

Análise dos atributos do solo e da produtividade da cultura de cana-de-açúcar com o uso da geostatística e árvore de decisão

Analyze the soil attributes and sugarcane yield culture with the use of geostatistics and decision trees

Zigomar Menezes de Souza^{I*} Domingos Guilherme Pellegrino Cerri^{II} Marcelo José Colet^{II}
Luiz Henrique Antunes Rodrigues^I Paulo Sérgio Graziano Magalhães^I
Rafael Junqueira Araújo Mandoni^{III}

RESUMO

Um dos desafios da agricultura de precisão é oferecer subsídios para a definição de unidades de manejo para posteriores intervenções. Portanto, o objetivo deste trabalho foi avaliar os atributos químicos do solo e a produtividade da cultura de cana-de-açúcar por meio da geostatística e mineração de dados pela indução da árvore de decisão. A produtividade da cana-de-açúcar foi mapeada em uma área de aproximadamente 23ha, utilizando-se o critério de célula, por meio de um monitor de produtividade que permitiu a elaboração de um mapa digital que representa a superfície de produção para a área em estudo. Para determinar os atributos de um Argissolo Vermelho-Amarelo, foram coletadas as amostras no início da safra 2006/2007, utilizando-se uma grade regular de 50 x 50m, nas profundidades de 0,0-0,2m e 0,2-0,4m. Os dados dos atributos do solo e da produtividade foram analisados por meio da técnica de geostatística e classificados em três níveis de produção para indução de árvore de decisão. A árvore de decisão foi induzida no programa SAS Enterprise Miner, sendo utilizado algoritmo baseado na redução de entropia. As variáveis altitude e potássio apresentaram os maiores valores de correlação com a produtividade de cana-de-açúcar. A indução de árvores de decisão permitiu verificar que a altitude é a variável com maior potencial para interpretar os mapas de produtividade de cana-de-açúcar, auxiliando na agricultura de precisão e mostrando-se uma ferramenta adequada para o estudo de definição de zonas de manejo em área cultivada com essa cultura.

Palavras-chave: agricultura de precisão, variabilidade espacial, mineração de dados, monitor de produtividade.

ABSTRACT

One of the challenges of precision agriculture is to offer subsidies for the definition of management units for

posterior interventions. Therefore, the objective of this work was to evaluate soil chemical attributes and sugarcane yield with the use of geostatistics and data mining by decision tree induction. Sugarcane yield was mapped in a 23ha field, applying the cell criterion, by using a yield monitor that allowed the elaboration of a digital map representing the surface of production of the studied area. To determine the soil attributes, soil samples were collected at the beginning of the harvest in 2006/2007 using a regular grid of 50 x 50m, in the depths of 0.0-0.2m and 0.2-0.4m. Soil attributes and sugarcane yield data were analyzed by using geostatistics techniques and were classified into three yield levels for the elaboration of the decision tree. The decision tree was induced in the software SAS Enterprise Miner, using an algorithm based on entropy reduction. Altitude and potassium presented the highest values of correlation with sugarcane yield. The induction of decision trees showed that the altitude is the variable with the greatest potential to interpret the sugarcane yield maps, then assisting in precision agriculture and, revealing an adjusted tool for the study of management definition zones in area cropped with sugarcane.

Key words: precision agriculture, spatial variability, data mining, yield monitor.

INTRODUÇÃO

A adoção das técnicas relacionadas ao conceito de agricultura de precisão é um desafio recente no Brasil. Questões e dúvidas de toda ordem são levantadas e são semelhantes às apontadas pelos países mais desenvolvidos. Os aspectos que suscitam maiores dúvidas estão relacionados a como interpretar tanta informação e transformá-la em ferramentas que

^IFaculdade de Engenharia Agrícola (FEAGRI), Universidade Estadual de Campinas (UNICAMP), CP 6011, 13083-970, Campinas, SP, Brasil. E-mail: zigomarms@agr.unicamp.br. *Autor para correspondência.

^{II}Programa de Pós-graduação, FEAGRI, UNICAMP, Campinas, SP, Brasil.

^{III}Curso de Engenharia Agrícola, UNICAMP, Campinas, SP, Brasil

auxiliem na tomada de decisões para o correto tratamento da variabilidade espacial inegável de nossas lavouras. Essa variabilidade que se expressa na colheita, em termos de diferenças de produtividade ao longo da lavoura, tem causas das mais variadas e de maneira alguma pode ser atribuída a um fator isoladamente.

Uma alternativa do setor sucro-alcooleiro na redução de custos é a adoção da agricultura de precisão, que promove o conhecimento da variabilidade dos solos. A variabilidade dos solos tem sido abordada pela classificação numérica, por métodos de estatística multivariada, classificação contínua (*fuzzy*), geoestatística, métodos fractais, morfologia matemática e teoria do caos (BURROUGH et al., 1994). Embora esses métodos estatísticos permitam inferir sobre a variabilidade espacial do solo (vertical e horizontal), a dependência espacial entre as amostras somente pode ser modelada por meio da geoestatística (WEBSTER, 2001). Análises estatísticas clássicas que consideram a independência entre as amostras, baseadas na média, vêm sendo substituídas por análises geoestatísticas fundamentadas na teoria das variáveis regionalizadas (ISAAKS & SRIVASTAVA, 1989), por intermédio do semivariograma e da dependência espacial.

O desenvolvimento de sistemas computacionais para auxiliar a tomada de decisão sobre o uso e manejo do solo em áreas agrícolas é importante, visto que possibilita avaliar grande quantidade de informações a respeito do solo e da planta, as quais podem produzir estratégias mais adequadas para o aumento da produtividade e para a proteção ambiental (DE LA ROSA et al., 1993; YANG et al., 2002). Nesse contexto, apesar de ser uma técnica relativamente nova, a utilização da mineração de dados por meio da indução de árvore de decisão é relatada como promissora na análise de dados em estudos de agricultura de precisão.

O planejamento e o gerenciamento de áreas agrícolas envolvem práticas de manejo das culturas que, por sua vez, exigem um conhecimento simultâneo de atributos do solo, do relevo e do rendimento dos talhões. As técnicas da geoestatística e da estatística clássica e a indução por árvore de decisão possibilitam o armazenamento, o tratamento, a sobreposição, a análise e a visualização, em conjunto, dessas informações espaciais, sob a forma de mapas digitais, permitindo tomadas de decisões rápidas, ágeis e com maior eficiência, minimizando custos, otimizando os recursos e as atividades produtivas e acarretando ganho de produtividade e maximização dos lucros.

Nesse sentido, sistemas computacionais têm que ser desenvolvidos para auxiliar na tomada de

decisão em diferentes áreas das ciências agrárias. Portanto, o objetivo deste trabalho foi avaliar os atributos químicos do solo e a produtividade da cultura de cana-de-açúcar por meio da geoestatística e mineração de dados pela indução da árvore de decisão.

MATERIAL E MÉTODOS

O experimento foi conduzido em uma área de cana-de-açúcar pertencente à Usina São Domingos, localizada no Município de Catanduva, São Paulo (SP). As coordenadas geográficas são de 24°20'33" de latitude Sul e 50°17'40" de longitude Oeste. O clima da região é do tipo subtropical CWa e mesotérmico de inverno seco, segundo a classificação de Köppen. A distribuição das chuvas segue o regime típico das zonas tropicais de baixa altitude, verão chuvoso e inverno seco. A área experimental está sob cultivo de cana-de-açúcar há mais de 35 anos. O solo da área foi classificado como um Argissolo Vermelho-Amarelo.

Em uma área de 23ha, os solos foram amostrados nos pontos de cruzamento de uma malha com intervalos regulares de 50m contendo 118 pontos amostrais, na safra de 2006/2007. Em cada ponto dessa malha, foi levantada a sua cota, com o auxílio de um teodolito e DGPS. Para efetuar a amostragem do solo, foi utilizado um trado para coleta do solo. Os atributos químicos foram estimados a partir da coleta de amostras de solos nas profundidades de 0,0-0,2m e 0,2-0,4m. No mapeamento da produtividade da cultura da cana-de-açúcar, utilizou-se o monitor de produtividade embarcado na colhedora, denominado Simprocana. Para correlacionar a produtividade da cana-de-açúcar com os atributos químicos do solo, utilizou-se o coeficiente de correlação linear (nível de 5% de significância).

Na caracterização química do solo, foi determinada potenciometricamente a acidez ativa (pH em CaCl₂). As bases trocáveis cálcio (Ca), magnésio (Mg) e potássio (K) e fósforo disponível (P) foram extraídas com o método da resina trocadora de íons proposto por RAIJ et al. (2001). Com base nos resultados das análises químicas, foi calculada a porcentagem de saturação por bases do solo (V%). O teor de matéria orgânica foi obtido pelo método descrito pela EMBRAPA (1997).

Os atributos químicos e a produtividade da cultura foram avaliados por meio da análise estatística descritiva, calculando-se a média, a mediana, o desvio padrão, o coeficiente de variação, o coeficiente de assimetria, o coeficiente de curtose e os valores máximos e mínimos. A hipótese de normalidade dos dados foi testada pelo teste de Kolmogorov-Smirnov,

por meio do programa computacional SAS (SCHLOTZHAVER & LITTELL, 1997). A dependência espacial foi avaliada por meio de ajuste de semivariogramas pelo programa GS+ (ROBERTSON, 1998). A análise do grau de dependência espacial dos atributos foi realizada segundo CAMBARDELLA et al. (1994), em que são considerados de dependência espacial forte os semivariogramas que têm um efeito pepita menor ou igual a 25% do patamar, moderada entre 25% e 75%, e fraca quando for maior que 75%.

O algoritmo básico de indução de árvores de decisão constrói a árvore de forma recursiva, de cima para baixo (HAN & KAMBER, 2001). A indução da árvore de decisão inicia-se com um conjunto de treinamento, que é dividido de acordo com um teste sobre uma das variáveis independentes, sendo formados subconjuntos mais homogêneos em relação à variável dependente. Esse procedimento é repetido até que sejam obtidos conjuntos de exemplos bem homogêneos, para os quais seja possível atribuir um único valor para a variável dependente. O critério utilizado para escolher a variável independente que divide o conjunto de exemplos em cada repetição é o aspecto principal do processo de indução.

Na análise dos dados para a indução da árvore de decisão, inicialmente a variável produção foi classificada, de forma a converter os valores de produção medidos (variável contínua) em níveis de produção (variável discreta). Foram testadas classificações com três e cinco níveis, inicialmente pela distribuição de frequência, que, na indução da árvore de decisão, não apresentou resultados satisfatórios. Adotou-se então o desvio padrão como critério para classificação, sendo definidos, como nível médio de produção, os dados no intervalo entre a média mais ou menos um desvio padrão. Os dados inferiores a esse intervalo foram definidos como nível baixo, e os superiores foram definidos como nível alto.

A indução da árvore de decisão foi realizada por meio da ferramenta *Tree (Tree node)* do SAS *Enterprise Miner*, sendo utilizado o algoritmo de redução de entropia (QUINLAN, 1993). A árvore de decisão foi escolhida para ser binária, com dois ramos a partir de cada nó interno. Para evitar que o modelo ficasse muito específico para o conjunto de treinamento, o que comprometeria a sua generalização e o desempenho com novos exemplos, foram adotadas duas regras de parada do algoritmo de indução. A primeira regra limitou a profundidade da árvore, permitindo que esta tivesse no máximo seis níveis. A segunda regra limitou a fragmentação do conjunto de treinamento, requerendo um mínimo de 10 exemplos em cada nó para a busca de uma nova divisão e pelo menos cinco exemplos em cada nó folha.

Além das regras de parada, denominadas de pré-poda, foi realizado um procedimento de pós-poda, após a indução da árvore completa. Junto com essa árvore completa, foram avaliadas todas as suas possíveis subárvores e foi escolhida a menor subárvore (menor complexidade).

RESULTADOS E DISCUSSÃO

Os valores da média e mediana de todos os atributos químicos do solo e a produtividade da cultura de cana-de-açúcar, nas profundidades de 0,0-0,2m e 0,2-0,4m, são próximos, caracterizando distribuição simétrica, porém somente a saturação por bases apresentou normalidade pelo teste Kolmogorov-Smirnov (Tabela 1). RACHID JÚNIOR et al. (2006), estudando as variabilidades espacial e temporal de atributos químicos do solo e da produtividade da soja, não encontraram normalidade para os atributos do solo estudados. A normalidade dos dados não é uma exigência da geoestatística, é conveniente apenas que a distribuição não apresente caudas muito alongadas, o que poderia comprometer as estimativas da krigagem, as quais são baseadas nos valores médios (ISAAKS & SRIVASTAVA, 1989).

De acordo com a classificação do coeficiente de variação (CV), proposta por WARRICK & NIELSEN (1980), o pH, nas profundidades de 0,0-0,2m e 0,2-0,4m, apresentou baixo CV (< 12%), os demais atributos químicos do solo estudados tiveram CV médio (12 a 24%) (Tabela 1). A produtividade da cultura de cana-de-açúcar também apresentou CV médio. Segundo VANNI (1998), coeficiente de variação maior que 35% revela que a série é heterogênea, e a média tem pouco significado. Portanto, observa-se, para todos os atributos químicos do solo e para a produtividade da cana, que a série de dados é homogênea.

Os resultados da análise geoestatística mostraram que os atributos químicos do solo e a produtividade da cana-de-açúcar apresentaram dependência espacial nas duas profundidades, e o mesmo foi observado para produtividade da cultura da cana-de-açúcar (Tabela 2). Na seleção dos modelos dos semivariogramas, foram considerados os menores coeficientes da validação cruzada. Todos os dados dos atributos químicos do solo e da produtividade da cultura da cana-de-açúcar ajustaram-se ao modelo esférico, concordando com os resultados de várias pesquisas que indicam o modelo esférico como o de maior ocorrência para os atributos químicos do solo (VIEIRA, 2000; CORÁ et al., 2004; SOUZA et al., 2004; GREGO & VIEIRA, 2005).

Tabela 1 - Estatística descritiva dos atributos pH, cálcio (mmolc dm^{-3}), magnésio (mmolc dm^{-3}), potássio (mmolc dm^{-3}), fósforo (mg dm^{-3}), teor de matéria orgânica (g dm^{-3}) e saturação por bases (V%) de um Argissolo Vermelho-Amarelo nas profundidades de 0,0-0,2m e 0,2-0,4m e da produtividade da cultura de cana-de-açúcar (t ha^{-1}).

Parâmetros	pH	Ca	Mg	K	P	MO	V%	Produtividade
-----Profundidade de 0,0-0,2m-----								
Média	4,96	10,96	5,00	2,12	6,34	15,45	47,39	67,31
Mediana	5,00	11,00	5,00	2,08	6,15	15,50	47,40	67,52
¹ DP	0,38	2,37	1,16	0,72	1,08	4,12	6,76	12,46
Curtose	-1,10	-0,54	3,22	0,61	-0,16	0,60	-0,38	3,21
Assimetria	-0,03	-0,02	1,00	0,75	0,48	0,51	-0,23	1,15
Máximo	5,50	17,00	9,00	4,90	10,00	29,00	60,70	128,16
Mínimo	4,20	5,00	3,00	1,10	5,00	7,00	31,10	40,18
² CV	7,69	21,68	24,6	33,98	17,15	26,64	14,26	18,51
³ d	0,13	0,10	0,26	0,11	0,19	0,08	0,05	0,28
-----Profundidade de 0,2-0,4m-----								
Média	4,80	9,45	4,16	2,37	5,94	15,18	43,35	–
Mediana	4,85	9,14	4,06	2,45	6,68	15,08	43,60	–
DP	0,34	2,46	1,11	0,64	1,00	4,00	6,80	–
Curtose	-0,69	-0,21	1,61	-0,98	0,28	0,57	-0,33	–
Assimetria	0,40	0,51	0,46	-0,12	0,79	0,46	0,21	–
Máximo	5,50	17,00	9,00	3,80	10,00	30,00	60,8	–
Mínimo	4,30	5,00	2,00	1,20	4,00	6,00	29,3	–
CV	7,05	24,02	24,84	27,01	20,32	29,00	15,69	–
d	0,14	0,17	0,22	0,09	0,24	0,07	0,04 ^{ns}	–

¹ DP = desvio padrão; ² CV = coeficiente de variação; ³ d = estatística do teste de Kolmogorov-Smirnov, não significativo a 5% de probabilidade (^{ns}).

O efeito pepita é um parâmetro importante do semivariograma e indica a variabilidade não explicada, considerando a distância de amostragem utilizada. Quanto maior for a diferença do efeito pepita em relação ao patamar do semivariograma, maior a continuidade do fenômeno, menor a variância da estimativa ou maior a confiança que se pode ter na estimativa (VIEIRA, 2000). A análise da relação $C_0/(C_0+C_1)$ dos dados dos atributos químicos do solo e da produtividade da cultura da cana-de-açúcar mostrou grau de dependência espacial forte para todos os atributos químicos e a produtividade da cana, com exceção das variáveis potássio e matéria orgânica nas profundidades estudadas, que apresentaram grau de dependência espacial moderado (Tabela 2). Isso demonstra que os semivariogramas explicam a maior parte da variância dos dados experimentais com confiabilidade na estimativa. Com relação ao alcance da dependência espacial, os atributos químicos do solo apresentaram valores próximos nas profundidades de 0,0-0,2m e 0,2-0,4m, com exceção das variáveis potássio e produtividade da cultura da cana-de-açúcar, que apresentaram valores superiores às demais variáveis estudadas.

A correlação linear entre a produtividade da cana-de-açúcar com os atributos químicos do solo nas profundidades de 0,0-0,2m e 0,2-0,4m mostrou coeficientes de correlação baixos para todos os atributos do solo estudados, com exceção da variável potássio, que apresentou significância (Tabela 3). Resultados semelhantes foram observados por RACHID JÚNIOR et al. (2006), os quais verificaram baixa correlação entre a produtividade da soja com os atributos do solo e alta correlação com o potássio. MEGDA et al. (2008) e ROSA FILHO et al. (2009), estudando a correlação linear e espacial entre a produtividade do feijão e da soja, respectivamente, com atributos do solo, observaram correlação linear baixa, porém significativa.

Baixa correlação da produtividade de culturas com atributos do solo vem sendo observada em diversas pesquisas, o que dificulta a implantação da agricultura de precisão (MATA et al., 1999; YANAI et al., 2001; VIEIRA & MOLIN, 2001). Porém, KITAMURA et al. (2007) afirmam que a baixa correlação, entre dois atributos quaisquer, não invalidam a hipótese de ocorrência de apreciável correlação espacial entre eles. A altitude do terreno revelou uma correlação positiva com a produtividade, indicando que, nas áreas

Tabela 2 - Modelos e parâmetros dos semivariogramas experimentais para a variável pH, cálcio ($\text{mmol}_e \text{dm}^{-3}$), magnésio ($\text{mmol}_e \text{dm}^{-3}$), potássio ($\text{mmol}_e \text{dm}^{-3}$), fósforo (mg dm^{-3}), teor de matéria orgânica (g dm^{-3}) e saturação por bases (V%) de um Argissolo Vermelho-Amarelo e da produtividade da cultura de cana-de-açúcar (t ha^{-1}).

Parâmetros	pH	Ca	Mg	K	P	MO	V%	Produtividade
-----Profundidade de 0,0-0,2m-----								
Modelo	Esférico							
Efeito pepita	0,017	0,34	0,26	0,23	0,22	6,94	7,21	51,12
Patamar	0,08	4,69	1,21	0,59	1,18	14,30	38,92	324,01
Alcance	87	102	83	399	90	170	105	471
¹ GDE	21	7	21	39	19	48	18	16
² R ²	67	73	64	83	68	80	92	86
³ VC	62	71	65	78	64	77	72	76
-----Profundidade de 0,2-0,4m-----								
Modelo	Esférico	-						
Efeito pepita	0,02	0,99	0,18	0,11	0,21	9,08	8,40	-
Patamar	0,09	5,54	0,93	0,41	1,20	18,17	41,19	-
Alcance	96	97	99	268	111	241	89	-
GDE	22	18	19	27	18	50	20	-
R ²	85	60	70	94	80	93	65	-
VC	67	60	65	74	68	78	70	-

¹ GDE = grau de dependência espacial - $[C_0/(C_0+C_1)] \times 100$; ² R² = coeficiente de determinação; ³ VC = coeficiente de determinação do teste de validação cruzada.

menos elevadas do talhão, a produtividade é menor. Resultados semelhantes foram observados por MONTEZANO et al. (2008), os quais avaliaram a variabilidade de nutrientes em plantas de milho cultivado em talhão manejado homogêneo e verificaram que a altitude proporcionou correlação

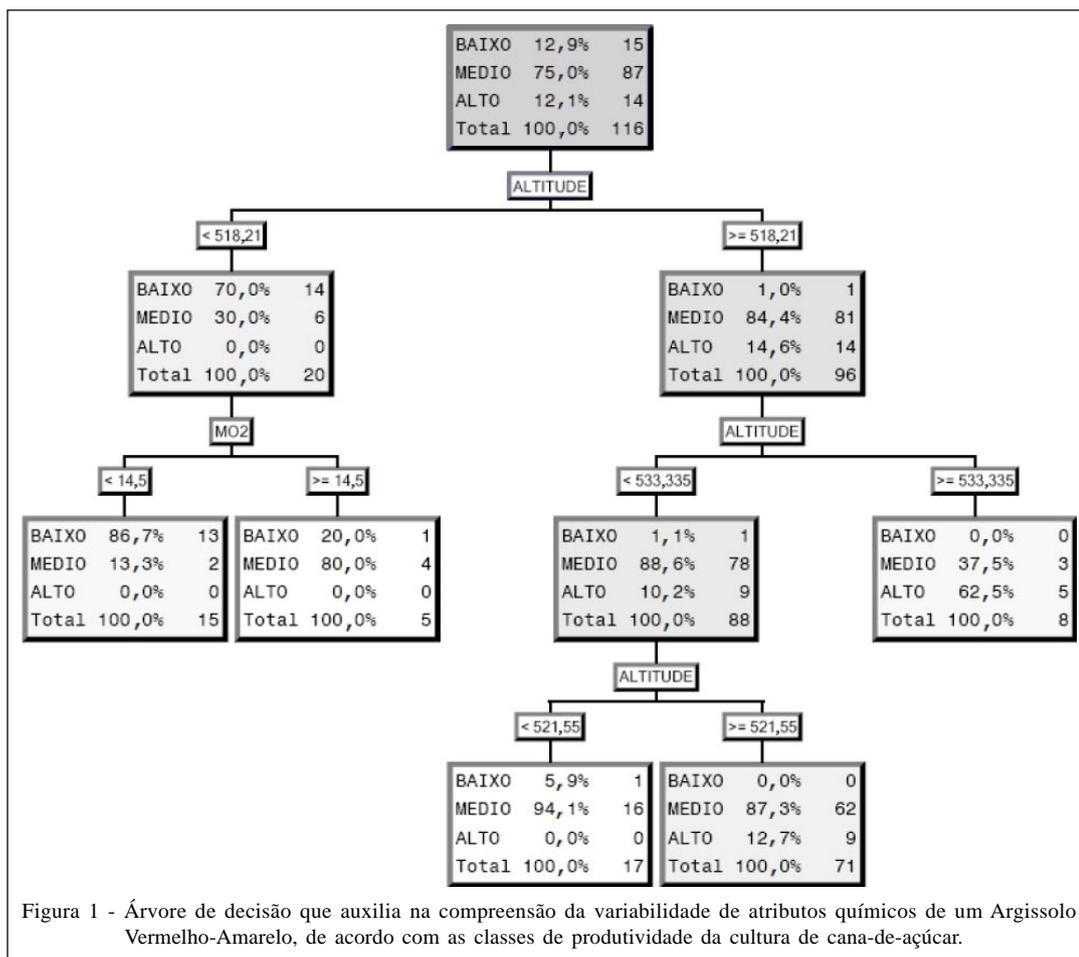
positiva com a produtividade de milho, demonstrando que essa variável influencia o fluxo da água no solo, diretamente associado à produtividade das culturas.

A árvore de decisão induzida demonstrou a influência da altitude na variável nível de produção da cana-de-açúcar, para o talhão estudado (Figura 1). Entre

Tabela 3 – Matriz de correlação linear entre a produtividade da cana-de-açúcar e os atributos de um Argissolo Vermelho-Amarelo.

¹ Altitude	pH	Ca	Mg	K	P	MO	V%	Produtividade
-----Profundidade de 0,0-0,2m-----								
Altitude	-0,22	-0,04	-0,06	0,50**	-0,02	0,20	0,32	0,60**
pH		0,48	0,27	0,23	0,04	-0,05	0,68**	0,17
Ca			0,59**	0,09	-0,05	0,02	0,74**	-0,07
Mg				0,06	0,03	0,09	0,54**	-0,02
K					-0,03	0,13	0,34	0,53**
P						-0,09	-0,10	0,03
MO							0,11	0,12
V%								0,18
-----Profundidade de 0,2-0,4m-----								
Altitude	0,18	-0,09	0,08	0,33	-0,01	-0,02	0,34	0,58**
pH		0,30	0,27	0,23	0,23	-0,01	0,57**	0,15
Ca			0,64**	-0,01	-0,09	0,16	0,73**	-0,15
Mg				-0,04	-0,09	-0,09	0,61**	-0,13
K					-0,16	0,30	0,19	0,52**
P						-0,06	-0,08	-0,03
MO							0,07	0,14
V%								0,12

¹ = significativo a 1%.



um conjunto de atributos químicos do solo, associados com a altitude, a árvore induzida estabelece limites de altitude com alta correlação com os níveis de produção observados. Nos pontos amostrais de menor altitude (< 518,21m), foram observados os menores níveis de produção. Para os locais de baixa altitude, a definição dos níveis de produção se deu pelo teor de matéria orgânica na camada de 0,2-0,4m. Quando o teor de matéria orgânica foi de < 14,5g dm⁻³, o nível de produção foi baixo, com 86,7% de acerto; para um teor de MO ≥ 14,5g dm⁻³, o nível de produção observado passou de baixo para médio, com acerto de 80%.

Para o intervalo de altitude entre 518,21m e 521,55m (altitude média), a árvore demonstra ocorrência de valores médios de produção, com 94,1% de acertos. Já nos pontos amostrais de maior altitude (≥521,55m) foram observados níveis de produção médios e altos, com sua localização definida por intervalos de altitude. Nesse caso, a árvore de decisão não foi robusta na identificação, podendo a definição exata dos níveis mais altos de produção ocorrer em função de alguma outra variável não contemplada no modelo.

Observa-se, por meio da matriz de correlação e pela árvore de decisão, que a altitude da área em estudo foi a variável que apresentou melhor correlação com a produtividade (Tabela 3 e Figura 1). Estudos têm demonstrado que o relevo tem sido um atributo que define a distribuição de atributos químicos do solo ao longo da paisagem e tem apresentado altas correlações com a produtividade de culturas e os atributos do solo (KRAVCHENKO & BULLOCK, 2000; SOUZA et al., 2003; SOUZA et al., 2004; MONTEZANO et al., 2008).

CONCLUSÕES

As variáveis altitude e potássio apresentaram os maiores valores de correlação com a produtividade de cana-de-açúcar.

A indução de árvores de decisão permitiu verificar que a altitude é a variável com maior potencial para interpretar os mapas de produtividade de cana-de-açúcar, auxiliando na agricultura de precisão e mostrando-se uma ferramenta adequada para o estudo

de definição de zonas de manejo em área cultivada com essa cultura.

REFERÊNCIAS

- BURROUGH, P.A. et al. The state of the art in pedometrics. **Geoderma**, Amsterdam, v.62, n.1-3, p.311-326, 1994. Disponível em: <http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6V67-48B0MYK-88&_user=972058&_rdoc=1&_fmt=&_orig=search&_sort=d&_docanchor=&view=c&_acct=C000049648&_version=1&_urlVersion=0&_userid=972058&md5=a9197027531de4960615e2e39b8d3ef7>. Acesso em: 22 de nov. 2008. doi: 10.1016/0016-7061(94)90043-4.
- CAMBARDELLA, C.A. et al. Field-scale variability of soil properties in Central Iowa. **Soil Science Society of American Journal**, Madison, v.58, n.5, p.1501-1511, 1994.
- CORÁ, J.E. et al. Variabilidade espacial de atributos do solo para adoção do sistema de agricultura de precisão na cultura de cana-de-açúcar. **Revista Brasileira de Ciência do Solo**, Viçosa, v.28, n.6, p.1013-1021, 2004. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-06832004000600010&lng=en&nrm=iso&tlng=pt>. Acesso em: 22 out. 2008. doi: 10.1590/S0100-06832004000600010.
- DE LA ROSA, D. et al. Expert evaluation system for assessing field vulnerability to agrochemical compounds in Mediterranean regions. **Journal of Agricultural Engineering Research**, London, v.56, n.2, p.153-164, 1993. Disponível em: <http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6WH1-45P12Y2-S&_user=972058&_rdoc=1&_fmt=&_orig=search&_sort=d&_docanchor=&view=c&_acct=C000049648&_version=1&_urlVersion=0&_userid=972058&md5=30e198bd86ceed800a9ca256312ce655>. Acesso em: 20 de nov. 2008. doi: 10.1006/jaer.1993.1068.
- EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA – EMBRAPA. **Manual de métodos de análise de solo**. 2.ed. Rio de Janeiro: Ministério da Agricultura e do Abastecimento, 1997. 212p.
- GREGO, C.R.; VIEIRA, S.R. Variabilidade espacial de propriedades físicas do solo em uma parcela experimental. **Revista Brasileira de Ciência do Solo**, Viçosa, v.29, n.2, p.169-177, 2005. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-06832005000200002&lng=pt&nrm=iso&tlng=pt>. Acesso em: 19 de out. 2008. doi: 10.1590/S0100-06832005000200002.
- HAN J.; KAMBER M. **Data mining: concepts and techniques**. San Francisco: Morgan Kaufmann/CA, 2001. 550p.
- ISAACS, E.H.; SRIVASTAVA, R.M. **An introduction to applied geostatistics**. New York: Oxford University, 1989. 561p.
- KITAMURA, A.E. et al. Relação entre a variabilidade espacial das frações granulométricas do solo e a produtividade do feijoeiro sob plantio direto. **Revista Brasileira de Ciência do Solo**, Viçosa, v.31, n.2, p.361-369, 2007. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-06832007000200018&lng=pt&nrm=iso&tlng=pt>. Acesso em: 08 de nov. 2008. doi: 10.1590/S0100-06832007000200018.
- KRAVCHENKO, A.N.; BULLOCK, D.G. Correlation of corn and soybean yield with topography and soil properties. **Agronomy Journal**, Madison, v.75, n.1, p.75-83, 2000. Disponível em: <<http://agron.sci journals.org/cgi/content/full/92/1/75>>. Acesso em: 13 de out. 2007.
- MATA, J.V.D. et al. Relação entre produtividade e resistência à penetração em área irrigada por pivô central, sob dois sistemas de preparo. **Acta Scientiarum**, Maringá, v.21, n.3, p.519-525, 1999. Disponível em: <<http://www.periodicos.uem.br/ojs/index.php/ActaSciAgron/article/view/4264/2942>>. Acesso em: 20 de fev. 2008.
- MEGDA, M.M. et al. Correlação linear e espacial entre a produtividade de feijão e a porosidade de um Latossolo Vermelho de Selvíria (MS). **Revista Brasileira de Ciência do Solo**, Viçosa, v.32, n.2, p.781-788, 2008. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-06832008000200032&lng=pt&nrm=iso&tlng=pt>. Acesso em: 18 de Nov. 2007. doi: 10.1590/S0100-06832008000200032.
- MONTEZANO, Z.F. et al. Variabilidade de nutrientes em plantas de milho cultivado em talhão manejado homogêneo. **Bragantia**, Campinas, v.67, n.4, p.969-976, 2008. Disponível em: <http://www.scielo.br/scielo.php?pid=S0006-87052008000400020&script=sci_arttext&tlng=target=_blank>. Acesso em: 01 de nov. 2007. doi: 10.1590/S0006-87052008000400020.
- QUINLAN, J.R. **C4.5: programs for machine learning**. San Francisco: Morgan Kaufmann/CA, 1993. p.234-278.
- RACHID JÚNIOR, A. et al. Variabilidade espacial e temporal de atributos químicos do solo e da produtividade da soja num sistema de agricultura de precisão. **Engenharia na Agricultura**, Viçosa, v.14, n.3, p.156-169, 2006. Disponível em: <<http://www.seer.ufv.br/seer/index.php/reveng/article/viewFile/116/51>>. Acesso em: 22 de out. 2008.
- RAIJ, B. van et al. (Ed). **Análise química para avaliação da fertilidade de solos tropicais**. Campinas: Instituto Agrônomo, 2001. 285p.
- ROBERTSON, G.P. **GS+: Geostatistics for the environmental sciences** (version 5.1 for windows). Plainwell: Gamma Design Software, 1998. 152p.
- ROSA FILHO, G. et al. Variabilidade da produtividade da soja em função de atributos físicos de um Latossolo Vermelho distroférrico sob plantio direto. **Revista Brasileira de Ciência do Solo**, Viçosa, v.33, n.2, p.275-283, 2009. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-06832009000200006&lng=pt&nrm=iso&tlng=pt>. Acesso em: 25 de nov. 2008. doi: 10.1590/S0100-06832009000200006.
- SCHLOTZHAVER, S.D.; LITTELL, R.C. **SAS: system for elementary statistical analysis**. 2.ed. Cary: SAS, 1997. 905p.

- SOUZA, C.K. et al. Influência do relevo na variação anisotrópica dos atributos químicos e granulométricos de uma latossolo em Jaboticabal-SP. **Engenharia Agrícola**, Jaboticabal, v.23, n.3, p.486-495, 2003.
- SOUZA, Z.M. et al. Variabilidade espacial do pH, Ca, Mg e V% do solo em diferentes formas do relevo sob cultivo de cana-de-açúcar. **Ciência Rural**, Santa Maria, v.34, n.6, p.1763-1771, 2004. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-84782004000600015&lng=pt&nrm=iso&tlng=pt>. Acesso em: 01 de dez. 2007. doi: 10.1590/S0103-84782004000600015.
- VANNI, S.M. **Modelos de regressão: estatística aplicada**. São Paulo: Legmar Informática & Editora, 1998. 177p.
- VIEIRA, S.R. Geoestatística em estudos de variabilidade espacial do solo. In: NOVAIS, R.F. et al. (Eds). **Tópicos em ciência do solo**. Viçosa: Sociedade Brasileira de Ciência do Solo, 2000. V.1, p.1-53.
- VIEIRA S.R.; MOLIN J.P. Spatial variability of soil fertility for precision agriculture. In: EUROPEAN CONFERENCE ON PRECISION AGRICULTURE, 3., 2001, Montpellier. **Proceedings...** Montpellier: Agro Montpellier, 2001. V.1, n.3, p.491-496.
- WARRICK, A.W.; NIELSEN, D.R. Spatial variability of soil physical properties in the field. In: HILLEL, D. (Ed.). **Applications of soil physics**. New York: Academic, 1980. Cap.2, p.319-344.
- WEBSTER, R. Statistics to support soil research and their presentation. **European Journal of Soil Science**, Oxford, v.52, n.2, p.331-340, 2001. Disponível em: <<http://www3.interscience.wiley.com/cgi-bin/fulltext/122371407/HTMLSTART>>. Acesso em: 12 de out. 2007.
- YANAI, J. et al. Geostatistical analysis of soil chemical properties and rice yield in a paddy field and application to the analysis of yield-determining factors. **Soil Science and Plant Nutrition**, Tokyo, v.47, n.2, p.291-301, 2001.
- YANG C. et al. Use of hyperspectral imagery for identification of different fertilization methods with decision-tree technology. **Biosystems Engineering**, Amsterdam, v.83, n.3, p.291-298, 2002.