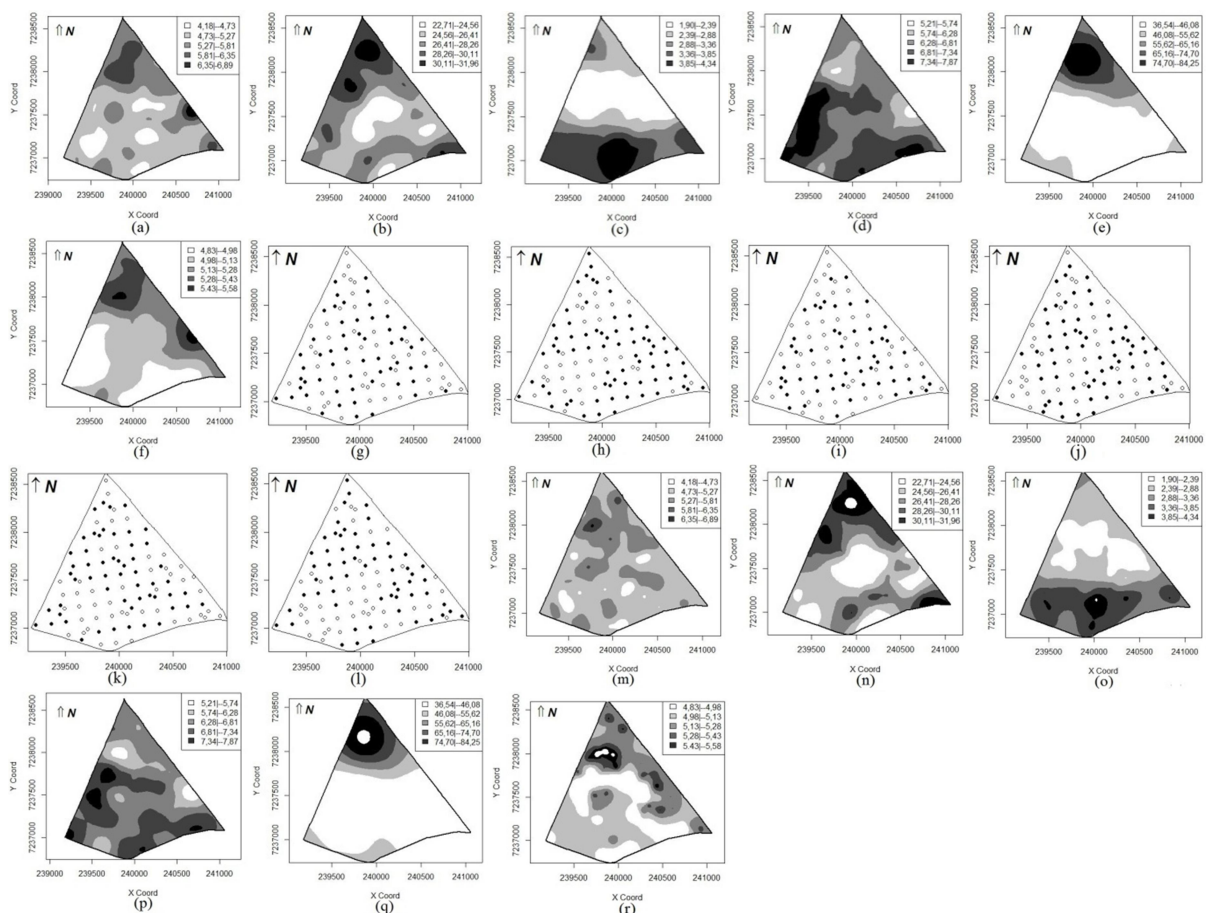


Figura 1 Mapa temático dos atributos com configuração original ((a) á (f)); disposição dos pontos da configuração amostral reduzida ((g) á (l)) em que: (● pontos selecionados; ○ pontos não selecionados); e mapa temático dos atributos com configuração amostral reduzida ((m) á (r)), para respectivamente os atributos: Ca, C, Cu, H+Al³, Mn e pH.

Tabela 1 Estatística descritiva e parâmetros dos modelos da configuração amostral original e da configuração amostral de tamanho reduzido dos atributos Ca, C, Cu, H+Al³, Mn e pH.

Atributo		Média	C.V (%)	Modelo	$\hat{\phi}_1$	$\hat{\phi}_1 + \hat{\phi}_2$	\hat{a}	EPR (%)
Ca	Configuração	5,20	26,41	Gaussiano	1,14	1,56	223,25	73,08
C	amostral	26,93	12,06	Matérn k=2	4,50	9,08	383,47	49,53
Cu	original	2,95	27,86	Esférico	0,13	0,53	849,99	25,14
H+Al ³		6,78	22,50	Exponencial	1,62	2,32	512,97	69,82
Mn		49,32	39,50	Esférico	149,20	255,7	699,99	58,35
pH		5,10	7,58	Exponencial	0,10	0,15	875,47	64,78
Ca	Configuração	5,09	27,68	Esférico	1,29	1,32	44,67	97,42
C	amostral de	26,47	12,75	Esférico	4,71	12,37	620,09	38,09
Cu	tamanho	2,95	29,03	Exponencial	0,11	0,40	503,77	27,99
H+Al ³	reduzido	6,65	24,60	Esférico	1,82	2,67	322,01	68,28
Mn		48,40	42,31	Matérn K=2	140,80	279	821,75	50,47
pH		5,10	7,31	Exponencial	0,01	0,11	214,35	9,30

C.V: coeficiente de variação, $\hat{\phi}_1$: efeito pepita estimado, $\hat{\phi}_2$: contribuição estimada, \hat{a} : alcance prático estimado (m), EPR: efeito pepita relativo ($EPR = \hat{\phi}_1 / (\hat{\phi}_1 + \hat{\phi}_2)$) (%), $\hat{\phi}_1 + \hat{\phi}_2$: patamar estimado.

CONCLUSÕES: Todos os atributos químicos do solo apresentaram dependência espacial. O AG foi eficiente quando a redução da configuração amostral para todos os atributos. Não foi evidenciado uma similaridade entre os mapas temáticos elaborados com a malha original e a otimizada. Alguns atributos com tamanho amostral reduzido apresentaram diferenças quanto à escolha do modelo geoestatístico estimado e quanto às estimativas dos parâmetros deste modelo, quando comparados com os resultados obtidos pela configuração original.

AGRADECIMENTOS: A CAPES, CNPq e Fundação Araucária pelo apoio financeiro.

REFERÊNCIAS: ANDERSON, J.F.; HARDY, E.E.; ROACH, J.T.; WITMER, R. E. A land use and land cover classification system for use with remote sensor data, U.S. Washington, DC: U.S. **Geologic Survey**, 2001. 41 p.

CÂMARA, G.; MEDEIROS, J. S. **GIS para meio ambiente**. In: Sistema de informações geográficas, 1998, Curitiba. Anais. Curitiba: Sages, 1998. p. 140-180.

CAMBARDELLA, C.A.; MOORMAN, T.B.; PARKIN, T.B.; NOVACK, J.M.; KARLEN, D.L.; TURCO, R.F.; KNOPKA, A.E. Field-scale variability of soil properties in Central Iowa Soils. **Soil Science Society America Journal**, Madison, v.58, n.5, p. 1501-1511, 1994.

CRESSIE, N. A. C. **Statistics for Spatial Data** (Rev. Ed.). John Wiley & Sons, 2015. p.928.

GUEDES, L. P. C.; RIBEIRO JR. P.J.; URIBE-OPAZO, M.A.; BASTIANI, F. Soybean yield maps using regular and optimized sample with different configurations by simulated annealing. **Engenharia Agrícola**, Jaboticabal, v.36, n.1, p.114-125, 2016.

LINDEN, R. **Algoritmos Genéticos**. Rio de Janeiro: Ciência Moderna Ltda, 2012. 496 p.

R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Disponível em: <<http://www.R-project.org>> Acesso em: 15 mar. 2017.

RIBEIRO JR., P. J.; DIGGLE, P. J. **geoR: A package for geostatistical analysis**. R-NEWS, vol.1, n. 2, ISSN 1609-3631, 2001.