



APPLICATION OF FUZZY SOIL CLASSIFICATION IN VISUALIZING 3-D SOIL STRATA

Jing-Wen Chen

Professor, Department of Civil Engineering, National Cheng Kung University, 1, Da-Shine RD., Tainan, Taiwan 701,
R.O.C, geothen@mail.ncku.edu.tw

Follow this and additional works at: <https://jmstt.ntou.edu.tw/journal>



Part of the [Civil and Environmental Engineering Commons](#)

Recommended Citation

Chen, Jing-Wen (2005) "APPLICATION OF FUZZY SOIL CLASSIFICATION IN VISUALIZING 3-D SOIL STRATA," *Journal of Marine Science and Technology*. Vol. 13: Iss. 4, Article 4.

DOI: 10.51400/2709-6998.2120

Available at: <https://jmstt.ntou.edu.tw/journal/vol13/iss4/4>

This Research Article is brought to you for free and open access by Journal of Marine Science and Technology. It has been accepted for inclusion in Journal of Marine Science and Technology by an authorized editor of Journal of Marine Science and Technology.

APPLICATION OF FUZZY SOIL CLASSIFICATION IN VISUALIZING 3-D SOIL STRATA

Jing-Wen Chen* and Chang-Hua Chen**

Key words: fuzzy similarity measure, principal component analysis, index of soil classification.

ABSTRACT

The distribution of soil classes is an important factor in engineering design. In order to generate the soil class map, fuzzy soil classifications were developed to provide the means to characterize and quantify the soil classes, and then the spatial distribution of soil class can be interpolated by geostatistical method. For establishing continuous class soil maps, this paper present an index of fuzzy soil classification generated from principal component analysis (PCA). The ability of classification of the index is tested with a soil database which including 9,260 soil samples, and the accurate rate of simulation was 91.7%. There are two procedures presented to establish the 3-D soil class map. The first method is to integrate 2-D membership degree maps into 3-D maps; while the second method is to interpolate the spatial distribution of index of fuzzy soil classification directly by 3-D ordinary kriging. These procedures are illustrated by a demonstrated area.

INTRODUCTION

Geotechnical engineering design is strongly influenced by the distribution of soil classes. For visualizing the three-dimensional distribution of soil classes, the quantitative system of soil classification is required. The traditional systems of soil classification are deterministic and symbolic and cannot to be used to conduct numerical calculation or to predict the spatial distribution of soil classes. McBratney and de Gruijter[8] proposed fuzzy k -means (FKM) for fuzzy quantitative soil classification. Fuzzy soil classification uses the membership degrees to correspond to the degrees of classification. Hence, the fuzzy information is able to express the gradual transition of soil classes [7] and can be interpolated by geostatistical methods [5] to obtain

the spatial prediction of soil classes [3, 4, 6, 11].

FKM is an unsupervised cluster system; it needs to define the number of cluster in advance but not to define the rule of experience. It can obtain the optimal centers of groups that minimize the total distance from sample to cluster center. Chen *et al.* [2] showed that the fuzzy similarity measure could improve the ability of FKM's identification in soil classification. This system is referred to the fuzzy similarity of grain-size distribution (FSGSD). FSGSD is based on the similarity degree between curves of grain-size distribution to identify the soil classes. The membership degrees are obtained by FKM, and the degrees of similarity are evaluated by FSGSD for each class. Then the soil class maps are generated by geotechnical method and illustrated individually for each class; not an integrated soil class map. Therefore this paper uses principal component analysis (PCA) to establish an index of fuzzy soil classification. Through the index, the spatial distribution of soil classes can be estimated easily, and soil class map can be made further. In this paper the procedure of analysis includes four parts: the quantifying soil classification system, using PCA to obtain the index of fuzzy soil classification from the degree of fuzzy similarity, applying geostatistics to predict the distribution of index of fuzzy soil classification, and visualizing the 3-D soil class map.

FUZZY SIMILARITY OF GRAIN-SIZE DISTRIBUTION

The fuzzy similarity measure which first presented by Pappis and Karacapilidis [12] is mainly applied to determine the degree of similarity between two fuzzy numbers. FSGSD, which presented by Chen *et al.* [2], identifies the soil classes based on the similarity between patterns of grain-size distribution curve. The similarity degree is defined as the ratio of the area that is confined by minimum curve to the area that is confined by maximum curve as shown in Figure 1. In the figure, the x -axis is grain size and y -axis is percent finer. The x -axis is assumed nominal scale for simplifying

Paper Submitted 03/13/05, Accepted 07/12/05. Author for Correspondence: Jing-Wen Chen. E-mail: geochen@mail.ncku.edu.tw.

*Professor, Department of Civil Engineering, National Cheng Kung University, 1, Da-Shine RD., Tainan, Taiwan 701, R.O.C.

**Doctoral student, Department of Civil Engineering, National Cheng Kung University, 1, Da-Shine RD., Tainan, Taiwan 701, R.O.C.

calculation. Since the area measure is a trapezoid area, the measure equation can be written as follows:

$$S(A, B) = \frac{(G_L + 2S_L + 2M_L + C_L)}{(G_U + 2S_U + 2M_U + C_U)} \quad (1)$$

where $G_L, S_L, M_L,$ and C_L represent the lower accumulated content of gravel, sand, silt, and clay between two curves, respectively; $G_U, S_U, M_U,$ and C_U are the higher accumulated content.

According to the Unified Soil Classification System (USCS), the grain size of gravel is over 4.75 mm (sieve no. 4), sand is between 0.075 mm (sieve no. 200) and 4.75 mm, and fine grain is under 0.075 mm. Then according to the American Association of State Highway and Transportation Officials (AASHTO), the fine grain size is divided into silt size (between 0.002 mm and 0.075 mm) and clay size (under 0.002 mm). The simplified distribution curves of fuzzy soil classes were obtained by FKM from Chen *et al.* [2]. The accumulated contents of grain size of fuzzy soil class are listed in Table 1. The procedure of FSGSD is illustrated with an example as follows. If a soil sample is defined SM by USCS and the accumulated grain-size contents of gravel, sand, silt, and clay are 100%, 100%, 19%, and 1%, respectively. In the fuzzy gravel class, the accumulated grain-size contents of gravel, sand, silt, and clay are 100%, 64.36%, 26.58%, and 11.43%, respectively. The transformed distribution curves can be obtained by comparing the two curves. The curve of maximum value

consists of the higher accumulated content, and the $G_U, S_U, M_U,$ and C_U are 100%, 100%, 26.58%, and 11.43%, respectively. The curve of minimum value consists of the lower accumulated content, and the $G_L, S_L, M_L,$ and C_L are 100%, 64.36%, 19%, and 1%, respectively. Then the degree of similarity between this sample and fuzzy gravel class can be evaluated by Eq. (1), and the degree is 0.7343. The other degrees of similarity for fuzzy sand, fuzzy silt, and fuzzy clay class are 0.9554, 0.7335, and 0.6507, respectively. Finally this sample will be classified to the class of the highest degree of similarity, *i.e.* fuzzy sand class.

PRINCIPAL COMPONENT ANALYSIS

PCA is concerned with explaining the variance-covariance structure through a few linear combinations of the original variables. PCA can conduct the orthogonal transformation of a coordinate system and removes linear correlations among the data. The new coordinate axes are called as principal components. PCA can efficiently represent data distributed on a linear hyper plane and retain as much information as possible in fewer components instead of the original variables. The PCA need to evaluate the eigenvalues and eigenvectors from the correlation matrix of data. The eigenvalues represent the ability of explained variance. The eigenvectors are the coefficient of transformation of the coordinate system. Therefore, the eigenvector in maxi-

Table 1. The accumulated contents of grain size of fuzzy soil class form FKM

Fuzzy soil class	Accumulated contents of grain size			
	Gravel size	Sand size	Silt size	Clay size
Gravel	100%	64.36%	26.58%	11.43%
Sand	100%	99.75%	25.10%	4.12%
Silt	100%	99.89%	75.79%	10.27%
Clay	100%	99.67%	89.63%	40.66%

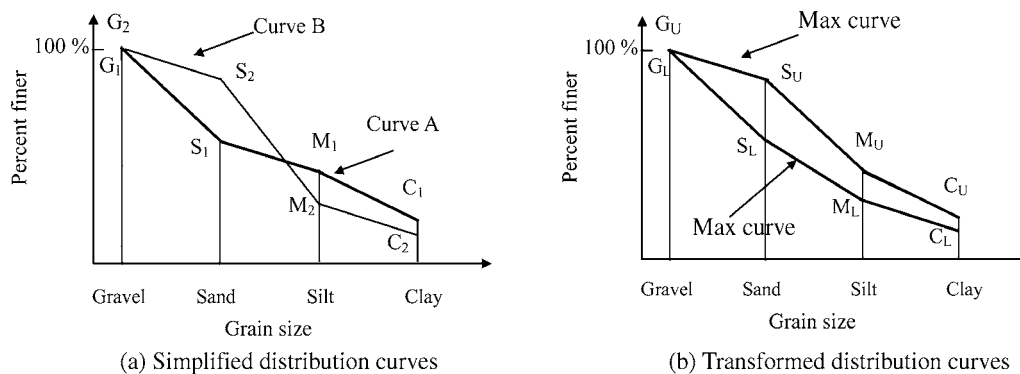


Fig. 1. The max. and min. value curves of grain-size distribution.

imum eigenvalue is the coefficient of transformation of first principal component. In this paper, the PCA with correlation matrix is conducted by MINITAB software [10]. The PCA is applied in FSGSD and in FKM, and the first two components are required.

ANALYSIS OF A DATABASE

For establishing the index of fuzzy soil classification, the PCA is performed with a database. The database includes 650 boreholes, 9,260 samples, in Kaohsiung city. The depth was from 0 to 30 meters and the statistic of soil classes are listed in Table 1. From Table 1, the number of gravel, sand, silt, and clay soil are 103 (1.11%), 5,892 (63.63%), 1,534 (16.57%) and 1,731 (18.69%), respectively. The soil classes in the soil database mainly are sandy soil (63.63%); the gravel is relatively little (1.11%) and in some sub-groups there is no sample.

In FSGSD, the four eigenvalues are 3.3907, 0.4577, 0.1095, and 0.0421, respectively. The ratio of an eigenvalue to the sum of eigenvalues is called as the ratio of explained variance. Therefore, the first component is able to explain the variance up to 84.7%. The values in the new coordination system are called as the component scores. The scatter plot of the first and the second component scores in FSGSD is illustrated in Figure 2; the soil classes are obviously clustered.

For establishing the index of fuzzy soil classification, the distribution of the first component scores is shown in Figure 3, and the difference soil classes can be readily identified. Therefore the first component is regarded as the index of fuzzy soil classification (I_{fc}). The formula of first component (or called as the index of fuzzy soil classification) is as follows.

$$I_{FC} = 0.50206 \times \frac{S_G - 0.69251}{0.07939} + 0.50124 \times \frac{S_S - 0.86416}{0.11174} - 0.46345 \times \frac{S_M - 0.81305}{0.08546} - 0.53096 \times \frac{S_C - 0.75734}{0.12091} \quad (2)$$

where S_G , S_S , S_M , and S_C are the degrees of similarity of a soil sample belonging to fuzzy gravel, fuzzy sand, fuzzy silt, and fuzzy clay class, respectively. The coefficient of Eq. (2) is the eigenvector of first component. The following, $(S_G - 0.69251)/0.07939$, $(S_S - 0.86416)/0.11174$, $(S_M - 0.81305)/0.08546$, and $(S_C - 0.75734)/0.12091$, are the normalized scores. According to the distribution of the first principal component score, the intervals of the index of fuzzy soil classification can be properly defined for each soil class: gravel soil is greater than 1.81, sand soil between 1.81 and

-0.4, silt soil between -0.4 and -2.71, and clay soil is less than -2.71. The result of classification by the index is listed in Table 2 and the rate of accuracy of gravel, sand, silt, and clay soil are 75.73%, 98.23%, 80.70%, and 80.13%, respectively. The overall accuracy is 91.70% on the average.

In FKM the scatter plot of the first and the second component scores is illustrated in Figure 4; the soil classes are not obviously clustered. The distribution of the first component score is quite extreme as shown in Figure 5 and the soil classes cannot be identified.

RESULTS AND DISCUSSION

This study selects a regular 100 m × 100 m area as demonstrated case and collects a total of 15 boreholes (including 195 soil samples, at 1.5 m depth intervals between 0 to 20 m) as shown in Figure 6. Among these, 11 boreholes (143 soil samples) are used as the trained

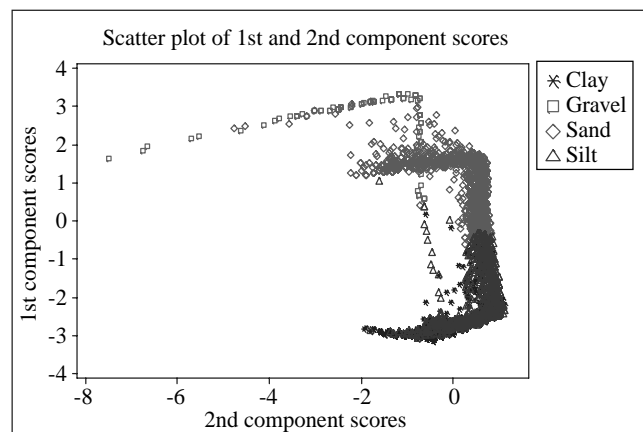


Fig. 2. The scatter plots of the first and second component scores in FSGSD.

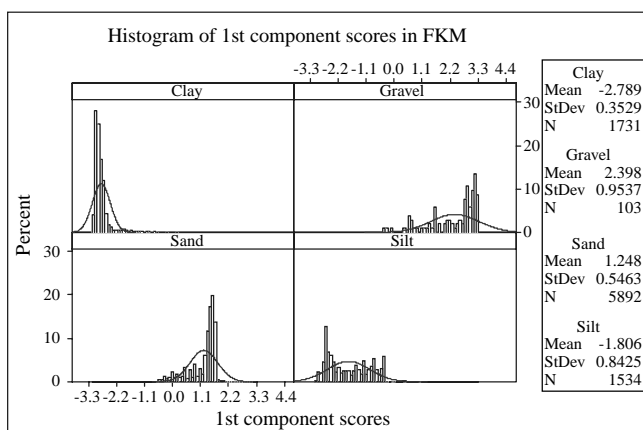


Fig. 3. The distributions of the first component score in FSGSD.

data and 4 boreholes (52 soil samples) as the verified data. Two procedures are presented to establish 3-D soil class map. The details described as following paragraphs.

1. The 3-D soil class map integrated from 2-D membership degree maps

In this demonstrated site, the soil horizontal layers are internally uniform corresponding the strata units but the profiles in vertical are discontinuous relatively. Hence, 2-D ordinary kriging is used to interpolate the spatial distribution of membership degree of fuzzy soil classes in each horizontal layer [4]. The global variograms with the exponential model are taken in the geostatistics and the number of 2-D soil maps is 52, four fuzzy classes at 13 depths. The class memberships are integrated into an index of fuzzy soil classification. The geostatistical analysis was conducted by Vesper software [9] for each grid-sampling interval. Vesper is developed by the Australian Center of Precision Agri-

culture (ACPA) for spatial prediction that is capable of performing ordinary kriging with local variograms. The 3-D soil class map is visually illustrated by GridSTAT software [1] as shown in Figure 6. The coordinate system in the figure is the 2-degree transverse Mercator projection with Taiwan Datum 1997.

The results of verified data are listed in Table 3. The accurate rates of sand, silt and clay are 95.83%, 85.71%, and 61.9%, respectively. The overall accurate rate is 80.77%.

2. The 3-D soil class map generated by 3-D ordinary kriging

The procedure that described above is complicated since the degrees of class membership need to be interpolated individually in each layer and to be transformed into the index of fuzzy soil classification.

Hence, a comparatively easy procedure is presented. The 3-D kriging is used directly to estimate the index of fuzzy soil classification in 3-D geological

Table 2. The result of classification identified by the first principal component in FSGSD

Number (percent)	Classification by USCS				Sum
	Predict groups	Gravel	Sand	Silt	
Gravel	78 (75.73%)	66 (1.12%)	0 (0%)	0 (0%)	144
Sand	24 (23.30%)	5788 (98.23%)	84 (5.48%)	2 (0.12%)	5898
Silt	1 (0.97%)	38 (0.64%)	1238 (80.70%)	342 (19.76%)	1619
Clay	0 (0%)	0 (0%)	212 (13.82%)	1387 (80.13%)	1599
Sum	103	5892	1534	1731	9260

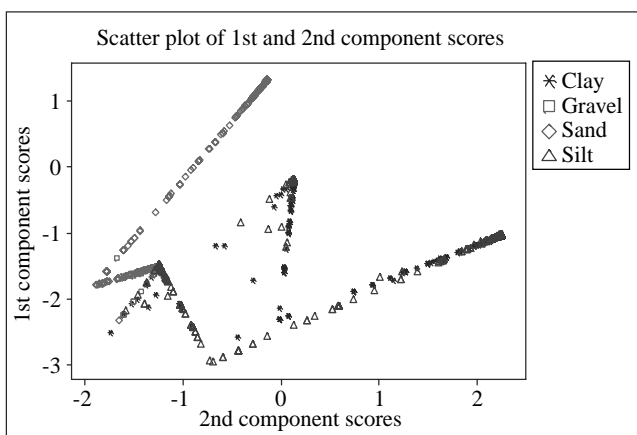


Fig. 4. The scatter plots of the first and second component scores in FKM.

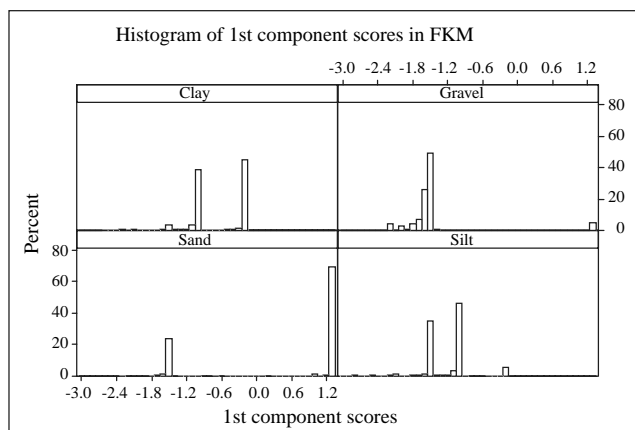


Fig. 5. The histograms of the first component score in FKM.

space. The procedure is conducted by GridSTAT Pro software, and the parameters of anisotropy ellipse are used in the variograms. After calculating by GridSTAT Pro, the major correlation length is 231 m, the minor is 192 m, vertical is 18 m, and the direction of greatest continuity is 53.906 degrees in clockwise from the north. Since the 3-D kriging consider the correlation of

vertical variation, it is more suitable to interpolate the non-homogeneous layers in vertical direction. The 3-D soil class map estimated by 3-D ordinary kriging is shown in Figure 7. This soil map is different from the map integrated by 2-D membership degree maps (see Figure 6), and its the vertical variation is stronger.

The results of verified data are listed in Table 4.

Table 3. The result of classification predict by 2-D ordinary kriging

Number (percent)	Classification by USCS			Sum
	Predict groups	Sand	Silt	
Sand	23 (95.83%)	1 (14.29%)	3 (14.29%)	27
Silt	1 (4.17%)	6 (85.71%)	5 (23.81%)	12
Clay	0 (0.00%)	0 (0.00%)	13 (61.90%)	13
Sum	24	7	21	52

Table 4. The result of classification predicted by 3-D ordinary kriging

Number (percent)	Classification by USCS			Sum
	Predict groups	Sand	Silt	
Sand	23 (95.83%)	5 (71.43%)	4 (19.05%)	32
Silt	1 (4.17%)	1 (14.29%)	4 (19.05%)	6
Clay	0 (0.00%)	1 (14.29%)	13 (61.90%)	14
Sum	24	7	21	52

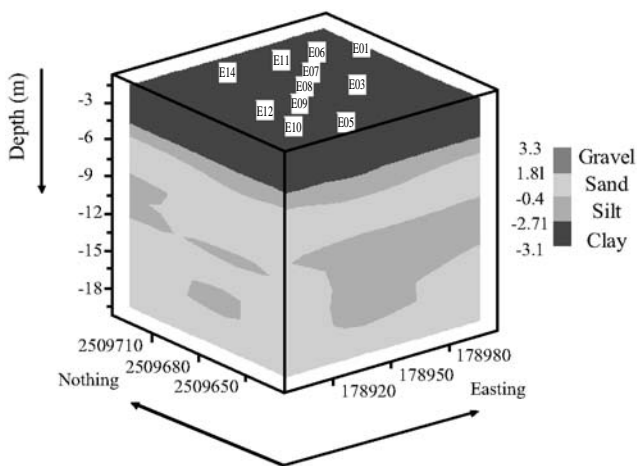


Fig. 6. The 3-D soil map integrated from 2-D memberships (vertical exaggeration is 5).

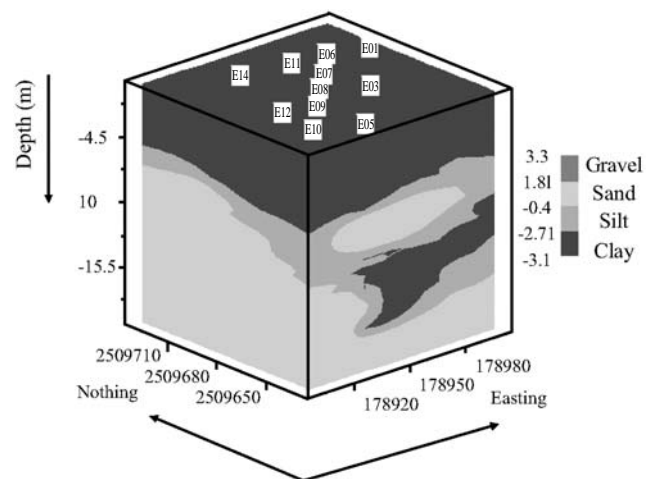


Fig. 7. The 3-D soil maps interpolated by 3-D ordinary kriging (vertical exaggeration is 5).

The accurate rates of sand, silt and clay are 95.83%, 14.29%, and 61.00%, respectively. The small accurate rate of silt maybe results from the small number of silt samples, the error of spatial interpolation, and the error the identification by index. The overall accurate rate is 71.15%, which is inferior to the result of the 3-D soil map integrated by 2-D membership degree maps. The demonstrated area in the study is internally uniform in horizons; thereby the 2-D kriging model is the better. However, 3-D kriging model is comparatively easy and considers the correlation of vertical variation; it is more suitable to the strata relatively continuous in vertical.

CONCLUSION

Fuzzy soil classification can express quantitative information of soil properties. In order to establish continuous class soil maps, the index of fuzzy soil classification is presented and the accurate rate is up to 91.7% in a soil database from Kaohsiung city.

Two procedures are presented to establish the 3-D soil class map. One is to integrate 2-D membership maps into 3-D soil class map. The other is to interpolate the index of fuzzy soil classification directly by 3-D ordinary kriging and comparatively easy to be conducted. Since the 3-D kriging model considers the correlation of vertical variation, it is more proper in the strata that are relatively continuous in vertical. But in the demonstrated area the strata are internally uniform in horizontal, the accurate rate of the 3-D kriging model is only 71.15%, which is inferior the result of 2-D kriging model, 80.77%. The both two procedures are quite effective, however they should be chosen properly for the different strata situation. The 2-D kriging model is proper for alluvium, and 3-D kriging model could be proper for diluvium.

REFERENCES

1. Applied Computer Engineering Inc., *GridSTAT Pro Tutorial*, (1999). (<http://www.gridstat.com>)
2. Chen, J.W., Chen, C.H., and Chen, S.C., "Application of Fuzzy k -mean Cluster and Fuzzy Similarity in Soil Classification," *Proceeding of The Fifteenth International Offshore and Polar Engineering Conference, Seoul Korea*, pp. 459-465 (2005).
3. De Gruijter, J.J., Walvoort, D.J.J., and van Gaans, P.F.M., "Continuous Soil Maps-a Fuzzy Set Approach to Bridge the Gap Between Aggregation Levels of Process and Distribution Models," *Geoderma*, Vol. 77, pp. 169-195(1997).
4. Grunwald, S., Lowery, B., Rooney, D.J., and McSweeney, K., "Profile Cone Penetrometer Data used to Distinguish Between Soil Materials," *Soil Till. Res.*, Vol. 62, pp. 27-40 (2001).
5. Hendricks Franssen, J.J.W.M., van Eijnsbergen, A.C., and Stein, A., "Use of Spatial Prediction Techniques and Fuzzy Classification for Mapping Soil Pollutants," *Geoderma*, Vol. 77, pp. 243-262 (1997).
6. Lagacherie, P., Cazemier, D.R., van Gaans, P.F.M., and Burrough, P.A., "Fuzzy k -means Clustering of Fields in an Elementary Catchment and Extrapolation to a Larger Area," *Geoderma*, Vol. 77, pp. 197-216 (1997).
7. McBratney, A.B. and De Gruijter, J.J., "A Continuum Approach to Soil Classification by Modified Fuzzy k -means with Extragrades," *J. Soil Sci.*, Vol. 43, pp. 159-175 (1992).
8. McBratney, A.B., De Gruijter, J.J., and Brus, D.J., "Spatial Prediction and Mapping of Continuous Soil Classes," *Geoderma*, Vol. 54, pp. 39-64 (1992).
9. Minasny, B., McBratney, A.B., and Whelan, B.M., VESPER version 1.6, Australian Centre for Precision Agriculture, University of Sydney, Australia (2002). (<http://www.usyd.edu.au/su/agric/acpa>)
10. Minitab Inc., MINITAB Release 14, computer software, PA, USA (2004).
11. Odeh, I.O.A., McBratney, A.B., and Chittleborough, D.J., "Soil Pattern Recognition with Fuzzy c -Means: Applications to Classification and Soil-Landform Interrelationships," *Soil Sci. Soc. Am. J.*, Vol. 56, pp. 505-516 (1992).
12. Pappis, C.P. and Karacapilidis, N.I., "A Comparative Assessment of Measures of Similarity of Fuzzy Values," *Fuzzy Set. Syst.*, Vol. 56, pp. 171-174 (1993).