

POSICIONAMENTO ÓTIMO PARA INSTALAÇÃO DE SENSORES USANDO O ALGORITMO FUZZY C-MEANS

CLAUDIO L. BAZZI¹, KELYN SCHENATTO², SHRINIVASA UPADHYAYA³, FRANCISCO ROJO⁴, ERIN KIZER⁵

¹ Doutor em Eng. Agrícola, Docente do Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio, Universidade Tecnológica Federal do Paraná - Medianeira, Paraná, Brasil (45)3240-8000, bazzi@utfpr.edu.br.

² Doutora em Eng. Agrícola, Professora do Departamento de Ciência da Computação, Universidade Tecnológica Federal do Paraná - Santa Helena.

³ Professor Doutor, Universidade da Califórnia - Davis, Califórnia, Estados Unidos.

⁴ Mestre em Engenharia Agrícola, Universidade da Califórnia - Davis, Califórnia, Estados Unidos.

⁵ Mestranda em Engenharia Agrícola, Universidade da Califórnia - Davis, Califórnia, Estados Unidos.

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

RESUMO: Devido ao aumento na demanda de água agrícola e urbana, tornou-se necessário o uso racional dos recursos hídricos. Neste sentido, no meio agrícola, o uso de sensores e controladores para aplicar apenas a quantidade de água necessária, onde e quando é necessário (irrigação de precisão), está crescendo em importância. Para alcançar esse objetivo, os pesquisadores têm desenvolvido tecnologias que medem o teor de umidade do solo e o estado da água da planta para implementar irrigação de precisão para melhorar a eficiência do uso da água. Este estudo utilizou o algoritmo Fuzzy C-Means para gerar zonas de manejo e determinar a colocação ideal de apenas alguns sensores para que uma estimativa confiável do estado da água da planta pudesse ser obtida para implementar irrigação de precisão em amendoeiras e pés de uva. Os resultados mostraram que o algoritmo permite não somente a geração de zonas de manejo eficientes com base nas características do solo e da planta, mas também a colocação de um número limitado de sensores dentro de cada zona de manejo para capturar a variabilidade espacial no estado da água da planta.

PALAVRAS-CHAVE: monitor de folha; irrigação de precisão; zonas de manejo.

OPTIMAL PLACEMENT OF INSTALLATION SENSORS USING FUZZY C-MEANS ALGORITHM

ABSTRACT: Due to increases in agricultural and urban water demand, it has become necessary to use water resources wisely. In agriculture, use of sensors and controllers to apply only the quantity of water required, where and when it is needed (i.e., precision irrigation), is growing in importance. To achieve this goal, researchers have been developing technologies that measure soil moisture content and plant water status to implement precision irrigation to improve water use efficiency. This study used the Fuzzy C-Means algorithm to generate management zones and determine optimal placement of just a few sensors so that reliable estimation of plant water status could be obtained to implement precision irrigation in almond and grape crops. The results showed that the algorithm allows not only the generation of efficient management zones based on soil and plant characteristics but also the placement of a limited number of sensors within each management zone to capture spatial variability in plant water status.

KEYWORDS: leaf monitor; precision irrigation; management zone.

INTRODUÇÃO

A demanda por água para o uso doméstico e industrial está constantemente aumentando e as mudanças do clima em condições extremas estão causando a redução da disponibilidade dos recursos hídricos (MOUAZEN et al., 2014). Desta forma, é necessário o desenvolvimento de pesquisas com o objetivo de otimizar o uso da água (NIJBROEK et al., 2003), incluindo a agricultura que tem uma alta demanda por água, pois é o principal elemento que contribui para variabilidade espacial da produtividade das lavouras (JAYNES; COLVIN, 1997; CAMP; SADLER, 1998). Para Mouazen et al., 2014, o melhor manejo da água em sistemas irrigados podem prover uma economia significativa nestes sistemas de produção.

Zona de manejo (ZM) é uma sub-área onde há um potencial equivalente de produção, tais como níveis similares de nutrientes e de capacidade de retenção de água no solo, relativos a aspectos naturais de relevo, tipo de solo e manejo utilizado (YAN et al., 2007). Uma ZM mostra praticamente as mesmas características de solo e requerem quantidades similares de insumos, tais como água e nutrientes (MORAL et al., 2010; SCHEPERS et al., 2004).

A definição de ZMs pode ser realizada considerando várias características e a capacidade de retenção de solo para otimização do uso da água (ISLAM et al., 2011). Sistemas de controle para irrigação podem ser acionados ou desligados considerando a disponibilidade de água de cada zona ou por características das próprias plantas, como por meio de índices de estresse hídrico destas. O algoritmo Fuzzy C-Means é comumente utilizado para definição de zonas de manejo, tendo-se bons resultados (ILIADIS et al., 2010; ARNO et al., 2011; LI et al., 2013; JIPKATE; GOHOKAR, 2012; MINGOTI; LIMA, 2006), criando zonas de manejo de forma automática e não subjetiva, permitindo a divisão em C-clusters, com referência para o centro de massa ou centroide para cada grupo do cluster (FRIDGEN et al., 2004). Ainda, as estatísticas de MPE e FPI ajudam a definir o número ideal de zonas necessárias para cada talhão.

Quando sistemas de irrigação são baseados em sensores, outro problema que ocorre é sobre o número ideal de sensores necessários e em quais locais estes sensores devem ser adicionados dentro de cada zona. Para responder estas questões, o presente trabalho objetivou criar uma metodologia para definir os melhores locais para instalação de sensores no campo usando zonas de manejo. Duas áreas experimentais (cultivadas com amêndoas e uvas) foram utilizadas para definir zonas de manejo e selecionar o melhor local para instalação de sensores em cada zona. Dados de textura e condutividade elétrica do solo foram utilizadas para definir as zonas de manejo com o algoritmo Fuzzy C-Means. Dados de capacidade de retenção de água no solo foram utilizados para definir as melhores posições para a instalação dos sensores.

MATERIAL E MÉTODOS

Para este estudo, foram utilizados dados de duas áreas agrícolas localizadas na Califórnia, EUA, sendo uma cultivada com amêndoas (2,1 ha) e outra com uvas (4,2 ha), conforme apresentada na Figura 1. Seguindo a recomendação de Doerge (2000) foram utilizadas somente atributos estáveis para geração das zonas de manejo, fazendo uso de dados de textura (areia, argila e silte), elevação e condutividade elétrica (superficial e profunda). Para a área com uva, foram ainda utilizados dados de NDVI. A condutividade elétrica foi obtida por meio do Veris EC superyor 3150.

Para representação das variáveis para cada área fazendo uso de técnicas multivariada de análise de componentes principais (PCA), foram definidas três componentes principais, representando mais que 85% da variabilidade (JOHNSON; WICHERN 2007; MARDIA et al., 1989) que foram utilizadas para criar as zonas de manejo.



FIGURA 1. Talhão 1 - amêndoas (esquerda) e talhão 2 – uva (direita)

Os dados das componentes principais (PCA 1, PCA 2 e PCA 3) foram analisados fazendo uso da análise exploratória e geoestatística (JOURNAL; HUIJBREGTS, 1978), sendo considerados 252 pontos amostrais no talhão 1 e 240 no talhão 2. Foram ajustados os modelos exponencial, gaussiano e esférico ao semivariograma experimental e então selecionado o melhor modelo considerando as estatísticas de validação cruzada (SUN et al., 2009; ARSLAN, 2012). Os dados então foram interpolados por krigagem ordinária com pixel de 2m x 2m.

O software para definição de unidades de manejo (SDUM) (BAZZI et al., 2013) foi utilizado para criar e avaliar as zonas de manejo criadas fazendo uso do algoritmo Fuzzy C-Means (BEDZEK, 1981). Para avaliação do número ótimo de classes, fazendo uso do mesmo software foram utilizados os seguintes métodos:

1. Redução da variância – RV (DOBERMANN et al., 2003; XIANG et al., 2007):

$$RV = 1 - \frac{\sum_{i=1}^n W_i * V_{um_i}}{V_{talhão}} * 100 \quad (1)$$

em que, n – é o tamanho da amostra; W_i – é a proporção de área de cada zona de manejo; V_{um_i} – variância dos dados em cada zona de manejo; $V_{talhão}$ – variância dos dados para todo o talhão.

2. Índice de performance Fuzzy – FPI (FRIDGEN et al., 2004):

$$FPI = 1 - \frac{c}{(c-1)} \left[1 - \sum_{j=1}^n \sum_{i=1}^c (u_{ij})^2 / n \right] \quad (2)$$

em que, c – número de clusteres, n – número de amostras de toda a área; u_{ij} – elemento ij de relevância da matriz Fuzzy.

3. Índice de entropia de partição modificado – MPE (BOYDELL; MCBRATNEY, 2002):

$$MPE = \frac{- \sum_{j=1}^n \sum_{i=1}^c u_{ij} \log(u_{ij}) / n}{\log c} \quad (3)$$

4. em que, c – número de clusters; n – número de amostras de toda a área; u_{ij} – elemento ij de relevância da matriz Fuzzy.

5. Teste de comparação de médias fazendo uso da Anova.

Os locais para instalação dos sensores no campo foram definidos utilizando o algoritmo Fuzzy C-Means com algumas alterações necessárias (Figura 2). Como os locais para instalação dos sensores são fixos foram utilizados os dados dos valores do atributo de retenção de água do solo como centroides no algoritmo, conforme o número prévio de sensores estipulados. Para definir os melhores locais foram calculados os valores de FPI e MPE para a combinação de locais disponíveis para instalação dos sensores e o número de sensores previamente definidos. Foram simuladas as combinações de uso de 2, 3 e 4 sensores para cada talhão, considerando a área como um todo e a divisão por zona de manejo.

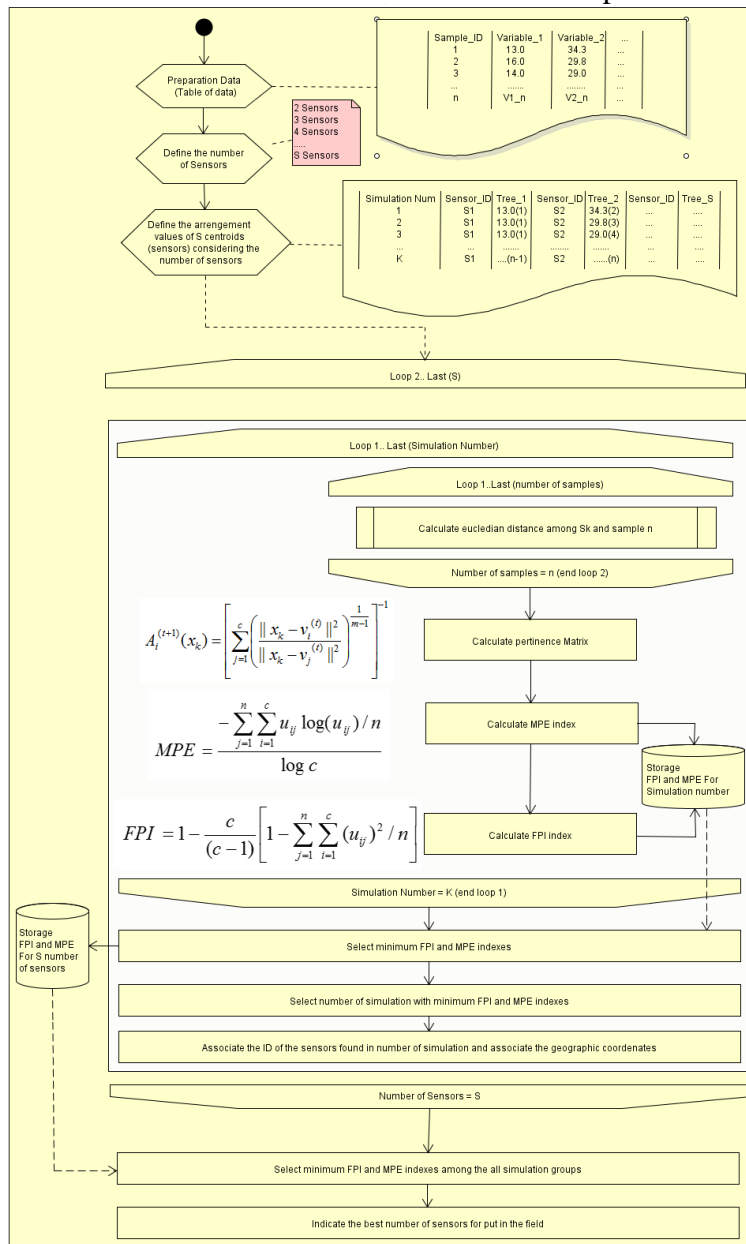


FIGURA 2. Fluxograma da sequencia de passos para definir as melhores posições para instalação dos sensores.

RESULTADOS E DISCUSSÃO

O índice de correlação espacial (Tabela 1), indicou que houve autocorrelação espacial para todos os atributos, indicando dependência espacial. A argila apresentou-se negativamente correlacionada com a elevação indicando que em locais mais altos do talhão o percentual de argila é menor, assim como ocorreu com os dados de condutividade elétrica, indicando correlação positiva com a argila, mas negativa com a elevação, concordando com os resultados apresentados por Sudduth et al. (2005).

TABELA 1. Matriz de correlação espacial (Índice de Moran) para todas as variáveis

		Talhão 1							
	Argila	Areia	Silte	Elevação	CE_superficial	CE_profunda	NDVI	Produtividade	
Argila	0,212								
Areia	-0,137	0,213							
Silte	0,090	-0,209	0,222						
Elevação	-0,137	0,027	0,017*	0,399					
CE_superficial	0,084	-0,143	0,144	-0,055	0,175				
CE_profunda	0,089	-0,130	0,126	-0,080	0,172	0,176			

		Talhão 2							
	Argila	Areia	Silte	Elevação	CE_superficial	CE_profunda	NDVI	Produtividade	
Argila	0,3154								
Areia	-0,2468	0,2917							
Silte	-0,2005	0,046	0,2596						
Elevação	-0,2314	0,3083	-0,0017	0,4885					
CE_superficial	0,1582	-0,1804	-0,0391	-0,2266	0,2137				
CE_profunda	0,2812	-0,2594	-0,1363	-0,312	0,1979	0,3228			
NDVI	-0,1329	-0,0219*	0,2331	-0,1085	0,0228	-0,0782	0,2928		
Produtividade	-0,0815	0,0117*	0,1027	0,0153*	0,0586	-0,0535	0,1477	0,2068	

* Não significativa a 5%.

Para o talhão de amêndoas, a variável tida como mais influente foi a areia (45,7%) e para as videiras, foi a argila (48,4%). Seguindo a recomendação de Johnson e Wichern (2007) e Mardia et al. (1989), foram utilizadas as três primeiras componentes principais em cada talhão, considerando que conseguiram juntas representar pelo menos 85% da variabilidade dos dados.

TABELA 2. Análise de componentes principais (PCA)

	Talhão 1						Talhão 2					
	PC1	PC2	PC3	PC4	PC5	PC6	PC1	PC2	PC3	PC4	PC5	PC6
Proporção (%)	0,662	0,151	0,114	0,069	0,004	0	0,601	0,186	0,114	0,067	0,032	0
Acumulativo (%)	0,662	0,813	0,927	0,996	1	1	0,601	0,787	0,901	0,968	1	1
Variável	PC1	PC2	PC3	PC4	PC5	PC6	PC1	PC2	PC3	PC4	PC5	PC6
Argila	-0,380	-0,182	0,543	-0,69	-0,004	-0,227	-0,484	-0,261	0,195	0,318	0,207	0,718
Areia	0,457	-0,337	-0,296	-0,147	-0,023	-0,754	0,431	-0,345	-0,254	-0,582	-0,117	0,527
Silte	-0,419	0,478	0,162	0,434	0,032	-0,617	0,258	0,816	0,014	0,151	-0,192	0,455
Elevação	0,278	0,781	-0,203	-0,519	0,043	-0,001	0,397	-0,350	-0,264	0,706	-0,391	0,012
CE_superficial	-0,447	-0,035	-0,518	-0,148	-0,713	-0,001	-0,350	0,154	-0,909	0,023	0,162	0,010
CE_profunda	-0,441	-0,118	-0,53	-0,148	0,699	0,001	-0,483	-0,01	0,027	-0,196	-0,853	-0,004

Para cada uma das PCAs, foi realizada a análise geoestatística ajustando os modelos e obtendo-se os parâmetros de efeito pepita, patamar e alcance. Exceto a PCA2 no talhão 1 e a PCA3, no talhão 2, que apresentaram moderada dependência (CAMBARDELA et al., 1994), as demais PCAs apresentaram forte dependência espacial. Os mapas então foram gerados utilizando a krigagem ordinária.

Com os mapas gerados, foram geradas as zonas de manejo em 2, 3, 4 e 5 classes, fazendo uso do SDUM (BAZZI et al., 2013) e o algoritmo Fuzzy C-Means (Figura 3) e por meio dos índices de FPI e MPE verificou-se que o número ideal de zonas para ambos os

talhões foi de 2 zonas (Figura 4).

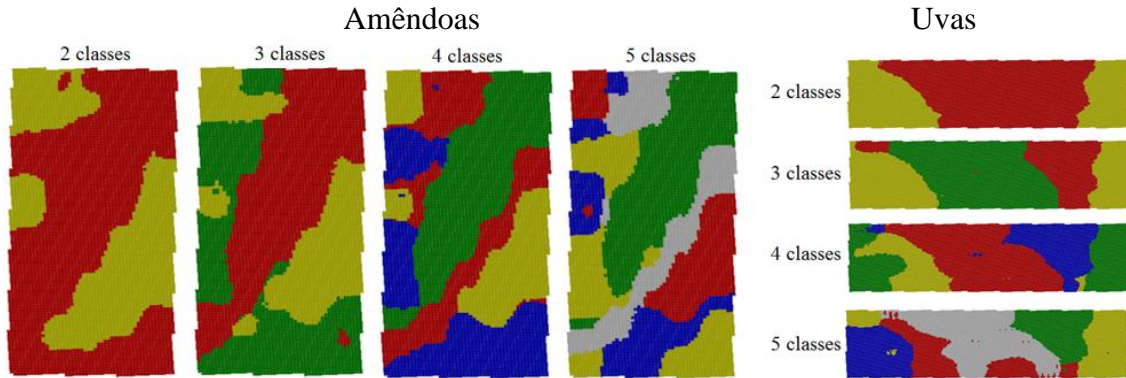


FIGURA 3. Zonas de manejo criadas com Fuzzy C-Means nos talhão 1 e 2

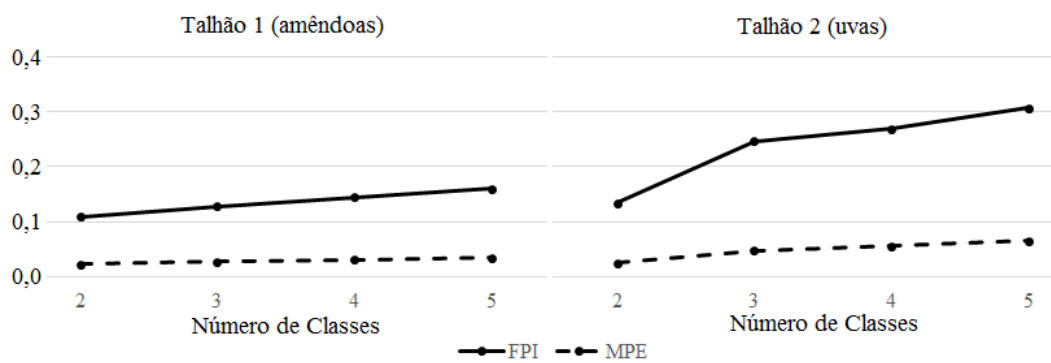


FIGURA 4. Índices FPI e MPE para divisão das zonas de manejo em 2, 3, 4 e 5 classes.

Executados os procedimentos para seleção dos melhores locais para adição dos sensores fazendo uso de dados de capacidade de retenção de água no solo, pode-se verificar que em todos os casos, o sensor foi selecionado representou de forma adequada a variável, conforme pode ser visualizado nas Figura 5.

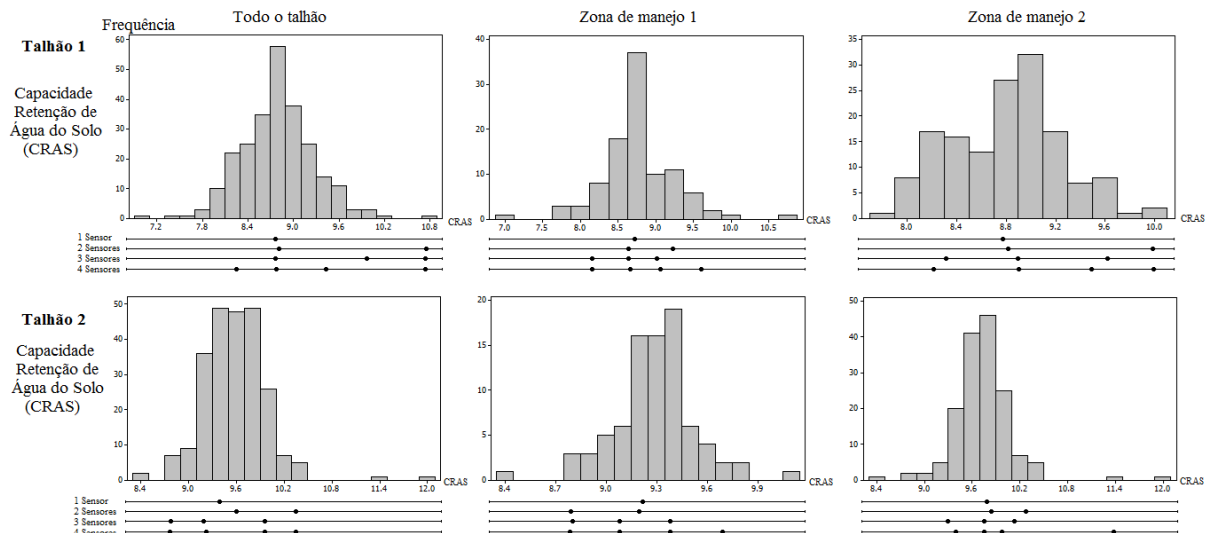


FIGURA 5. Histograma da distribuição da capacidade de retenção de água no solo e os valores deste atributo nos locais selecionados para cada talhão e para divisão em duas zonas de manejo.

A Figura 6 permite verificar a localização geográfica de onde os sensores devem ser

instalados no campo, considerando a execução do algoritmo desenvolvido.

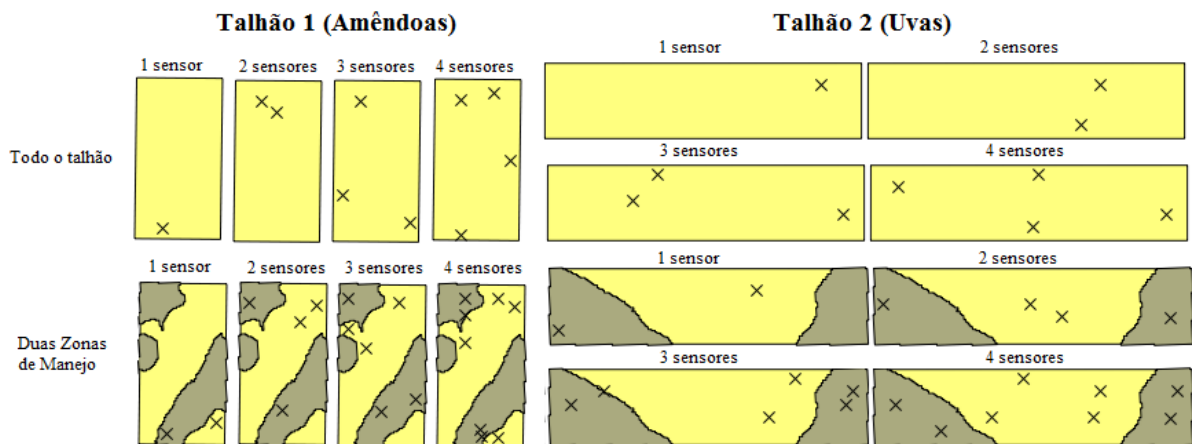


FIGURA 6. Localização dos sensores para cada simulação realizada (1, 2, 3, 4 sensores) para a área total e considerando a divisão em zonas de manejo.

CONCLUSÕES

A técnica de análise de componentes principais (PCA) permitiu a representação satisfatória das variáveis estudadas e permitiu a geração de zonas de manejo de forma satisfatória.

As modificações realizadas no algoritmo Fuzzy C-Means permitiu obter o número ótimo de sensores para serem adicionados em cada talhão ou zona de manejo, bem como identificar quais árvores (plantas) estes sensores deveriam ser instalados para melhor representar a capacidade de retenção de água do solo para servir de base para instalação de sensores de estresse hídrico das plantas.

AGRADECIMENTOS

À Universidade Tecnológica Federal do Paraná, à Universidade da Califórnia e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pelo apoio financeiro.

REFERÊNCIAS

- ARNO, J.; MARTINEZ-CASASNOVAS, J. A.; RIBES-DASI, M.; ROSELL, J. R. Clustering of grape yield maps to delineate site-specific management zones. **Spanish Journal of Agricultural Research**, v.9, n.3, p.721-729, 2011.
- ARSLAN, H. Spatial and temporal mapping of groundwater salinity using ordinary kriging and indicator kriging: The case of Bafra Plain, Turkey. **Agriculture Water Management**, v.113, n.1, p. 57-63, 2012.
- BAZZI, C. L.; SOUZA, E. G.; OPAZZO, M. O.; NOBREGA, L. H.; ROCHA, D. M. Management Zones definition using soil chemical and physical attributes in a soybean area. **Engenharia Agrícola**, v.33, n.5, p. 952-964, 2013.
- BEZDEK, C. J. **Pattern recognition with fuzzy objective function algorithms**. 1. ed. New York: Plenum Press, 1981. 272 p.

BOYDELL, B.; MCBRATNEY, A. B. Identifying potential within-field management zones from cotton yield estimates. **Precision Agriculture**, v.3, n.1, p. 9-23, 2002.

CAMBARDELLA C. A.; MOORMAN T. B.; NOVAK J. M.; PARKIN, T. B.; KARLEN, D. L.; TURCO, R. F.; KONOPKA A. E. Field-scale variability of soil properties in central Iowa soils. **Soil Science Society American Journal**, v.58, n.1, p. 1501-1511, 1994.

CAMP, C. R.; SADLER, E. J. Site-specific crop management with a center pivot. **Journal of soil and water conservation**, v.53, n.4, p. 312-314, 1998.

DOBERMANN, A.; PING, J. L.; ADAMCHUK, V. I.; SIMBAHAN. G. C.; FERGUSON, R. B. Classification of crop yield variability in irrigated production fields. **Agronomy Journal**, v.95, n.1, p. 1105-1120, 2003.

DOERGE, T.A. Site-Specific Management Guidelines. **Potash & Phosphate Institute**, Norcross, 2000.

ILIADIS, L. S.; VANGELOUDH, M.; SPARTALIS, S. An intelligent system employing an enhanced fuzzy C-Means clustering model: Application in the case of forest fires. **Computers and Electronics in Agriculture**, v.70, n.1, p. 276-284, 2010.

JAYNES, D. B.; COLVIN, T. S.; KASPAR, T. C. Identifying potential soybean management zones from multi-year yield data. **Computer and electronics in agriculture**, v.46, n.3, p. 309-327, 2005.

JIPKATE, B. R.; GOHOKAR, V. V. A. Comparative analysis of Fuzzy C-Means clustering and K-Means clustering algorithms. **International Journal of Computational Engineering**, v.2, n.3, p. 737-739, 2012.

JOHNSON, R. A.; WICHERN, D. W. **Applied Multivariate Statistical Analysis**. 6. ed. New Jersey: Pearson, 2007. 800 p.

JOURNEL, A. G.; HUIJBREGTS, C. H. J. **Mining Geostatistics**. 1. ed. New York: The Blackburn Press, 1978. 600 p.

FRIDGEN, J. J.; KITCHEN, N. R.; SUDDUTH, K. A.; DRUMMOND, S. T.; WIEBOLD, W. J.; FRAISSE, C. W. Management zone analyst (MZA): software for subfield management zone delineation. **Agronomy Journal**, v.96, n.1, p. 100-108, 2004.

LI, Y.; SHI, Z.; WU, H.; LI, F.; LI, H. Definition of management zones for enhancing cultivated land conservation using combined Spatial Data. **Environmental Management**, v.52, n.1, p. 792-806, 2013.

MARDIA, K. W.; KENT, J. T.; BIBBY, J. M. **Multivariate Analysis**. London: Academic Press, 1989.

MINGOTI, S. A.; LIMA, J. O. Comparing SOM neural network with Fuzzy C-means, K-means and traditional hierarchical clustering algorithms. **European Journal of Operational Research**, v.174, n.3, p. 1742-1759, 2006.

MORAL, F. J.; TERRÓN, J. M.; SILVA, J. R. M.; Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. **Soil Tillage Research**, v.106, n.1, p. 335-343, 2010.

MOUAZEN, A. M.; ALHWAIMEL, S. A.; KUANG, B.; WAINE, T. Multiple on-line soil sensors and data fusion approach for delineation of water holding capacity zones for site specific irrigation. **Soil & Tillage Research**, v.143, n.1, p. 95-105, 2014.

NIJBROEK, R.; HOOGENBOOM, G.; JONES, J. W. Optimizing irrigation management for a spatially variable soybean Field. **Agricultural Systems**, v.76, n.1, p. 359-377, 2003.

SCHEPERS, A. R.; SHANAHAN, F. J.; LIEBIG, M. A.; SCHEPERS, J. S.; JOHNSON, S. H.; LUCHIARI, J. A. Appropriateness of management zones for characterizing spatial variability of soil properties and irrigated corn yields across years. **Agronomy Journal**, v.96, n.1, p. 195-203, 2004.

SUDDUTH, K. A.; KITCHEN, N. R.; WIEBOLD, W. J.; BATCHELOR, W. D.; BOLLERO, G. A.; BULLOCK, D. G.; CLAY, D. E.; PALM, H. L.; PIERCE, F. J.; SCHULER, R. T.;

THELEN, K. D. Relating apparent electrical conductivity to soil properties across the north-central USA. **Computers and Electronics in Agriculture**. v.46, n.1, p. 263-283, 2005.

SUN, Y.; KANG, S.; LI, F.; ZHANG, L. Comparison of interpolation methods for depth to groundwater and its temporal and spatial variations in the Minqin oasis of northwest China. **Environmental Modelling & Software**, v.24, n.10, p. 1163-1170, 2009.

XIANG, L.; PAN, Y.; GE, Z.; ZHAO, C. Delineation and Scale Effect of Precision Agriculture Management Zones Using Yield Monitor Data Over Four Years. **Agriculture Sciences in China**, v.6, n.2, p. 180-188, 2007.

YAN, L.; ZHOU, S.; CIFANG, W.; HONGYI L.; FENG, L. Classification of management zones for precision farming in saline soil based on multi-data sources to characterize spatial variability of soil properties. **Transactions of the Chinese Society of Agricultural Engineering**, v.23, n.8, p. 84–89, 2007.