Establishment of homogeneous zones in a soil of alluvial origin

Determinação de zonas homogêneas num solo de origem fluvial

Camilo Ignacio Jaramillo-Barrios[1]*, Andrés Felipe Ortiz-Rico[2],

Gustavo Alfonso Araujo Carrillo[3], Viviana Marcela Varón-Ramírez[3]

[1]Corporación Colombiana de Investigación Agropecuaria-AGROSAVIA, Centro de Investigación Nataima, vía Espinal-Ibagué, Tolima, Colombia
[2]Universidad Santo Tomás, Facultad de Estadística, Bogotá, Colombia
[3]Corporación Colombiana de Investigación Agropecuaria-AGROSAVIA, Centro de Investigación Tibaitatá, Cundinamarca, Colombia
*Corresponding author: camilojaramillo@usantotomas.edu.co, cijaramillo@agrosavia.co
Received in November 11, 2018 and approved in April 16, 2019

ABSTRACT

Soils have the ability to maintain plant growth and biological activity due to their physical and chemical properties. The aim of this study was to observe the spatial distribution of some chemical properties of the soil, such as pH, organic matter (OM), electrical conductivity (EC), effective cation exchange capacity (ECEC), sulfur (S) and aluminum (Al) content and to establish zones with homogeneous chemical characteristics using the MULTISPATI-PCA technique and the fuzzy c-means algorithm. The study area was located in the Tundama and Sugamuxi Valleys (Boyacá, Colombia) with an area of 8,017 ha. Chemical properties such as pH, OM, EC, S, Al, and ECEC were indicators of the chemical degradation of these soils. Four homogeneous zones were identified. The first zone represents areas with acidity and excessive sulfur, with a pH of 4.54, 15.88% OM, 3.19 dS m^{-1} EC, 2.47 meq 100 g^{-1} Al and 365.59 meq 100 g^{-1} S. In contrast, the second zone represents areas with a high self-neutralizing capacity, with a pH of 5.98, 4.22% OM, 0.75 dS m^{-1} EC, 0.20 meq 100 g^{-1} Al and 44.64 meq 100 g^{-1} S. Zone three showed a high similarity with the first two, except for its EC and S contents. Finally, zone four showed similarity with the first, except in OM, EC and S contents. These data show that S and EC influenced the homogeneous zones because the soils in this area are called acid sulfate soils.

Index terms: Geostatistics; semivariogram; soil degradation; Multispati-PCA; spatial distribution.

RESUMO

A capacidade do solo para manter o crescimento da planta e a actividade biológica reside nas suas propriedades físicas e químicas. O objectivo deste estudo foi observar a distribuição espacial das propriedades químicas do solo como pH, matéria orgânica (MO), condutividade eléctrica (CE), capacidade de troca catiônica eficaz (CTCE) e conteúdo de S e Al e determinar áreas com características químicas homogêneos por meio da técnica MULTISPATI-PCA e c-meios algoritmo. A área de estudo está localizada nos vales de Tundama e Sugamuxi (Boyacá-Colômbia) com uma área de 8,017 ha. O pH, MO, CE, S, Al, e as propriedades de CTCE eram indicativos de degradação química destes solos. Identificaram-se quatro zonas homogêneas que representam as primeiras áreas com excesso de acidez e de enxofre, com pH 4,54, 15,88 % de MO, 3,19 dS m^{-1} de EC, 2,47 meq 100g^{-1} de AL e 365,59 meq 100g^{-1} de S; em contraste, a segunda zona representa áreas com elevada capacidade de auto-neutralização, com um pH de 5,98, 4,22% de MO, 0,75 dS m^{-1} de CE, 0,20 meq 100g^{-1} de Al g e 44,64 meq 100g^{-1} de S. A zona três apresentou a maior similaridade com os dois primeitos parâmetros, exceto CE e S. Finalmente, a zona quatro mostrou semelhança com a zona um, exceto MO, CE e S. Portanto, conclui-se que as zonas de manejo foram influenciadas pelo enxofre e condutividade elétrica, e, portanto, os solos nesta área são denominados sulfatados ácidos.

Termos para indexação: Geoestatística; semivariograma; degradação do solo; Multispati-PCA; Distribuição espacial.

INTRODUCTION

The ability of soils to maintain plant growth and biological activity lies in their physical and chemical properties (Lal, 2002). These properties are the result of specific interactions among the five formation factors in a given place (McGraw, 1994) and of dozens of pedogenetic processes, thus generating spatial variability in their chemical, physical, biological, and mineralogical properties, among others (Jaramillo, 2014).

The study of the spatial variability of agricultural soil properties is important to make appropriate management decisions and to improve soil quality (Rosemary et al., 2017). The variability has a strong relationship with the soil use (Wang; Shao, 2013), i.e., a soil without human intervention shows less variation than one under agricultural use. Moreover, in the latter, management practices affect the change of soil properties (Jaramillo, 2012). Once the source of variation is known, a higher

efficiency in establishing homogeneous zones is achieved (Mzuku et al., 2005), allowing the implementation of differential management. This management improves efficiency and sustainability in production through fertilization, irrigation and tillage practices, among others, that are specific to each place (Jaramillo; Sadeghian; Lince, 2013).

The use of several statistical tools to study the spatial variability of soil properties has been recorded (Reichardt; Timm, 2008). Classical statistics is the discipline that started these studies. However, classical statistical evaluations only include generalizations regarding the magnitude of the variation and cannot be used to evaluate autocorrelated data (Stoyan et al., 2000). As an alternative, Bachmeier and Bufa (1992) mention that the theory of regionalized variables allows the measurement of the spatial dependence of edaphic properties and that geostatistics provide a means to define an autocorrelation with semivariograms (Stoyan et al., 2000; Díaz, 2002).

The use of geostatistics to establish the variability of the chemical properties of soils has been widely studied in various types of soils and production systems worldwide (Rahal, 2015; Aghasi et al., 2017). Within geostatistical techniques, interpolation with ordinary kriging (OK) is the most widely used method, and it assumes intrinsic stationarity. However, if the spatial stochastic process shows a tendency, the assumption is unsustainable, and a more robust model that can better explain the variation is needed (Li et al., 2015). As a solution, Matheron (1969) introduced so-called universal kriging (UK), where the trend is removed using surface models that turn out to be linear combinations of spatial coordinates (Díaz, 2002).

Principal component analysis (PCA) is commonly used at the multivariate level to construct linear combinations with study variables (Moral; Terrón; Da Silva, 2010); however, this method does not consider the dependency structures expected in spatial data. As an alternative to this analysis, a method called MULTISPATI-PCA (Dray; Saïd; Débias, 2008) has been designed with good results in the study of soils. This method incorporates spatial information and uses the Moran index (MI) to measure the dependence or spatial correlation between observations at a site and the average in their neighborhood (Arrouays et al., 2011). Thus, this analysis considers spatial structure (autocorrelation) in the original variables to produce synthetic variables and the relationships between the variables is measured (covariability analysis) (Córdoba et al., 2013).

MULTISPATI-PCA has been used in the delineation of management zones based on soil and terrain variables (Peralta et al., 2015). In Córdoba, Argentina, this technique was used for multivariate zoning at the regional scale, involving edaphic and climatic data (Giannini et al., 2018). In France, it was used across the entire country to study the main soil characteristics of the topsoil and to assess their multivariate spatial patterns (Arrouays et al., 2011). Regarding the definition of management zones of the different cluster analyses, the fuzzy c-means algorithm has been widely used (Rodrigues; Corá, 2015; Tripathi et al., 2015; Gavioli et al., 2016). This algorithm has been determined to be preferable for grouping properties in the soil continuum (Odeh; McBratney; Chittleborough, 1992) and has been used in the grouping of similar points using the main spatial components as inputs for site classification (Córdoba et al., 2013).

Accordingly, the aim of this study was to observe the spatial distribution of some chemical properties of the soil, such as pH, organic matter (OM), electrical conductivity (EC), effective cation exchange capacity (ECEC), sulfur and aluminum contents. Furthermore, the study aims to establish zones with chemical homogeneous characteristics using the MULTISPATI-PCA technique and the fuzzy c-means algorithm in a soil of alluvial origin.

MATERIAL AND METHODS

Description and location of the study area

This study was carried out in an area of 8,017 ha in the department of Boyacá (Colombia), specifically in the region known as Valles del Tundama and Sugamuxi. These areas are located geographically between the parallels 5º 43' 28.8228" N and 5º 50' 32.4162" N, and the meridians 73º 6' 38.4798" W and 72º 56' 0.3264" W of Greenwich, at an average altitude of 2,500 m a.s.l. (Figure 1). Soils within the study area are mostly used for livestock grazing, and only approximately one thousand hectares are used for agricultural production, where potato, wheat, corn, beans, peas, beans, carrots, and bulb onions, among others, are the main crops produced (Forero; Castillo, 2016).

The study area remains flooded most of the time because it is located in a "valley" geomorphological position. These soils are characterized by continuous chemical degradation due to various forms of sulfur, high concentrations of soluble aluminum and iron, production of sulfuric acid, low availability of phosphorus and low base saturation. For this reason, they are known as acid sulfate soils (Bernal; Forero, 2014).
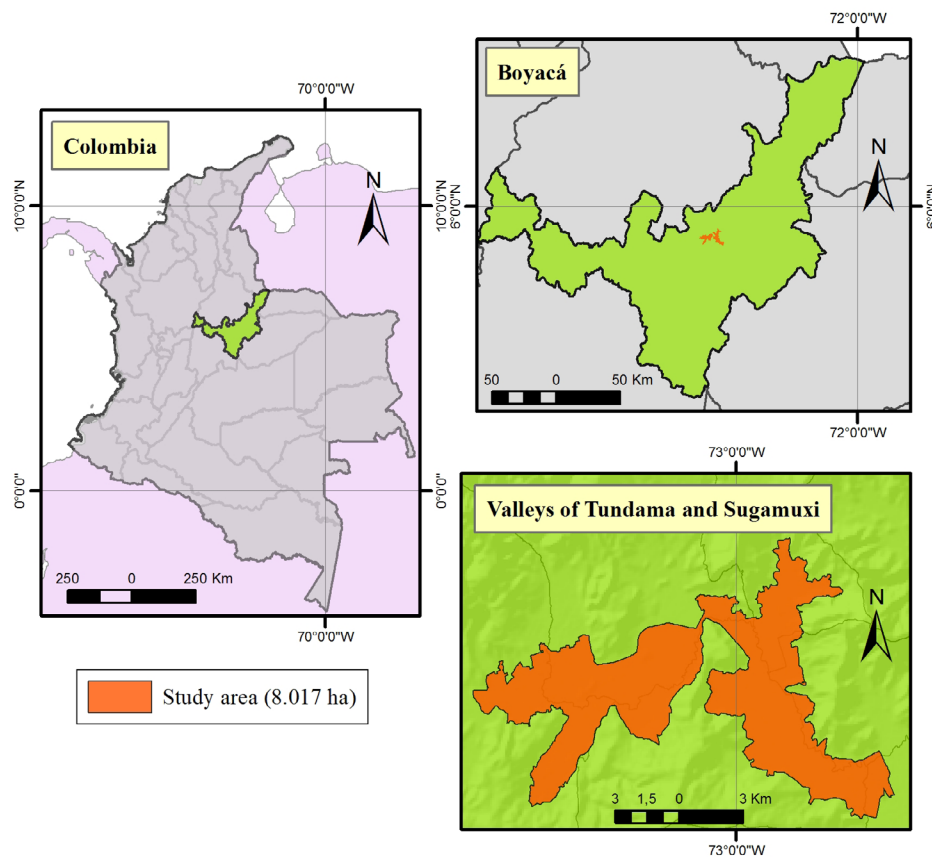
Figure 1: Study area to establish the spatial distribution of chemical soil properties.

Soil sampling and laboratory analyses

The data set used in this study was obtained under the Agreement 20110060 (Internal Code 1723) between Ministerio de Agricultura y Desarrollo Rural de Colombia (MADR) and Corporación Colombiana de Investigación Agropecuaria (Agrosavia). The sampling design was stratified according to the type of soil, land uses, origin and timing of floodwaters and ponding. Two hundred ninety-five observation points were sampled at a depth between zero (0) and 20 cm, and their allocation was randomized in each unit. The chemical soil properties measured were pH (soil:water ratio of 1:2.5 determined by potentiometry) and organic matter (OM) measured by the Walkley and Black method. Available phosphorus (P) was measured by Bray II through VIS spectrophotometry, and calcium (Ca), magnesium (Mg), potassium (K) and sodium (Na) were measured using 1N ammonium acetate at pH 7.0, employing atomic absorption spectrophotometry. Exchangeable aluminum (Al) was measured by extraction with KCl when the pH was <5.5 through volumetry,

available sulfur (S) was determined by calcium monobasic phosphate through spectrophotometry, and Fe, Mn, Cu and Zn contents were determined by the modified Olsen method through atomic absorption spectrophotometry. Available boron (B) was measured using calcium monobasic phosphate through VIS spectrophotometry, and electrical conductivity (EC) and effective cationic exchange capacity (ECEC) were calculated by the sum of the cations of Ca, Mg, Na and K.

Statistical analysis

Initially, a descriptive and exploratory analysis was carried out with all the study variables to calculate central tendency and dispersion measures (Rodríguez; Camacho; Rubiano, 2016). Additionally, the significance of Pearson's product-moment correlation coefficients was evaluated using the Clifford, Richardson and Hémon (1989) procedure and Moran's index to estimate the spatial autocorrelation. The coefficient of variation (CV) was analyzed using the Warrick and Nielsen (1980) criterion,

in which values lower than 12% are considered to have low variability, those between 12% and 60% to have medium variability, and those higher than 60% to have high variability.

Subsequently, an exploratory analysis was carried out through visual diagnostic tests to evaluate the assumptions of normality using box diagrams, histograms and qq-plots. Stationarity was evaluated with scatter plots of the variable versus latitude and longitude.

Stationarity was verified with a spatial trend analysis estimating a polynomial model through a multiple regression. In this regression, the variable under study was the dependent variable, and the sampling point coordinates were the independent variables (Jaramillo; Sadeghian; Lince, 2013). Once a significant regression model was obtained (p<0.05), residuals were extracted to carry out semivariance analysis.

A normality test was performed with the residuals employing the Shapiro-Wilks test at 5% significance; when this assumption was not met, the variable was transformed into a natural logarithm (Li; Webster; Shi, 2015). In this case, the analysis described above was performed again with respective trend removal according to each case (Jaramillo, 2009). Semivariograms were created, adjusting each one according to the model found for each variable. The first step was to determine the initial parameters with the eyefit function using the interactive Tcl-Tk interface. In these, different models were tested (exponential, spherical, Gaussian, Matern, among others).

The values obtained from the initial parameters were used to estimate the semivariance models, adjusting these according to the model found for each spatial stochastic process, establishing the sill (C), range (A) and nugget effect (C0) (Hernández et al., 2018). The estimated parameters were calculated through ordinary least squares, weighted "n pairs", weighted "cressie", maximum likelihood and maximum restricted likelihood (Cressie, 1993; Selby; Kockelman, 2013; Li; Webster; Shi, 2015).

To establish the goodness of the predictions made by different methods, cross-validations were made, and the best model was selected by the highest cross-validation coefficient (CVC), the smallest root of the mean square error (RMSE), the reduced error (RE) value closer to zero, the value of the standard deviation of the reduced errors (SDRE) closer to one (Faraco et al., 2008; Johann et al., 2010; Cortés; Camacho-Tamayo; Giraldo, 2016) and the best degree of spatial dependence (DSD) of each of the chemical soil properties, according to the classification proposed by Cambardella et al. (1994). These authors consider the degree of spatial dependence as strong when DSD ≤ 25%, moderate when 25 < DSD ≤ 75%, and weak when DSD > 75%.

For interpolations, the polygon of the zone was used, and this was calculated from the sample size and total number of pixels with the equation of inspection density suggested by Hengl et al. (2006) used when predictions are created. Each map should have approximately an equal sample density per area. Therefore, an approximate pixel size (TP) of 40 was established for a total of 50,106 pixels used in the maps per variable, as well as the main interpolated spatial components. The predictions were carried out by implementing UK, and for variables transformed by natural logarithm, the inverse transformation was applied with the correction of Laurent (1963) to estimate the original scale. The correction consists of adding a value of 0.5 to the estimated regionalized variable and multiplying it by the estimated variance in the exponential function.

For the principal spatial components (sPC), the MULTISPATI-PCA method was used. A weighting matrix *Wnxn*, which is a mathematical representation of the geographical distribution of the study sites, was created. The neighborhood network was defined based on the Euclidean distance between adjacent neighboring points. Subsequently, sPCs were calculated, and the associated eigenvalues equivalent to the spatially structured variance were obtained. The presence of autocorrelation in sPCs was analyzed with the Moran index (MI). UK interpolation was applied based on the MULTISPATI-PCA sPC1 and sPC2 semivariograms to obtain multivariate spatial variability maps. (Dray; Said; Debias, 2008, Córdoba et al., 2012).

The fuzzy c-means cluster algorithm was employed with the interpolations of the first two sPCs to classify the homogeneous zones using the Euclidean distance, a fuzziness exponent of 1.3 (Córdoba et al., 2016) and a maximum number of iterations of 500. To validate the results, internal measures of four indexes, partition density (PD), Xie and Beni (XB), Fukuyama and Sugeno (FS) and partition coefficient (PC) values were used. Between two and eight homogeneous zones were tested, which were compared using the lowest values of the first three indexes and the value closest to one for the PC (Meyer et al., 2017). The analyses described above were performed with the statistical package R version 3.4.4 (R Core Team, 2018) and Pearson's product-moment correlation with the software PASSaGE (v.2) (Rosenberg; Anderson, 2011).

RESULTS AND DISCUSSION

Spatial analysis

In Table 1, descriptive statistics for the study properties can be observed. The CV showed a high

variability in all the evaluated properties, except for pH and Fe, which showed medium variability; for the latter, its interpretation agrees with Hernández et al. (2018). Regarding pH, studies have shown that its variability range fluctuates between 2 and 15% (Cox; Gerard; Abshire, 2006); however, in the current study, a 21.43% range was found. This can be explained by the physical, mineralogical and biochemical changes in the acidity-related processes of this type of soil, which directly affect this property (Rosicky; Sullivan; Slavich, 2004).

Six chemical properties were selected: pH, OM, EC, S, Al and ECEC, for the univariate interpolation by UK, and their significant spatial correlation ($p < 0.001$) was verified through the Moran index (MI) as follows: pH (MI = 0.295), OM (MI = 0.468), EC (MI = 0.311), S (MI = 0.281), Al (MI = 0.267) and ECEC (MI = 0.329). The above agrees with studies of the area where changes in these properties have been highlighted, establishing that pH, Al and S are important indicators in the diagnosis of acidity and sulfation status of this type of soil (Castro et al., 2006; Rincón; Castro; Gómez, 2008). pH showed a lower value of skewness (0.17) in comparison with the values between 1.53 and 2.91 that OM, EC, S, Al and

ECEC presented; these last variables were transformed by natural logarithm. Regarding stationarity, all the variables presented a tendency, and it was necessary to remove it through a model as a function of the coordinates. The semivariograms obtained for the chemical properties studied were adjusted to exponential models, with ranges above 345.9 m, being higher for S with 1013.5 m (Table 2). The restricted maximum likelihood (REML) method was used to adjust the OM, S and ECEC, ordinary least squares (OLS) was used for pH and Al, and maximum likelihood (ML) for EC (Figure 2).

The degree of spatial dependence (DSD) was classified as strong (DSD ≤ 25%) for four properties (OM, EC, Al and ECEC) and moderate (25 < DSD ≤ 75%) for the two remaining properties; however, all of them were less than or equal to 37.0%. According to Cambardella et al. (1994), the variables with strong spatial dependence are more influenced by soil formation factors (Table 2). The evaluation of UK interpolation was intermediate, with a range of cross-validation coefficients between 0.53 and 0.73, i.e., lower than those found by Varón-Ramírez, Camacho-Tamayo and González, (2018), and similar to those found by Cortés, Camacho-Tamayo and Giraldo (2016).

**Table 1:** Descriptive statistics of the chemical properties evaluated.

Soil parameter	Mean	Min	Max	CV	Skewness
pH	5.42	3.00	7.90	22	0.17
OM (%)	7.79	1.02	36.08	72	1.59
EC (dS m^{-1})	1.53	0.16	6.57	87	1.53
S (meq 100 g^{-1})	144.50	2.41	1200.00	78	2.91
Al (meq 100 g^{-1})	1.11	0.00	9.40	164	1.94
ECEC (meq 100 g^{-1})	19.74	2.22	97.23	75	2.00
P (ppm)	13.84	0.52	56.50	87	1.64
K (meq 100 g^{-1})	0.94	0.04	4.81	75	2.15
Na (meq 100 g^{-1})	1.42	0.03	12.40	128	2.94
Ca (meq 100 g^{-1})	14.14	0.44	87.30	91	2.07
Mg (meq 100 g^{-1})	2.14	0.12	11.80	78	2.66
Fe (meq 100 g^{-1})	155.36	15.00	380.00	35	0.01
Cu (meq 100 g^{-1})	0.90	0.01	7.26	104	2.02
Mn (meq 100 g^{-1})	16.01	0.24	66.40	70	1.20
Zn (meq 100 g^{-1})	4.99	0.39	17.40	64	0.83
B (meq 100 g^{-1})	0.51	0.06	2.01	61	1.86

pH= Potential of hydrogen, OM= Organic matter, EC=Electrical conductivity, S= Sulfur, Al=Aluminum, ECEC= Effective cation exchange capacity, P=Phosphorus, K=Potassium, Na= Sodium, Ca= Calcium, Mg= Magnesium, Fe= Iron, Cu= Copper, Mn= Manganese, Zn= Zinc, B= Boron, CV= Coefficient of variation.

Table 2: Estimated parameters of the semivariogram models of the soil properties studied.

Variable	Model	Method	Co	Range (m)	Co + C	DSD	CVC	RMSE	RE	SDRE
pH	Exponential	OLS	0.394	566.82	1.037	27.0	0.54	0.99	0.007	0.87
OM	Exponential	REML	0.051	608.11	0.401	11.0	0.73	0.19	-0.046	0.86
EC	Exponential	ML	0.113	406.42	0.449	20.0	0.57	0.02	-0.022	1.04
S	Exponential	REML	0.630	1013.50	1.089	37.0	0.54	19.0	-0.171	0.98
Al	Exponential	REML	0.055	759.21	0.568	9.0	0.57	0.59	-0.071	0.92
ECEC	Exponential	REML	0.059	345.90	0.310	16.0	0.62	0.15	-0.020	1.05

OM= Organic matter, EC= Electrical conductivity, ECEC=Effective cation exchange capacity, Co=nugget, Co+C1=Sill, DSD= Degree of spatial dependence, CVC= cross validation coefficient, RMSE: root mean square error, RE= mean reduced error, SDRE= standard deviation of the reduced error.
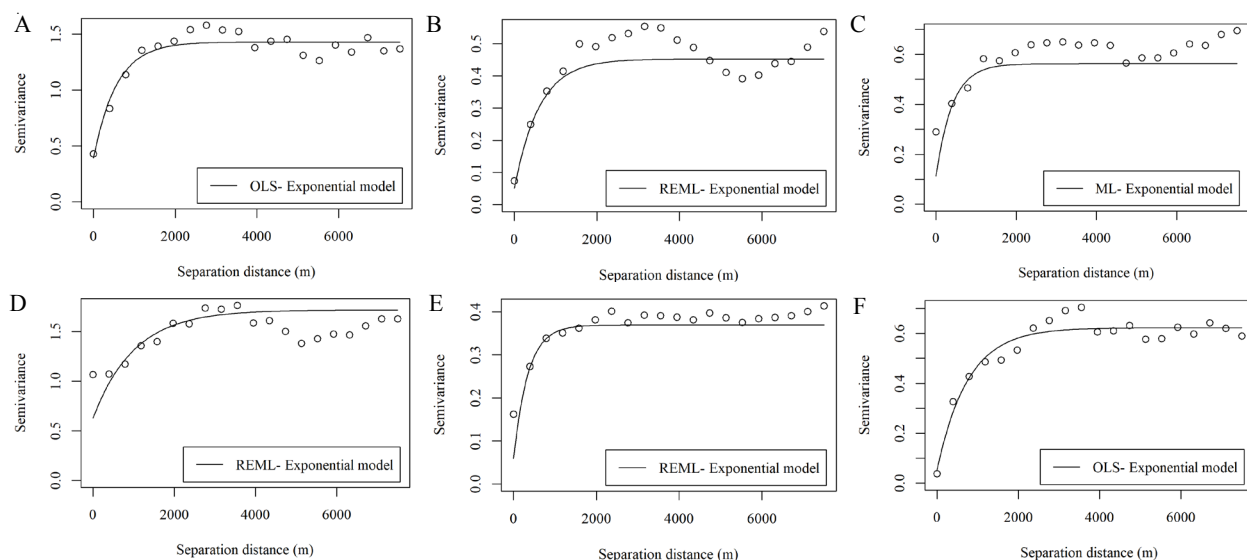
Figure 2: Semivariogram models of the chemical soil properties selected. a) pH, b) OM, c) EC, d) S, e) Al and f) ECEC.

The pH interpolation range fluctuated between 3.5 and 7.4, with a mean of 5.49 and a standard deviation of 0.98. Figure 3A shows that some areas of the Tundama and Sugamuxi Valleys have soils with ultra-acidic to extremely acidic reactions (pH 3.1-4.0) (Castro et al., 2006). The Pearson's product-moment correlations were significant, namely, low pH values are related to high levels of Al (r = -0.75, p-value= <0.0001) (Figure 3E), increases in OM (r = -0.52, p-value= 0.0002) (Figure 3B), increases in EC (r = -0.29, p-value= 0.0995) (Figure 3C) and excessive levels of S (r = -0.45, p-value=0.0079) (Figure 3D), which generate undesirable conditions for vegetative development.

In contrast, areas with pH > 5.5 and lower contents of Al, S, OM, EC and ECEC were observed. In these areas, according to Welch et al. (2009), corrections have generally been used with $CaCO_3$, with the purpose of neutralizing the acidity caused by hydrogen, aluminum, manganese, iron and organic matter acids, to allow the normal growth of plants. This implies that as the $CaCO_3$ doses increase, the pH stabilizes to 7.0, indicating that the high load of sulfates and iron exerts self-neutralization with joint formation of gypsum and Fe and Al hydroxides, a specific phenomenon of acid sulfate soils (Dent, 1986).

Homogeneous zones

A spatial principal component analysis (sPCA) was performed, which showed significant spatial autocorrelation values (p-values <0.001) and MI (0.49 and 0.26) for the first and second spatial components,

respectively. Figure 4A shows eigenvalues that suggest two main structures. The eigenvalues are the product between the variance and the spatial autocorrelation of the components. These first two axes collected 78.89% of the cumulative spatial variance. SPC1 includes properties related to soil acidity pH, showing pH, Cu, Zn and Mn opposed to OM, Al, S, Ca, B, EC and ECEC, which explains 61.68% of the cumulative spatial variance. SPC2 is represented by the Al content of the soil, contrasted with pH, Ca, Mg, ECEC, which explains the 17.21% value (Figure 4A). Finally, Figure 4B shows the sPC3 that is explained by Mn, Fe and Mg as opposed to the pH, explaining the 8.37% value.

After carrying out the sPCA, the semivariograms obtained were adjusted to exponential models, with the maximum likelihood (MV) and maximum restricted likelihood (REML) methods (Figures 4C-D). The estimated parameters obtained and the validation measures are presented in Table 3. SPC1 showed a strong DSD and a 0.74 CVC, in contrast to sPC2, which exhibited a moderate DSD and a 0.46 CVC.

The spatial distribution of sPC1 (Figure 4E) showed that areas with low values (blue scale) have acidic pH, low levels of Cu, Mn and Zn and high values of OM, Al, S, Ca, B, EC and ECEC.
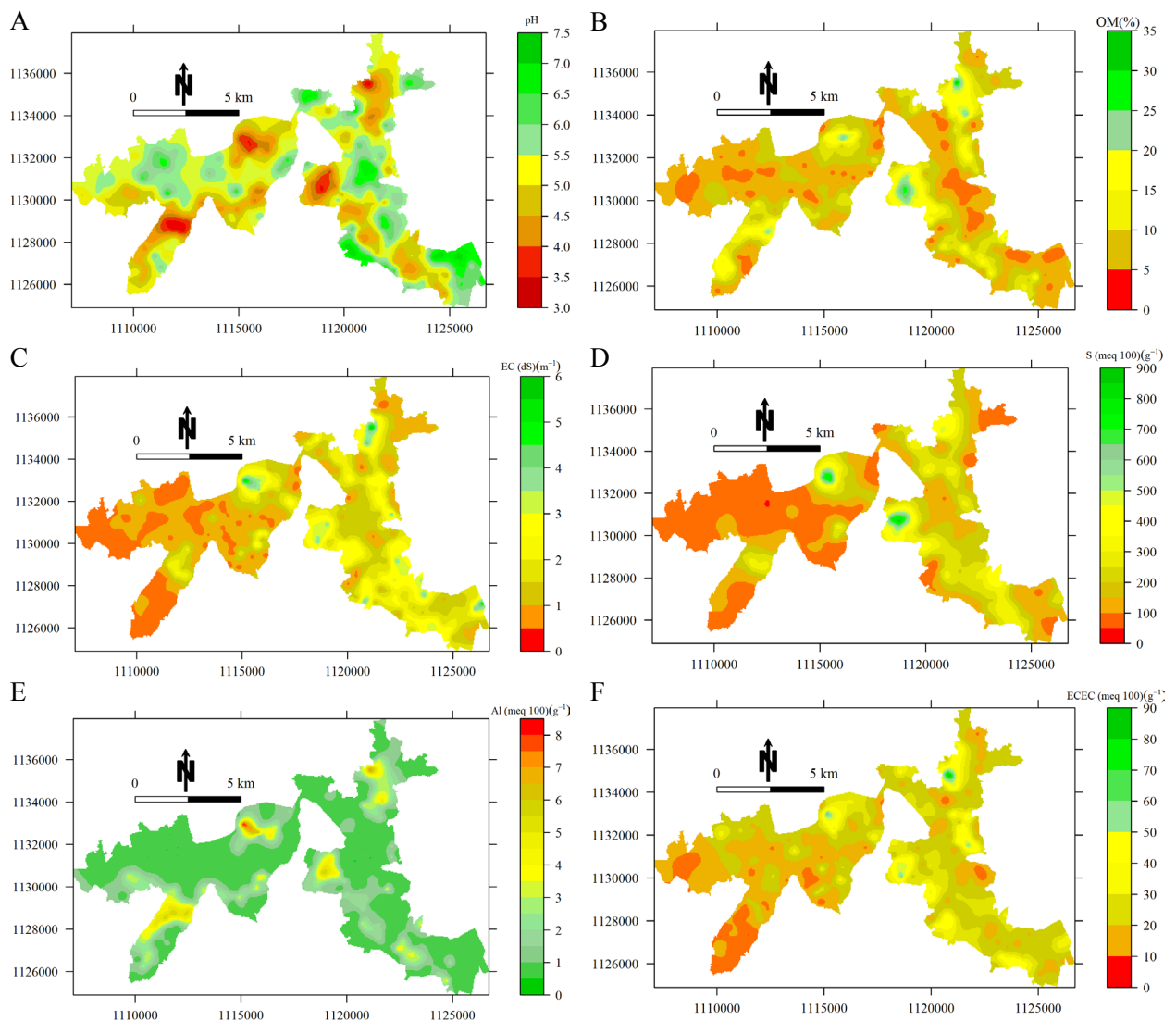

Figure 3. Spatial distribution of the chemical properties studied. a) pH, b) OM, c) EC, d) S, e) Al, and f) ECEC.
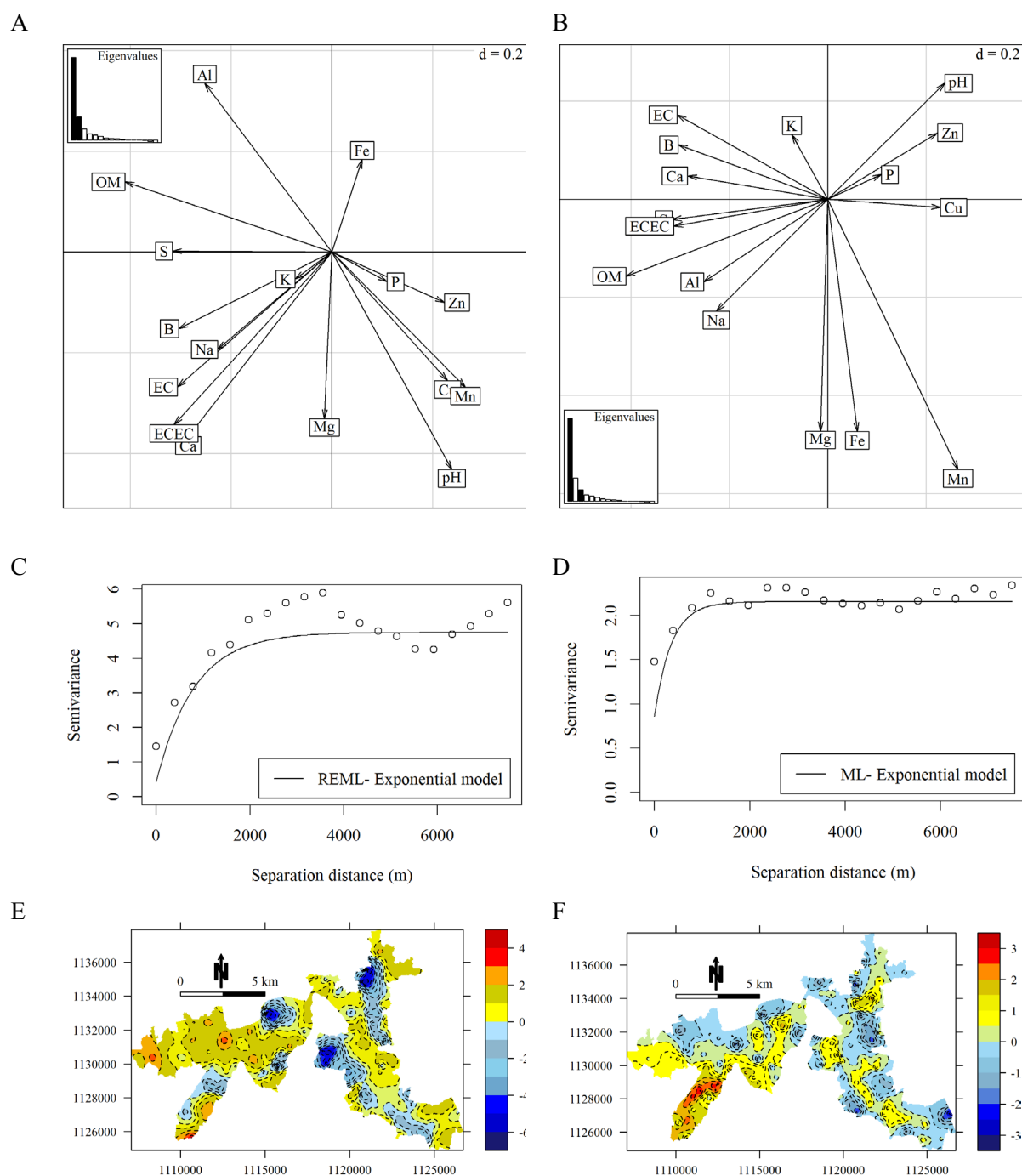
Figure 4: Spatial principal component analysis. a) Graphical representation of the first two axes and eigenvalues associated with the sPC, b) Graphic representation of the first and third axes and eigenvalues associated with the sPC, c) sPC1 semivariogram, d) sPC2 Semivariogram, e) sPC1 spatial distribution, and f) sPC2 spatial distribution.

In these areas, the excessive Ca content makes the absorption of all metallic micronutrients (Mn, Zn) difficult due to a decrease in their solubility (Gómez et al., 2007). SPC2 (Figure 4F) recorded areas with negative values (blue scale) with slightly acidic pH, low Al content, and high Ca, Mg, ECEC values, contrary to areas with positive values (yellow and red scale) that showed high Al content, acidic pH, and low Ca, Mg and ECEC values.

Homogeneous zones for the Tundama and Sugamuxi valleys were defined through the fuzzy c-means cluster algorithm for the interpolations of the first two spatial principal components. The number of optimal homogeneous zones (HZ) was established according to the registered indexes for the algorithm, where a number of classes of four were selected due to the lowest values found in the first three indexes and the value closest to one of the partition coefficients (PC) (Figure 5).

From this point, four homogeneous zones (HZs) were established (Figure 6). Subsequently, the sampling points of the properties were located on the map of these areas, where the HZ1 concentrated 15%, HZ2 39%, HZ3 22% and HZ4 concentrated 25%. Table 4 shows the main group characteristics. The chemical property with the highest discrimination among the HZs was S, which agrees with what was reported by Combatt, Palencia and Marin (2003).

HZ1 was named *"Areas with excessive acidity and sulfur levels"*. This type of area is characterized by having an acid pH (> 4.0), high Al content and excessive sulfur levels that appear neutralized by high Ca and ECEC concentrations. The highest limitations for this area are the high EC values and Na levels due to the effect of groundwater loaded with calcium and sodium sulfates in depressed areas (Castro et al., 2006). These zones are composed of materials rich in sulfides whose aeration

Table 3: Estimated parameters of the semivariogram models of the first two spatial principal components.

Variable	Model	Method	Co	Range (m)	Co + C	DSD	CVC	RMSE	RE	SDRE
sPC1	Exponential	REML	0.436	810.81	4.317	0.09	0.74	1.49	0.024	0.69
sPC2	Exponential	ML	0.861	353.68	1.295	0.40	0.46	1.39	0.001	0.74

Co=nugget, Co+C1=Sill, DSD= Degree of spatial dependence, CVC= cross validation coefficient, RMSE: root mean square error, RE= reduced error, SDRE= standard deviation of reduced error.
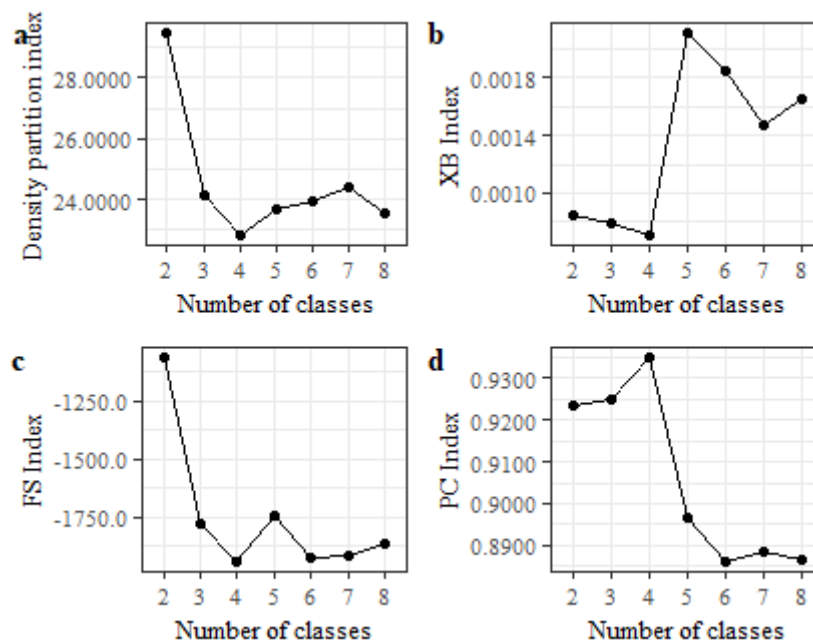

Figure 5: Fuzzy c-means algorithm method indexes used to establish the optimal number of classes or groups. a) Partition density index, b) Xie and Beni index, c) Fukuyama and Sugeno index, and d) Partition coefficient index.
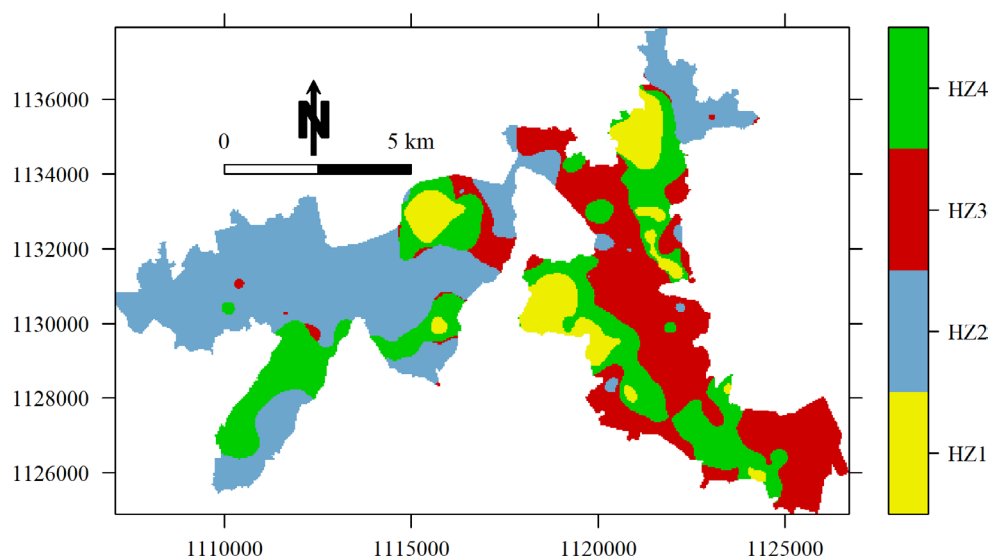
Figure 6: Homogeneous zones (HZ) defined by the fuzzy c-means algorithm in the first two axes of the spatial components of the study area.

Table 4: Chemical properties of the homogeneous zones defined by the fuzzy c-means algorithm in the first two spatial components of the study area.

Chemical soil properties	Homogeneous zones			
	1-2sPC	2-2sPC	3-2sPC	4-2sPC
	Mean ± SE	Mean ± SE	Mean ± SE	Mean ± SE
pH	4.54±0.15	5.98±0.08	6.11±0.13	4.47±0.09
OM (%)	15.88±0.99	4.22±0.21	6.17±0.44	10.02±0.51
EC (dS m^{-1})	3.19±0.21	0.75±0.07	1.78±0.15	1.54±0.12
S (meq 100 g^{-1})	365.59±44.2	44.64±4.38	124.31±14.62	183.98±18.50
Al (meq 100 g^{-1})	2.47±0.41	0.20±0.05	0.28±0.09	2.45±0.22
ECEC (meq 100 g^{-1})	41.29±3.06	11.99±0.57	22.50±1.21	16.77±1.26
P (ppm)	10.15±1.50	16.20±1.23	15.48±1.58	10.93±1.15
K (meq 100 g^{-1})	0.96±0.09	0.89±0.07	1.10±0.09	0.85±0.08
Na (meq 100 g^{-1})	3.39±0.44	0.69±0.07	1.72±0.22	1.13±0.15
Ca (meq 100 g^{-1})	31.81±2.75	8.05±0.47	16.85±1.14	10.89±1.20
Mg (meq 100 g^{-1})	2.66±0.37	2.16±0.12	2.54±0.24	1.46±0.12
Fe (meq 100 g^{-1})	153.50±6.71	164.11±5.69	142.09±6.58	154.42±5.47
Cu (meq 100 g^{-1})	0.27±0.08	1.37±0.09	1.12±0.09	0.37±0.08
Mn (meq 100 g^{-1})	10.87±1.08	23.34±1.09	16.11±1.19	7.61±0.55
Zn (meq 100 g^{-1})	3.38±0.42	6.27±0.28	5.50±0.39	3.51±0.31
B (meq 100 g^{-1})	0.84±0.05	0.37±0.01	0.52±0.03	0.53±0.04

pH= Potential of hydrogen, OM= Organic matter, EC=Electrical conductivity, S= Sulfur, Al=Aluminum, ECEC= Effective cation exchange capacity, P=Phosphorus, K=Potassium, Na= Sodium, Ca= Calcium, Mg= Magnesium, Fe= Iron, Cu= Copper, Mn= Manganese, Zn= Zinc, B= Boron, SE= Standard error.

produces a decrease in the pH to levels not tolerable by plants. The extreme acidity has determinant implications for soil toxicity, including an increase in aluminum and iron solubility. For this reason, the nutritional quality of this soil can become severely deficient. Similarly, in the root zone, acidity causes a loss in plant productivity, and acid runoff causes adverse environmental impacts (Rincón; Castro; Gómez, 2008).

HZ2 was established as *"Areas with a high capacity for self-neutralization"*. This zone differed because its average pH was 5.98, that is, slightly acid, its sulfur levels were lower, as well as its contents of all variables, such as Al, Ca, Na and OM, in contrast with higher values in P, Zn, Mn, Cu and Fe.

HZ3 was designated as *"Sulfated areas with slightly acidic pH"*. Its main characteristics are high levels of S and an average pH of 6.11. Moreover, Ca levels are higher compared to those of HZ2, with a consequent decrease in Cu, Zn and Mn.

Finally, HZ4, *"Areas with high acidity and sulfur levels"* is characterized by having an average pH of 4.47, high sulfur levels, although lower compared to those of HZ1, high Al content and lower EC and ECEC values in relation to those of HZ1.

CONCLUSIONS

The homogeneous zones (HZs), considering the soil index of the two sPCs, were spatially related to the behavior of the chemical properties. The Moran index (MI) and the incorporation of a distance matrix determined the analysis of the spatial variability in the components generated. Because the soils in this area are called acid sulfate soils, the homogeneous zones were strongly influenced by S, as well as by properties associated with constant crop management, such as acidity (pH), Al contents, % OM and ECEC. In the case of the cultivable areas of HZs1 and HZs4, practices such as washing before liming, liming and green coverings are recommended.

ACKNOWLEDGMENTS

REFERENCES

AGHASI, B. et al. Sub-basin scale spatial variability of soil properties in Central Iran. **Arabian Journal of Geosciences**, 10(6):136. 2017.

ARROUAYS, D. Large trends in French topsoil characteristics are revealed by spatially constrained multivariate analysis. **Geoderma**, 161(3-4):107-114, 2011.

BACHMEIER, O. A.; BUFFA, E. Variabilidad espacial de un suelo bajo vegetación de *Prosopis* sp. **Turrialba**, 42:365-370. 1992.

BERNAL, A. A.; FORERO, F. E. Evaluación de especies vegetales para el manejo de la acidez en suelos sulfatados ácidos de Paipa, Boyacá. **Revista Corpoica: Ciencia y Tecnología Agropecuaria**, 15(2):229-236, 2014.

CAMBARDELLA, C. A. et al. Field-scale variability of soil properties in central Iowa soils. **Soil Science Society of America Journal**, 58:1501-1511, 1994.

CASTRO, H. E. et al. Diagnostic and controlling acidity in acid sulfate soils by using incubation tests at the Upper Chicamocha basin, Boyaca. **Agronomía Colombiana**, 24(1): 122-130, 2006.

CLIFFORD, P.; RICHARDSON, S.; HÉMON, D. Assessing the significance of the correlation between two spatial processes. **Biometrics**, 45:123-134, 1989.

COMBATT, E.; PALENCIA, G.; MARTÍN, N. Clasificación de suelos sulfatados acidos según azufre extraible en los municipios del Medio y Bajo Sínu. **Temas Agrarios**, 8(2):22-29, 2003.

CÓRDOBA, M. et al. Identificación de zonas de manejo sitio-específico a partir de la combinación de variables de suelo. **Revista Corpoica: Ciencia y Tecnología Agropecuaria**, 13(1):47-54, 2012.

CORDOBA, M. et al. Subfield management class delineation using cluster analysis from spatial principal components of soil variables. **Computers and Electronics in Agriculture**, 97:6-14, 2013.

CÓRDOBA, M. et al. Protocol for multivariate homogeneous zone delineation in precision agriculture. **Biosystems Engineering**, 143(1):95-107, 2016.

CORTÉS, D. L.; CAMACHO-TAMAYO, J. H.; GIRALDO, R. Spatial prediction of soil penetration resistance using functional geostatistics. **Scientia Agricola**, 73(5):455-461, 2016.

COX, M. S.; GERARD, P. D.; ABSHIRE, M. J. Selected soil properties variability and their relationships with yield in three Mississippi fields. **Soil science**, 171(7):541-551, 2006.

CRESSIE, N. **Statistics for Spatial Data**. John Wiley & Sons, New York, 1993. p.928.

DENT, D. Acid sulfate soils: A baseline for research and development. **International institute of land reclamation and improvement**. Publication 39. Wageningen. 1986. 204p.

DÍAZ, V. M. **Geoestadística aplicada**. Instituto de Geofísica, UNAM e Instituto de Geofísica y Astronomía, CITMA, Cuba, 2002. 131p.

DRAY, S.; SAÏD, S.; DÉBIAS, F. Spatial ordination of vegetation data using a generalization of Wartenberg's multivariate spatial correlation. **Journal of Vegetation Science**, 19(1):45-56, 2008.

FARACO, M. A. et al. Seleção de modelos de variabilidade espacial para elaboração de mapas temáticos de atributos físicos do solo e produtividade da soja. **Revista Brasileira de Ciência do Solo**, 32(2):463-476, 2008.

FORERO, S.; CASTILLO, M. Diagnóstico de emprendimiento en el fortalecimiento empresarial en la provincia del Tundama, Colombia. **Ingeniería Industrial. Actualidad y Nuevas Tendencias**, 9(5):139-152, 2016.

GAVIOLI, A. et al. Optimization of management zone delineation by using spatial principal components. **Computers and Electronics in Agriculture**, 127:302-310, 2016.

GIANNINI, F. et al. Enhancing edaphoclimatic zoning by adding multivariate spatial statistics to regional data. **Geoderma**, 310:170-177, 2018.

GÓMEZ, M. et al. Optimización de la producción y calidad en cebolla cabezona (*Allium cepa*) mediante el balance nutricional con magnesio y micronutrientes (B, Zn y Mn), Valle Alto del Río Chicamocha, Boyacá. **Agronomía Colombiana**, 25(2):339-348, 2007.

HENGL, T. Finding the right pixel size. **Computers & geosciences,** 32(9): 1283-1298, 2006.

HERNÁNDEZ, W. et al. Variabilidad espacial del pH y del contenido de Fe_2O_3 en suelos de la cuenca del río Tabure del Estado Lara. **Ciencia y Tecnología**, 11(1):19-27, 2018.

JARAMILLO, D. F. Variabilidad espacial de las propiedades ándicas de un Andisol hidromórfico del Oriente Antioqueño (Colombia). **Revista Facultad Nacional de Agronomía Medellín**, 62(1),4907-4921, 2009.

JARAMILLO, D. F. Variabilidad espacial del suelo: Bases para su estudio. **Revista de La Facultad de Ciencias**, 1:73-87, 2012.

JARAMILLO, D. F.; SADEGHIAN, S.; LINCE, L. A. Variabilidad espacial de las bases en un Andisol de la zona cafetera central colombiana. **Boletín de Ciencias de la Tierra**, 33: 111-124, 2013.

JARAMILLO, D. F. **El suelo: Origen, Propiedades, Espacialidad**. Universidad Nacional de Colombia. Medellín, 2014. 553p.

JOHANN, J. A. et al. Variabilidade espacial da rentabilidade, perdas na colheita e produtividade do feijoeiro Spatial variability of profitability, harvest losses and productivity of beans. **Engenharia Agrícola**, 30(4):700-714, 2010.

LAURENT, A. G. The lognormal distribution and the translation method: Description and estimation problems. **Journal of the American Statistical Association**, 58(301):231-235. 1963.

LAL, R. Soil carbon dynamics in cropland and rangeland. **Environmental Pollution**, 116(3): 353-362, 2002.

LI, H. Y.; WEBSTER, R.; SHI, Z. Mapping soil salinity in the Yangtze delta: REML and universal kriging (E-BLUP) revisited. **Geoderma**, 237:71-77, 2015.

MATHERON, G. **Le krigeage universel**. vol 1. Cahiers du Centre de Morphologie Mathématique, École des Mines de Paris, Fontainebleau, 1969. 82p.

McGRAW, T. Soil test level variability in Southern Minnesota. **Better Crops Plant Foods**, 78:24-25, 1994.

MEYER, D. et al. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. **R package version 1.6-8**, 2017.

MORAL, F. J.; TERRÓN, J. M.; DA SILVA, J. M. Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. **Soil and Tillage Research**, 106(2):335-343, 2010.

MZUKU, M. et al. Spatial variability of measured soil properties across site-specific management zones. **Soil Science Society of America Journal**, 69:1572-1579, 2005.

ODEH, I. O. A.; McBRATNEY, A. B.; CHITTLEBOROUGH, D. J. Soil pattern recognition with fuzzy c-means: Application to classification and soil-landform interrelationships. **Soil Science Society of America Journal**, 56:505-516, 1992.

PERALTA, N. R. et al. Delineation of management zones to improve nitrogen management of wheat. **Computers and Electronics in Agriculture,** 110:103-113, 2015.

RAHAL, N. S. Characterization of spatial variability of some physiochemical soil properties of mesopotamian plain soils. **European Journal of Agriculture and Forestry Research**, 3(3):1-16, 2015.

R CORE TEAM. **R**: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Versión 3.4.4. 2018.

REICHARDT, K.; TIMM, L. C. **Solo, planta e atmosfera: Conceitos, processos e aplicações**. São Paulo, Manole, 2008. 478p.

RINCÓN, Á.; CASTRO, H.; GÓMEZ, M. Caracterización física de los suelos sulfatados ácidos del Distrito de Riego del Alto Chicamocha (Boyacá) y su aplicación al manejo. **Agronomía Colombiana,** 26(1):134-145. 2008.

RODRÍGUEZ, F. A.; CAMACHO, J. H.; RUBIANO, Y. Variabilidad espacial de los atributos químicos del suelo en el rendimiento y calidad de café. **Corpoica Ciencia y Tecnología Agropecuaria**, 17(2):237-254, 2016.

RODRIGUES, M. S.; CORÁ, J. E. Management zones using fuzzy clustering based on spatial-temporal variability of soil and corn yield. **Engenharia Agrícola**, 35(3):470-483, 2015.

ROSEMARY, F. et al. Exploring the spatial variability of soil properties in an Alfisol soil catena. **Catena**, 150:53-61, 2017.

SELBY, B.; KOCKELMAN, K. M. Spatial prediction of traffic levels in unmeasured locations: Applications of universal kriging and geographically weighted regression. **Journal of Transport Geography**, 29:24-32, 2013.

STOYAN, H. et al. Spatial heterogeneity of soil respiration and related properties at the plant scale. **Plant and Soil**, 222(1-2):203-214, 2000.

ROSENBERG, M. S.; ANDERSON, C. D. PASSaGE: Pattern Analysis, Spatial Statistics and Geographic Exegesis. Version 2. **Methods in Ecology and Evolution**, 2(3):229-232, 2011.

TRIPATHI, R. et al. Delineation of soil management zones for a rice cultivated area in eastern India using fuzzy clustering. **Catena**, 133:128-136, 2015.

ROSICKY, M. A.; SULLIVAN, L. A.; SLAVICH, P. G. Factor contributing to the acid soil scalding process in coastal floodplain of New South Wales. **Australian Journal of Soil Research**, 42:587-594, 2004.

VARÓN-RAMÍREZ, V. M.; CAMACHO-TAMAYO, J. H.; GONZÁLEZ, J. Management zones determination based on physical properties of the soil. **Ciência e Agrotecnologia,** 42(3): 248-260, 2018.

WARRICK, A. W.; NIELSEN, D. R. Spatial variability of soil physical properties in the field. In: HILEEL, D, (ed.). **Applications of soil physics**. Nueva York: Academic Press, 1980, p.319-344.

WANG, Y. Q.; SHAO, M. A. Spatial variability of soil physical properties in a region of the Loess Plateau of PR China subject to wind and water erosion. **Land Degradation & Development**, 24:296-304. 2013.

WELCH, S. et al. Mineralogical control of rare earth elements in acid sulfate soils. **Geochimica et Cosmochimica Acta**, 73:44-64, 2009.