Defining management zones for a papaya culture considering soil chemical attributes by fuzzy c-means

Fabricia Benda de Oliveira[1]* [ID]**; Julião Soares de Souza Lima[2]** [ID]**;**
Carlos Henrique Rodrigues de Oliveira[3] [ID]**; Walas Permanhane Sturião[4]** [ID]**;**
Jacyelli Sgranci Angelos[5] [ID]

[1]Departamento de Geologia. Universidade Federal do Espírito Santo (UFES), Alto Universitário, s/n, Caixa Postal 16, CEP: 29500-000, Alegre, ES, Brazil.
[2]Departamento de Engenharia Rural. Universidade Federal do Espírito Santo (UFES), Alto Universitário, s/n, Caixa Postal 16, CEP: 29500-000, Alegre, ES, Brazil. E-mail: limajss@yahoo.com.br
[3]Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo (IFES), Rodovia ES-482 (Cachoeiro-Alegre), Km 72, CEP: 29500-000, Alegre, ES, Brazil. E-mail: carlos.oliveira@ifes.edu.br
[4]Departamento de Agronomia. Universidade Federal de Viçosa (UFV), Avenida Peter Henry Rolfs, s/n, CEP: 36570-000, Viçosa, MG, Brazil. E-mail: wsturiao@gmail.com
[5]Programa de Pós-graduação em Agroquímica. Universidade Federal do Espírito Santo (UFES), Alto Universitário, s/n, Caixa Postal 16, CEP: 29500-000, Alegre, ES, Brazil. E-mail: jacysgranci@gmail.com
*Corresponding author. E-mail: fabricia.oliveira@ufes.br

ABSTRACT

In precision agriculture, one of the major problems is the need for a dense sampling grid. A solution to this problem has been the generation of management zones and the correlation between them. This work aims at defining management zones for a papaya culture using the fuzzy C-means grouping method, based on the following soil attributes: phosphorus, calcium, magnesium and potassium; and base saturation. The work was conducted in an approximately 1.2 ha area planted with a commercial culture of *Carica papaya* L. cultivated in typical Argisol found on the coastal terraces, in the municipality of São Mateus, ES, Brazil. The management zones were defined by the proposed method, based on the following analyses: soil chemical parameters P, Ca, Mg, and base saturation (V), analyzed together; soil chemical parameters P, Ca, Mg, and V, analyzed separately; and, the following combinations of soil chemical attributes Ca-Mg, Ca-V, Ca-Mg-V, P-Ca, P-Mg, P-V, P-Ca-Mg, P-Ca-V, Mg-V, and P-Ca-Mg-V. The generated management zones facilitated an understanding of the spatial variability of the attributes in the study area and the observed similarities between management zones generated from different attributes. In addition, the data were influenced especially by the P and V% attributes.

Keywords: *Carica papaya* L, geostatistics, grouping analysis, precision agriculture.

Definição de zonas de manejo para a cultura do mamoeiro considerando atributos químicos do solo por meio de c-means fuzzy

RESUMO

Na agricultura de precisão, um dos maiores problemas é a necessidade de uma grade amostral densa. Uma solução para este impasse tem sido a geração de zonas de gestão e a

correlação entre elas. Este trabalho tem como objetivo definir zonas de manejo para a cultura do mamoeiro utilizando o método de agrupamento fuzzy C-means, baseado nos seguintes atributos do solo: fósforo, cálcio, magnésio e potássio; e saturação de bases. O trabalho foi conduzido em uma área de aproximadamente 1,2 ha plantada com cultura comercial de Carica papaya L. cultivada em Argissolo típico encontrado em terraços costeiros, no município de São Mateus, ES, Brasil. As zonas de manejo foram definidas pelo método proposto, com base nas seguintes análises: parâmetros químicos do solo P, Ca, Mg e saturação por bases (V), analisados em conjunto; parâmetros químicos do solo P, Ca, Mg e V, analisados separadamente; e, as seguintes combinações de atributos químicos do solo Ca-Mg, Ca-V, Ca-Mg-V, P-Ca, P-Mg, P-V, P-Ca-Mg, P-Ca-V, Mg-V, e P-Ca-Mg-V. As zonas de manejo geradas permitiram explicar a variabilidade espacial dos atributos na área de estudo e as semelhanças observadas entre as zonas de manejo geradas a partir de diferentes atributos, além disso, os dados foram influenciados especialmente pelos atributos P e V%.

Palavras-chave: agricultura de precisão, análise de agrupamento, *Carica papaya* L, geoestatítica.

1. INTRODUCTION

Agriculture has been undergoing major technological transformations, among them the use of precision agriculture, which can benefit the crop since it identifies areas with potential for producing better quality fruits and assists in understanding the factors determining that. It provides careful and detailed soil and crop management, recognizing differences in soils and plants, allowing adequate management practices for each planting area (Saiz-Rubio and Rovira-Más, 2020). Understanding the spatial distribution of soil attributes is the first step to establish adequate management practices to optimize agricultural productivity and minimize environmental damage. Geostatistics enables us to verify spatial dependence and visualize the attributes of the studied areas (Rodrigues *et al.*, 2012; Delalibera *et al.*, 2012).

In order to determine different management practices with variable rates of input, it is necessary to generate management zones that aim to reduce sampling, resulting in maps representative of the area, with a combination of factors limiting productivity and quality, to which a uniform dose of inputs can be applied (Ramos *et al.*, 2017).

Data-grouping techniques that allow using a set of prominent factors for crop development and for identifying crop variabilities have been used to generate these management areas. Among these mathematical methods, fuzzy C-means for the use of a set of attributes correlated with the nutritional variability of the study area (Rodrigues Junior *et al.*, 2011; Alves *et al.*, 2013).

Rodrigues Junior *et al.* (2011) defined management zones in a 2.1 ha area cultivated with Coffea arabica L. cv. Catuaí, in Paula Cândido, MG, based on foliar analysis performed with a chlorophyll sensor using the C-means and fuzzy C-means methods, and concluded that both grouping methods had similar results regarding management areas. Alves *et al.* (2013) established management zones in no-tillage corn and soybean areas based on the spatial variability of soil apparent electrical conductivity and organic matter, using fuzzy C-means. The best results for defining the management zones according to soil texture attributes were obtained from organic matter or electrical conductivity maps and, for the chemical attributes, from organic matter or altitude and organic matter maps.

Taylor *et al.* (2003) used soil productivity and electromagnetic induction data for defining management zones by the C-means algorithm and reported on obtaining coinciding zones. They also proposed a more targeted sampling method for reducing the number of soil samples from 27 to 4.

The essence of the fuzzy methodology is the acceptance of uncertainty that must be

modeled and treated by mathematical methods and was introduced by Zadeh (1965). It has been used in several study areas due to its ability to detect similarities among individual data from a dataset to aid in the decision-making process. The fuzzy set expresses the degree to which an element belongs to a set using a pertinence function.

Thus, considering that management zones guide area sampling, the spatial distribution of soil and plant attributes, and the varying dosage of inputs, the correlation of the management zones with soil chemical attributes can lead to a better understanding of the spatial distribution of soil attributes and sampling optimization.

Therefore, the objective of this work was the preliminary mapping of management zones in an area planted with a papaya crop based on the P, K, Ca, Mg, and base saturation (V%) chemical attributes determined by chemical analysis of soil samples, using the Fuzzy C-means (FCM) method.

2. MATERIAL AND METHODS

The data used in this work originated from the experimental work for the master thesis conducted by Sturião *et al.* (2018) at the Vovô Délio farm, in São Mateus, ES, located at the approximate coordinates E = 407241 m and N = 7905948 m, with 40 m average altitude, and spindle 24. According to Embrapa (2013), the soil in the area is typical of the coastal terraces of this Brazilian region, characterized as Argissolo Amarelo Distrófico of low activity, with smooth undulating relief, and approximately 3.0% slope.

The largest extensions of coastal terraces in the country are found in the Brazilian states of Bahia and Espírito Santo (Jacomine, 1996), which are responsible for more than 80% of the papaya produced. However, the presence of cohesive layers in the soils acts as a counterpoint to some of the main potentialities (Cintra, 2005), meaning that it is important to select more suitable areas and adjust management practices to soils with cohesive horizons subsurface to guarantee the longevity and productivity of the papaya (Salgado and Costa, 2003). These soils generally have a very dense cohesive B horizon, due to the presence of a higher content of dispersed microclay (Corrêa *et al.*, 2008; Lima Neto *et al.*, 2010).

The cohesive sandy texture soil was conventionally prepared and cultivated with papaya of the Golden THB variety. After plant sexing, soil sampling was performed. The deformed samples were collected using an auger at 0 - 0.20 m depth, in the crown projection, from two points equidistant approximately 0.20 m, in the fertilizer application point in the planting line direction. The two samples were placed in plastic bags, forming a single sample per sampling point, identified and sealed.

Soil samples were collected from a 1.2 ha plot, in which a sampling point of approximately 5.0 m² was defined, totaling 129 sample points (Figure 1). All samples were georeferenced using a GPS receiver (TechGeo®, Model GTR G2 geodesic) and post-processed.

The chemical variables phosphorus (P), calcium (Ca), magnesium (Mg), and potassium (K) from soil samples collected between 0 - 0.20 m depth were analyzed and the base saturation (V%) value was calculated. The data analysis first step was to perform a descriptive statistical analysis of the stored data to evaluate the distribution of the variables throughout the experiment by calculating the position and dispersion parameters such as mean, median, maximum and minimum values, upper and lower quartiles, standard deviation, and the coefficients of variation, asymmetry, and kurtosis.

Subsequently, the Pearson correlation ($p < 0.05$) coefficients were determined for soil chemical attributes following the classification of Sturião *et al.* (2018): extra-high ($80\% \leq r \leq 100\%$); high ($60\% \leq r \leq 80\%$); moderate ($40\% \leq r \leq 60\%$); low ($20\% \leq r \leq 40\%$); and zero ($0\% \leq r \leq 20\%$) (likewise for negative r values).

The data underwent geostatistical analysis to check for spatial dependence and, when present, quantify the degree, using the adjustment of Matheron classical semivariogram (Félix

et al., 2016). For each variable, the best fit for the experimental model (Santos *et al.,* 2019) among the semivariogram models was determined by geostatistical analysis. The Vesper® software was used and the model was chosen based on the lowest residual sum of the squares (RSS) and the highest coefficient of determination (R^2). Additionally, the spatial dependency index (SDI) was calculated by $[C1/(C0 + C1)]*100$, where C1 is the structural variance, and $C0 + C1$ is the threshold. The SDI was classified as follows: weak (SDI $\leq 25\%$), moderate ($25\% <$ SDI $\leq 75\%$) and strong (SDI $> 75\%$) (Silva *et al.*, 2013).

The spatial dependence was verified for all variables, and the spatial variability map was elaborated using kriging interpolation, which according to Ferreira *et al.* (2017), is a set of regression analysis techniques that minimizes the estimated variance based on a previous model, taking into account the stochastic dependence between data spatial distribution.

The spatial variability maps were generated using interpolations with a 0.5 x 0.5 m pixel size for the coordinates E and N and the same contour polygon, defined in Figure 1. This standardization guarantees that all generated maps have the same number of pixels and can overlap.
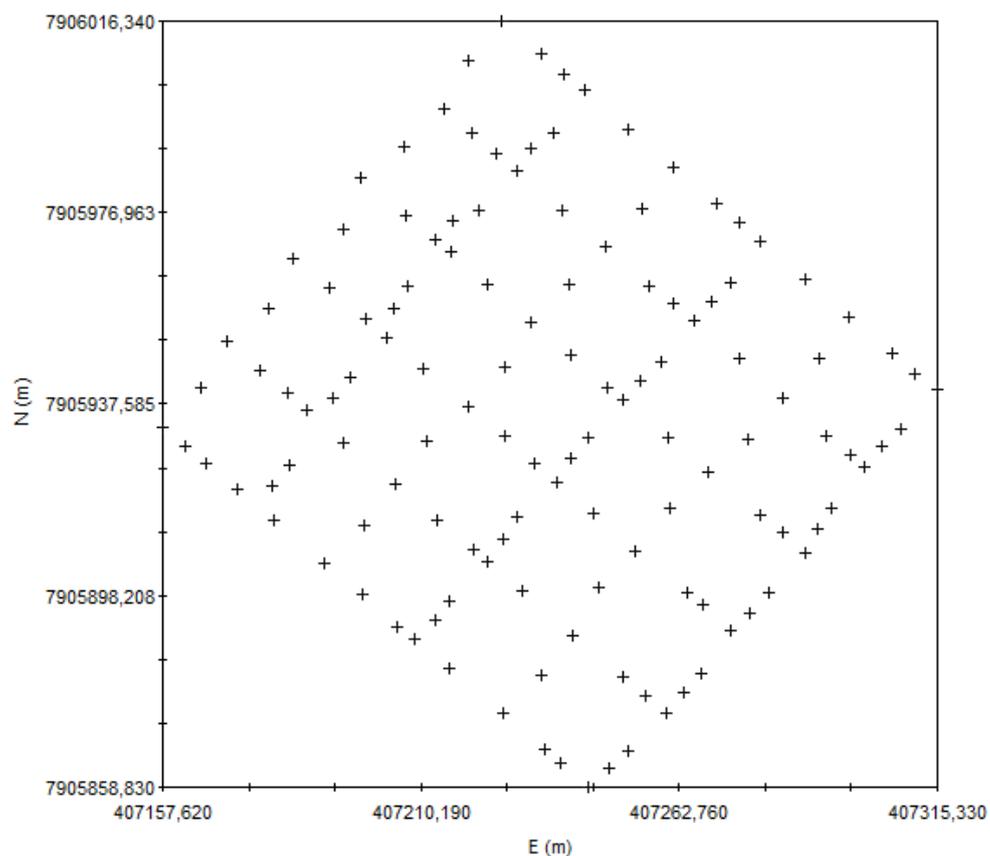
Figure 1. Sample mesh elaborated for soil sampling to determine its attributes.

According to Linden (2009), grouping analysis, or clustering, has the purpose of separating objects into groups, based on the characteristics that the objects have. The basic idea is to place objects that are similar according to some predetermined criteria in the same group. There is the possibility of generating the grouping, considering the relevance of each value associated with each class, with diffuse clustering being recommended for this use (Yang, 1993; Bezdek and Pal, 1992). This approach allows an individual to be associated with all clusters using a membership function (Zadeh, 1965). Thus, fuzzy sets offer the advantage of manifesting each individual partially to all groups. Among fuzzy algorithms, the most popular is Fuzzy C-means (FCM). This approach to FCM classification minimizes the within-class sum of squares.

In the study of management zones, among the analyzed variables, those that presented spatial dependence were analyzed by the FCM algorithm of the FuzMe® software. The FCM algorithm was applied to group the data to minimize intraclass variance and maximize interclass variance, thus minimizing the objective function Jm of a characteristic function μij (μijª[0,1]) and a class center ci.

This method was used to divide the data groups used to determine the management zones into two and three classes (Ricardo *et al.*, 2016). The objective function used by FCM is described as Equation 1:

$$\text{Minimize J}_m = \sum_{i=1}^{k} \sum_{j=1}^{n} \left(\mu_{ij}\right)^m d^2\left(x_j, c_i\right) \tag{1}$$

Where:

$$d^2\left(x_j, c_i\right) = \|x_j - c_i\|;$$

m is a classification index of the fuzzy degree.

The variable m > 1 defines the allowed distance between the considered point and the centroid being calculated. Thus, the greater the value, the more elements are assigned to this group (Rodrigues Junior *et al.*, 2011).

A preliminary test with five different m values (1.1, 1.3, 2, 5, and 10) was carried out. The results showed that the defined management zones remained constant with the changing m values, as already observed by Rodrigues Junior *et al.* (2011) and, therefore, m = 1.1 was adopted.

The iteration was recalculated until the difference between Jm actual and Jm anterior was equal to or less than 0.0001. Ten random initiations were made with different initial centroids.

The FCM algorithm was implemented following the steps:

1) Initialize the values of the elements μij of the membership matrix with random numbers between 0 and 1, satisfying the conditions described in Equations 1;

2) Calculate the centroids ci (i = 1, 2,3 ... c);

3) Update the pertinence matrix values;

4) If the number of maximum iterations initially defined has been reached, then the algorithm ends, otherwise, return to Step 2.

Further, the Euclidean distance was used for calculating the distance d between the i-th center of the class and the j-th point of the sample. To optimize the number of classes, the value of two indexes, the Fuzzy Performance Index (FPI) and Modified Partition Entropy (MPE), were determined according to Equations 2 and 3 (Metwally *et al.*, 2019), respectively.

$$\text{FPI} = 1 - \frac{c}{c-1}\left[1 - \frac{\sum_{j=1}^{n} \sum_{i=1}^{c} (\mu_{ij})^2}{n}\right] \tag{2}$$

$$\text{MPE} = \left[\frac{\sum_{j=1}^{n} \sum_{i=1}^{c} \mu_{ij}.\log\mu_{ij}}{n.\log(c)}\right] \tag{3}$$

After delimiting the groups by the FCM method, the ArcGis® 10.2 software was used for spatializing the data and generating thematic maps. A visual analysis of the maps generated allowed us to choose the maps with the most similar management zones for the different soil attributes used.

IPABHi

3. RESULTS

3.1. Descriptive and geostatistical analysis

Table 1 shows the descriptive statistics of the following soil attributes P, Ca, Mg, K and V% and their range (SR) for papaya culture.

The Pearson correlation coefficients (p < 0.05) between soil chemical attributes (Table 2) were null ($0.0 \leq r \leq 20.0\%$) for "P and Mg", "P and K", "P and V", "K and Ca", "V and K"; low ($20.0 \leq r \leq 40.0\%$) for "Ca and P", "K and Mg", and moderate ($40.0 \leq r \leq 60.0\%$) for "Mg and Ca", "V and Ca", "V and Mg". Further, the soil base saturation (V) indicates how many of the negative charges are occupied by Ca^{2+}, Mg^{2+} and K^+, in relation to the exchange points of the acidic cations H^+ and Al^{3+}, being directly correlated with the soil bases.

Table 1. Analysis of descriptive statistics of P, Ca, Mg, K, and V%.

Attribute	Min.	Q1	Md.	Q3	Max.	Mean	s	CV (%)	Ck	Cs	SR
P ($mg\ dm^{-3}$)	2.00	9.00	15.00	27.00	66.0	20.12	14.81	73.61	0.93	1.24	10.00 – 20.00
Ca ($cmol_c\ dm^{-3}$)	0.80	1.40	1.90	2.40	7.80	2.06	0.88	42.75	12.86	2.35	1.60 – 4.00
Mg ($cmol_c\ dm^{-3}$)	0.20	0.40	0.60	0.70	1.30	0.59	0.19	32.86	0.88	0.59	0.60 – 1.00
K ($mg\ dm^{-3}$)	20.00	50.00	60.00	80.00	140.00	67.44	24.15	35.80	-0.05	0.37	30.00 – 60.00
V (%)	46.97	63.60	71.04	77.70	94.71	70.69	10.09	14.27	-0.20	-0.09	70.00 – 80.00

Min.: minimum value; Q1 = first quartile; Md = median; Q3 = third quartile; Max = maximum value; s = standard deviation; CV = coefficient of variation; Ck = kurtosis coefficient; Cs = coefficient of asymmetry; SR = Range, adapted from Costa and Costa (2003) and Prezotti *et al.* (2007).

Table 2. Pearson correlation coefficient (p<0.05), %, for chemical variables of soil sampled at 0-0.20 m depth.

	P	Ca	Mg	K	V
P	1				
Ca	22*	1			
Mg	11	48*	1		
K	16	8	31*	1	
V	13	41*	49*	13	1

P: Phosphorus; Ca: Calcium; Mg: Magnesium, K: Potassium; and V: base saturation.
* Significant at 5%.

According to Sturião (2018), the soil sampling location affects the correlations found in the results, mainly due to the sampling location and application and homogenization of fertilizers and the quantitative relationship between attributes.

All attributes were submitted to geostatistical analysis. The parameters of the adjusted semivariograms and the coefficient of determination of the cross-validation for each attribute are shown in Table 3, indicating that K did not show spatial dependence, since the semivariogram presented pure nugget effect (PNE).

Table 3. Semivariogram parameters adjusted for the studied attributes.

Attribute	Model	C_0	C_0+C_1	a (m)	SDI (%)	R_1^2 (%)	R_2^2 (%)
P	EXP	59.40	135.50	78.30	56	58	21
Ca	ESF	0.027	0.91	21.80	97	63.6	24
Mg	ESF	0.01	0.04	22.04	70	99.7	14
K	PNE	-	-	-	-	-	-
V	ESF	4.90	65.66	22.70	93	97.5	20

EXP: theoretical exponential model; ESF: theoretical spherical model; PNE: Pure nugget effect; C0: nugget effect; C0 + C1: threshold; a: scope of spatial dependence; SDI: spatial dependency index [(C1/C0 + C1)*100]; R_1^2: coefficient of determination of semivariogram; and R_2^2: coefficient of determination of cross-validation (significant at p<0.05).

The exponential model was the best fit for the P semivariogram based on the highest coefficient of determination (R_1^2) while the spherical model was the best for the Ca, V and Mg variables, similar to the result reported by Menezes *et al.* (2004), for the latter variable.

3.2. Management areas by the FCM method

Figures 2 and 3 show the management areas generated by the FCM method from the P, Ca, Mg and V attributes and their combinations. Tables 4 and 5 show the mean values of the three data classes determined for all analyzed soil chemical attributes and their combinations.
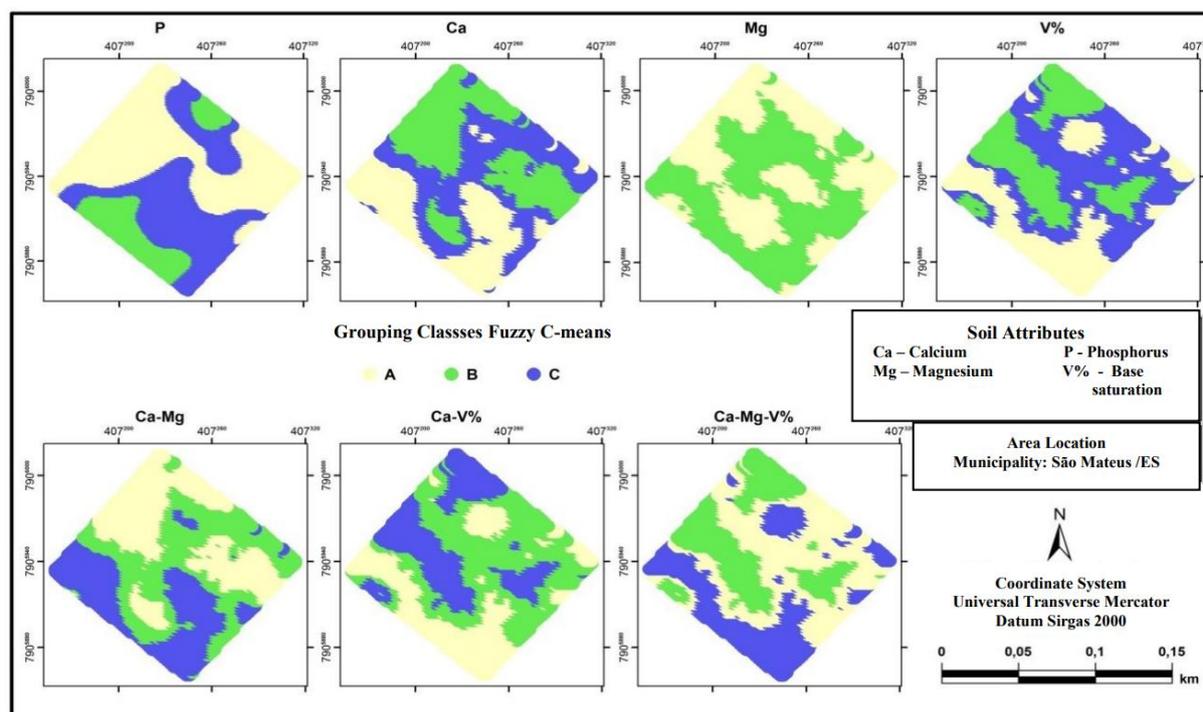
Figure 2. Grouping map according to soil attributes P, Ca, Mg, V%, Ca-Mg, Ca-V%, and Ca-Mg-V%, by the Fuzzy C-means algorithm.
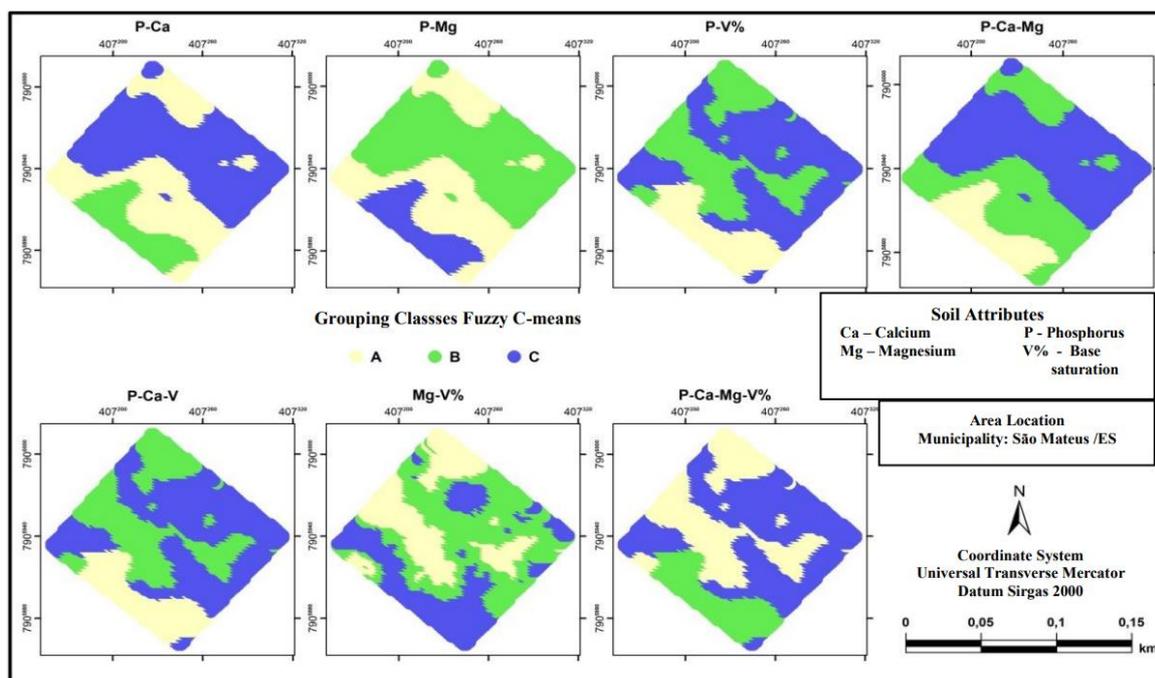
Figure 3. Grouping map according to soil attributes P-Ca, P-Mg, P-V%, P-Ca-Mg, P-Ca-V%, Mg-V%, and P-Ca-Mg-V%, using the Fuzzy C-means algorithm.

Table 4. Means of the three data classes for the P, Ca, Mg and V% attributes, by the FCM algorithm.

Classes	Means			
	P	Ca	Mg	V
3a	16.60	2.45	0.48	76.02
3b	20.20	1.66	0.69	61.98
3c	24.50	2.02	0	69.54

Table 5. Means of the three data classes for the Ca-Mg, Ca-V, P-Ca, P-Mg, P-V, and Mg-V attributes, by the FCM algorithm.

Classes	Means											
	Ca-Mg		Ca-V		P-Ca		P-Mg		P-V		Mg-V	
	Ca	Mg	Ca	V	P	Ca	P	Mg	P	V	Mg	V
3a	1.67	0.50	2.31	76.02	19.42	2.12	19.41	0.54	29.12	75.64	0.47	61.98
3b	2.02	0.58	1.97	69.54	29.70	2.40	13.26	0.57	16.17	62.51	0.56	69.54
3c	2.46	0.65	1.87	61.98	13.27	1.87	29.70	0.64	15.14	71.43	0.70	76.02

4. DISCUSSION

4.1. Descriptive and geostatistical analysis

A general analysis of the data shows that the levels of the attributes in the sampled soil were above the upper level of the ideal range, probably due to the fertilization done in the area, leaving the area with excess nutrients in the soil.

Table 1 shows that the mean and median values, measurements that characterize the center of the frequency distribution, are close for the Ca, Mg, K and V variables, indicating a trend

toward data normal distribution. Thus, we assumed normal data distribution for these variables.

Furthermore, the Mg, K and V data tend to symmetry, which is verified by the asymmetry and kurtosis coefficients. The asymmetry coefficient close to zero characterizes a symmetrical distribution, and the kurtosis coefficient close to zero for P, Mg, K, and V, characterizes a mesokurtic (normal) distribution; however, Ca did not have Cs and Ck near zero, characterizing a distribution that does not tend to normality. Montezano *et al.* (2006) and Montanari *et al.* (2005) studied the variability of chemical attributes and reported no normal distribution for most attributes studied, which contradicts the results found in this study.

Table 1 shows that 60% of the attributes studied present an asymmetric distribution to the right, with a mean greater than the median while the kurtosis coefficient (Ck) indicates platykurtic distribution for 60% of the attributes.

The data variability analyzed by the coefficient of variation (CV) and classified indicates that (Gomes *et al.*, 2018):
• mean CV for Ca, Mg, K and V, totaling 80% of the attributes studied ($12 < CV < 62\%$). Similar results were found by Menezes *et al.* (2004), Silveira *et al.* (2000) and Silva *et al.* (2003), in the 0.0 - 0.2 m depth and by Souza *et al.* (2003) at the 0.6 - 0.8 m depth;
• high CV for P ($CV > 62\%$), representing 20% of the attributes. Cavalcante *et al.* (2007) reported that higher CV is due to fertilization residual effects and application method, usually in the sowing line and without complete homogenization, unlike liming, which was distributed throughout the area.

Menezes *et al.* (2004) studied the spatial variability of pH, Ca, Mg and V% of the soil in different reliefs cultivated with sugarcane in Guariba, SP, and found mean CV values for Ca, Mg and V that corroborate the results of this work. Barbieri *et al.* (2008) investigated the spatial variability of chemical attributes of Argisol regarding variable application rate of inputs in different reliefs and reported a low coefficient of variation only for base saturation in the concave area, at the 0 - 0.20 m depth. The other attributes had average or high coefficients of variation in the two depths studied.

The mean soil fertility was analyzed based on the ideal range for the papaya crop in Espírito Santo (Table 1). Thus, based on these classes, we have mean values above the upper value of the ideal/satisfactory range (SR) for P (20.12 mg.dm^{-3}) and K (67.44 mg.dm^{-3}). The mean value of Mg is below the lower value of the ideal range while the other attributes are within the ideal range.

The greater mobility of P in the soil of the study area is possibly explained by its low clay content in the 0 - 0.20 m depth, while the predominant kaolinite and irrigation favored the diffusion and availability to the absorption by the plants (Sturião, 2018). Also, the samples were collected in the crown projection in the fertilization site.

When we study the geostatistics, spherical and exponential models have been the most suitable to describe the behavior of semivariograms of soil and plant attributes (Menezes *et al.*, 2004; Barbieri *et al.*, 2008; Rodrigues Junior *et al.*, 2011; Alves *et al.*, 2013). Analyzing the Semivariogram, the pure nugget effect (PNE) indicates a lack of spatial dependence for the shortest distance sampled (5.7 m), so K behaves randomly in the considered distance.

The range reflects the spatial continuity of the attribute that at distances greater than the range behaves randomly. Therefore, the range value is desired to be as large as possible, since it conveys greater precision to data interpolation, i.e., greater spatial continuity. The Ca, Mg and V attributes had a range of approximately 22 m and adjusted to the spherical model, allowing us to infer the same spatial distribution pattern because they have close ranges and adjusted to the same theoretical model.

Table 2 shows that the elements Ca and V had strong SDI. According to Silva *et al.* (2013), with moderate and high R_1^2, respectively, but presented low values of R_2^2. All adjustments presented R_1^2 higher than 50%, and 75% had R_1^2 greater than 75%.

IPABH

The cross-validation study identifies areas with no spatial dependence, indicating the need for new samplings (Rodrigues Junior *et al.*, 2011), showing the importance of cross-validation to indicate the robustness of the chosen model to estimate the values in places where no information is available.

From the adjusted semivariograms, the attributes with spatial dependence (P, Ca, Mg, V) were interpolated by the ordinary kriging method to determine their values in non-sampled locations. Subsequently, a visual analysis of the area using the calculated attributes allowed the selection of different management zones/areas according to their specific needs.

Based on the results of R_2^2, all chemical variables (P, Ca, Mg and V) that fit the theoretical semivariograms were used to define the management zones except for K, due to the pure nugget effect presented. Therefore, the management zones were defined using the FuzMe® software and FCM algorithm based on the following soil chemical parameters: P, Ca, Mg, and V, analyzed separately, and the following combinations Ca-Mg, Ca - V, Ca - Mg - V, P - Ca, P - Mg, P - V, P - Ca - Mg, P - Ca - V, Mg - V, P - Ca - Mg -V as well.

The ideal number in each proposed problem was analyzed to determine the number of classes into which the dataset was divided, and consideration was given as to whether the defined classes were technically and economically feasible. According to Máquina *et al.* (2020), the optimal number of classes for each management zone is determined by the minimum values of the Fuzziness Performance Index (FPI) and Modified Partition Entropy (MPE). Thus, the FPI and MPE indexes were calculated for 2 and 3 classes with similar results, but slightly lower FPI and MPE values were obtained for 3 classes. Likewise, Molin and Castro (2008), Yan *et al.* (2008), Xin-Zhong *et al.* (2009) and Morari *et al.* (2009) in similar studies also obtained an optimum number of classes equal to three. Thus, this study recommends using three classes to define management zones.

4.2. Management zones obtained by the FCM method

According to Loisel *et al.* (2018), small management areas are difficult to handle, considering subclasses due to technical and economic limitations. When increasing the number of classes, some pixels remain unchanged, while others are reclassified due to the weak Euclidean proximity with the classes allocated and the classification is independent of the variable coordinates. Pedroso *et al.* (2010) highlighted the disregard for the variable coordinates during the classification process as a limitation of the fuzzy C-means algorithm.

Tables 2, 3 and 4 show that, using the FCM method for grouping using more than one attribute and between different combinations, the mean values of the attribute centroids were maintained in most cases, including the classes. Another similarity verification criterion consisted of visually inspecting the maps (Figure 1 and 2) to analyze the location and composition of the management zone classes aiming at investigating their relationships in different attribute combinations. Similarity was observed among the following management zones "Ca, Ca - Mg"; "V, Ca - V, Ca - Mg - V"; "P - Ca, P - Mg, P - Ca - Mg, P - Ca - V"; and "P - Ca - Mg - V, P, Ca - V."

The results suggest that the P and V% attributes guide, for the most part, the delimitation of the management zones. V% was to be expected since it depends on the sum of bases (Ca, Mg and K).

Different data sources can be used for delimiting management zones that are economically more feasible than the total sampling of the area, and easier for applications at variable rates. As an example, it is possible to cite using farmer knowledge (Karydas *et al.*, 2020), productivity data (Ribeiro *et al.*, 2020), soil sampling (Demattê *et al.*, 2014), images (Rodrigues Junior *et al.*, 2011), electromagnetic induction sensors (Barrowes *et al.*, 2019) and the combination of these data (Taylor *et al.*, 2003; Fleming *et al.*, 2004; Koch *et al.*, 2004; Schepers, 2004).

It is noteworthy that although the studied area is relatively small (1.2 ha) for generating

management zones compared to other works found in the literature and has low variability due to adequate fertilization, the results showed that it was possible to generate management zones.

Several attributes can be used to define management zones in an area; it is suggested that isolated or combined attributes be used that are able to reflect the behavior of the variables which one desires to characterize. The higher the correlation between the attributes used to generate the management zones and the object to be characterized, the greater the reliability of the maps generated with the attributes. The management classes suggest areas that have similar behavior regarding the studied attributes, and are therefore markers of the sampling points for the fertility study of the area.

5. CONCLUSIONS

The generated management zones facilitate our understanding of the spatial variability of the attributes in the study area and to correlate similarities between management zones generated from different variables. However, it was observed that P and V% attributes particularly influenced the data, indicating a degree of coincidence between the generated zones that suggest a more targeted sampling, reducing the number of samples.

Additionally, comparing the results of this work with the work proposed by Sturião (2018), the sampling of the area becomes more directed, requiring fewer sampling points, besides decreasing the number of attributes studied, confirming the effectiveness of the method.

6. ACKNOWLEDGMENTS

7. REFERENCES

ALVES, S. M de F.; ALCÂNTARA, G. R.; REIS, E. F dos.; QUEIROZ, D. M de.; VALENTE, D. S. M. Definição de zonas de manejo a partir de mapas de condutividade. **Revista Biociências**, v. 1, p. 104–114, 2013.

BARBIERI, D. M.; MARQUES JÚNIOR, J.; PEREIRA, G. T. Variabilidade espacial de atributos químicos de um argissolo para aplicação de insumos à taxa variável em diferentes formas de relevo. **Engenharia Agrícola**, v. 28, p. 645–653, 2008. https://doi.org/10.1590/S0100-69162008000400004

BARROWES, B.; PRISHVIN, M.; JUTRAS, G., SHUBITIDZE, F. Frequency Electromagnetic Induction (HFEMI) Sensor Results from IED Constituent Parts. **Remote Sensing**, v. 11, n. 20, p. 2355, 2019. https://doi.org/10.3390/rs11202355

BEZDEK, J. C.; PAL, S. K. (eds). *Fuzzy* **models for pattern recognition – methods that search for structure in data**. New York, IEEE Pres, 1992.

CAVALCANTE, E. G. S.; ALVES, M. C.; SOUZA, Z. M.; PEREIRA, G. T. Variabilidade espacial de atributos químicos do solo sob diferentes usos e manejos. **Revista Brasileira de Ciência do Solo**, v. 31, p. 1329–1339, 2007. https://doi.org/10.1590/S0100-06832007000600012

CINTRA, F. L. D. Reflexões sobre o efeito dos horizontes coesos no movimento de água no solo e na distribuição do sistema radicular. *In*: MARTINS, D. dos S. (ed.). **Papaya Brasil**: mercado e inovações tecnológicas para o mamão. Vitória: INCAPER, 2005. p.103-107.

CORRÊA, M. M.; KER, J. C.; BARRÓN, V.; TORRENT, J.; CURI, N.; TORRES, T. C. P. Caracterização física, química, mineralógica e micromorfológica de horizontes coesos e fragipãs de solos vermelhos e amarelos do ambiente tabuleiros costeiros. **Revista Brasileira de Ciência do Solo**, v. 32, p. 297-313, 2008.

DELALIBERA, H. C.; WEIRICH NETO, P. H.; NAGATA, N. Management zones in agriculture according to the soil and landscape variables. **Engenharia Agrícola,** v. 32, n. 6, p. 1197–1204, 2012. 10.1590/S0100-69162012000600021

DEMATTÊ, J. A. M.; DEMATTÊ, J. L. I.; ALVES, E. R.; NEGRÃO, R.; MORELLI J. L. Precision agriculture for sugarcane management: a strategy applied for brazilian conditions. **Acta Scientiarum Agronomy,** v. 36, n.1, p. 111, 2014. https://doi.org/10.4025/actasciagron.v36i1.17664

EMBRAPA. **Sistema brasileiro de classificação de solos**. 3. ed. Rio de Janeiro: Embrapa/CNPSo, 2013. 353 p.

FÉLIX, V. B.; GONZATTO JÚNIOR, O. A.; ROSSONI, D. F.; HENRIQUES, M. J. Estimadores de semivariância: uma revisão. **Ciência e Natura,** v. 38, n. 3, p. 1157, 28 set. 2016. https://doi.org/10.5902/2179460X21326

FERREIRA, I. O.; RODRIGUES, D. D.; SANTOS, G. R.; ROSA, L. M. F. In Bathymetric Surfaces: idw or kriging? **Boletim de Ciências Geodésicas**, v. 23, p. 493–508, 2017. https://doi.org/10.1590/S1982-21702017000300033

FLEMING, K. L.; HEERMANN, D. F.; WESTFALL, D. G. Evaluating Soil Color with Farmer Input and Apparent Soil Electrical Conductivity for Management Zone Delineation. **Agronomy Journal**, v. 96, n. 6, p. 1581, 2004. https://doi.org/10.2134/agronj2004.1581

GOMES, R. P.; CAMPOS, M. C. C; BRITO, W. B. M.; CUNHA, J. M.; MUNIZ, A. W.; SILVA, L. S. *et al.* Variability and spatial correlation of aggregates and organic carbon in indian dark earth in Apuí region, AM. **Bioscience Journal**, v. 34, n. 5, p. 1188–1199, 2018. https://doi.org/10.14393/BJ-v34n5a2018-39414

JACOMINE, P. K. T. Distribuição geográfica, características e classificação dos solos coesos dos tabuleiros costeiros. *In*: REUNIÃO TÉCNICA SOBRE SOLOS COESOS DOS TABULEIROS COSTEIROS, 1996, Cruz das Almas. **Anais[...]** Aracaju : Embrapa-CPATC, 1996.

KARYDAS, C.; LATROU, M.; LATROU, G.; MOURELATOS, S. Management Zone Delineation for Site-Specific Fertilization in Rice Crop Using Multi-Temporal RapidEye Imagery. **Remote Sensing**, v. 12, n. 16, p. 2604, 2020. https://doi.org/10.3390/rs12162604

KOCH, B.; KHOSLA, R.; FRASIER, W. M.; WESTFALL, D. G; INMAN, D. Economic Feasibility of Variable-Rate Nitrogen Application Utilizing Site-Specific Management Zones. **Agronomy Journal**, v. 96, n. 6, p. 1572–1580, 2004. https://doi.org/10.2134/agronj2004.1572

LIMA NETO, J. A.; RIBEIRO, M. R.; CORRÊA, M. M.; SOUZA-JÚNIOR, V. S.; ARAÚJO FILHO, J. C. *et al.* Atributos químicos, mineralógicos e micromorfológicos de horizontes coesos de Latossolos e Argissolos dos tabuleiros costeiros do estado de Alagoas. **Revista Brasileira de Ciência do Solo**, v. 34, p.473-486, 2010.

LINDEN, R. Técnicas de Agrupamento. **Revista de Sistemas de Informação da FSMA**, n. 4, p. 18-36, 2009.

LOISEL, P.; CHARNOMORDIC, B.; JONES, H.; TISSEYRE, B. An optimisation-based approach to generate interpretable within-field zones. **Precision Agriculture,** v. 20, n. 1, p. 101–117, 2018. 10.1007/s11119-018-9584-3

MÁQUINA, C. M.; SANTOS, N. T.; COSTA, M. M.; SILVA, S. A. Using multivariate analysis to design management zones. **Científica**, v. 48, n. 1, p. 25–35, 2020. http://dx.doi.org/10.15361/1984-5529

MENEZES, Z. S.; MARQUES, J.; TADEU, G. P.; MOREIRA, L. F. Variabilidade espacial do pH, Ca, Mg e V% do solo em diferentes formas do relevo sob cultivo de cana-de-açúcar. **Ciência Rural**, v. 34, p. 1763–1771, 2004. https://doi.org/10.1590/S0103-84782004000600015

METWALLY, M. S.; SHADDAD, S. M.; LIU, M.; YAO, R.; ABDO, A. I.; LI, P. *et al.* Soil Properties Spatial Variability and Delineation of Site-Specific Management Zones Based on Soil Fertility Using Fuzzy Clustering in a Hilly Field in Jianyang, Sichuan, China. **Sustainability**, v. 11, n. 24, p. 7084, 2019. https://doi.org/10.3390/su11247084

MOLIN, J. P.; CASTRO, C. N. de. Establishing management zones using soil electrical conductivity and other soil properties by the fuzzy clustering technique. **Scientia Agricola**, v. 65, n. 6, p. 567–573, 2008. https://doi.org/10.1590/S0103-90162008000600001

MONTANARI, R. *et al*. Landscape form as a criterion for sampling optimization of an oxisol under cultivation of sugar cane. **Pesquisa Agropecuária Brasileira**, v. 40, p. 69-77, 2005. https://doi.org/10.1590/S0100-204X2005000100010

MONTEZANO, Z. F.; CORAZZA, E. J.; MURAOKA, T. Variabilidade espacial da fertilidade do solo em área cultivada e manejada homogeneamente. **Revista Brasileira de Ciência do Solo**, v. 30, n. 5, p. 839–847, 2006. https://doi.org/10.1590/S0100-06832006000500010

MORARI, F.; CASTRIGNANÒ, A.; C. PAGLIARIN. Application of multivariate geostatistics in delineating management zones within a gravelly vineyard using geo-electrical sensors. **Computers and Electronics in Agriculture,** v. 68, n. 1, p. 97–107, 2009. https://doi.org/10.1016/j.compag.2009.05.003

PEDROSO, M.; TAYLOR, J.; TISSEYRE, B.; CHARNOMORDIC, B.; GUILLAUME, S. A. Segmentation algorithm for the delineation of agricultural management zones. **Computers and Electronics in Agriculture**, v. 70, n. 1, p. 199–208, 2010. https://doi.org/10.1016/j.compag.2009.10.007

RAMOS, F. T.; SANTOS, R. T.; CAMPELO JUNIOR, J. H.; MAIA, J. C. S. Defining management zones based on soil attributes and soybean productivity. **Revista Caatinga**, v. 30, p. 427–436, 2017. https://doi.org/10.1590/1983-21252017v30n218rc

RIBEIRO, F. O.; FERNANDES, A. R, MATOS, G. S. B.; LINDOLFO, M. M.; GUEDES, R. S.; RODRIGUES, G. R. Spatial variability in leaf analysis and productivity of fertirrigated açaí. **Engenharia Agrícola**, v. 40, p. 800–808, 2020. https://doi.org/10.1590/1809-4430

IPABHi

RICARDO, S.; SOUZA, E. G.; BAZZI, C. L.; URIBE-OPAZO, M. A.; BETZEK, N. M. Redundant variables and the quality of management zones. **Engenharia Agrícola**, v. 36, n. 1, p. 78–93, 2016. https://doi.org/10.1590/1809-4430

RODRIGUES JUNIOR, F. A.; VIEIRA, L. B.; QUEIROZ, D. M.; SANTOS, N. T. Geração de zonas de manejo para cafeicultura empregando-se sensor SPAD e análise foliar. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 15, p. 778–787, 2011. https://doi.org/10.1590/S1415-43662011000800003

RODRIGUES, M. S.; CORÁ, J. E.; FERNANDES, C. Soil sampling intensity and spatial distribution pattern of soils attributes and corn yield in no-tillage system. **Engenharia Agrícola**, v. 32, n. 5, p. 852–865, 2012. https://doi.org/10.1590/S0100-69162012000500005

SAIZ-RUBIO, V.; ROVIRA-MÁS, F. From Smart Farming towards Agriculture 5.0: A Review on Crop Data Management. **Agronomy**, v. 10, n. 2, p. 207, 2020. https://doi.org/10.3390/agronomy10020207

SALGADO, J. S.; COSTA, A. N. Solos cultivados com mamão. *In*: MARTINS, D. dos S.; COSTA, A. de F. S. (eds.). **A cultura do mamoeiro**: Tecnologias de Produção. Vitória: INCAPER, 2003. p.117- 125.

SANTOS, J. A.; SOUSA, K. N. S.; SANTOS, P. R. B.; LIMA, J. L.; BACELAR, R. J. Habitat, limnological signatures and spatial modeling: a zoning proposal for the Curuá-Una hydroelectric reservoir, Pará, Brazil. **Acta Limnologica Brasiliensia**, v. 31, 2019. https://doi.org/10.1590/S2179-975X12017

SCHEPERS, A. R.; SHANAHAN, J. F.; LIEBIG, M. A; SCHEPERS, J. S.; JOHNSON, S. H.; LUCHIARO, 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. https://doi.org/10.2134/agronj2004.1950

SILVA, V. R.; REICHERT, J. M.; STORCK, L.; FEIJÓ, S. Variabilidade espacial das características químicas do solo e produtividade de milho em um Argissolo Vermelho-Amarelo distrófico arênico. **Revista Brasileira de Ciência do Solo**, v. 27, n. 6, p. 1013–1020, 2003. https://doi.org/10.1590/S0100-06832003000600005

SILVA, A. F.; BARBOSA, A. P.; ZIMBACK, C. R. L.; LANDIM, P. M. B. Geostatistics and remote sensing methods in the classification of images of areas cultivated with citrus. **Engenharia Agrícola,** v. 33, p. 1245–1256, 2013. https://doi.org/10.1590/S0100-69162013000600017

SILVEIRA, P. M.; ZIMMERMANN, F. J. P.; SILVA, S. C.; CUNHA, A. A. Amostragem e variabilidade espacial de características químicas de um latossolo submetido a diferentes sistemas de preparo. **Pesquisa Agropecuária Brasileira**, v. 35, n. 10, p. 2057–2064, 2000. https://doi.org/10.1590/S0100-204X2000001000018

SOUZA, C. K.; MARQUES, J. R. J, MARTINS FILHO, M. V.; PEREIRA, G. T. Influência do relevo e erosão na variabilidade espacial de um Latossolo em Jaboticabal (SP). **Revista Brasileira de Ciência do Solo**, v. 27, n. 6, p. 1067–1074, 2003. https://doi.org/10.1590/S0100-06832003000600011

STURIÃO, W. P.; GONTIJO, I.; FONSECA, A. S.; LIMA, J. S.S. Variability and Spatial Correlation between Phenotypic Attributes and Productivity of Papaya. **Journal of Experimental Agriculture International**, p. 1–13, 2018. 10.9734/JEAI/2018/41819

TAYLOR, J. C.; WOOD, G. A.; EARL, R.; GODWIN, R. J. Soil Factors and their Influence on Within-field Crop Variability, Part II: Spatial Analysis and Determination of Management Zones. **Biosystems Engineering**, v. 84, n. 4, p. 441–453, 2003. https://doi.org/10.1016/S1537-5110(03)00005-9

XIN-ZHONG, W.; GUO-SHUNA, L.; HONG-CHAO, H.; ZHEN-HAIB, W.; QINGHUAB, L.; XU-FENGC, L. *et al.* Determination of management zones for a tobacco field based on soil fertility. **Computers and Electronics in Agriculture**, v. 65, n. 2, p. 168–175, 2009. https://doi.org/10.1016/j.compag.2008.08.008

YAN, L.; ZHOU, S.; CI-FANG, W.; HONG-YI, L.; FANG, L. Determination of potential management zones from soil electrical conductivity, yield and crop data. **Journal of Zhejiang University-SCIENCE B,** v. 9, n. 1, p. 68–76, 2008. 10.1631/jzus.B071379

YANG, M. S. A Survey of *Fuzzy* Clustering. **Mathematical and Computer Modelling**, v. 18, n. 11, p. 1-16, 1993. https://doi.org/10.1016/0895-7177(93)90202-A

ZADEH, L. A. Fuzzy sets. **Information and Control**, v. 8, n. 3, p. 338–353, 1965. https://doi.org/10.1016/S0019-9958(65)90241-X