



Revista Ciência Agronômica

ISSN: 0045-6888

ISSN: 1806-6690

Universidade Federal do Ceará

Mattos, Amanda Guimarães de; Lima, Erica Souto Abreu; Amaral Sobrinho, Nelson Moura Brasil do; Zoffoli, Hugo José de Oliveira; Pérez, Daniel Vidal  
Multivariate analyses to establish reference values for  
soils in Médio Paraíba, state of Rio de Janeiro, Brazil<sup>1</sup>

Revista Ciência Agronômica, vol. 49, no. 1, January-March, 2018, pp. 1-10  
Universidade Federal do Ceará

DOI: 10.5935/1806-6690.20180001

Available in: <http://www.redalyc.org/articulo.oa?id=195358115001>

- How to cite
- Complete issue
- More information about this article
- Journal's homepage in redalyc.org



Scientific Information System Redalyc

Network of Scientific Journals from Latin America and the Caribbean, Spain and Portugal

Project academic non-profit, developed under the open access initiative

## Multivariate analyses to establish reference values for soils in Médio Paraíba, state of Rio de Janeiro, Brazil<sup>1</sup>

Análises multivariadas para o estabelecimento dos valores de referência em solos do Médio Paraíba, RJ

Amanda Guimarães de Mattos<sup>2</sup>, Erica Souto Abreu Lima<sup>3\*</sup>, Nelson Moura Brasil do Amaral Sobrinho<sup>3</sup>, Hugo José de Oliveira Zoffoli<sup>3</sup> and Daniel Vidal Pérez<sup>4</sup>

**ABSTRACT** - Quality Reference Values (QRVs) for potentially toxic elements are obtained from background levels in soils. However, this determination from mean values or percentiles is not appropriate given the variability in the natural distribution of these elements in soils. Therefore, the objective of this study was to propose a new methodology to establish the QRVs, using the Médio Paraíba region (RJ, Brazil) as an example, from groups of soils defined based on the pseudo-total levels of B, Ba, Co, Cr, Cu, Ni, Pb, and Zn through the use of multivariate analyses and discriminant functions. A total of 40 points, collected at depths of 0-20 and 20-40 cm, were used for the determination of pseudo-total contents, according to the EPA 3051A methodology. The samples were separated into three groups to better represent the variability of the soils of the region. The classification functions were obtained based on the variables Mn, Fe, and Mg. In general, the groups G1 and G2 presented lower values than the ones obtained when the sample universe was used, whereas G3 presented higher values. The QRVs obtained from the soil groups presented substantial differences that translate into advantages for the management of the contaminated areas of the region.

**Key words:** Guiding values. Heavy metals. Cluster analysis.

**RESUMO** - Valores de Referência de Qualidade (VRQ) para elementos potencialmente tóxicos são obtidos através dos teores naturais nos solos. No entanto, esta determinação a partir de valores médios ou de percentis não é adequada, diante da variabilidade existente na distribuição natural desses elementos nos solos. Diante disso, o objetivo do estudo foi propor uma nova metodologia de estabelecimento dos VRQs, usando como exemplo a Região do Médio Paraíba - RJ, a partir de grupos de solos definidos com base nos teores pseudo-totais de B, Ba, Co, Cr, Cu, Ni, Pb e Zn, mediante a utilização de análises multivariadas e de funções discriminantes. Foram utilizados 40 pontos, coletados nas profundidades 0-20 e 20-40 cm, para a determinação dos teores pseudo-totais, conforme a metodologia EPA 3051A. As amostras foram separadas em três grupos, para melhor representar a variabilidade existente nos solos da região. As funções de classificação foram obtidas com base nas variáveis Mn, Fe e Mg. Em geral, os grupos G1 e G2 apresentaram valores inferiores, em relação aos obtidos quando se utilizou o universo amostral. Enquanto que, G3 apresentou valores superiores. Os VRQs obtidos a partir de grupos de solos apresentaram diferenças substanciais que traduzem vantagens para o gerenciamento de áreas contaminadas da região.

**Palavras-chave:** Valores orientadores. Metais pesados. Análise de agrupamento.

DOI: 10.5935/1806-6690.20180001

\*Autor para correspondência

Recebido para publicação em 13/01/2016; aprovado em 06/02/2017

<sup>1</sup>Parte da Dissertação do primeiro autor, apresentado ao Curso de Pós-Graduação em Agronomia - Ciência do Solo da Universidade Federal do Rio de Janeiro/UFRRJ

<sup>2</sup>Universidade Federal Rural do Rio de Janeiro/UFRRJ, aamandagmattos@gmail.com

<sup>3</sup>Universidade Federal Rural do Rio de Janeiro/UFRRJ, ericaabreulima@gmail.com, nelmoura@ufrj.br, zoffoliho@yahoo.com.br

<sup>4</sup>Embrapa Solos, daniel.perez@embrapa.br

## INTRODUCTION

The determination of the background levels of potentially toxic elements in soils, as well as the proposition of Quality Reference Values (QRVs), is essential for the construction of legislation that serves as the basis for the monitoring and remediation of these areas. The use of lists with guiding values has been common practice in countries that have developed an efficient environmental policy for the protection of natural resources. In these cases, contamination of resources is recorded when concentrations of the elements of environmental interest reach a level above a set threshold, called the guiding value. Internationally, although no uniformity exists regarding the nomenclature used (background values, baseline values, guiding values, reference values, trigger, etc.), the guiding values represent the basis of the soil and groundwater protection policy (CETESB, 2001). In Brazil, among the guiding values proposed by Resolution 420 of the National Council for the Environment (Conselho Nacional do Meio Ambiente - CONAMA) of 2009 are the Quality Reference Values, which indicate the natural concentration of a substance in soils that have not been influenced by anthropic actions (CONAMA, 2009).

In recent years, several research groups, from Brazil and from other countries, have directed their research toward obtaining reference values specific to each study region (ALFARO *et al.*, 2015; ALMEIDA JÚNIOR *et al.*, 2016; BINI *et al.*, 2011; BIONDI *et al.*, 2011; COSTA *et al.*, 2014; PRESTON *et al.*, 2014; RÉKÁSI; FILEP, 2012; SANTOS; ALLEONI, 2013). However, an evaluation of the obtained values allows clear distinctions to be made between them. This is because these values depend on the composition of the source material, the pedogenetic processes, and the degree of soil development, which are specific to each environment.

The state of Rio de Janeiro does not yet have soil QRVs, and values determined for other Brazilian states are being used. However, the use of QRVs from another locality can lead to errors in evaluations due to the diversity of geomorphoclimatic conditions.

The region of Médio Paraíba, located in the south of the state of Rio de Janeiro, has a very intense industrial activity, where several cases of contamination have occurred, harming the quality of life in the surrounding population. Therefore, a decision was made to study the background levels of the potentially toxic elements of this region, thus providing support for understanding the relation between the soil characteristics and the distribution of these elements in these areas.

To establish the QRVs of inorganic substances, the Brazilian legislation recommends the use of the 75<sup>th</sup> or 90

th percentile of the sample universe, with the anomalous values (outliers) being previously withdrawn (CONAMA, 2009). However, the determination of QRVs from simple statistical criteria, whether based on the mean or percentiles of the frequency distribution of the data, has limitations. According to Paye, Mello, and Melo (2012), the main limitations are related to the variability of the physical and chemical attributes of the soils, which determine the distribution of the potentially toxic elements. In addition, the removal of values considered outliers may represent another error. Therefore, particularities can occur in an area that could generate natural values of potentially toxic elements above those observed for the other areas under study. Thus, further in-depth individual investigations into the origin of such values are needed (PRESTON *et al.*, 2014).

The establishment of reference concentration ranges has been performed through different statistical procedures with the aim of minimizing this problem, where multivariate statistical methods, which allow the simultaneous interpretation of a large number of variables, have been the most used (FADIGAS *et al.*, 2006; PAYE; MELLO; MELO, 2012).

The objective of this work was to propose the use of multivariate analyses and discriminant functions as a new methodology for the establishment of QRVs, using as an example the Médio Paraíba region - RJ, for groups of soils defined based on the pseudo-total levels of B, Ba, Co, Cr, Cu, Ni, Pb, and Zn.

## MATERIALS AND METHODS

### Soil selection

Soil samples were selected from the Médio Paraíba region - RJ, whose collection points were obtained from a joint analysis of the soil, geology, use, and land cover maps of the state of Rio de Janeiro, all on a 1:500,000 scale. The maps were overlaid using ArcGIS Desktop 10, produced by ESRI; to obtain the sampling points, the cLHS - Conditioned Latin Hypercube System was used. In this way, 40 points, considered sufficient to obtain all the variability of environmental covariates, were chosen for the collection of soils in areas of minimal anthropic interference in the Médio Paraíba region - RJ. The samples were collected at a minimum distance of 200 m from the roads, thus avoiding interference from particulates emitted by vehicles.

All points were georeferenced using GARMIN® GPS navigation equipment, and the information (geographical coordinates datum WGS84, and municipality) for each point is presented in Table 1.

**Table 1** - Information of the selected points for soil collection in the Médio Paraíba region - RJ in areas of lower anthropogenic interference

Points	Coordinates		Municipality
	W	S	
1	44°05'48.2"	22°35'06.64"	Barra Mansa
2	43°45'9.6"	22°05'17.1"	Valença
3	43°51'15.9"	22°36'06.2"	Piraí
4	43°33'44.2"	22°07'10.5"	Rio das Flores
5	44°33'29.3"	22°19'56.5"	Itatiaia
6	44°03'01.5"	22°38'49.1"	Rio Claro
7	44°23'18.4"	22°27'18.1"	Resende
8	44°11'54.5"	22°36'22.2"	Barra Mansa
9	44°33'29.3"	22°19'56.5"	Itatiaia
10	43°55'05.1"	22°37'27.7"	Piraí
11	44°48'15.9"	22°21'02.4"	Resende
12	43°55'54.4"	22°36'04.3"	Piraí
13	43°59'26.9"	22°21'24.9"	Barra do Piraí
14	44°07'31.1"	22°16'17.2"	Valença
15	44°24'02.8"	22°19'26.4"	Resende
16	43°57'19.2"	22°18'21.7"	Valença
17	43°50'19.0"	22°18'15.2"	Valença
18	43°52'08.6"	22°24'29.8"	Barra do Piraí
19	44°16'30.1"	22°28'24.8"	Barra Mansa
20	43°46'19.0"	22°08'57.5"	Valença
21	43°57'49.4"	22°08'24.5"	Valença
22	43°45'9.6"	22°05'17.1"	Valença
23	44°10'28.2"	22°23'54.1"	Barra Mansa
24	44°07'31.1"	22°16'17.2"	Valença
25	44°14'49.4"	22°21'21.7"	Quatis
26	44°35'27.3"	22°19'19.5"	Itatiaia
27	44°10'28.2"	22°23'54.1"	Barra Mansa
28	43°45'9.6"	22°05'17.1"	Valença
29	43°46'50.9"	22°19'29.5"	Valença
30	43°59'12.0"	22°40'32.0"	Rio Claro
31	43°57'49.4"	22°08'24.5"	Valença
32	44°14'49.4"	22°21'21.7"	Quatis
33	43°37'21.9"	22°04'58.2"	Rio da Flores
34	43°55'6.1"	22°37'4.9"	Piraí
35	43°54'8.6"	22°21'8.5"	Barra do Piraí
36	43°55'13.0"	22°19'27.9"	Valença
37	44°11'45.3"	22°35'56.7"	Barra Mansa
38	44°18'50.1"	22°25'27.1"	Porto Real
39	44°42'50.4"	22°28'18.7"	Resende
40	44°07'49.4"	22°23'51.6"	Barra Mansa

## Soil collection and preparation

The collection and preparation of soil samples were performed based on procedures described in Embrapa (2009). For each sampling point, small trenches were opened with the aid of a straight blade, where samples of approximately 2 kg were collected at depths of 0-20 and 20-40 cm. To avoid possible contamination, samples of the lower depth were collected first.

The soil samples were air dried using white paper and a veil-type screen to avoid contamination. After drying, they were crushed with an agate mortar, homogenized, and passed through a stainless steel sieve with a 2-mm opening (10 mesh) in preparation for the physical and chemical analyses and through a sieve with a 0.149-mm opening (100 mesh) for the soil digestion procedure.

## Soil analysis

The chemical and physical analyses were performed on the 80 collected samples and consisted of organic carbon, pH in H<sub>2</sub>O (1:2.5), exchangeable cations, potential acidity, base sum calculation, cation exchange capacity (CEC), and quantification of the grain size fractions (sand, silt, and clay fraction) according to the methods proposed by Donagemma *et al.* (2011).

For the determination of the B, Ba, Co, Cr, Cu, Ni, Pb, and Zn in the soil, the samples (0.500 ± 0.001 g) underwent wet digestion in a closed system in a MARS Xpress® digester for 8 min and 40 s until reaching 175 °C according to the SW-846 3051A method (USEPA, 2007); a v/v ratio of 3 HNO<sub>3</sub>:1 HCl was used. For the determination of the element concentrations, the extracts were analyzed in an Inductively Coupled Plasma Optical Emission Spectrometry (ICP-OES) device.

The quality control of the analysis was carried out with SRM 2709 *San Joaquin soil* certified material, from the National Institute of Standards and Technology (NIST). The levels obtained were compared to the leachable concentrations, as well as to those concentrations recommended by NIST for pseudo-total digestion methods, i.e., partial sample digestion (NIST, 2003). When compared to the leached values, the recovery rates of the certified samples were satisfactory, above 60%, indicating the efficiency of the method.

## Statistical analyses

The descriptive statistical analysis of the contents of the potentially toxic elements was performed in a Microsoft® Office Excel® spreadsheet. The Pearson correlation coefficients were determined in SAS (2010) software, with a statistical significance level of 1% probability.

The Euclidian distance and the Ward method were used as a measure of similarity in the cluster analysis (CA). For this, the mean pseudo-total values of eight elements (B, Ba, Co, Cr, Cu, Ni, Pb, and Zn) were used as clustering variables (standardized data for mean 0 and variance 1). For selection of the most appropriate number of groups, a cross-validation was performed in the discriminant analysis, using the same discriminatory variables used in the clustering, assuming an equal covariance matrix and equal classification probabilities for the groups.

After the groups were defined, successive discriminant analyses were performed to reallocate the observations within the groups until reaching a 0% classification error rate by the cross-validation method. This procedure is defined in this work as Reallocation with Discriminant Analysis (RDA).

After the groups were obtained, the soil variables that best discriminated them were selected. For this purpose, the *proc stepdisc* of the SAS statistical program (2010) was used, with *backward* and *stepwise* selection criteria at a 5% probability for the model variables' input or output. Subsequently, the stability of each set of variables selected by the *proc stepdisc* was verified, based on the coefficient of variation of the multiplier constants ( $\beta$ ) of the discriminant model and on the error rate by cross-validation. After the soil variables were defined, the classification functions of each group were then elaborated.

The QRV of each element was established based on the general mean values, mean of each group, and percentiles (75th and 90th) of the frequency distribution of the results, as suggested by Conama (2009), with the percentiles also being determined for each group formed.

## RESULTS AND DISCUSSION

### Correlation analysis between potentially toxic element contents and soil attributes

Pearson's correlations among the B, Ba, Co, Cr, Cu, Ni, Pb, and Zn contents and of these contents with the chemical attributes (pH, P, K<sup>+</sup>, Ca<sup>2+</sup>, Mg<sup>2+</sup>, Al<sup>3+</sup>, Na<sup>+</sup>, Corg, CEC, Fe, and Mn) and the granulometric (silt and clay) levels at a significance level of 1% are presented in Table 2.

According to Pearson's correlation analysis, the B, Ba, Co, Cr, Cu, Ni, Pb, and Zn elements showed high correlations with each other (Table 2). According to Sheng *et al.* (2012), these elements possibly were associated with the same soil source material and were released as weathering products of these rocks.

**Table 2** - Pearson correlation coefficients among the background levels of B, Ba, Co, Cr, Cu, Ni, Pb, and Zn and the soil attributes

	B	Cr	Pb	Cu	Co	Zn	Ni	Ba	Fe	Mn
B	1									
Cr	0.72*	1								
Pb	0.19	0.16	1							
Cu	0.33*	0.39*	0.24	1						
Co	0.09	0.25	0.41*	0.58*	1					
Zn	0.02	0.22	0.41*	0.55*	0.75*	1				
Ni	0.28	0.56*	0.30*	0.68*	0.74*	0.75*	1			
Ba	-0.09	0.17	0.21	0.56*	0.80*	0.79*	0.80*	1		
Fe	0.95*	0.73*	0.23	0.36*	0.16	0.08	0.32*	-0.05	1	
Mn	0.03	0.23	0.31*	0.51*	0.73*	0.48*	0.47*	0.51*	0.33	1
pH	0.30*	0.37*	0.15	0.24	0.43*	0.36*	0.39*	0.44*	0.33*	0.33*
P	-0.15	-0.11	-0.03	0.07	0.17	0.28	0.08	0.27	-0.14	0.13
K <sup>+</sup>	0.02	0.05	0.11	0.18	0.11	0.23	0.11	0.17	0.02	0.08
Ca <sup>2+</sup>	0.16	0.28	0.15	0.22	0.45*	0.48*	0.34*	0.43*	0.18	0.41*
Mg <sup>2+</sup>	0.08	0.20	0.16	0.53*	0.62*	0.68*	0.60*	0.77*	0.13	0.40*
Al <sup>3+</sup>	0.51*	0.39*	0.04	0.07	-0.12	-0.01	0.32*	-0.13	-0.24	-0.3
Na <sup>+</sup>	-0.13	-0.11	0.32*	0.17	0.17	0.21	0.16	0.27	0.46*	-0.21
Corg <sup>1</sup>	0.02	0.03	0.01	-0.02	-0.10	0.02	-0.02	-0.17	-0.1	0.08
CEC <sup>2</sup>	-0.09	-0.04	0.01	-0.07	0.01	0.16	0.04	0.02	0.02	0.06
Silt	0.26	0.18	0.48*	0.09	0.25	0.21	0.18	0.09	-0.11	0.04
Clay	0.43*	0.19	-0.08	-0.09	-0.38*	-0.45*	-0.36*	-0.52*	0.27	0.23

\*significant values at 1% probability highlighted <sup>1</sup> Organic carbon<sup>2</sup> Cation exchange capacity at pH 7.0

Positive and significant correlations with Fe and/or Mn were observed for all elements, suggesting a high geochemical affinity between them. In this study, the Fe and Mn contents were considered soil attributes, as they were directly related to the content of the Fe, Mn oxides, hydroxides, and oxyhydroxides of the soils. Correlations of Fe and Mn with potentially toxic elements are associated with the formation of secondary oxides that have high adsorption capacity (BURAK *et al.*, 2010; SIPOS *et al.*, 2008). According to Friedrich, Hasenmueller and Catalano (2011), Fe and Mn oxides control the distribution and speciation of potentially toxic elements in soils weathered by adsorption, occlusion, and coprecipitation reactions.

B and Cr presented similar behavior because the correlation was high among them (72%), and therefore, they were similarly influenced by Al<sup>3+</sup>, pH, and Fe.

Because the elements Ba, Co, Ni, and Zn had a very close correlation with each other ( $r > 70\%$ ), they showed similar associations and were positively correlated with pH, Ca<sup>2+</sup>, Mg<sup>2+</sup>, and Mn, and they showed a negative correlation with clay. Ba presented a high

correlation coefficient with Mg (77%). Kabata-Pendias and Mukherjee (2007) reported that Ba behaves similarly to Ca<sup>2+</sup> and Mg<sup>2+</sup> and may even replace these elements in the structure of carbonate minerals.

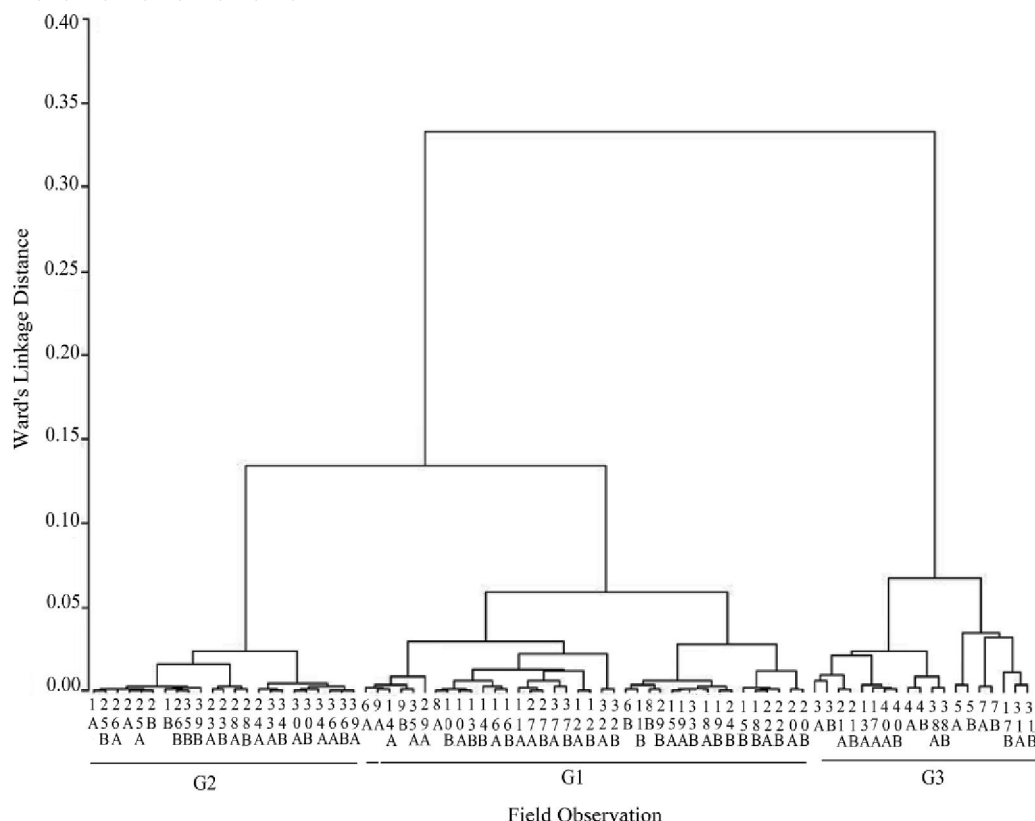
Pb correlated with Na<sup>+</sup>, Mn, and silt, and Cu correlated with Mg<sup>2+</sup>, Fe, and Mn.

### Classification of samples into groups

Figure 1 shows the dendrogram resulting from the cluster analysis, which was performed using the B, Ba, Co, Cr, Cu, Ni, Pb, and Zn contents of the 80 soil samples. The bond distance greater than 1.25 times the standard deviation of the bond distance among all the observations was adopted as a cutoff point (MILLIGAN; COOPER, 1985). This criterion suggested the formation of two or three groups.

To better elucidate the choice of the number of groups, the error rate of the cross-validation in the discriminant analysis was used separately for each analytical repetition. By means of this error rate, the formation of two groups and of three groups that presented

**Figure 1** - Dendrogram formed by Ward's Cluster Analysis and Euclidian distance as a function of the mean contents of three analytical replicates of B, B, Ba, Co, Cr, Cu, Ni, Pb, and Zn



low values (less than 20%) was verified. We then chose three groups, as a greater number of groups gives a greater distinction of variability in the samples. The number of samples composing the groups were as follows: group 1 (G1), composed of 39 observations; group 2 (G2), composed of 26 observations; and group 3 (G3), composed of 15 observations, totaling the 80 observations.

To select soil variables that classify the groups formed and provide a better measure of the discriminant function, the procedure was performed in two stages. The first step consisted of using the *proc stepdisc* procedure (SAS, 2010) through two selection criteria—*backward* and *stepwise*—in three training samples, which resulted in eight sets of variables (Table 3). The most common variables sets were verified as Mg-Fe-Mn-CEC-Corg, with 12 observations, followed by Mg-Fe-Mn, with eight observations.

The second step consisted of verifying the stability of the variables, based on the coefficient of variation (CV%) of the multiplier constants ( $\beta$ ), by using the variables of the discriminant function for all sets of variables obtained in the first step (Table 3). Only the Mg-Fe-Mn and Mg-Fe-Mn-Clay models presented  $\beta$  with CV% less than 50% in the two clustering procedures (CA and RDA), but the Mg-

Fe-Mn-Clay model presented a higher classification error rate relative to the Mg-Fe-Mn model, indicating that it is therefore less appropriate than the Mg-Fe-Mn model.

By selecting the soil variables Mg, Fe, and Mn, the discriminant analysis allowed the elaboration of a function for each group, generating a classification model (Table 4), where new samples can be allocated within the established groups. The framing of the group to which the observation belongs will be defined by the highest value generated between the discriminant functions (SAS, 2010).

### Characterization of groups

Figure 2 presents the standardized data (mean equal to 0 and variance 1) of the contents of B, Ba, Co, Cr, Cu, Ni, Zn, and Pb and of the soil attributes for each group. Group 1 (G1) is verified as being the one with the lowest levels of potentially toxic elements; it also has the lowest values of Fe, Mn, Mg, Ca, and pH, with means of 23,331.1 mg kg<sup>-1</sup>; 97.0 mg kg<sup>-1</sup>; 0.2 cmol<sub>c</sub> dm<sup>-3</sup>; 0.7 cmol<sub>c</sub> dm<sup>-3</sup>; and 5.1, respectively.

The highest values of B and Cr were obtained in group 2 (G2), where the highest values of Fe, Al, and

**Table 3** - Characteristics of discriminant models with soil variables for training subsamples

Model	Criterion	n*	Coefficient of variation (%)							Overall error rate
			$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	
Mg-Fe-Mn	CA <sup>1</sup>	6	16	42	18	24	-	-	-	25%
	RDA <sup>2</sup>	2	14	36	19	31	-	-	-	24%
Mg-Fe-Mn-CEC-Corg	CA <sup>1</sup>	4	17	44	19	36	26	8	-	24%
	RDA <sup>2</sup>	8	12	6	17	31	18	128	-	23%
Mg-Fe-Mn-CEC-Corg-Al	CA <sup>1</sup>	3	17	22	16	50	25	118	198	23%
	RDA <sup>2</sup>	0	11	59	16	40	17	2	36	24%
Mg-Fe-Mn-Al	CA <sup>1</sup>	3	19	45	16	27	116	-	-	26%
	RDA <sup>2</sup>	0	14	37	18	29	70	-	-	25%
Mg-Fe-Mn-Ca	CA <sup>1</sup>	1	16	36	16	19	58	-	-	26%
	RDA <sup>2</sup>	2	14	34	17	32	47	-	-	25%
Mg-Fe-Mn-Clay	CA <sup>1</sup>	0	16	30	29	21	17	-	-	27%
	RDA <sup>2</sup>	2	20	43	37	33	21	-	-	27%
Mg-Fe-Mn-Corg-Silt	CA <sup>1</sup>	2	18	326	24	39	36	116	-	25%
	RDA <sup>2</sup>	0	14	353	21	48	31	69	-	23%
Mg-Fe-Mn-Corg-Silt-Clay	CA <sup>1</sup>	0	17	45	43	24	44	76	13	26%
	RDA <sup>2</sup>	1	18	57	43	36	57	49	17	24%
Mg-Fe-Mn-CEC-Corg-Ca	CA <sup>1</sup>	1	17	68	17	23	59	15	25	23%
	RDA <sup>2</sup>	0	12	42	14	29	30	12	25	24%

\*Number of occurrences in the selection of variables by the STEPDISC procedure; <sup>1</sup>groups formed by cluster analysis, <sup>2</sup>groups formed by reallocation with discriminant analysis

**Table 4** - Discriminant functions of classification with the variables Mg, Fe, and Mn for the formation of three groups of observations

Group	Discriminant function
1	$-6.6688028 + 1.6765472 \cdot \text{Mg} + 0.0005336 \cdot \text{Fe} + 0.0037539 \cdot \text{Mn}$
2	$-15.9871706 + 3.7745883 \cdot \text{Mg} + 0.0007932 \cdot \text{Fe} + 0.0090506 \cdot \text{Mn}$
3	$-20.7115228 + 9.0556961 \cdot \text{Mg} + 0.0006968 \cdot \text{Fe} + 0.0163894 \cdot \text{Mn}$

clay were also observed, with means of 36,385.1 mg kg<sup>-1</sup>; 36,380.0 cmol<sub>c</sub> dm<sup>-3</sup>; and 372.5 g kg<sup>-1</sup>, respectively.

On the other hand, the highest background levels of Ba, Co, Cu, Ni, Pb, and Zn were represented by group 3 (G3), which also presented the highest levels of Mn, Mg, Ca, and pH, whose averages were 601.2 mg kg<sup>-1</sup>; 1.3 cmol<sub>c</sub> dm<sup>-3</sup>; 2.5 cmol<sub>c</sub> dm<sup>-3</sup>; and 5.8, respectively. This confirms the Pearson correlation analysis (Table 1), where a positive and significant correlation was observed between the Mg and Mn contents and these six elements, except for Pb, which had no significant correlation with Mg. Clay presented a negative correlation with the G3 elements, explaining its low values in this group (mean of 278.2 g kg<sup>-1</sup>).

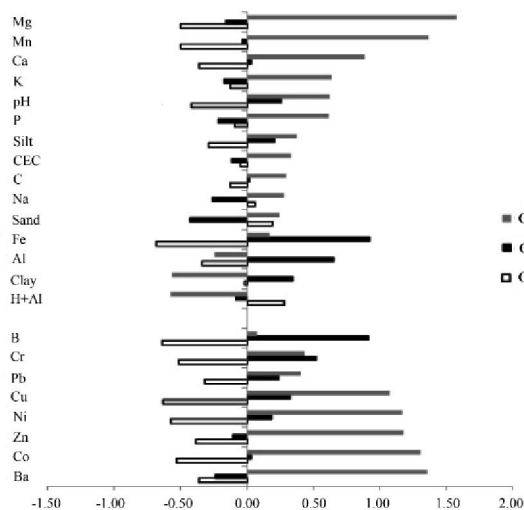
Thus, the relations between the soil attributes and the potentially toxic elements are evident; these relations corroborate the results obtained with the Pearson's correlation analysis (Table 1), and they confirm that the selection of the variables Mn, Fe, and Mg to compose the discriminant function classification is appropriate.

#### Establishment of Quality Reference Values (QRVs)

Table 5 shows the background levels of the B, Ba, Co, Cr, Cu, Ni, Pb, and Zn elements in the soils of the Médio Paraíba region - RJ; these levels are represented by the overall means and by the 75 th and 90 th percentiles for the total set of samples and for the three classification groups.



**Figure 2** - Mean of the groups standardized for a total mean equal to zero and variance 1 after classification with the discriminant model Mg-Fe-Mn



In this study, the 75 th percentile (P75), or the upper quartile, was considered in the establishment of the QRVs because it presents an intermediate condition between the mean levels and the 90 th percentile. The elimination of 25% of the data becomes more appropriate because it ensures a greater and better environmental assessment, as well as contributing to human health safety. Since the areas without anthropic interference in the Médio Paraíba region are limited, this means that the lithological units are not all represented by the points sampled; thus, the adoption of the 75 th percentile becomes more adequate.

However, the adoption of groups better contemplates the variability in soil samples and presents greater homogeneity within the groups. This behavior can be verified by the coefficients of variation, which were

smaller in the groups than in the total sample (Table 5). In this way, the use of the 75 th percentile of the groups, instead of the 75 th percentile of the sample set, is recommended since both the more restrictive and the more permissive situations would be represented. Thus, the acquisition of a range of variation is possible for the levels of potentially toxic elements in the soils of the region.

The B levels found in the soils of the Médio Paraíba region - RJ, both in the 75 th percentile (P75) and in the 90 th percentile (P90), are higher than the reference values established by Copam (2011) for the soils of Minas Gerais state (11.5 mg kg<sup>-1</sup>), and the lowest value was 74 mg kg<sup>-1</sup> in the G1 mean.

Ba presented great variation in the overall contents, according to the coefficient of variation; this was also observed between the groups. The 75 th percentile ranged from 28 to 135 mg kg<sup>-1</sup> in the groups. However, G1 and G2 presented levels below the value determined by Copam (2011), which is 93 mg kg<sup>-1</sup>. The 90 th percentile of G3 presented a Ba value of 198 mg kg<sup>-1</sup>, which is higher than the prevention value (150 mg kg<sup>-1</sup>) adopted by Conama (2009).

Co levels, in general, approach those levels found in other Brazilian states, varying in the P75 of the groups from 3 to 15 mg kg<sup>-1</sup>, within the range established by Fadigas *et al.* (2006), which was 2-20 mg kg<sup>-1</sup>. In G3, the values found were 15 mg kg<sup>-1</sup> in P75 and 16 mg kg<sup>-1</sup> in P90, which were higher but very close to the value of Cetesb (2014) of 13 mg kg<sup>-1</sup>.

In general, the elements Cr and B varied little between the groups (40 to 58 mg kg<sup>-1</sup> in P75), as they presented low coefficients of variation. In addition, throughout the study, B and Cr presented similar dynamics, proving the high correlation between them.

**Table 5** - Background levels of B, Ba, Co, Cr, Cu, Ni, Pb, and Zn (mg kg<sup>-1</sup>) in the soils of the Médio Paraíba region - RJ. Overall means, coefficients of variation, and 75th and 90th percentiles of the sample universe and groups

Element	Total Sample			G1			G2			G3			CV (%)			
	μ	P75	P90	μ	P75	P90	μ	P75	P90	μ	P75	P90	T	G1	G2	G3
B	94	117	140	74	88	102	123	139	148	97	103	133	34	28	19	27
Ba	41	40	112	23	28	46	28	32	82	108	135	198	123	90	109	66
Co	5	6	13	2	3	4	5	6	11	11	15	16	101	86	86	48
Cr	41	52	64	32	40	51	50	58	77	49	53	65	43	42	35	35
Cu	7	10	16	3	4	8	9	12	16	13	16	18	83	80	69	32
Ni	6	7	12	4	5	6	7	9	13	11	14	18	70	57	56	39
Pb	12	15	19	10	14	16	14	15	22	15	18	20	48	48	50	33
Zn	25	31	39	20	24	31	24	28	33	41	48	65	53	40	45	42

μ- mean; CV- coefficient of variation; T- total sample

The values of Cr in P75 in the three groups were lower than those proposed by COPAM (2011) for the soils of Minas Gerais ( $75 \text{ mg kg}^{-1}$ ) and Martinez-Lladó *et al.* (2008) for the soils of Catalonia (Spain) ( $83.6 \text{ mg kg}^{-1}$ ). However, they were higher than the values obtained for the Rio Grande do Norte state ( $26.5 \text{ mg kg}^{-1}$ ) (COSTA *et al.*, 2014).

Cu concentrations in the Médio Paraíba region - RJ were generally low and were lower than the values determined by Cetesb (2014) ( $35 \text{ mg kg}^{-1}$ ) and Copam (2011) ( $49 \text{ mg kg}^{-1}$ ).

The contents of Ni are close to those found in soils of other Brazilian states. The P75 values of the three groups were 5, 9, and  $14 \text{ mg kg}^{-1}$ , respectively, close to the reference values of  $13 \text{ mg kg}^{-1}$  established by Cetesb (2014) for the state of São Paulo and of  $6.6 \text{ mg kg}^{-1}$  for Espírito Santo state (PAYE *et al.*, 2010) and below the reference value of  $21.5 \text{ mg kg}^{-1}$ , established by Copam (2011) for Minas Gerais state. However, the lowest value verified for Ni ( $5 \text{ mg kg}^{-1}$  in G1) is above the value  $1.3 \text{ mg kg}^{-1}$  adopted for Mato Grosso state and Rondônia state (SANTOS; ALLEONI, 2013).

Pb contents showed little variation in the soils of the Médio Paraíba region - RJ, probably because 71% of the geology of the samples was gneiss since there is a relation between Pb and gneiss, as observed by Biondi *et al.* (2011), due to the presence of Pb in the composition of feldspars, which are constituents of gneiss.

The Pb contents, as well as those of Co and Ni, were similar to the values obtained for other Brazilian states. The highest Pb content obtained in this study was  $22 \text{ mg kg}^{-1}$  in the P90 of G2. In the P75 of the groups, the values were 14, 15, and  $18 \text{ mg kg}^{-1}$  in G1, G2, and G3, respectively. These values were lower than those of Minas Gerais state ( $19.5 \text{ mg kg}^{-1}$ ) (COPAM, 2011) and lower than those established by Rékási and Filep (2012) for the soils of Hungary ( $29.7 \text{ mg kg}^{-1}$ ) and by Alfaro *et al.* (2015) for the soils of Cuba ( $50 \text{ mg kg}^{-1}$ ).

The Zn levels established in P75 for groups G1, G2, and G3 were 24, 28, and  $48 \text{ mg kg}^{-1}$ , which were lower than the reference value established by Cetesb (2014) of  $60 \text{ mg kg}^{-1}$ , and the superior condition was slightly higher than that of Copam (2011) ( $46.5 \text{ mg kg}^{-1}$ ). The highest value of Zn was  $65 \text{ mg kg}^{-1}$ , found in the P90 of G3. In general, the values are within the range proposed by Fadigas *et al.* (2006) for this element ( $6\text{-}79 \text{ mg kg}^{-1}$ ) and below the value of Mediterranean Europe ( $83 \text{ mg kg}^{-1}$ ) (MICÓ *et al.*, 2007).

Considering the presented results, the QRVs obtained from the soil groups presented substantial

differences compared to the procedure regulated by Resolution 420/2009 of Conama. Although G1 and G2 present values lower than or close to those obtained when using the entire sample universe, the existence of G3, which stands out for the values being higher than the others, indicates the risk of disregarding it as an eventual outlier. The importance of group separation thus becomes an advantage for the management of contaminated areas and should be considered in a forthcoming re-evaluation of said Resolution.

## CONCLUSIONS

1. Soils showed a wide variation in the contents of B, Ba, Co, Cr, Cu, Ni, Pb, and Zn and were greatly influenced by the source material, the formation factors, and the physical and chemical characteristics of the soils, reflecting the lithological complexity of the region;
2. The use of multivariate techniques was an adequate and promising tool for obtaining QRVs, highlighting the separation of the samples into three groups, which best expressed the variability of the sample set, in addition to the adoption of P75 in the groups, which proved to be more pertinent to the study conditions;
3. The soil variables Mn, Fe, and Mg, which were used in the group classification function, had the greatest effect on the levels of the elements B, Ba, Co, Cr, Cu, Ni, Zn, Pb, and Cr in the soils;
4. The classification functions were shown to be adequate tools for the allocation of new samples in previously established groups;
5. The P75-based QRVs for the Médio Paraíba region for each group formed, in  $\text{mg kg}^{-1}$ , are as follows: G1 – B (88), Ba (28), Co (3), Cr (40), Cu (4), Ni (5), Pb (14), and Zn (24); G2 – B (139), Ba (32), Co (6), Cr (58), Cu (12), Ni (9), Pb (15), and Zn (28); and G3 and G2 – B (103), Ba (135), Co (15), Cr (53), Cu (16), Ni (14), Pb (18), and Zn (48).

## REFERENCES

- ALFARO, M. R. *et al.* Background concentrations and reference values for heavy metals in soils of Cuba. **Environmental Monitoring and Assessment**, v. 187, p. 4198-4207, 2015.
- ALMEIDA JÚNIOR, A. B. *et al.* Background and reference values of metals in soils from Paraíba State, Brazil. **Revista Brasileira de Ciências do Solo**, v. 40, p. 1-13, 2016.
- BINI, C. *et al.* Background levels of trace elements and soil geochemistry at regional level in NE Italy. **Journal of Geochemical Exploration**, v. 109, p. 125-133, 2011.

- BIONDI, C. M. *et al.* Teores de Fe, Mn, Zn, Cu, Ni e Co em solos de referência de Pernambuco. **Revista Brasileira de Ciências do Solo**, v. 35, p. 1057-1066, 2011.
- BURAK, D. L. *et al.* Geochemistry and spatial distribution of heavy metals in Oxisols in a mineralized region of the Brazilian Central Plateau. **Geoderma**, v. 160, p. 131-142, 2010.
- COMPANHIA DE TECNOLOGIA DE SANEAMENTO AMBIENTAL. Decisão de diretoria nº 045/2014/E/C/I, de 20 de fevereiro de 2014. Dispõe sobre a aprovação dos valores orientadores para solos e águas subterrâneas no Estado de São Paulo, em substituição aos valores orientadores de 2005 e dá outras providências. **Diário Oficial [do] Estado de São Paulo**, Poder Executivo, São Paulo, v. 124, n. 36, 21 fev. 2014. Seção 1, Caderno Executivo I, 4 p.
- COMPANHIA DE TECNOLOGIA DE SANEAMENTO AMBIENTAL. **Relatório de estabelecimento de valores orientadores para solos e águas subterrâneas no Estado de São Paulo**. São Paulo, 2001. 73 p. (Série Relatórios Ambientais).
- CONSELHO ESTADUAL DE POLÍTICA AMBIENTAL. Deliberação Normativa nº 166, de 29 de junho de 2011. Altera o Anexo I da Deliberação Normativa Conjunta COPAM/CERH nº 2 de 6 set. 2010, estabelecendo os valores de referência de qualidade dos solos. **Diário do Executivo de Minas Gerais**, 2011. Disponível em: <<http://www.siam.mg.gov.br/sla/download.pdf?idNorma=18414>>, Acesso em: 29 jun. 2014.
- CONSELHO NACIONAL DO MEIO AMBIENTE. Resolução nº 420, de 28 de dezembro de 2009. Dispõe sobre critérios e valores orientadores de qualidade do solo quanto à presença de substâncias químicas e estabelece diretrizes para o gerenciamento ambiental de áreas contaminadas por essas substâncias em decorrência de atividades antrópicas. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, nº 249, p. 81-84, 30 dez. 2009. Disponível em: <<http://www.mma.gov.br/port/conama/legiabre.cfm.htm>>, Acesso em: dia maio 2014.
- COSTA, W. P. *et al.* Valores de referência de qualidade para metais pesados em solos do Rio Grande do Norte. **Revista Brasileira de Ciência do Solo**, v. 38, p. 1028-1037, 2014.
- DONAGEMMA, G. K. *et al.* **Manual de métodos de análise de solos**. 2. ed. Rio de Janeiro: Embrapa Solos, 2011. 230 p. (Embrapa Solos. Documentos, 132).
- EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Manual de análises químicas de solos, plantas e fertilizantes**. 2. ed. rev. ampl. Brasília, DF: Embrapa Informação Tecnológica, 2009. 627 p.
- FADIGAS, F. S. *et al.* Proposição de valores de referência para a concentração natural de metais pesados em solos brasileiros. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 10, p. 699-705, 2006.
- FRIERDICH, A. J.; HASENMUELLER, E. A.; CATALANO, J. G. Composition and structure of nanocrystalline Fe and Mn oxide cave deposits: implications for trace element mobility in karst systems. **Chemical Geology**, v. 284, p. 82-96, 2011.
- KABATA-PENDIAS, A.; MUKHERJEE, A. B. **Trace elements from soil to human**. New York: Springer, 2007. 450 p.
- MARTINEZ-LLADÓ, X. *et al.* Trace element distribution in topsoils in Catalonia: background and reference values and relationship with regional geology. **Environmental Engineering Science**, v. 25, n. 6, p. 863-878, 2008.
- MICÓ, C. *et al.* Baseline values for heavy metals in agricultural soils in a European Mediterranean region. **Science of The Total Environment**, v. 378, p. 13-17, 2007.
- MILLIGAN, G. W., COOPER, M. C. An examination of procedures for determining the number of cluster in a data set. **Psychometrika**, v. 50, p. 159-179, 1985.
- NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. **Standard Reference Materials SRM 2709**: San Joaquin soil: baseline trace element concentrations. Certificate Issue Date: 18 jul. 2003. 13 p.
- PAYE, H. de S. *et al.* Valores de referência de qualidade para metais pesados em solos no Estado do Espírito Santo. **Revista Brasileira de Ciência do Solo**, v. 34, p. 2041-2051, 2010.
- PAYE, H. de S.; MELLO, J. W. V. de; MELO, S. B. de. Métodos de análise multivariada no estabelecimento de valores de referência de qualidade para elementos-traço em solos. **Revista Brasileira de Ciência do Solo**, v. 36, p. 1031-1041, 2012.
- PRESTON, W. *et al.* Valores de referência de qualidade para metais pesados em solos do Rio Grande do Norte. **Revista Brasileira de Ciência do Solo**, v. 38, p. 1028-1037, 2014.
- RÉKÁSI, M.; FILEP, T. Fractions and background concentrations of potentially toxic elements in Hungarian surface soils. **Environmental Monitoring and Assessment**, v. 184, p. 7461-7471, 2012.
- SANTOS, S. N. dos; ALLEONI, L. R. F. Reference values for heavy metals in soils of the Brazilian agricultural frontier in Southwestern Amazônia. **Environmental Monitoring and Assessment**, v. 185, n. 7, p. 5737-5748, 2013.
- SAS INSTITUTE Inc. **SAS Online Doc® 9.2**. Cary, NC, 2010.
- SHENG, J. *et al.* Heavy metals of the Tibetan top soils: level, source, spatial distribution, temporal variation and risk assessment. **Environmental Science and Pollution Research**, v. 19, p. 3362-3370, 2012.
- SIPOS, P. *et al.* Sorption of copper, zinc and lead on soil mineral phases. **Chemosphere**, v. 73, p. 461-469, 2008.
- UNITED STATES ENVIRONMENTAL PROTECTION AGENCY. **Method 3051A**: microwave assisted acid digestion of sediments, sludges, soils, and oils. Washington, DC, 2007. 30 p. Disponível em: <<http://www.epa.gov/osw/hazard/testmethods/sw846/pdfs/3051a.pdf>>. Acesso em: 20 jun. 2013.