

Management units based on the physical properties of an Oxisol

J.H. Camacho-Tamayo^{1*}, Y. Rubiano Sanabria², L. M. Santana³

¹*Facultad de Ingeniería. Programa de Ingeniería Agrícola. Universidad Nacional de Colombia, Sede Bogotá, Colombia. Ciudad Universitaria.* ²*Facultad de Agronomía. Universidad Nacional de Colombia, Sede Bogotá, Colombia.* ³*Departamento de Geografía, Universidad del Valle. Cali, Colombia.* *Corresponding author: jhcamachot@unal.edu.co

Abstract

In Colombia, the zone of the greatest agricultural expansion includes soils within the native savanna system. These soils are Oxisols, which have a stable structure but are susceptible to physical degradation. The objective of this paper was to study the spatial variability of the soil physical properties with the aim of determining management units. This study was carried out in the municipality of Puerto Gaitán, Meta, Colombia, in an area of 5100 ha where nodes were placed every 320 m in a grid. The samples were taken from the surface (A_H) and subsurface (B_H) horizons in which the texture, pore distribution, saturated hydraulic conductivity (K_s), bulk density (B_D) and soil particle density (P_D) were determined. The data analysis was conducted using descriptive statistics, geostatistics and multivariate analysis. The predominant textures of the soil were loam and sandy loam. High values of B_D , high values of microporosity, low presence of mesopores and slow K_s were found, which result in physical limitations for adequate development of crops. The B_D and the total porosity were the properties that best characterized the soil and showed the greatest number of significant correlations with the other physical properties. The K_s was the only property that did not show spatial dependence. Geostatistics allowed the identification of spatial variability patterns, estimation of the properties at non-sampled sites, and identification of management units of the properties that fit the semivariogram models. Multivariate analysis identified the most representative properties of the soil and permitted the establishment of management units.

Keywords: Geostatistics, multivariate analysis, pedometrics, spatial correlation, soil management

1. Introduction

Oxisols are the seventh most common soils in the world and occupy an area of approximately 9.61 million km². They are found principally in the tropics in flat or slightly undulating zones (USDA, 1999), and they represent 24.5% of soils. In Latin America, Oxisols cover approximately 2.43 million km² and occupy large areas in Brazil, Colombia and Venezuela,

specifically in zones occupied by natural savannas, which are the second largest biome in South America (Thomas and Ayarza, 1999).

In Colombia, Oxisols appear in the so-called Llanos Orientales (Eastern Plains), which cover more than 20 million hectares (Phiri *et al.*, 2001) of which more than

3 million are subjected to mechanized crop production. This is the agricultural zone of the greatest development during the last decade, and its use is passing from extensive cattle raising to technified crops, such as rice, corn, soy, palm oil, sugar cane, rubber, fruit trees and timber. The nutritional requirements for these crops contrast with the low fertility of the soil, which makes the use of correctives and fertilizers necessary because of its low content of interchangeable organic matter, P, macronutrients and micronutrients in addition to its high acidity (Camacho-Tamayo *et al.*, 2008).

From the physical point of view, these Oxisols have a stable structure and natural densification. Nevertheless, Oxisols are susceptible to degradation once the original condition of the soil is altered through farming and planting of crops. Oxisols properties can degrade over a period of 5 years if conventional tillage is utilized, which in conjunction with the high and intense precipitation in the zone affects the size of the aggregates, diminishes the porous volume and reduces the water flow in the soil (Phiri *et al.*, 2001).

The intrinsic variability of soil is due to edaphic and climatic factors that influence the biological, physical and chemical processes, which act simultaneously with differing intensity and are associated with its genesis (Malagón, 2003). To determine and evaluate the characteristics and processes as well as their variability, numerous techniques of soil survey and classification have been used. The resulting information is represented in maps that show the geographic distribution of the soils where regions of units are defined using the similar characteristics of texture, topography, horizon thickness, nutrient level, and slope (Moral *et al.*, 2011). However, the properties of soil can vary significantly even in small inside areas for the same unit of soil (Nyamangara *et al.*, 2013).

A first estimation of the variability of soil can be established by means of univariate techniques through

descriptive measures, such as the mean, maximum value, minimum value, mode, or coefficient of variation (CV). The use of these techniques has limitations because it analyzes the properties of the soil individually and does not permit visualizing the different interactions between them (Cruz *et al.*, 2011). The CV is sensitive to the distance of the sampling and to the number of samples with greater distances of the sampling resulting in greater values of CV (Garten *et al.*, 2007). Moreover, an increasing number of samples results in diminished CV values (Dong-Sheng *et al.*, 2011), which can erroneously lead to declaring low variability. Furthermore, CV does not permit defining the spatial variability because this technique does not provide information about the pattern of spatial distribution of a particular property, which makes it impossible to establish management units.

Knowing the spatial variability of soil permits a better understanding of the complex relationships between its characteristics and environmental factors (Goovaerts, 1998), and it also helps in determining specific management practices for adequate use of the soil (Orjuela Matta *et al.*, 2012). Recognition of the variability leads to reevaluating the conventional soil management in agricultural production, which assumes a homogeneous behavior of soil properties and uses general recommendations (Tesfahunegn *et al.*, 2011). However, it is necessary to consider that the determination of the variability requires an accurate and reliable sampling to avoid large errors of assessment, and it is necessary to carry out sampling by means of the identification of groups of interrelated soil properties (Peña *et al.*, 2009).

Multivariate analysis allows the identification of the relationships among a group of properties or samples studied. Principal component and cluster analyses are widely used for classification, modeling, and evaluation in environmental studies. In the study of soil properties, multivariate analysis is a tool that allows the characterization of an area and the establishment of the interaction with the surroundings of each resulting property and group. While it is true

that these techniques do not in themselves allow the identification of the spatial distribution of a group of samples, they help to classify and group them. With this grouping, it is possible to determine areas and establish management units allowing appropriate uniform management to be carried out within the units.

Physical properties play a key role in fertility and are a basic component of soil quality (Amézquita *et al.*, 2004), and their evaluation has been proposed to define specific management units (Cucunubá-Melo *et al.*, 2011). In particular, soil bulk density (B_D) influences the porosity and, therefore, the movement and storage of fluids (water and air) in soil. Particle density (P_D) depends on the mineral and organic composition of soil because the specific crop management and organic materials used can lead to noticeable spatial variability. Texture is a fundamental physical property with great influence on the B_D and porosity, and it also affects the heating capacity and conductance of soil. Furthermore, clays contribute to the stabilization of organic carbon and the retention of anions and cations in soil, which affect the processes of percolation and the stripping/retention of nutrients as well as the agrochemicals in soil (Godwin and Miller, 2003).

Considering the importance that soils have in the eastern plains of Colombia for the immediate future for Colombian agricultural production, the aim of the present research was to study the spatial variability of the physical properties of the dominant soils (Oxisols) through the use of geostatistical and multivariate analyses to determine soil management units.

2. Materials and Methods

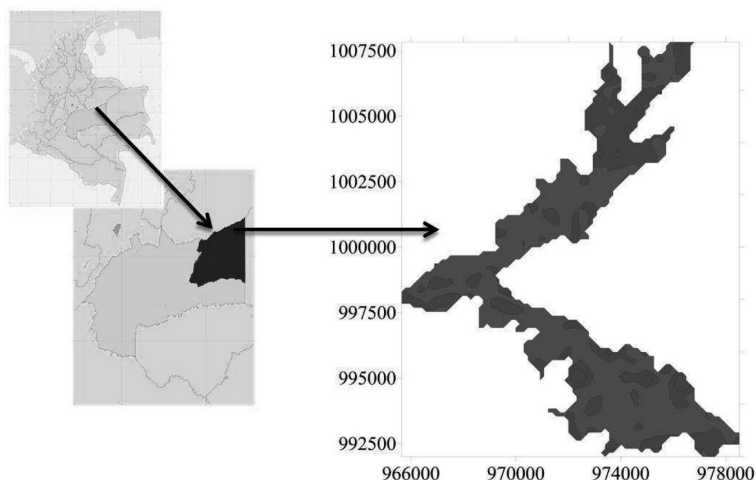
2.1. Locality and characterization of the study area

The study was carried out in the Carimagua Experimental Station located in the municipality of

Puerto Gaitán, Meta, Colombia with geographical coordinates of 4° 37' N and 71° 19' W and at an average altitude of 175 m (Figure 1). The zone has a slightly undulating relief with slopes between 2 and 5%, and the zone is covered with native savanna (used for more than 30 years for extensive cattle raising). Moreover, the zone has a sub-humid tropical climate with an average temperature of 27.8 °C and an average annual precipitation of 2,240 mm, which is concentrated between the months of April and November. The predominant soils of the zone are highly fertilized Oxisols, which are strongly acidic ($pH < 5$) and have low organic matter contents. The soil under study belongs to the Carimagua-Tomo complex with taxonomic components of Typic Haplustox and Tropectic Haplustox.

2.2. Field sampling and laboratory analysis

The sampling was carried out in an area of approximately 5,100 ha between the months of February and June in 2010. A grid design was used with sampling points distributed perpendicularly every 320 m for a total of 470 points. Each point was georeferenced with a GPS (approximation of ± 1 m), and the sampling of the soil was carried out on surface horizon (A_H) and subsurface horizon (B_H) for a total of 940 samples. At each point and horizon, disturbed samples were taken for the determination of soil particle density (P_D) using the pycnometer method and sand content using the Bouyoucos method. Additionally, an undisturbed sample was taken with a ring of 50 mm in height and 50 mm in diameter to determine the saturated hydraulic conductivity (K_S ; using the constant head well permeameter method), bulk density (B_D), saturation points (PS; 0 bar), field capacity (FC; 0.1 bar) and permanent wilting point (PWP; 15 bar) on ceramic pressure plate extractors. On the basis of these determinations, the total porosity (T_p) and the distribution of the macropores (M_{AP}), mesopores (M_{EP}) and micropores (M_{IP}) were established.



2.3. Data processing and statistical analysis

With the results of the laboratory analysis for the properties mentioned, an initial exploratory study of the data was carried out to calculate measures of placement (mean, median, minimum, and maximum), variability (coefficient of variation; CV) and form (skewness and kurtosis). Thus, the supposition of normality was empirically evaluated, which, while not indispensable, gives better predictions when geostatistical (Diggle and Ribeiro, 2000) and multivariate techniques are used. For the CV analysis, the Warrick and Nielsen (1980) classification was considered, which indicates low variability for values less than 12%, medium variability for values between 12 and 60%, and high variability for values greater than 60%.

To identify the properties that best characterize the horizons of the soil, principal components (PC) analysis was carried out. PC analysis reduces the dimensionality and shows the different basic components, which are called principal components. Each principal component is described in terms of new components, which are defined based on a

linear combination of the original variables. The first component, which is associated with the greatest eigenvalue, represents the maximum value of the total of the variance, and the second component is the second linear combination and is not correlated with the first component, which represents the maximum residual variance, and so on until the total variance is accounted. In general, a small number of components explain a high percentage of the total variance, so the entire set of data can be described in a smaller dimensional space. To facilitate the interpretation of each component in this study, Varimax rotation was applied after previously standardizing the values with a mean zero (0) and variance unit (1). The exploratory study of the data and the multivariate statistical analysis were carried out using the SPSS v.20 program (SPSS Inc., Chicago, IL, USA).

The experimental semivariograms of the properties under consideration were then calculated. This analysis assumes that the values of a property separated by small distances are more similar than those located at greater distances. The function of the experimental semivariogram $\gamma(h)$ is defined by Equation (1) as

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x) - Z(x+h)]^2 \quad (1)$$

where $Z(x)$ is the value of the variable at site x ; $Z(x+h)$ is an additional sampling value separated from the first by a distance h ; and $N(h)$ is the number of pairs that are separated by h . The function of the semivariogram is calculated for various distances (h). In practice, due to the irregularity of the sampling, distance intervals ($[0, h]$, $[h, 2h]$, $[2h, 3h]$, ...) are taken, and the experimental semivariogram corresponds to an average distance between pairs of sites within each interval and not to a specific distance (h). The solution of the problem of prediction in geostatistics requires knowledge of the structure of spatial autocorrelation for any possible distance between sites within the area under study. Thus, the fit to the models that determine the value of the semivariogram for any distance is necessary.

There are diverse theoretical models of semivariance that can be fit to the experimental semivariogram. Webster and Oliver (2007) discussed the characteristics and conditions with which theoretical models must satisfy. In general, such models can be divided into unbounded (linear, logarithmic and potential) and bounded (spherical, exponential and Gaussian). All of these models have three parameters in common, which are the nugget (C_0), sill ($C_0 + C_1$) and range (a). The nugget effect indicates the discontinuity between the samples, so the spatial variability is not detected during the sampling process. The sill is the value of the semivariance where the model stabilizes, thereby exhibiting a constant value. The range represents the distance to which a spatial correlation exists, thereby indicating that there is no longer any correlation between the samples. In the present study, bounded models were employed to fit the experimental semivariograms. The spherical model is defined as follows:

$\gamma(h) = C_0 + C_1 * [1.5*(h/a) - 0.5*(h/a)^3]$ for $0 < h < a$
 $\gamma(h) = C_0 + C_1$ for $h > a$. The exponential model is defined as follows: $\gamma(h) = C_0 + C_1 * [1 - e^{-(3h/a)}]$ for $0 < h < d$; where d is the maximum distance at which

the semivariogram is defined. The Gaussian model is defined as follows: $\gamma(h) = C_0 + C_1 * [1 - e^{-(3h^2/a^2)}]$.

Once the model of best fit for each property was established, the degree of spatial dependence (DSD) was verified through the relationship between the nugget and sill ($C_1/C_0 + C_1$). The DSD is classified as follows: strong if it is greater than 75%; moderate if it is between 25 and 75%; and weak if it is less than 25% (Cambardella, 1994). It is important that the nugget is not greater than 50% of the value of the sill so that the spatial correlation model will describe reality well (Cressie, 1993). In the other cases, the noise in the measurements would explain to a greater extent the spatial variability than the correlation of the phenomenon. In these cases, the model fit to the experimental semivariogram is called the nugget effect (Goovaerts, 1998) and is defined as follows: $\gamma(h) = C_0$ for $h > 0$.

For the geostatistical analysis, GS+ v.9 software (Gamma Design Software, LLC, Plainwell, MI, USA) was used. This software allowed the selection of the theoretical semivariogram models on the basis of the least value of the sum of the residual squares, the coefficient of determination (R^2) of the fitting equation and the similar values obtained between the real value and the estimated value obtained in the cross validation.

On the basis of the semivariogram models of the properties that expressed spatial dependence, the interpolation by the ordinary kriging method was performed, which is considered to be the best linear unbiased estimator of minimum variance (Diggle and Ribeiro, 2000) for making predictions at non-sampled sites. The results are shown by means of contour maps. This procedure was carried out with the Surfer v.10 program (Golden Software Inc., Golden, CO, USA). Furthermore, a linear correlation was established to determine the relationship between the properties, which can be visually verified on contour maps when they are significant.

To identify the management units, three methods were used. The first method was cluster analysis (CA), and the second and third methods were estimated on the basis of a soil index (SI) by means of principal component analysis (PC) and coefficient of variation (CV), respectively. In the CA, the results obtained from each property were included, and all sites of the sampling (observations of cases) were considered. The CA allowed the identification of observation groups that had some similarities between them. The difference among the groups is a distance measure. In the present study, the Euclidian distance was used to separate a set of sampled points into groups, and the results are shown in the form of a graph (dendrogram) to facilitate the identification of the groups formed by the analyzed samples to ultimately establish homogeneous units. The results were obtained using the Ward algorithm, which calculates the distance of the sum of the squares between two groups, which extends along all points. At each estimate, the sum of the squares between the groups is minimized along all segmentations by fusion of the two groups previously estimated. In general, this method is considered to be efficient, but it creates groups of small dimensions (Ferreira, 2011).

According to Ortega and Santibañez (2007), a SI corresponds to a linear combination of the properties of the selected soil where the weight of each property is chosen, so the difference (variation) among the sampling points is maximized according to Equation (2) as follows:

$$SI_z = \sum w_i P_{iz} \quad (2)$$

where SI_z is the soil index in position z of the sampling and w_i is the weight of the standardized variable P_{iz} in position z of the sampling. This SI is a continuous variable formed by the properties of the soil that can be classified and delimited with a representation on a map. The weights for each variable can be arbitrary or can be determined mathematically.

For the PC, an SI can be obtained on the basis of the variance of each component (w_i) and the factor score

(P_{iz}). In this case, the definition of the management units of the soils can be conducted through the consideration of percentiles, quartiles, or the average and its relation to the standard deviation for which it is convenient to know the zone under study and the spatial behavior of the soil properties. These management units will show a high correspondence of correlation with several of the properties, principally with those that are found in the first PC. According to the variance of the PCs, it is possible to also show a high correspondence with the properties immersed in the other principal components.

For the SI based on the CV, it is assumed that the properties that show a greater variability in the field, as expressed through the coefficient of variation (CV), have greater weight in a linear model that combines all of the soil properties for estimating a soil index (Equation 2). For this analysis, the standardized values in the CP and the values of CV obtained in the exploratory data analysis were taken. The relative weight of each variable was obtained using Equation (3) as follows:

$$w_i = \frac{CV_i}{\sum CV_i} \quad (3)$$

where w_i is the weight of variable i ; and CV_i is the coefficient of variation of variable i .

Once the management units for each method were identified, maps of the management units for each horizon were drawn through triangulation with linear interpolation. Finally, correlations between the units and each property were estimated to establish their correspondence and influence on the management units.

3. Results and Discussion

3.1. Descriptive analysis

According to the sand content, loam textures prevailed with the predominance of fine particles in the two horizons (Table 1) as was also reported for the zone by

Molina *et al.* (2003). The bulk density (B_D) and soil particle density (P_D) showed values and ranges typical for these Oxisols in the native savanna (Phiri *et al.*, 2001) and indicated natural processes of densification, which affected the presence of macropores (M_{AP}) and mesopores (M_{EP}) as well as the saturated hydraulic conductivity (K_S).

The K_S showed values less than that of loamy soils according to Clapp and Homberger (1978). These K_S values indicated that the water flow was slow in the two horizons. However, there were some places in the two horizons with moderate K_S values but with some restrictions of water flow. Despite observing that the T_p was less in the B_H , the K_S was slightly greater in this horizon than in the A_H due to the greater presence of M_{AP} and M_{EP} , which indicated a greater water storage capacity in this horizon as well as a greater range of water available to the plants. The behavior of the K_S and the greater presence of M_{IP} on the surface suggested that the natural process of densification in addition to being stepped on by cattle affected

this property. The T_p oscillated between 37.85 and 56.25% for the A_H with an average of 47.58%, which were values similar to those reported by Phiri *et al.* (2001) for native savanna on recent intervened soils. These ranges were similar for the B_H with less predominance of M_{IP} .

According to Warrick and Nielsen (1980), the K_S showed high variability for the two horizons, which is a property that does not fit with a symmetric distribution, as verified by comparing the mean and median as well as by the skewness and kurtosis values (Table 1). In general, K_S is a property that shows high variability (Cucunubá-Melo *et al.*, 2011), especially at the surface level or in highly intervened soils (Zimmermann and Elsenbeer, 2008). Sand, M_{AP} and M_{EP} showed medium variability for the two horizons. The M_{IP} and T_p showed low variability in the two horizons similar to the B_D and P_D , which are properties that help to define the porous space of soil. In general, soil particles, B_D and P_D present low or medium variability (Peña *et al.*, 2009).

Table 1. Descriptive analysis of soil properties in the Carimagua Experimental Station.

Property	Mean	Median	CV, %	Minimum	Maximum	Skewness	Kurtosis
A horizon							
K_S , cm h ⁻¹	0.20	0	186.32	0	2.04	2.70	7.94
M_{AP} , %	4.27	4.35	54.17	0.34	10.67	0.33	-0.59
M_{EP} , %	6.63	6.64	45.90	1.14	15.72	0.27	-0.52
M_{IP} , %	36.33	36.58	8.42	26.54	44.37	-0.56	0.66
T_p , %	47.58	47.67	7.10	37.85	56.25	-0.14	0.17
B_D , g cm ⁻³	1.45	1.45	5.36	1.24	1.68	0.21	0.11
P_D , g cm ⁻³	2.65	2.66	1.96	2.51	2.79	-0.12	0.24
Sand, %	26.27	26.00	19.94	13.00	43.00	0.20	0.25
B horizon							
K_S , cm h ⁻¹	0.26	0.13	127.04	0	1.53	1.56	2.22
M_{AP} , %	4.84	4.74	57.99	0.43	13.15	0.47	-0.47
M_{EP} , %	7.54	7.31	51.45	0.00	18.38	0.50	-0.27
M_{IP} , %	32.99	33.37	10.83	23.2	41.89	-0.30	0.22
T_p , %	45.52	45.39	8.59	35.44	54.86	-0.05	-0.31
B_D , g cm ⁻³	1.53	1.53	7.54	1.22	1.83	-0.06	-0.20
P_D , g cm ⁻³	2.70	2.70	1.69	2.57	2.82	0.09	-0.04
Sand, %	22.77	23.00	18.05	13.00	35.00	0.32	0.22

With the exception of K_s , the mean and median values were similar for the other properties, which indicated symmetric distributions, as confirmed by the values close to zero for skewness and kurtosis, thereby empirically corroborating the precept of normality. More importantly than normality, one should verify that the normal distribution does not show any elongated tails because they can compromise the results that are

obtained based on the observed data (Cressie, 1993), especially when interpolation with kriging is carried out, which is a method that bases the estimations on mean values (Warrick and Nielsen, 1980). Another important element to verify is the occurrence of the proportional effect between the mean and variance of the samples in space because this effect allows the estimation of sills defined in semivariogram models.

Table 2. Coefficients of the first four principal components (PCs) obtained using Varimax rotation for the soil physical properties.

Property	PC 1	PC 2	PC 3	PC 4	Communality
A horizon					
K_s	0.03354	0.05966	0.86754	0.13411	0.77530
M_{AP}	0.64130	0.15549	-0.00425	0.31013	0.53164
M_{EP}	0.09860	-0.95858	0.02048	-0.06118	0.93276
M_{IP}	0.55785	0.69361	0.08823	-0.20984	0.84410
T_p	0.94922	0.07494	0.09597	-0.07335	0.92122
B_D	-0.84271	0.13277	0.08899	0.13151	0.75301
P_D	0.18138	0.20264	0.62497	0.39310	0.61908
Sand	-0.14761	-0.03799	-0.04380	0.85905	0.76311
Eigenvalue	2.59	1.40	1.15	1.00	
Variance, %	32.31	17.54	14.38	12.53	
Cum. Variance, %	32.31	49.85	64.23	76.76	
B horizon					
K_s	0.06845	-0.10798	0.06068	0.864385	0.76719
M_{AP}	0.69536	0.08259	-0.32329	-0.142065	0.61505
M_{EP}	0.24569	-0.85297	0.12405	0.169219	0.83194
M_{IP}	0.28943	0.85125	0.19592	0.067290	0.85131
T_p	0.93187	0.11993	0.08282	0.117489	0.90343
B_D	-0.88499	0.15204	-0.11615	-0.016217	0.82008
P_D	-0.03396	0.38017	-0.34046	0.449046	0.46324
Sand	-0.06710	-0.06215	-0.87818	-0.032562	0.78062
Eigenvalue	2.32	1.66	1.05	1.00	
Variance, %	28.98	20.74	13.16	12.53	
Cum. Variance, %	28.98	49.72	62.88	75.41	

Loading factors higher than 0.6 (absolute value) are shown in bold.

3.2. Analysis of the properties by principal components

In the two horizons, the principal components (PCs) were analyzed with an eigenvalue greater than one (1) to include an adequate interval. As reflected in the accumulated variance, the first four PC explained more than 75% of the total variance (Table 2). For the A_H and B_H , the coefficients of the M_{AP} and P_D components as well as the communality value were verified as being low, which indicated that these properties had scarce correlation with the other attributes as well as low representation in the principal components analyzed. Thus, these properties were considered to not be representative for characterizing the soil under study.

The PC analysis was similar for the two horizons. The properties that stood out most for each component were the same for PC1 and PC2, and the components showed accumulated variances with similar values in the two horizons. The first PC incorporated 32.31 and 28.98% of the variance for the A_H and B_H , respectively, and the M_{AP} , T_p and B_D were the properties that best represented this component. The coefficient values of these properties showed that B_D was inversely related to T_p .

The PC2 showed a variance of 17.54 and 20.74% for the A_H and B_H , respectively, and the properties that best defined this component were M_{EP} and M_{IP} . The third and fourth components showed the properties that had the least influence on the variance, thus indicating that the properties that should be least considered for characterizing the soil, including K_s and the sand content for the two horizons and B_D for the A_H . These results indicated that it is convenient to consider first the T_p and B_D for the two horizons followed by M_{EP} and M_{IP} for the discrimination of the management units of the horizons.

3.3. Spatial behavior of the properties

The only property that did not fit to the theoretical semivariogram models in the two horizons was K_s , which is a property that showed a random distribution and independence in its spatial behavior (Table 3), and

this result correlated with the high variability observed in the CV. Various reports have demonstrated the spatial independence of K_s (Cucunubá-Melo *et al.*, 2011; Jiménez *et al.*, 2011) because it is affected by the B_D , soil particles, pore size and pore continuity. For the other properties, the predominant model was the exponential model followed by the spherical model for the two horizons. Sand showed the lowest coefficient of determination (R^2) in the two horizons with values of 0.68 for the A_H and 0.62 for the B_H . In turn, M_{IP} was the property with the highest R^2 for the two horizons. These R^2 values together with the cross validation coefficient (CVC) values close to one (1) for the properties that showed a fit to the theoretical semivariogram models indicated an adequate reliability in the estimation of the data in the non-sampled zones as estimated through kriging (Cucunubá *et al.*, 2011).

The properties that showed a strong degree of spatial dependence (DSD) fit the exponential model where the B_D was the only property that showed this behavior in the two horizons and the only one in the A_H . For the B_H , the T_p and P_D also followed this behavior. These properties in common showed a low CV (Table 1), which although it is not a necessary condition, upon carrying out the correlation between CV and DSD of the set of properties that showed a fit with the theoretical semivariogram models, values of -0.13 for the A_H and -0.50 for the B_H were obtained, thereby indicating an inverse correlation. With the exception of the M_{EP} and M_{IP} , greater values of DSD were observed for the properties in the B_H , which was due to the lesser effect of the external anthropic agents. According to Cambardella *et al.* (1994), a strong DSD in soil properties can be attributed to intrinsic factors (parental material, relief, weather, organisms and time), and a weak DSD can be attributed to extrinsic factors (fertilization, soil intervention, machinery traffic, pasturing and irrigation).

The property that showed the largest range in the two horizons was M_{AP} with values greater than 8,000 m, and the attribute with the smallest range was B_D for the A_H (Table 3).

Table 3. Parameters of the theoretical semivariogram models of soil physical properties.

Property	Model	C ₀	C ₀ + C ₁	Range, m	DSD, %	R ²	CVC
A horizon							
K _S	NE	0.58	0.58				
M _{AP}	Spherical	2.09	5.34	11600	0.61	0.95	1.03
M _{EP}	Exponential	1.82	9.26	4530	0.50	0.86	1.01
M _{IP}	Exponential	9.81	21.10	6300	0.54	0.94	1.03
T _P	Exponential	8.39	17.34	11400	0.52	0.83	1.04
B _D	Exponential	4.00*10 ⁻⁴	7.04*10 ⁻⁴	1290	0.94	0.77	0.90
P _D	Exponential	1.82*10 ⁻³	3.65*10 ⁻³	3750	0.50	0.90	0.91
Sand	Spherical	17.35	36.10	3360	0.52	0.68	0.95
B horizon							
K _S	NE	0.61	0.61				
M _{AP}	Exponential	3.97	10.04	8490	0.61	0.91	1.00
M _{EP}	Spherical	9.25	18.51	5550	0.50	0.87	1.02
M _{IP}	Exponential	10.55	21.11	5550	0.50	0.92	1.06
T _P	Exponential	0.04	17.71	1230	0.99	0.74	0.93
B _D	Exponential	3.60*10 ⁻⁴	1.26*10 ⁻²	1050	0.97	0.85	0.83
P _D	Exponential	3.23*10 ⁻⁴	2.46*10 ⁻³	960	0.87	0.86	0.81
Sand	Spherical	10.84	23.81	3380	0.55	0.62	0.97

CVC, cross validation coefficient; NE, nugget effect.

The other properties showed ranges greater than 1,000 m. This behavior verified that the distance chosen for the sampling was adequate (320 m) because it allowed the definition of the spatial dependence of the properties with the exception of K_S, which is a parameter for which it would be convenient to use a lesser distance or a nested system of sampling to verify whether a defined spatial dependency exists for properties that show random distribution when the sampling is carried out with a grid as proposed by Lin *et al.* (2005). In contrast, the ranges observed were smaller at the maximum distance between sampling points (close to 16,300 m), thereby fulfilling the criteria of the regionalized variables theory (Webster and Oliver, 2007). This condition is of special importance because the range provides information for the crop management and evaluation of experiments as well as information about the genesis and evolution of the soils in the landscape and limits between textural classes.

3.4. Linear correlation and kriging

The K_S showed a positive correlation with the T_P and an inverse one with the B_D in the two horizons, which indicated the influence of these two properties on the movement of the water in the soil (Table 4). The B_D and T_P properties showed the greatest number of significant correlations with the other properties for the two horizons, which correlated with the principal component analysis (Table 2), in turn, allowing the visualization of the importance of these properties in the characterization of the soils.

The M_{AP} showed a positive correlation with the T_P and a negative one with the B_D. Similarly, a negative correlation existed between B_D and T_P. These results were verified visually with the contour maps obtained by kriging where zones with high T_P or M_{AP} corresponded to zones of low B_D (Figure 2). Similarly, the negative correlation found in the analysis of the

correlation between the M_{EP} and M_{IP} in each horizon was reflected in the contour maps where zones of high M_{EP} corresponded to zones of low M_{IP} (Figure 2). Another important factor to consider based on the spatial variability of the B_D , T_P and M_{EP} is the relationship with water storage in the soil and the availability of water for plants. This relationship highlights the need to establish management units that are conducive to soil preparation, drainage works, or localized risk according

to the crop requirements and the time of year the farming is carried out. The P_D showed few significant correlations with the other properties in the two horizons, which affirmed the results found in the PC analysis because this property did not have any representation in the components analyzed. The greatest influence of the P_D was on the M_{EP} , which had a low negative correlation that was barely perceptible in the contour maps (Figures 2 and 3).

Table 4. Pearson correlation for soil physical properties.

	K_S	M_{AP}	M_{EP}	M_{IP}	T_P	B_D	P_D	Sand
A horizon								
K_S	1	ns	ns	ns	*	**	ns	ns
M_{AP}	0.053	1	**	ns	**	**	**	ns
M_{EP}	0.023	-0.314	1	**	ns	**	*	ns
M_{IP}	0.022	0.093	-0.505	1	**	**	ns	**
T_P	0.103	0.344	0.068	0.245	1	**	ns	**
B_D	-0.137	-0.268	-0.160	-0.174	-0.335	1	ns	ns
P_D	-0.080	0.167	-0.103	0.006	0.049	0.048	1	ns
Sand	0.032	-0.008	0.019	-0.188	-0.212	0.072	-0.006	1
B horizon								
K_S	1	ns	**	ns	**	**	ns	ns
M_{AP}	0.035	1	**	ns	**	**	ns	*
M_{EP}	0.144	-0.215	1	**	**	**	*	ns
M_{IP}	-0.020	-0.012	-0.503	1	**	**	ns	*
T_P	0.223	0.417	0.251	0.345	1	**	**	*
B_D	-0.139	-0.350	-0.335	-0.144	-0.725	1	**	ns
P_D	-0.013	-0.021	-0.111	0.044	-0.130	0.215	1	ns
Sand	-0.018	0.114	-0.075	-0.103	-0.120	0.015	0.018	1

* $p \leq 0.05$; ** $p \leq 0.01$.; ns not significant.

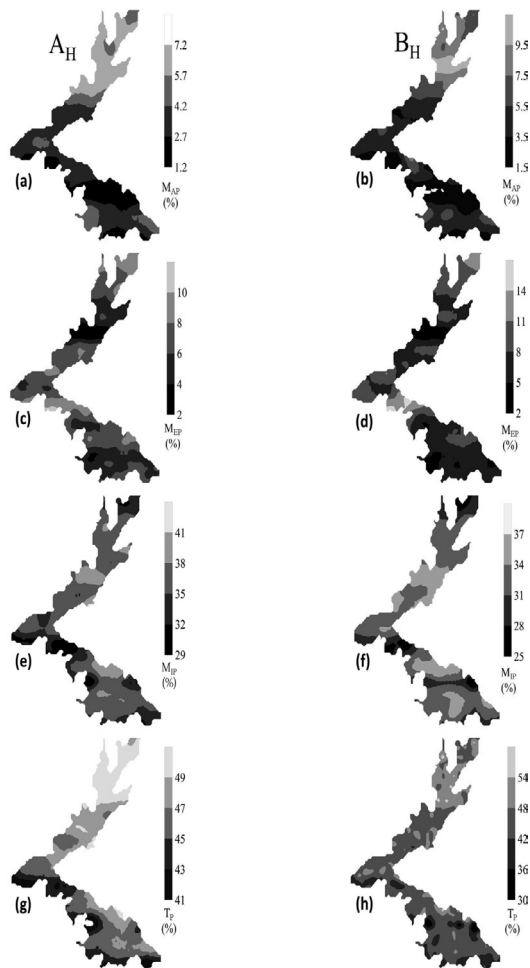


Figure 2. Contour maps obtained by ordinary kriging for macropores (a-b), mesopores (c-d), micropores (e-f) and total porosity (g-h) in the A and B horizons (left and right, respectively).

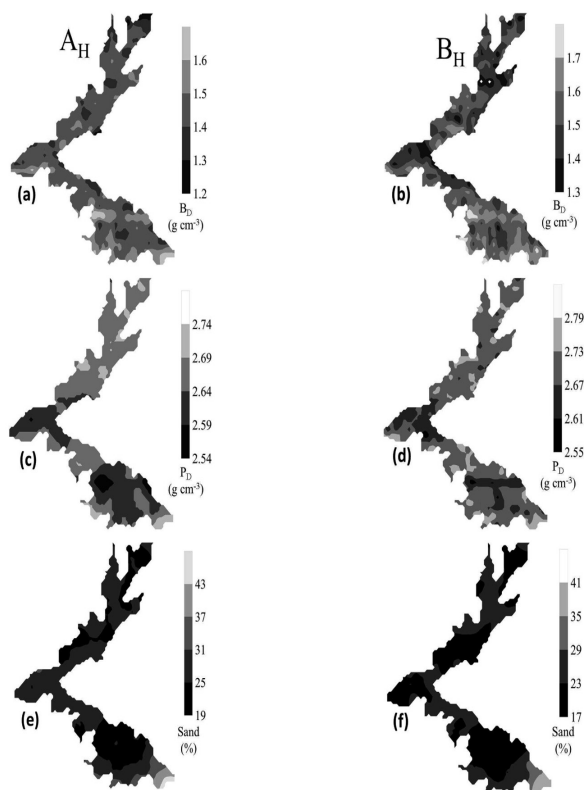


Figure 3. Contour maps obtained by ordinary kriging for bulk density (a-b), particle density (c-d) and sand (e-f) in the A and B horizons (left and right, respectively).

3.5. Management units and its relation to soil physical properties

Cluster analysis (CA) allowed definition of the units through the grouping of the sampling points that showed similarities based on the physical properties analyzed for each horizon (Figure 4). In this analysis, four homogenous units were defined for the A_H and B_H . With the aim of defining an equal number of units for the maps obtained using the soil indexes determined based on principal components (SIPC) and the coefficient of variation (SICV), the units were differentiated using the mean and standard deviation,

which is a methodology considered appropriate by Ortega and Santibáñez (2007).

When comparing the maps obtained by the different proposed techniques, a greater similarity existed between the maps generated on the basis of cluster analysis and SIPC with respect to the SICV (Figures 5 and 6). This difference was because the SICV was principally affected by the properties that showed a larger coefficient of variation, such as K_s , which is a property that did not show an important contribution in the first two principal components in addition to low correlations with the other properties.

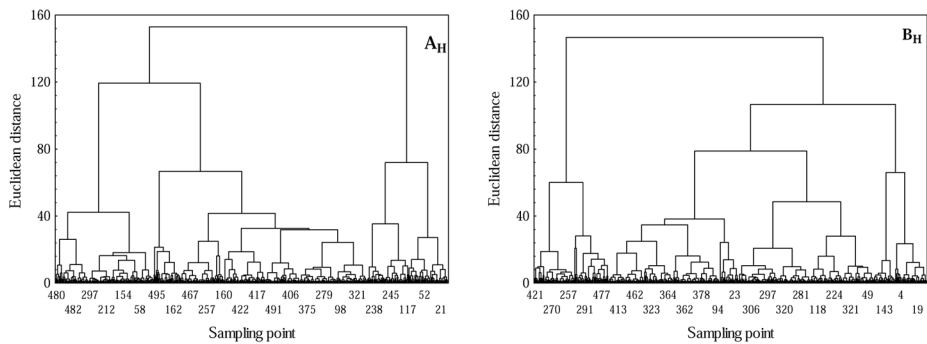


Figure 4. Cluster dendrograms for identifying management units in the AH and BH horizons.

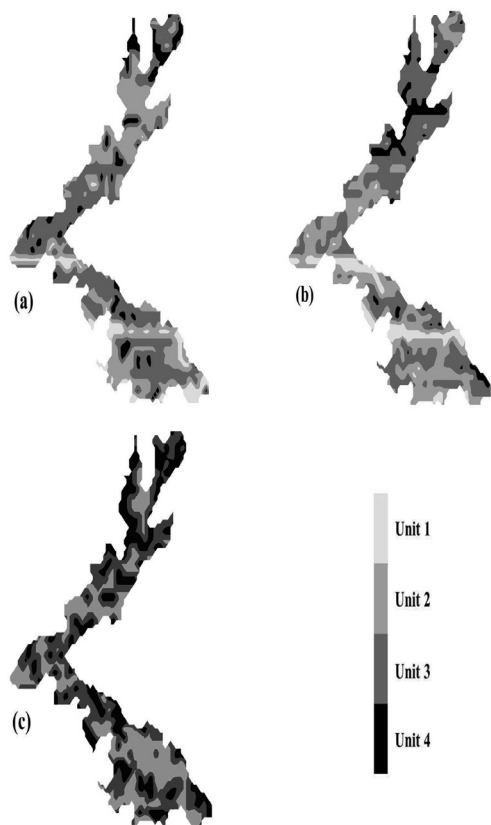


Figure 5. Management units defined for the A horizon from (a) cluster analysis and soil indices obtained from (b) principal components and (c) coefficient of variation.

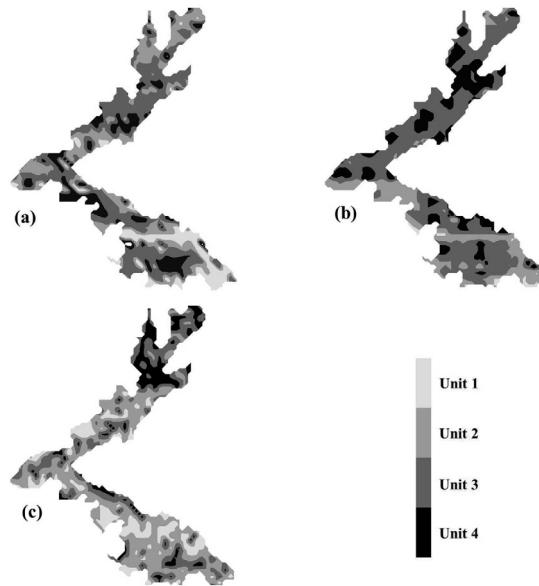


Figure 6. Management units defined for the B horizon from (a) cluster analysis and soil indices obtained from (b) principal components and (c) coefficient of variation.

The correlations among the management units and soil properties showed a high relationship with the principal component analysis results where the properties of the greatest representation in the first two PCs corresponded to the greatest correlations (Table 5), and the largest coefficients of correlation between the properties and management units were observed using the SIPC method for both horizons. The T_p , which was the property that showed the largest coefficient in CP1, was the property that had the greatest correlation with the management units, except for SICV in the

A_H , where K_S showed the largest coefficient due to its high CV value analyzed above. In contrast, the least representative property in the management units was P_D , except in the SICV method.

When the maps of the horizon management units were visually compared, the four units were similar in the A_H and B_H (Figures 5 and 6). This similarity allowed recommendations for each unit to be formulated considering the management of the two horizons conjunctly according to the crop that is to

be established, especially in the labor of farming where the physical properties of the soil are markedly affected, thereby affecting movement of water and air.

While it is true that the units showed different areas according to the method through which they were determined, characteristics directly related to the soil properties existed, which was observed by comparing the contour maps (Figures 2 and 3) with the management units. Unit 1 in both horizons showed the greatest sand content, the lowest value of T_p , the lowest value of M_{AP} , and the highest B_D . According to these conditions, Unit 1 had the greatest limitation for plant root development and water flow in the soil. In general, Unit 1 was the unit with the greatest physical limitations.

Unit 2 was characterized by a lower quantity of M_{EP} and a higher number of M_{IP} , which are properties that limit the quantity of water available to plants.

In this unit, the lowest K_s was found in the A_H , which is a behavior that restricts the ingress and movement of water in the soil, especially toward the B_H where the largest values of K_s existed. While it is true that Unit 2 presented fewer restrictions than Unit 1, its principal limitation was the scarce presence of M_{EP} . Unit 3 showed fewer physical restrictions for the two horizons with low values of K_s , greater presence of M_{EP} , higher values of T_p as well as lower B_D and intermediate sand content. With respect to its physical characteristics, Unit 3 was only surpassed by Unit 4 in both horizons, which was the unit with the lowest area in the estimations of the management units obtained by the three methods. Unit 4 had the greatest T_p with values greater than 50% for the A_H and B_H . With respect to the other units, Unit 4 showed a larger quantity of M_{EP} in the B_H , which also indicated that Unit 4 was the zone with the greatest water storage capacity and the greatest percentage of water available to plants. Similarly, this zone had the lowest values for B_D and P_D .

Table 5. Coefficients of correlation between management units and physical properties on the basis of cluster analysis (CA) in addition to the soil indexes obtained through principal components (PCs) and the coefficient of variation (CV).

Property	A horizon			B horizon		
	CA	PC	CV	CA	PC	CV
K_s	-0.343	0.193	-0.523	0.018	-0.158	-0.372
M_{AP}	-0.110	-0.616	-0.428	-0.382	-0.520	-0.411
M_{EP}	-0.254	0.349	-0.251	-0.265	0.145	-0.381
M_{IP}	-0.320	-0.635	0.093	-0.180	-0.670	0.158
T_p	-0.542	-0.696	-0.320	-0.643	-0.765	-0.512
B_D	0.485	0.507	0.272	0.593	0.612	0.436
P_D	0.080	0.012	-0.462	0.178	0.020	-0.164
Sand	0.284	-0.079	-0.098	0.324	0.366	-0.192

4. Conclusions

Soils in the native savanna of the eastern plains showed physical limitations because of natural processes of densification, as indicated by the high bulk density values and predominance of microporosity in the soil, which restrict the movement of water and air, in turn, affecting the development of plants. These physical limitations vary in space and can be identified through different properties, which are generally correlated among the properties and follow a spatial pattern where some have a greater influence or prevalence in the soil.

The properties that showed the greatest number of significant correlations were also those with the greatest expression in the principal component analysis and, in turn, those that showed the greatest spatial correspondence with the management units, as observed among the maps drawn for the management units and those obtained by kriging for the properties that were fit to semivariogram models.

The study of soil physical properties using geostatistics and multivariate analysis helps characterize soil. Geostatistics identifies the patterns of spatial variability, estimates the properties at non-sampled sites and identifies the management units with a reduced group of properties, thus allowing properties with less expensive and more economical sampling and laboratory analysis techniques to be chosen and optimizing the costs of sampling and definition of management units.

Acknowledgements

We would like to thank ECOPETROL for its financial support in the framing of the project and the Colombian Corporation for Agricultural Research (CORPOICA) for its logistic support in acquiring samples in the field and laboratory analysis. We would like to thank the Colombian Science, Technology and Innovation Department (COLCIENCIAS) for providing financial support for the first author.

References

- Amézquita, E., Thomas, R.J., Rao, I.M., Molina, D.L., Hoyos, P. 2004. Use of deep-rooted tropical pastures to build-up an arable layer through improved soil properties of an Oxisol in the Eastern Plains (Llanos Orientales) of Colombia. *Agric. Ecosyst. Env.* 103, 269-277.
- Camacho-Tamayo, J.H., Luengas, C.A., Leiva, FR. 2008. Effect of agricultural Intervention on the spatial variability of some soils chemical properties in the eastern plains of colombia. *Chilean J. Agric. Res.* 68, 42-55.
- Clapp, R.B., Homberger, G.M. 1978. Empirical equations for some soil hydraulic properties. *Water Resour. Res.* 14, 601-604.
- Cambardella, C.A., Moorman, T.B., Novak, J.M., Parkin, T.B., Karlen, D.L., Turco, R.F., Konopka, A.E. 1994. Field-scale variability of soil properties in Central Iowa Soils. *Soil Sci. Soc. Am. J.* 58, 1501-1511.
- Cressie, N. 1993. *Statistics for spatial data*. Revised Edition. John Wiley & Sons, New York, USA. 928 pp.
- Corrales, I.I., Amézquita, E., Rivera, M., Chávez, L.F. 2003. Condiciones físicas de un suelo bajo diferentes sistemas de labranza y su efecto sobre el comportamiento de cultivos y malezas en la Altillanura colombiana. En: D. Lobo Luján, D. Gabriels, G. Soto (eds.) *Evaluación de parámetros y procesos hidrológicos en el suelo*. VII Escuela Latinoamericana de Física de Suelos, Hidrología de Suelos, La Serena (Chile), pp: 101-106.
- Cruz, J.S., Assis, Júnior, R.N., Matias, S.S.R., Camacho-Tamayo, J.H. 2011. Spatial variability of an Alfisol cultivated with sugarcane. *Cien. Inv. Agr.* 38, 155-164.

- Cucunubá-Melo, J.L., Álvarez-Herrera, J.G., Camacho-Tamayo, J.H. 2011. Identification of agronomic management units based on physical attributes of soil. *J. Soil Sci. Plant Nutr.* 11, 87- 99.
- Diggle, P.J., Ribeiro, J.R. 2000. *Model Based Geostatistics*. São Paulo: Associação Brasileira de Estatística. 129 pp.
- Dong-Sheng, Y., Zhong-Qi, Z., Hao, Y., Xue-Zheng, S., Man-Zhi, T., Wei-Xia, W. 2011. Effect of Soil Sampling Density on Detected Spatial Variability of Soil Organic Carbon in a Red Soil Region of China. *Pedosphere*. 21, 207-213.
- Ferreira, D.F. 2011. *Estatística Multivariada*. 2nd. ed. Lavras: Ed. UFLA. 676 pp.
- Garten Jr., C.T., Kang, S., Brice, D.J., Schadt, C.W., Zhou, J. 2007. Variability in soil properties at different spatial scales (1 m - 1 km) in a deciduous forest ecosystem. *Soil Biol. Biochem.* 39, 2621-2627.
- Godwin, R.J., Miller, P.C.H. 2003. A review of the technologies for mapping within-field variability. *Biosyst. Eng.* 84, 393-407.
- Goovaerts, P. 1998. Geostatistical tools for characterizing the spatial variability of microbiological and physico-chemical soil properties. *Biol. Fertil. Soils*. 27, 315-334.
- Jiménez, J.J., Decaëns, T., Amézquita, E., Rao, I., Thomas, R.J., Lavelle, P. 2011. Short-range spatial variability of soil physico-chemical variables related to earthworm clustering in a neotropical gallery forest. *Soil Biol. Biochem.* 43, 1071-1080.
- Lin, H., Wheeler, D., Bell, J., Wilding, L. 2005. Assessment of soil spatial variability at multiple scales. *Ecol. Modell.* 182, 271-290.
- Malagón Castro, D. 2003. Ensayo sobre tipología de suelos colombianos - Énfasis en génesis y aspectos ambientales. *Rev. Acad. Colomb. Cienc.* 27, 319-341.
- Molina, D.L., Amézquita, E., Hoyos, P. 2003. Construcción de capas arables en suelos Oxisoles de la Altillanura colombiana. En: D. Lobo Luján, D. Gabriels, G. Soto (eds.) *Evaluación de parámetros y procesos hidrológicos en el suelo*. VII Escuela Latinoamericana de Física de Suelos, Hidrología de Suelos, La Serena (Chile), pp: 113-127.
- Moral, F.J., Terrón, J.M., Rebollo, F.J. 2011. Site-specific management zones based on the Rasch model and geostatistical techniques. *Comput. Electron. Agric.* 75, 223-230.
- Nyamangara, J., Masvaya, E.N., Tirivavi, R., Nyengerai, K. 2013. Effect of hand-hoe based conservation agriculture on soil fertility and maize yield in selected smallholder areas in Zimbabwe.. *Soil Till. Res.* 126, 19-25.
- Orjuela-Matta, H.M., Rubiano, Y., Camacho-Tamayo, J.H. 2012. Spatial analysis of infiltration in an Oxisol of the eastern plains of Colombia. *Chilean J. Agric. Res.* 72, 404-410.
- Ortega, R.A., Santibáñez, O.A. 2007. Determination of management zones in corn (*Zea mays* L.) based on soil fertility. *Comput. Electron. Agric.* 58, 49-59.
- Peña, R., Rubiano, Y., Peña, A., Chaves, B. 2009. Variabilidad especial de los atributos de la capa arable de un Inceptisol del piedemonte de la cordillera Oriental (Casanare, Colombia). *Agron. Colomb.* 27, 111-120.
- Phiri, S., Amézquita, E., Rao, I.M., Singh, B.R. 2001. Disc harrowing intensity and its impact on soil properties and plant growth of agropastoral systems in the Llanos of Colombia. *Soil Till. Res.* 62, 131-143.
- Tesfahunegn, G.B., Tamene, L., Vlek, P.L.G. 2011. Catchment-scale spatial variability of soil properties and implications on site-specific soil management in northern Ethiopia. *Soil Till. Res.* 117, 124-139.

- Thomas, R.J., Ayarza, M.A. 1999. Sustainable land management for the Oxisols of the Latin American savannas: dynamics of soil organic matter and indicators of soil quality. Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia. 231pp.
- USDA. 1999. Soil taxonomy. 2nd ed. Agriculture Handbook. Number 436. United States Department of Agriculture, Washington D.C., USA. 869 pp.
- Warrick, A.W., Nielsen, D.R. 1980. Spatial variability of soil physical properties in the field. In: D. Hillel (ed.). Applications of soil physics. Academic Press, New York, USA. pp. 319-344.
- Webster, R., Oliver, M.A. 2007. Geostatistics for environmental scientists. John Wiley & Sons Inc., Hoboken, NJ, USA. 330 pp.
- Zimmermann, B., Elsenbeer, H. 2008. Spatial and temporal variability of soil saturated hydraulic conductivity in gradients of disturbance. J. Hydrol. 361, 78-95.