TECHNICAL PAPER

MANAGEMENT ZONES IN AGRICULTURE ACORDING TO THE SOIL AND LANDSCAPE VARIABLES

HEVANDRO C. DELALIBERA[1], PEDRO H. WEIRICH NETO[2], NOEMI NAGATA[3]

ABSTRACT: The study of spatial variability of soil and plants attributes, or precision agriculture, a technique that aims the rational use of natural resources, is expanding commercially in Brazil. Nevertheless, there is a lack of mathematical analysis that supports the correlation of these independent variables and their interactions with the productivity, identifying scientific standards technologically applicable. The aim of this study was to identify patterns of soil variability according to the eleven physical and seven chemical indicators in an agricultural area. It was used two multivariate techniques: the hierarchical cluster analysis (HCA) and the principal component analysis (PCA). According to the HCA, the area was divided into five management zones: zone 1 with 2.87ha, zone 2 with 0.8ha, zone 3 with 1.84ha, zone 4 with 1.33ha and zone 5 with 2.76ha. By the PCA, it was identified the most important variables within each zone: V% for the zone 1, CTC in the zone 2, levels of H+Al in the zone 4 and sand content and altitude in the zone 5. The zone 3 was classified as an intermediate zone with characteristics of all others. According to the results it is concluded that it is possible to separate into groups (management zones) samples with the same patterns of variability by the multivariate statistical techniques.

KEYWORDS: multivariate analysis, precision agriculture, spatial variability.

ZONAS DE MANEJO AGRÍCOLAS DE ACORDO COM VARIÁVEIS DE SOLO E PAISAGEM

RESUMO: O estudo da variabilidade espacial de atributos do solo e plantas ou agricultura de precisão, técnica que visa ao uso racional dos recursos naturais, está expandindo-se comercialmente no Brasil. Porém ainda ressente-se de análise que dê respaldo matemático a correlação de variáveis independentes e suas interações com a variável produtividade, identificando padrões científicos tecnologicamente aplicáveis. Objetivou-se identificar padrões de variabilidade do solo com base em onze indicadores químicos e sete físicos em área agrícola. Utilizaram-se duas técnicas de estatística multivariada: análise hierárquica de agrupamentos (HCA) e análise por componentes principais (PCA). De acordo com a HCA, separou-se a área de estudo em cinco zonas de manejo, sendo a zona 1 com 2,87 ha; a zona 2 com 0,8 ha; a zona 3 com 1,84 ha; a zona 4 com 1,33 ha; e a zona 5 com 2,76ha. Por meio da PCA identificaram-se as variáveis mais importantes dentro de cada zona: na zona 1 o V%; na 2 a CTC; na 4 o H+Al, e na 5 o conteúdo de areia e a altitude. A zona 3 foi classificada como uma zona intermediária, apresentando características de todas as outras. Sendo assim, é possível separar em grupos (zonas de manejo) as amostras que possuem os mesmos padrões de variabilidade com técnicas de estatística multivariada.

PALAVRAS-CHAVE: análise multivariada, agricultura de precisão, variabilidade espacial.

[1] Pesquisador, Área de Engenharia Agrícola (AEA) - Instituto Agronômico do Paraná (IAPAR), hevandro@iapar.br.
[2] Professor doutor, Departamento de Ciência do Solo e Engenharia Agrícola; Laboratório de Mecanização Agrícola (Lama) – Universidade Estadual de Ponta Grossa (UEPG), lama1@uepg.br.
[3] Professora doutora, Departamento de Química; Laboratório de Química Ambiental e de Materiais (LabQAM) - Universidade Federal do Paraná (UFPR), nnagata@ufpr.br.
Recebido pelo Conselho Editorial em: 6-6-2011
Aprovado pelo Conselho Editorial em: 25-6-2012

INTRODUCTION

Among the extrinsic factors that affect crop yields, those related to soil as chemical fertility, position in the landscape and granulometry, show great variability. This variability and their interactions provide different bearer capabilities, and when subjected to the same farm management result in productivity variability. The use of so-called "information technology" in farming or precision agriculture, has become an important tool, optimizing the process of decision making in agricultural production (ORTEGA & SANTIBÁÑEZ, 2007).

The commercial techniques of use of machines to the reposition of chemical fertility in a variable rate do not consider the soil bearing capacity. Hence the division of farmland into subareas or management zones based on variables of importance for the factors of production must consider the soil formation, providing an agronomically, environmentally and economically rational management.

The recognition of patterns in attributes variability and soil and plant properties is crucial to the viability of different and localized management practices, mainly to an approach for crop management zones or classes (SANTOS et al., 2003). Understanding and classifying the variability of soil properties and landscape and its effects on crop productivity are important components for localized management of agricultural systems (JIANG & THELEN, 2004). However, as there is a multitude of variables that generate variability in crop yields, the most used techniques in data analysis for specific management site are the geomathematical and geostatistical models. These models, by themselves, cannot explain the changes in yield, because they do not allow the simultaneous analysis of a set of variables, given their interactions. The current concern is to study algorithms to identify patterns and magnitudes of variations (GUASTAFERRO et al., 2010).

There are several statistical methods that allow analysis of a set of variables and their relations with certain phenomena (HAIR JUNIOR et al., 2005). Termed as multivariate analysis, it aims to "summarize" a set of variables, considering the variance of the data set (MORRISON, 1976). With the improvement of the computing machines, such techniques have increased its use. In agriculture circle, because of the complexity of soil and plants, some papers use these studies to correlate with soil quality and/or productivity of crops (XIN-ZHONG et al., 2009 and JARADAT, 2009). YE & WRIGHT (2010) correlated the soil use with changes of chemical attributes.

The aim of this study was to use two qualitative techniques or exploratory multivariate analysis with some soil variables already identified in univariate analyzes, and importance of crop yields, to define management zones, as well as identify the most important for each zone.

MATERIAL AND METHODS

The study was conducted by the working group of the Laboratory of Agricultural Mechanization of the State University of Ponta Grossa (Lama/UEPG) in the area of commercial farming in the region of Ponta Grossa, in state of Paraná (PR), Brazil, with approximate coordinates of 25°05'S and 50°20'W and an average altitude of 950m.

Initially it was delimited regular mesh of 60 equidistant points of 40m, arranged in four transects of 15 points, approximately 9.6ha. In the study area the soil was classified as Dystrophic RED YELLOW LATOSOL. The field used had a total area of approximately 22ha, it was more than ten years under the Tillage System, the crop rotation being used in recent years were: soybean (*Glycine max*), triticale (*X Triticosecale* Widdmark), soybean, bristle oat (*Avena strigosa*), the experiment was performed in the new summer crop of maize (*Zea mays*). In the case of the bristle oat crop, the grain was not harvested, as it was handled mechanically (knife-roll) when in the filling phase, in the dough stage. For commercial, soybean, corn and triticale crops, it was performed chemical fertilization, according to routine chemical analyzes of soil and expected productivity. In all cases these were made up expecting high yields, common to the region.

At each point it was carried out the soil sampling (five subsamples), with Dutch auger at a depth of 0 to 20cm and subsequent chemical and physical routine analysis. Regarding the chemical one, it was determined the variables: pH, sum of hydrogen and aluminum (HAl), calcium (Ca), potassium (K), phosphorus (P), carbon (C), cation exchange capacity (CEC), base saturation (V%), the sum of calcium and magnesium (CaMg) calcium/magnesium ratio ($CaMg^{-1}$) and calcium, magnesium and potassium ratio (K basis). For the physical part, the granulometry was scaled (sand, silt and clay).

By sampling with Dutch auger, it was scaled the depth of horizon "A" (PHA) and the depth of the horizon "A" over the transition "AB" (PTHA). Besides these characteristics, topographical indicators were estimated for points, as the relative quota (Quota) to point 60, and topographic gradient, which would be the landscape positions relationship. The summary of the descriptive analysis of the variables are in Table 1.

TABLE 1. Descriptive analysis of variables.

Soil variables	Average	Deviation	Minimum	Maximum	CV (%)
sand (g kg^{-1})	356.65	72.55	213.00	538.00	20.34
silt (g kg^{-1})	185.52	37.42	80.00	307.00	20.17
clay (g kg^{-1})	457.83	59.66	320.00	560.00	13.03
PHA (cm)	19.02	2.35	15.00	25.00	12.34
PTHA (cm)	40.63	10.57	20.00	63.00	26.02
Quota (mm)	883.90	510.80	0.00	1949.00	57.79
Topographic gradient (%)	2.57	1.87	-2.68	6.73	72.81
pH	5.38	0.37	4.60	6.20	6.93
HAl ($cmol_C$ dm^{-3})	5.18	1.22	3.18	9.01	23.58
CaMg ($cmol_C$ dm^{-3})	7.88	1.17	5.50	10.50	14.89
Ca ($cmol_C dm^{-3}$)	4.81	0.78	3.20	6.30	16.22
K ($cmol_C dm^{-3}$)	0.22	0.11	0.08	0.55	50.36
P (mg dm^{-3})	6.09	2.35	1.80	13.10	38.53
C (g dm^{-3})	21.13	2.10	17.00	27.00	9.95
CTC ($cmol_C dm^{-3}$)	13.28	0.96	10.84	15.61	7.20
V %	61.10	8.14	39.00	77.00	13.32
Ca Mg^{-1}	1.58	0.18	1.00	2.10	11.52
K basis	44.95	22.39	14.70	101.10	49.81

With these values it was used two multivariate exploratory analyzes, the hierarchical cluster analysis (HCA) and principal component analysis (PCA).

The HCA model was used to define the number of groups of samples or management zones. This analysis, according to HAIR JUNIOR et al. (2005), groups the data by the Euclidean distance between samples, i.e., in "clusters" with similar attributes.

The PCA technique was then applied to identify the most important variables within each group found, because according to JIANG & THELEN (2004), it allows to separate in decreasing order of amount of variance that each component explains the total variance of the samples, thus to determine the intrinsic dimensionality of the set.

For both techniques it was carried out the pre-processing of data by self-scaling. This pre-processing was applied to compensate for the discrepancies between the magnitudes of the numerical values of the variables, avoiding minimize the effect of the lower magnitude variables, a factor that may influence the importance given to each variable in the analysis. The clustering technique or connection used was the incremental one, which improves the separation of close groups.

After the separation of the samples into groups with similar attributes for HCA, they were spatialized, defining the management zones, and, in these, the most important variables were defined by PCA. Later, there was again a descriptive analysis of the values for each management area, where it was gotten quantitative result that can be used for decision making.

In data spatialization and preparation of maps, for the graphic arrangement of the management zones it was used the interpolation by the method of the nearest neighbor and, for the representation of relief, the minimum curvature method.

RESULTS AND DISCUSSION

In the dendrogram obtained with the use of samples of HCA, it was separated five groups or management zones (Figure 1). This degree of similarity (0.53) and number of clusters was chosen because the area of study was relatively small and a larger number of zones, however constituted groups with a larger similarity, could prevent the management. Studying two agricultural areas and considering similar variables plus the yield variable, FROGBROOK & OLIVER (2007) used similarity of 0.60 and 0.53. The decision of the number of management zones becomes a case study, where the magnitude and spatial statistical aspects are crucial.

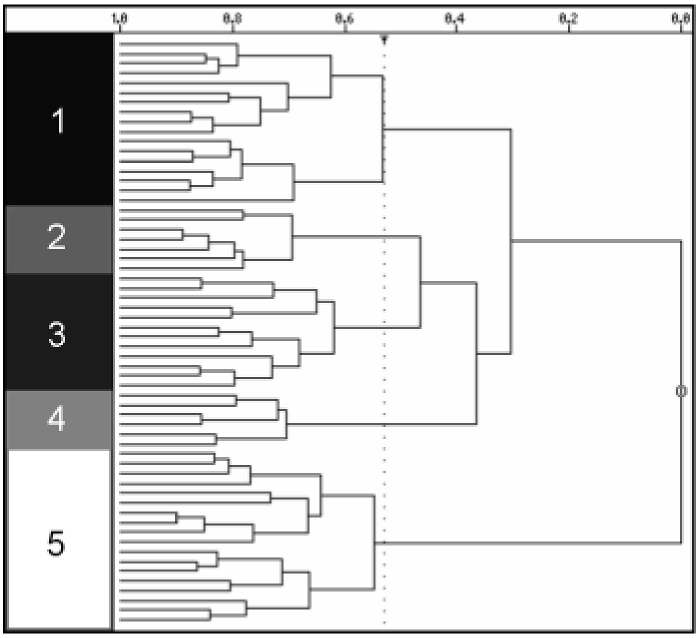
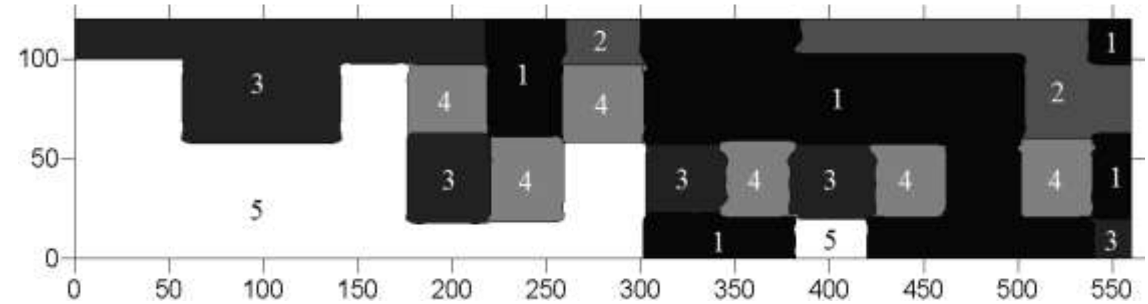
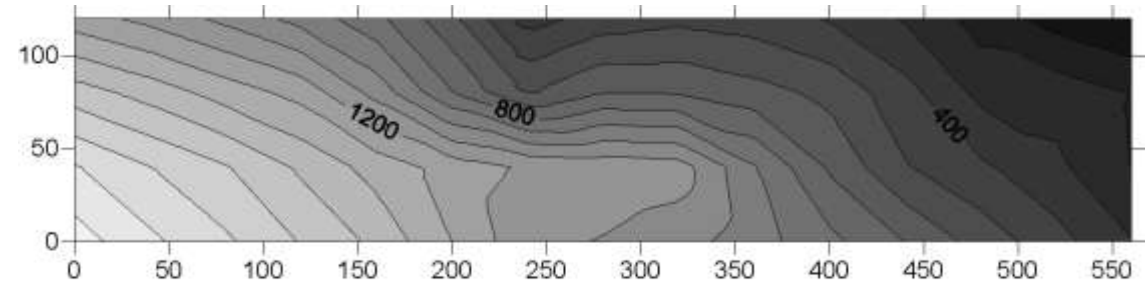


FIGURE 1. Dendrogram of the samples divided in five groups (similarity of 0.53).

With the groups, the defined handling zones were spatialized, where zone 1 corresponds to an area of 2.87ha, zone 2 with 0.8ha, 3 with 1.84ha, 4 with 1.33ha and zone 5 with 2.76ha. When comparing the map of management zones, as analysis, to relative isolines, it is possible to see a tendency of correlation. It visually explains the existence of influence of landscape position on the variability of the attributes studied (Figure 2).



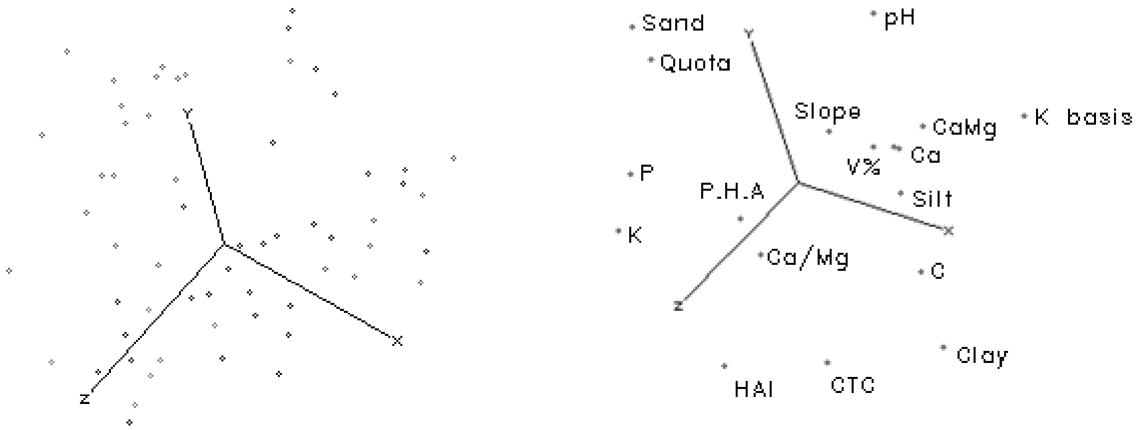
(a)



(b)

FIGURE 2. Spatial distribution of management zones (a) and relative isolines map (mm) (b).

In order to find out which are the most important variables within each management zone, it was performed a PCA analysis. This procedure allows visual differentiation of the groups, and allows the comparison of the scores graphs (representative of the samples) with the loadings (representatives of the variables), facilitating the identification of variables (Figure 3). To delineate the appropriate management zones, it is necessary to identify the most important variables within each zone, aiming to explain or intervene in the phenomenon of agricultural productivity (ORTEGA & SANTIBÁNHEZ, 2007).



(a) (b)

FIGURE 3. Arrangement of the principal components with the same angle to the "scores" (samples) (a) and to the "loadings" (variables) (b).

In the PCA analysis, the first component explained 29.2%, the first two components explained 46.1%, the first three principal components can explained 58.0%, the first four components explained 65.3% of the variances of the studied attributes. Thus, where a smaller number of components is better, however they should have statistical significance, so it was chosen to work with the first three components.

Relating the graphs of scores and loadings of PCA and evaluating the same for both main components, as shown in Figure 4, it is observed that for the group of samples comprising the first management area, the most important variables in the formation of this group were the pH, Ca, CaMg, V%, Ca Mg^{-1} and depth of horizon A. For the samples that form the management area 2, the most important variables are clay, CTC, carbon, K basis and silt. For the management zone 4, the most important variables were hydrogen and aluminum (HAl) and the slope. For the management zone 5, the most important variables are sand and Quota. The management area 3 is comprised of samples that are average values, and, for this reason, it appears on all scores graphs for overlap with other groups. In this case, as a way to improve operational efficiency, in a first year of adoption of management zones, the study could be performed to join the management zone 3 with some of its neighboring.

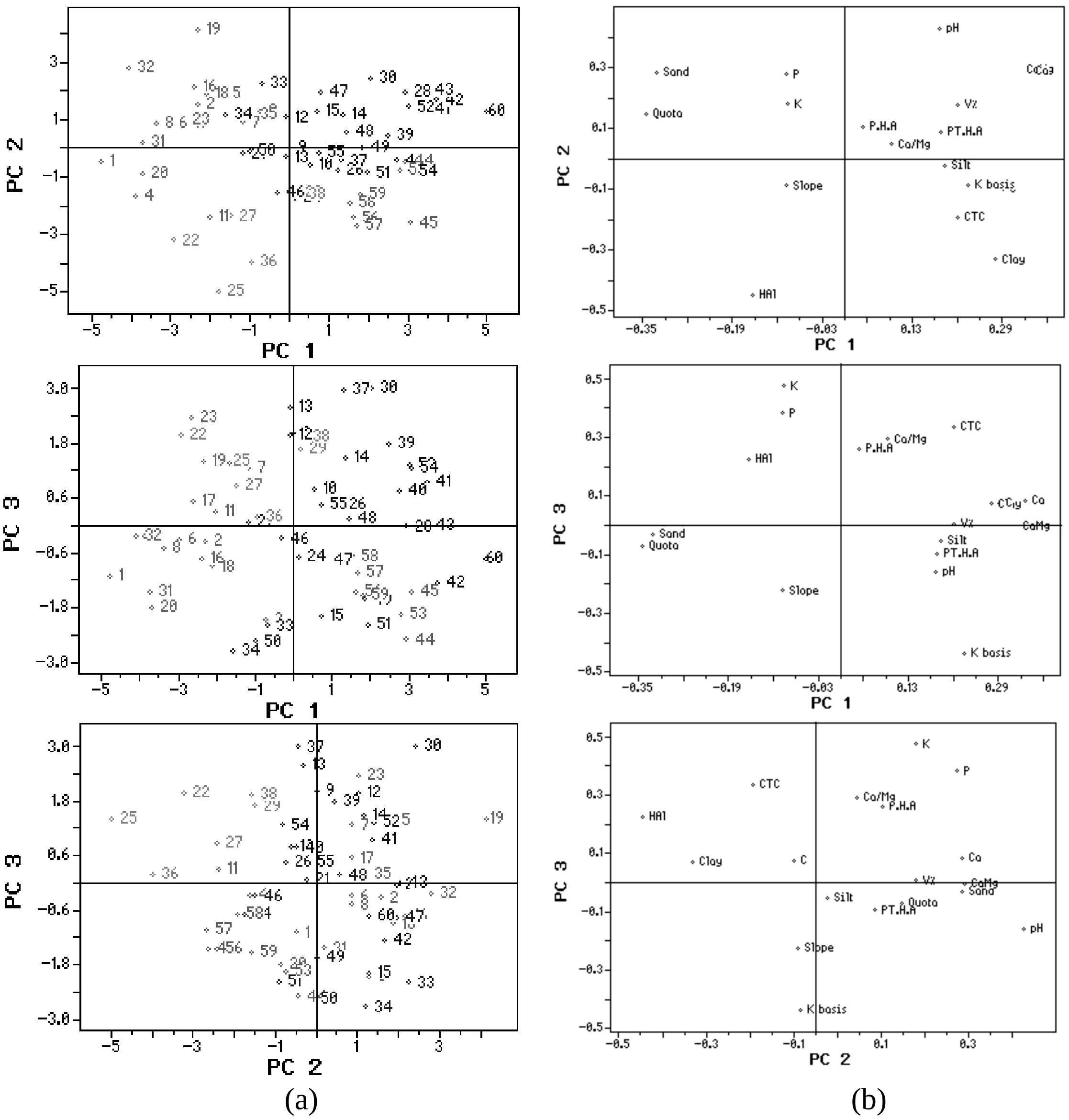


FIGURE 4. Differentiation model of the “scores” (a) and “loadings” (b).

It is important to highlight that in the relationships made between scores and loadings, it should be noted the provision of groups of samples for more than one two-dimensional graphical display, covering at least the possible combinations between the number of components chosen to explain the phenomenon variances. As in this case it was chosen three components, it was used these graphical representations to show the mentioned (Figure 4).

The application of the two methods of multivariate exploratory analysis showed similar results to those found by ORTEGA & SANTIBÁÑEZ (2007), with respect to efficiency in the separation of management zones.

After identifying the most important variables in the formation of management zones, there was new descriptive analysis for each group, yielding quantitative response based on qualitative analyzes generated (Table 2). It was observed that the variables listed as important in forming each group (area) by PCA have the highest averages when compared between them, even presenting, in some cases of coefficients of variation, that may be considered as medium to high.

Therefore, with the management zones mathematically demarcated, it can be performed studies and/or specific agronomic treatments. For example, where the analysis highlights organic matter or carbon (area 2) it is possible to see that, even though this area does not present the highest CTC due to low hydrogen, it has one of the highest averages of effective CTC, thus higher investment correction in chemical fertility can be interesting. Practically, this means that basic fertilization could be performed in the culture of economic interest, which would tend to greater productivity in relation to other management zones.

For the variable carbon, one of the most important regarding chemical, physical and biological fertility of the soil, zone 5 has the lowest average value, and thus a better investment in biomass may be recommended in winter, for example increasing the plant population of the culture of bristle oat. Another situation is the correlation of carbon (C) with relative dimensions. This trend can also be observed with respect to the contents of sand, where areas with larger quotas tend to have higher contents.

TABLE 2. Average and coefficient of variation of the variables for each management zone.

Soil variables	Management zones									
	1		2		3		4		5	
	A	CV	A	CV	A	CV	A	CV	A	CV
sand (g kg^{-1})	317.2	11	268.9	13	363.7	12	303.3	7	**441.2**	11
silt (g kg^{-1})	203.4	14	**216.9**	23	169.7	23	183.3	5	167.7	18
clay (g kg^{-1})	479.4	6	**514.3**	5	466.7	12	513.3	4	391.1	9
PHA (cm)	**20.0**	7	17.9	14	19.1	18	19.2	8	18.4	12
PTHA (cm)	43.6	20	40.4	18	**51.4**	19	32.3	32	33.5	17
Quota (mm)	517.2	49	256.7	52	993.0	32	738.2	41	**1450**	21
Topographic gradient (%)	2.0	184	-2.0	410	0.1	7127	**-3.2**	495	0.7	1134
pH	**5.5**	7	5.4	5	5.4	6	4.9	5	5.4	6
HAl (cmol$_C$ dm^{-3})	4.6	20	5.1	13	4.9	20	**7.4**	16	5.2	19
CaMg (cmol$_C$ dm^{-3})	**8.7**	11	8.0	7	8.3	7	6.4	15	7.1	13
Ca (cmol$_C$dm^{-3})	**5.6**	9	4.7	8	5.0	10	3.9	16	4.3	14
K (cmol$_C$dm^{-3})	0.27	43	0.10	16	0.15	21	0.25	38	0.25	46
P (mg dm^{-3})	6.8	27	3.7	17	4.7	18	5.7	48	7.4	36
C (g dm^{-3})	22.5	9	**23.0**	9	20.6	9	20.3	4	19.7	7
CTC (cmol$_C$dm^{-3})	13.8	4	13.2	3	13.3	7	14.1	4	12.6	8
V %	**66.6**	10	61.3	7	63.5	8	47.7	15	58.7	11
Ca Mg^{-1}	**1.7**	9	1.5	12	1.5	10	1.5	10	1.6	11
K basis	40.4	53	**78.6**	19	57.1	25	28.3	40	33.6	44

A – Average; VC – Coefficient of variation (%)

In the case of the area 4, it identifies the need for correction of acidity, thus raising the pH and decreasing the saturation by HAl, therefore increasing V%. This area also has the highest values of topographic gradient, maybe differentiated management of mechanization at the time of distribution, for example, without incorporating corrective movements could prevent unnecessary runoff and leaching, among others.

With table 2, it could be suggested countless different managements, as well as tested. This fact shows that the union of multivariate analysis with simpler analysis becomes more interesting. The multivariate statistical allows holistic analysis of the phenomena, considering multidimensional models where the interactions between variables are statistically estimated and considered for the position and/or importance of these. This fact is not possible when using only the three-dimensional models, as when only geostatistics is applied. And the simple analyzes applied on models facilitate and reduce errors of interpretation of the final results.

CONCLUSION

It was possible to perform the separation into management zones from soil variables employing multivariate statistical analysis by PCA and HCA. With the qualitative result of these analyzes, the application of quantitative analysis allowed the simple choice of the most significant variables in each one of the classified areas.

REFERENCES

FROGBROOK, Z. L.; OLIVER, M. A. Identifying management zones in agricultural fields using spatially constrained classification of soil and ancillary data. *Soil Use and Management*, Oxford, v.23, p.40–51, 2007.

GUASTAFERRO, F.; CASTRIGNANO, A.; DE BENEDETTO, D.; SOLLITTO, D.; TROCCOLI , A.; CAFARELLI, B. A comparison of different algorithms for the delineation of management zones. *Precision agriculture*, Dordrecht, v.11, n.6, p.600-620, 2010.

HAIR JUNIOR, F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. *Análise multivariada de dados.* 5. ed. Porto Alegre: Bookman, 2005. 593p.

JARADAT, A. A. Modeling biomass allocation and grain yield in bread and durum wheat under abiotic stress. *Australian Journal of Crop Science,* v.3, n.5, p.237-248, 2009.

JIANG, P.; THELEN, K. D. Effect of soil and topographic properties on crop yield in a north-central corn-soybean cropping system. *Agronomy Journal*, Madison, v.96, p.252-258, 2004.

MORRISON, D. E. *Multivariate statistical methods*. New York: McGrw-Hill, 1976. 338p.

ORTEGA, R.A.; SANTIBÁÑEZ, O.A.; Determination of management zones in corn (*Zea mays* L.) based on soil fertility. *Computers and Electronics in Agriculture,* New York, v.58, n.1, p.49–59, 2007.

SANTOS, A. O.; PALLONE FILHO, W. J.; UNGARO, M. R. G.; LINO, A. C. L.; RADIN, B. Prospecção de zonas potenciais para manejo diferenciado em agricultura de precisão utilizando-se padrões de solo-planta-clima1. *Revista Brasileira de Engenharia Agrícola e Ambiental*, Campina Grande, v.7, n.3, p.463-468, 2003.

XIN-ZHONG**,** W**.;** GUO-SHUN**,** L**.;** HONG-CHAO**,** H.; ZHEN-HAI**,** W.; QING-HUA**,** L.; XU-FENG**,** L.;WEI-HONG**,**H.;YAN-TAO**,** L. Determination of management zones for a tobacco field based on soil fertility. *Computers and Electronics in Agriculture,* New York, v. 65, n.2, p.168-175, 2009.

YE, R.; WRIGHT, A. L. Multivariate analysis of chemical and microbial propertiesin histosols as influenced by land-use types. *Soil & Tillage Research,* Amsterdam, v.110, p.94–100, 2010.