



Multivariate statistical analysis applied to assess the dispersion of contaminants in a mining tailings basin in the semiarid region of Bahia – Brazil

ARTICLES doi:10.4136/ambi-agua.2572

Received: 05 May 2020; Accepted: 28 Jul. 2020

Nelize Lima Santos^{1*}; Maria da Conceição Rabelo Gomes²
José Ângelo Sebastião Araújo dos Anjos²; Fernanda Gonçalves Cunha³

¹Instituto de Geociências. Programa de pós-graduação em Geologia. Universidade Federal da Bahia (UFBA), Rua Barão de Jeremoabo, s/n, CEP: 40170-290, Salvador, BA, Brazil.

²Departamento de Geologia. Instituto de Geociências. Universidade Federal da Bahia (UFBA), Rua Barão de Jeremoabo, s/n, CEP: 40170-020, Salvador, BA, Brazil.

E-mail: conceicaorabelo@yahoo.com.br, jose.anjos@ufba.br

³Consultora independente, Rua Paissandu, nº 370, CEP: 22210-080, Flamengo, RJ, Brazil.

E-mail: nandagcunha@gmail.com

*Corresponding author. E-mail: nelizelima.geo@gmail.com

ABSTRACT

This study employed multivariate analysis techniques to identify and evaluate the chemical variables responsible for the contamination of the urban area of Boquira, Bahia, due to the abandonment of the tailings basin of Pb-Zn mining, in order to assist in the environmental management of the area. Factor analysis was performed on main and grouping components. The factor analysis allowed grouping the variables into two main factors for street sediment samples, adding up to 72% of the total accumulated variance, and three factors for house dust samples, which explained 77% of the total variance. The variables have a strong correlation with the composition of the tailings basin. Cluster analysis classified the samples according to the concentration of metals in the area, where the influence of the tailings basin and the natural background of the region's rocks in the contamination distribution can be identified.

Keywords: cluster analysis, contamination, main component analysis.

Análise estatística multivariada aplicada na avaliação da dispersão de contaminantes de uma bacia de rejeitos de mineração no semiárido baiano - Brasil

RESUMO

As técnicas de análise multivariada utilizadas nesta pesquisa tiveram por objetivo identificar e avaliar as variáveis químicas responsáveis pela contaminação da área urbana de Boquira, Bahia, devido ao abandono da bacia de rejeitos da mineração de Pb-Zn, com a finalidade de auxiliar nas medidas de gerenciamento ambiental na área. Foram realizadas a análise fatorial em componentes principais e de agrupamento. A análise fatorial permitiu agrupar as variáveis em dois fatores principais para as amostras de sedimentos de rua, somando 72% da variância total acumulada e em três fatores para as amostras de poeira de casas, que



explicaram 77% do total da variância. As variáveis apresentam forte correlação com a composição da bacia de rejeitos. A análise de agrupamento classificou as amostras de acordo com a concentração dos metais na área, onde pode-se concluir a influência da bacia de rejeitos e do *background* natural das rochas da região na distribuição da contaminação.

Palavras-chave: análise de agrupamento, análise de componentes principais, contaminação.

1. INTRODUCTION

The municipality of Boquira, located in the south-central region of the state of Bahia, is located on the geological domain of the Paramirim deformation corridor, composed of, among other lithotypes, rocks from the Boquira Unit that provide this region with a high natural background for lead and zinc (Gomes *et al.*, 2020). Because of this characteristic, the municipality was the scene of mineral exploration from the late 1950s to 1992. The extensive production at the mine resulted in the generation of a tailings basin of approximately 3360 m² in area and 894x10³ m³ in volume, consisting of heavy metals associated with the mineralogical composition of the geological framework of the area.

The tailings basin, abandoned with the end of extraction activities, was the target of an inefficient revegetation program, and its fine-grained material is therefore constantly susceptible to wind erosion. Two aggravating factors in this situation are the location of the tailings basin, immediately next to the urban area and cultivation areas of the municipality, and the creation of the municipal dump over the basin. The atmospheric dispersion of material from the tailings basin exposes the municipality to a series of toxic metals that can cause damage to the health of the population.

Studies on the behavior of toxic metals in urban areas make it possible to interpret the potential risks of these elements; however, the assessment of the environmental quality of an area generally involves a wide range of variables, making it difficult to understand their interrelations (Gomes *et al.*, 2020). Multivariate analysis of the data makes it possible to establish variables essential to environmental management, as this method allows the selection of the variables with greater participation in the contamination that must be monitored and thus reduce costs with less important parameters.

This work therefore aims to define the most important variables in the contamination of the urban area of the municipality due to the presence of the tailings basin through the use of multivariate analysis techniques, in order to assist in the environmental management of the municipality.

2. METHODOLOGY

In the present study, the contamination data were obtained by the Low Density Geochemical Survey of the State of Bahia (Cunha *et al.*, 2016), which contains information on 110 house dust samples and 66 street sediment samples collected at different points of the urban area of the Boquira municipality.

The chemical variables for the home dust samples considered in this study were: aluminum (Al), arsenic (As), barium (Ba), bismuth (Bi), calcium (Ca), cadmium (Cd), cobalt (Co), chromium (Cr), copper (Cu), lead (Pb), sulfur (S), strontium (Sr), iron (Fe), phosphorus (P), lithium (Li), magnesium (Mg), manganese (Mn), molybdenum (Mo), nickel (Ni), potassium (K), sodium (Na), vanadium (V), zinc (Zn). The 110 samples analyzed and the 23 chemical analyses add up to 2530 results for house dust. In addition to the 23 cited elements above, 28 other elements were considered for the street sediment samples, namely, antimony (Sb), boron (B), beryllium (Be), cesium (Cs), scandium (Sc), tin (Sn), phosphorus (P), germanium (Ge), gallium (Ga), hafnium (Hf), indium (In), lanthanum (La), mercury (Hg), niobium (Nb), gold

(Au), palladium (Pd), platinum (Pt), silver (Ag), rhenium (Re), rubidium (Rb), selenium (Se), tungsten (W), thallium (Tl), tellurium (Te), thorium (Th), tantalum (Ta), titanium (Ti), uranium (U), in street sediment samples. The 66 samples analyzed and 51 chemical analyses add up to 3366 results for street sediment.

For statistical analysis in this study, the SPSS Statistics software, Version 17.0, was used, where the techniques of factor analysis and cluster analysis were performed. The mathematical model of factor analysis is controlled by the following Equation 1:

$$X_i = a_{i1} F_1 + a_{i2} F_2 + a_{i3} F_3 + \dots + a_{ij} F_j + e_i \quad (1)$$

Where: X_i are the standardized variables, a_i are the factor loads, F_j are the common factors not related to each other and e_i is an error that represents the variation portion of variable i that is unique to it and cannot be explained by a factor nor by another variable in the analyzed set (Bezerra, 2014).

The factorial analysis aimed to group information contained in a group of original variables into a smaller set of statistical variables; these variables are calculated by the linear combination of the original variables. The factor analysis was divided into the calculation of the correlation matrix, the extraction of factors and the rotation of the matrix.

An examination of the correlations between the variables was performed to obtain the correlation matrix that made it possible to identify the subsets of variables that were highly correlated with each other. The factor extraction method was based on Principal Component Analysis through a linear combination between the variables and the analysis of the variables was carried out based on the similarities between the variables (Mode R).

In order to observe whether the data were sufficiently linked to carry out the factor analysis, the correlation matrix was calculated using the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) method, which determines the adequacy of the data, where values between 0.5 and 0.9 indicate the degree of correlation between the variables, therefore a satisfactory factor analysis. Still, the Bartlett Test of Sphericity (BTS) was used, which tests the hypothesis that there is no correlation between the variables analyzed, thus indicating whether there was a sufficient relationship between the variables for the application of factor analysis.

In order to improve the explanatory potential of the variables, an orthogonal rotation of the factors was carried out using the Varimax Method (Johnson and Wichern, 2007), in order to minimize the number of variables with high loads in different factors, allowing the association of a variable with a single factor (Gomes *et al.*, 2020).

The analysis of hierarchical clustering of the samples was used to identify subgroups statistically different from each other, but composed of individuals or similar variables according to some criterion. Ward's method was used to measure the similarity given by the square Euclidean distance (Equation 2), from the total sum of the squared values of the deviations of each object in relation to the average value of the group in which it was inserted.

$$d_{ij}^2 = \sum_{k=1}^p (x_{ik} - x_{jk})^2 \quad (2)$$

Where: d_{ij} is the distance between two observations that corresponds to the sum of the squares of the differences between i and j for all p variables (Landim, 2011).

The environmental quality assessment was carried out using the clusters formed based on CONAMA Resolution No. 420 of 2009 (CONAMA, 2009) which establishes guidelines for the environmental management of areas contaminated by these substances as a result of anthropic activities. This resolution was used due to the lack of Brazilian legislation that determines guiding values for street sediments and house dust in relation to human health.

3. RESULTS AND DISCUSSION

The factor analysis aimed to order the variables so that it was possible to understand the distribution of concentrations of contaminants in the urban area of the municipality.

The factor analysis in street sediments was initially performed with the 51 chemical variables. Three simulations were necessary to obtain a satisfactory result, taking into account the criteria adopted for this analysis. The final simulation significantly reduced the number of variables, resulting in 6 variables (Al, Pb, Ca, Fe, Mg and Na). Table 1 shows the correlation matrix for the analyzed attributes, where one can observe that 40% of the pairs showed good correlation index (≥ 0.50), only 20% in the range of $0.6 < [r] < 0.9$ indicating a strong correlation, according to the classification by Callegari-Jaques (2003).

Table 1. Correlation matrix of the chemical variables of street sediment samples, in bold are the strong correlations and in italic, the weakest correlation. Sampling period: July / 2013

Variables	Al	Pb	Ca	Fe	Mg	Na
Aluminum (Al)	1.000					
Lead (Pb)	0.278	1.000				
Calcium (Ca)	0.497	0.123	1.000			
Iron (Fe)	0.678	0.197	0.371	1.000		
Magnesium (Mg)	0.661	0.295	0.739	.541	1.000	
Sodium (Na)	0.123	<i>-.083</i>	0.527	0.105	0.568	1.000

The results obtained through the correlation matrix allowed the identification of subsets of variables correlated strongly with each other, but little related to other variables. It was observed that there is a strong and positive correlation between the pairs of the variables Fe-Al, Mg-Al and Ca-Mg and the lowest correlation is between Na-Pb. The KMO index of the set of attributes analyzed was 0.706 and Bartlett's sphericity statistical test was significant with sig = 0.00, indicating that the factors can adequately describe the data variation (Table 2).

Table 2. Factor loads, commonality and variance explained in the factor analysis of the analyzed variables, after rotation by the Varimax method for street sediment samples. Sampling period: July / 2013.

Variables	Factor 1	Factor 2	Communalities
Aluminium (Al)	0.349	0.806	0.772
Lead (Pb)	-0.156	0.664	0.466
Calcium (Ca)	0.807	0.319	0.753
Iron (Fe)	0.269	0.759	0.648
Magnesium (Mg)	0.764	0.541	0.875
Sodium (Na)	0.888	-0.178	0.820
Variance explained by factors (%)	51.062	21.164	
Accumulated variance (%)	72.226		

Concentration units: mg L⁻¹, except EC (μS/cm). KMO = 0.706.

The factor analysis of the main component applied to street sediments condensed the analyzed variables into two ordered factors, which account for 72% of the total variance. Using

Varimax orthogonal rotation, it was observed that Factor 1 (F1) was responsible for 51% of this variance and Factor 2 (F2) was responsible for 21% of the total variance. The variables present in Factor 1 were Na (0.888), Ca (0.807) and Mg (0.764), whereas Factor 2 included the variables Al (0.806), Fe (0.759) and Pb (0.664).

Factor 1, represented by the variables Ca, Mg, Na, is strongly correlated to the mineralogical composition of the geological framework of the area. The Boquirá Unit, outcropping in the region, consists of a sedimentary chemical-terrestrial sequence, with associations of metacarbonates, quartzites, shales and banded iron formations (Carvalho, 2000; Garcia, 2011). The Boquirá Unit's iron formation can be subdivided into five different facies, one of which is composed of, among other types of carbonates, dolomites (rock formed by double calcium and magnesium carbonate, $[\text{CaMg}(\text{CO}_3)_2]$).

Factor 2, which comprises the variables Al, Pb, Fe, correlates strongly with the indicator attributes of clay mineral formation (Al and Fe) through chemical weathering in the soil. Factor 2 also indicated the presence of Pb, which is related to both concentration by natural processes in the Boquirá Unit (Carvalho, 2000; Garcia, 2011) and exposure due to the mineral exploration and abandonment process in the municipality. Despite this, the presence of a moderate form suggests its atmospheric dispersion due to Pb being generally associated with particles smaller than 53 μm in size (Alves *et al.*, 2018).

Factor analysis on house dust was performed initially with the 23 chemical variables. Four simulations were necessary to obtain a satisfactory result, taking into account the criteria adopted for this analysis. The final simulation was reduced to 13 variables (Al, Pb, Cr, Cd, Sr, Fe, P, Zn, Mn, Mo, Ni, K and V). Table 3 shows the correlation matrix for the analyzed attributes, where it can be seen that 23% of the total pairs had a good correlation index (≥ 0.50), with only 15% in the range of $0.6 < [r] < 0.9$ indicating a strong correlation, according to the Callegari-Jaques (2003) classification.

The results obtained through the correlation matrix allowed the identification of subsets of variables strongly correlated with each other, but little related to other variables. It was observed that there is a strong and positive correlation between the pairs of variables Cd-Pb, Cd-Zn and Pb-Zn. On the other hand, the correlations involving the variables Ni, Sr, K and Mo have a low to moderate correlation with the other variables. The KMO index of the set of attributes analyzed was 0.752 and Bartlett's sphericity statistical test was significant with $\text{sig} = 0.00$, indicating that the factors can adequately describe the data variation (Table 4).

The factor analysis of the main component applied to street sediments condensed the analyzed variables into three ordered factors, which account for 77% of the total variance. Using Varimax orthogonal rotation, it was observed that Factor 1 (F1) was responsible for 35% of this variance, Factor 2 (F2) was responsible for 24% and Factor 3 (F3) was responsible for 18% of total variance. The variables present in Factor 1 were Pb (0.954), Zn (0.917), Cd (0.892), Mn (0.815), Fe (0.789) and P (0.591), in Factor 2 were Al (0.851), V (0.847), Cr (0.800) and Ni (0.750) and in Factor 3 were K (0.970), Mo (0.969) and Sr (0.893).

Factor 1 is represented by the variables Cd, Pb, Zn, Mn, Fe and P, and maintains a strict relationship with the elements present in the tailings basin. Chemical analyses performed in the tailings basin show a correspondence with the mineralogical association of the primary ore (Carvalho, 2000; Garcia, 2011; Alves *et al.*, 2018) and point out Pb and Zn as the main metals that make up the basin (Cunha *et al.*, 2016; Alves *et al.*, 2018).

The granulometric-chemical analyses of the tailings basin indicate a concentration in the granulometry range below 105 μm , with more than half being smaller than 53 μm in size (Alves *et al.*, 2018). Because it is located immediately next to the urban area of the municipality and does not have any type of erosion control, the material from the tailings basin has been dispersed in the region, especially by air.

Table 3. Correlation matrix of the chemical variables of house dust samples, in bold are the strong correlations and in italic, the weakest correlations. Sampling period: July / 2013.

Variables	Al	Cd	Pb	Cr	Sr	Fe	P	Mn	Mo	Ni	K	V	Zn
Al	1.000												
Cd	-0.116	1.000											
Pb	-0.106	0.880	1.000										
Cr	0.548	0.056	0.138	1.000									
Sr	-0.338	0.067	0.032	-0.067	1.000								
Fe	0.211	0.559	0.692	0.453	-0.076	1.000							
P	0.213	0.525	0.491	0.216	-0.175	0.401	1.000						
Mn	-0.050	0.574	0.783	0.263	-0.034	0.684	0.349	1.000					
Mo	-0.225	<i>-0.038</i>	-0.060	-0.141	0.787	-0.155	-0.202	-0.085	1.000				
Ni	0.498	-0.246	-0.198	0.515	<i>-0.032</i>	0.094	0.037	-0.102	0.057	1.000			
K	-0.227	<i>-0.036</i>	-0.058	-0.142	0.788	-0.153	-0.202	-0.084	1.000	0.056	1.000		
V	.803	<i>-0.021</i>	0.094	0.623	-0.285	0.422	0.179	0.224	-0.169	0.392	-0.170	1.000	
Zn	<i>0.013</i>	0.836	0.846	0.263	0.053	0.686	0.487	0.651	-0.067	-0.178	-0.066	0.162	1.000

Subtitle: Al – Aluminium; Cd – Cadmium; Pb – Lead; Cr – Chrome; Sr – Strontium; Fe – Iron; P – Phosphor; Mn – Manganese; Mo – Molybdenum; Ni – Nickel; K – Potassium; V – Vanadium; Zn – Zinc.

Table 4. Factor loads, commonality and variance explained in the factor analysis of the analyzed variables, after rotation by the Varimax method for street sediment samples. Sampling period: July / 2013.

Variables	Factor 1	Factor 2	Factor 3	Communalities
Aluminium (Al)	-0.035	0.851	-0.237	0.781
Cadmium (Cd)	0.892	-0.176	0.024	0.827
Lead (Pb)	0.954	-0.089	0.013	0.918
Chrome (Cr)	0.251	0.800	-0.030	0.704
Strontium (Sr)	0.031	-0.145	0.893	0.820
Iron (Fe)	0.789	0.342	-0.062	0.743
Phosphor (P)	0.591	0.154	-0.201	0.413
Manganese (Mn)	0.815	0.064	-0.016	0.669
Molybdenum (Mo)	-0.082	-0.044	-0.969	0.948
Nickel (Ni)	-0.188	0.750	0.118	0.612
Potassium (K)	-0.080	-0.046	0.970	0.949
Vanadium (V)	0.168	0.847	-0.161	0.771
Zinc (Zn)	0.917	0.017	0.018	0.842
Variance explained by factors (%)	34.788	24.517	17.596	
Accumulated variance (%)	76.901			

Concentration units: mg L⁻¹, except EC (µS/cm). KMO = 0.706.

Particles larger than 100 µm can be suspended in the atmosphere for short periods, but tend to settle quickly, while particles ranging in size between 0.002 and 100 µm tend to remain in suspension longer (Finlayson-Pitts and Pitts, 2000) and are transported over greater distances (Järup, 2013). Thus, unlike the concentration in street sediments, the significant presence of these metals in house dust samples may indicate that the material disposed irregularly in the tailings basin has been transported by air, deposited in the open areas, but remobilized and redeposited in the environments urban areas.

The chemical composition of the tailings basin showed a moderate correlation with Factor 2, which included the variables Ni, Cr, V and Al, and an even lower correlation with Factor 3, represented by the variables Sr, Mo and K. The variables mentioned above were found in the basin of tailings, however at low concentrations (Cunha *et al.*, 2016).

The hierarchical cluster analysis applied to the analysis of toxic metal contamination in the urban area of the municipality allowed the classification of contaminants in different groups, but with similar chemical characteristics in each factor. In street sediment samples, the number of clusters was defined from cut-off point 5 for Factors 1 and 2, where the formation of three homogeneous groups for Factor 1 and two groups for Factor 2 was observed.

According to the Factor 1 variables (Ca, Mg and Na), three similar groups were generated, comprising 6%, 2% and 88% of the samples analyzed in clusters 1, 2 and 3, respectively (Figure 1A). This factor is composed of sediments without a concentration of toxic metals, which have no indication of maximum values allowed in CONAMA Resolution No. 420/2009.

Cluster 1 was characterized by sediments with high concentrations of essential, non-toxic metals. In this group, the concentration of calcium (Ca) varied between 6.85 and 13.10 mg L⁻¹, magnesium (Mg) varied between 0.61 and 4.52 mg L⁻¹ and sodium (Na) varied between 0.01 and 3.73 mg L⁻¹. Cluster 2 was represented by only one sample (60) and characterized by the highest concentration of Ca with 23.10 mg L⁻¹, while Mg and Na showed similar values to

Cluster 1, with 3.48 mg L^{-1} and $2,37 \text{ mg L}^{-1}$, respectively. Cluster 3 was characterized by sediments with low concentrations of essential metals. In this group, the Ca concentration varied between 0.01 and 4.33 mg L^{-1} , Mg varied between 0.01 and 2.20 mg L^{-1} and Na varied between 0.01 and 1.22 mg L^{-1} .

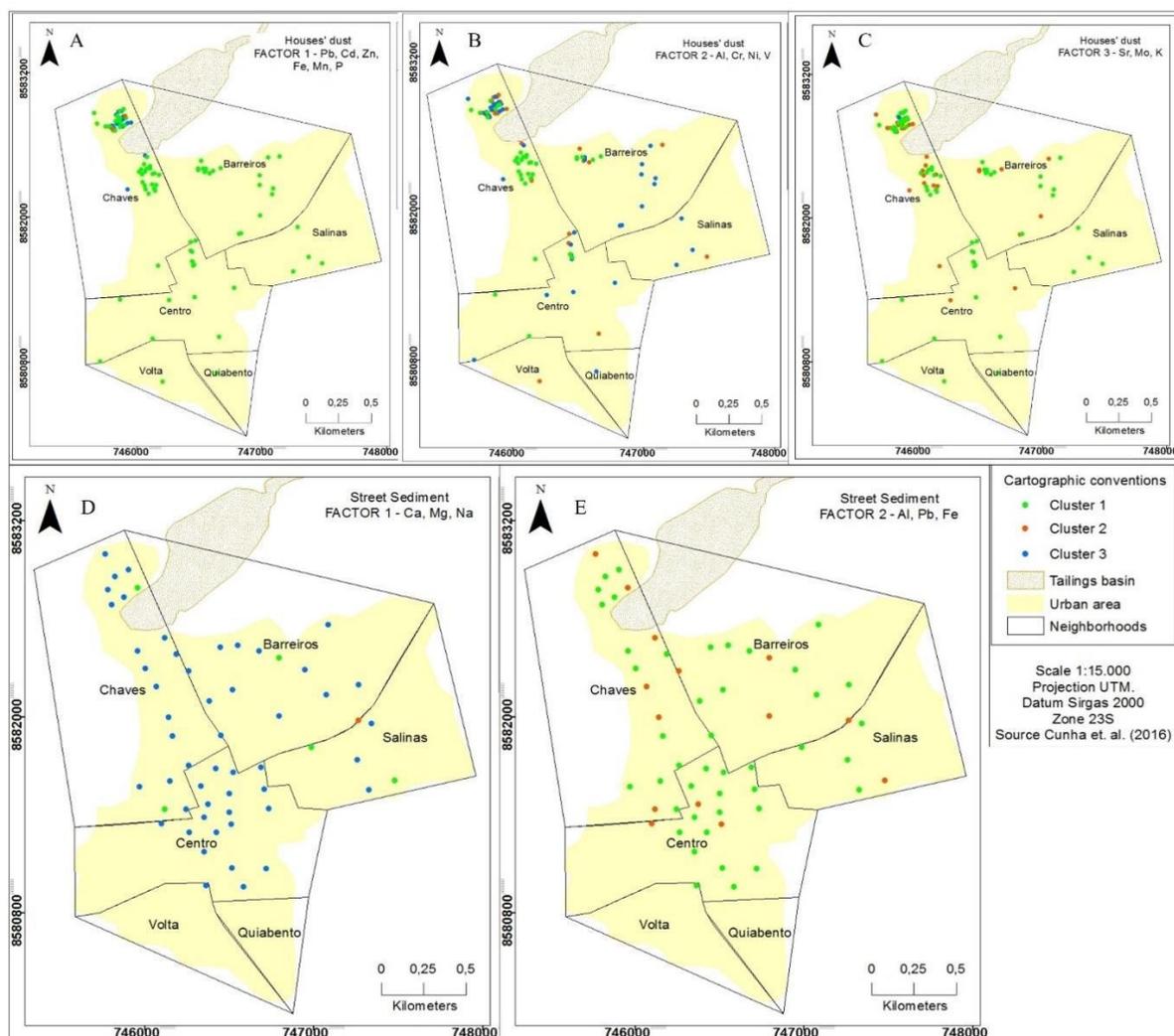


Figure 1. Distribution of variables by clusters. (A) Factor 1 and (B) Factor 2 for street sediment samples and (C) Factor 1, (D) Factor 2 and (E) Factor 3 for house dust samples. Sampling period: July / 2013. **Source of data:** Cunha *et al.* (2016).

According to the variables of Factor 2 (Al, Pb and Fe), two similar groups were generated, comprising 45% and 10% of the samples analyzed in Clusters 1 and 2, respectively (Figure 1B). This factor is composed of sediments with a concentration of contaminating metals, although Al and Fe do not have maximum permissible values determined by CONAMA Resolution No. 420/2009.

Cluster 1 was characterized by sediments with a higher concentration of contaminating metals. In this group, the concentration of aluminum (Al) varied between 0.80 and 4.65 mg L^{-1} , lead (Pb) varied between 0.10 and 1.10 mg L^{-1} and iron (Fe) varied between 5.58 and 12.40 mg L^{-1} . Cluster 2 was characterized by sediments with lower concentrations, where the Al concentration varied between 0.01 and 2.86 mg L^{-1} , Pb varied between 0.10 and 1.20 mg L^{-1} and Fe varied between 0.02 and 5.04 mg L^{-1} . Pb concentrations showed values below the maximum allowed values, according to CONAMA Resolution No. 420/2009, as limits for prevention (72 mg / L^{-1}) and investigation (300 mg / L^{-1}).

The samples of street sediment in both factors indicated a dispersion of the metals in a homogeneous manner throughout the entire urban area, due to the erosion and sedimentation, in the long term, of the metals. Studies carried out in Mariana (MG), after a disaster (Silva *et al.*, 2019), indicate that suspended material can remain in the air and reach a coverage area of up to 1.5 kilometers from its source. From this perspective, all locations in the urban area of Boquira (BA) are under the influence of contamination from the tailings basin. The low presence of metals indicates constant remobilization of street sediments due to the erosive action of winds, street cleaning procedures and vehicle traffic, as discussed by Vianna *et al.*, 2011; Pereira *et al.*, 2015; Abiye *et al.*, 2016.

For house dust samples, the number of clusters was defined from cut-off point 5 for Factors 1, 2 and 3, where the formation of three groups was observed for each factor. According to the Factor 1 variables (Cd, Pb, Zn, Mn, Fe and P), three similar groups were generated, comprising 81%, 2% and 17% of the samples analyzed in clusters 1, 2 and 3, respectively (Figure 1C). This factor is composed of dust from houses with a concentration of contaminating metals and essential metals, but which have a contaminating character in high concentrations.

Cluster 1 was characterized by dust from houses with lower concentrations of metals. In this group, the concentration of cadmium (Cd) varied between 0.01 and 8.0 mg L⁻¹, which means that 10 samples have a concentration above the maximum allowed values, as prevention limits (1.3 mg / L⁻¹) but below the investigation limits (8.0 mg / L⁻¹), according to CONAMA Resolution No. 420/2009. Lead (Pb) varied between 0.2 and 2080.0 mg L⁻¹, with 27% of the samples above the prevention limit (72 mg / L⁻¹) and 39% above the investigation limits (300 mg / L⁻¹), according to CONAMA Resolution No. 420/2009. Zinc (Zn) varied between 1.0 and 3340.0 mg L⁻¹, with 22% of the samples above the prevention limit (300 mg / L⁻¹) and 12% above the investigation limits (1000 mg / L⁻¹), according to CONAMA Resolution No. 420/2009. In turn, iron (Fe), manganese (Mn) and phosphorus (P) do not have maximum allowable values defined by CONAMA Resolution No. 420/2009, Fe varied between 0.6 and 11.3 mg L⁻¹, Mn between 99.0 and 2080.0 mg L⁻¹ and P between 26.0 and 1710.0 mg L⁻¹.

Cluster 2 was characterized by two samples with higher concentrations of contaminating metals, the concentrations of cadmium (Cd) were 54.0 and 95.0 mg L⁻¹, with the samples above the prevention limits (1.3 mg / L⁻¹) and research (8.0 mg / L⁻¹), according to CONAMA Resolution No. 420/2009. The lead (Pb) was 6360.0 and 9020.0 mg L⁻¹, above the prevention (72 mg / L⁻¹) and investigation (300 mg / L⁻¹) limits, according to CONAMA Resolution No. 420 / 2009. The zinc (Zn) was 4290.0 and 5790.0 mg L⁻¹, above the limit of prevention (300 mg / L⁻¹) and investigation (1000 mg / L⁻¹), according to CONAMA Resolution n° 420 / 2009. Iron (Fe) presented concentrations of 8.3 and 10.3 mg L⁻¹, manganese (Mn) 2020.0 and 3790.0 mg L⁻¹ and phosphorus (P) 1360.0 and 2770.0 mg L⁻¹.

Cluster 3 was characterized by dust samples from houses with intermediate values of concentration of contaminating metals. In this group, the concentration of cadmium (Cd) varied between 0.0 and 61.0 mg L⁻¹, with 84% of the samples above the investigation limits (8.0 mg / L⁻¹), according to CONAMA Resolution n° 420/2009. Lead (Pb) varied between 1590.0 and 5210.0 mg L⁻¹, with all samples showing values above the investigation limits (300 mg / L⁻¹), according to CONAMA Resolution No. 420/2009. The zinc (Zn) varied between 1180.0 and 4130.0 mg L⁻¹, with all samples showing values above the investigation limits (1000 mg / L⁻¹), according to CONAMA Resolution No. 420/2009. Iron (Fe), manganese (Mn) and phosphorus (P) varied between 4.1 and 9.1 mg L⁻¹, 774.0 and 4350.0 mg L⁻¹ and 569.0 and 1900.0 mg L⁻¹, respectively.

According to the Factor 2 variables (Al, Cr, Ni and V), three similar groups were generated, comprising 33%, 33% and 34% of the samples analyzed in Clusters 1, 2 and 3, respectively (Figure 1D). This factor is formed by contaminating metals and essential metals that, if in high concentrations, can damage human health.

Cluster 1 was characterized by dust samples from houses with higher concentrations. Aluminum (Al) and vanadium (V) do not have maximum allowable values defined by CONAMA Resolution No. 420/2009; Al varied between 1.9 and 4.1 mg L⁻¹ and V between 45.0 and 108, 0 mg L⁻¹. The concentration of chromium (Cr) varied between 37.0 and 91.0 mg L⁻¹, with two samples above the prevention limit (75.0 mg / L⁻¹) but below the investigation limit (300.0 mg / L⁻¹), in accordance with CONAMA Resolution No. 420/2009. Nickel (Ni) varied between 7.0 and 41.0 mg L⁻¹, with two samples above the prevention limit (30.0 mg / L⁻¹) but below the investigation limit (100.0 mg / L⁻¹), according to CONAMA Resolution No. 420/2009.

Cluster 2 is composed of samples with concentration of metals in intermediate values. Aluminum (Al) and vanadium (V) varied between 1.4 and 3.1 mg L⁻¹ and 39.0 and 65.0 mg L⁻¹, respectively. The concentration of chromium (Cr) varied between 24.0 and 68.0 mg L⁻¹, with all samples below the prevention limit (75.0 mg / L⁻¹), according to CONAMA Resolution No. 420/2009. Nickel (Ni) varied between 0.5 and 29.0 mg L⁻¹, with all samples below the prevention limit (30.0 mg / L⁻¹), according to CONAMA Resolution No. 420/2009.

Cluster 3 was characterized by house dust samples with lower concentrations of metals. Aluminum (Al) and vanadium (V) varied between 0.8 and 2.9 mg L⁻¹ and V between 1.0 and 47.0 mg L⁻¹, respectively. The concentration of chromium (Cr) varied between 9.0 and 63.0 mg L⁻¹, with all samples below the prevention limit (75.0 mg / L⁻¹), according to CONAMA Resolution No. 420/2009. Nickel (Ni) varied between 0.5 and 27.0 mg L⁻¹, with all samples below the prevention limit (30.0 mg / L⁻¹), according to CONAMA Resolution No. 420/2009.

According to the Factor 3 variables (Sr, Mo and K), three similar groups were generated, which comprised 66%, 33% and 1% of the samples analyzed in Clusters 1, 2 and 3, respectively (Figure 1E). Cluster 1 was characterized by low concentrations of strontium (Sr) and potassium (K), varying between 36.0 and 112.0 mg L⁻¹ and 0.1 and 2.8 mg L⁻¹, respectively, both do not have reference value in CONAMA Resolution No. 420/2009. The concentration of molybdenum (Mo) was the same for all samples, 0.1 mg L⁻¹, below the prevention limit (30.0 mg / L⁻¹) of CONAMA Resolution No. 420/2009. Cluster 2 was characterized by intermediate concentrations of strontium (Sr) and potassium (K), varying between 119.0 and 296.0 mg L⁻¹ and 0.1 and 3.4 mg L⁻¹, respectively. The concentration of molybdenum (Mo) did not vary between samples, remaining 0.1 mg L⁻¹. Cluster 3, on the other hand, presented a higher value for the three variables, with a concentration of 755.0 mg L⁻¹ for Sr, 1340.0 mg L⁻¹ for K and 11.0 mg L⁻¹ for Mo.

The dust samples from houses in Factor 1 showed higher values of metal concentration in places closer to the tailings basin, with higher concentration in the Chaves district, suggesting that this source of contamination still plays an important role in the contamination of the urban area of the municipality. This result agrees with the studies carried out by Machado et al. (2010) and Quiterio et al. (2001).

Although the results obtained by Quiterio *et al.* (2001) have indicated a higher concentration of these metals in the external area than inside the houses; this divergence can be explained by the methodological differences used in the two studies: the collection carried out by these authors was carried out in places subject to constant cleaning, such as carpets, furniture and curtains, while the research carried out by Cunha *et al.* (2016) sought to collect samples of dust located in places of difficult access, such as roof rafters, half-walls and behind the pictures. In areas more distant from the tailings basin, the presence of metals deposited inside the houses is accentuated due to the remobilization of sediments from the external area of the municipality.

The heavy metals found in the area have been associated with adverse effects on human health, being extensively studied by international bodies (Jarup, 2013). Pb, Cd and Zn presented the highest concentrations, well above that established by CONAMA Resolution 420/2009. These metals can cause renal, gastrointestinal effects and respiratory and neurological damage

in humans (Cetesb, 2018) Cr and Ni also showed values above the established in CONAMA Resolution 420/2009. These metals, although essential at low concentrations, can cause kidney disease, gastric irritation, dermatitis and allergic reactions in humans (Cetesb, 2018).

Heavy metals (K, V and Mn) do not have reference values, but in high concentrations they can result in respiratory, renal and neurological problems or even lead to death (Cetesb, 2018). In addition to the toxicity of these metals, the risk of effects on human health is related to long-time of exposure to these toxic metals, which can come into contact with humans through direct ingestion, dermal contact or inhalation of waste, street sediments and house dust.

4. CONCLUSION

The factor analysis allowed the classification of the most significant variables for the assessment of environmental quality, especially related to the dispersion of toxic metals that cause risk to human health. The cluster analysis classified the metals according to the main sources and allowed the analysis of environmental quality, according to CONAMA Resolution 420/2009, by different groups (Clusters). The street sediment samples resulted in three groups for Factor 1 and two groups for Factor 2, while the samples for house dust formed three groups for the three factors generated.

This allowed us to conclude that the region's natural background influences metal concentrations in the urban area, but that the main source of contamination by toxic metals is the tailings basin, abandoned since 1960. The concentration of these metals was above the maximum limits established for human exposure, such as cadmium, lead and zinc (Clusters 1, 2 and 3 of Factor 1) and chromium and nickel (Cluster 1 of Factor 2).

The results obtained through the multivariate analysis allow the formulation of a more assertive environmental management plan, able to assist in the risk assessment to human health due to the exposure to toxic metals in the municipality.

5. ACKNOWLEDGMENTS

The authors would like to thank the Coordination for the Improvement of Higher Education Personnel (CAPES) for financing this research, through the process: 88882.452960 / 2019-01.

6. REFERENCES

- ABIYE, O. E.; SUNMONU, L. A.; AJAO, A. I.; AKINOIA, O. E.; AYOOLA, M. A.; JEGEDE, O. O. Atmospheric dispersion modeling of uncontrolled gaseous pollutants (SO₂ and NO_x) emission from a scrap-iron recycling factory in Ile-Ife, Southwest Nigeria. **Cogent Environmental Science**, v. 2, p. 1275413, 2016. <https://doi.org/10.1080/23311843.2016.1275413>
- ALVES, F. E. A.; BERTOLINO, L. C.; MENDES, J. Mineralogical Characterization of Lead Mine Tailing in Boquira, Bahia State, Brazil. **Anuário do Instituto de Geociências**, v. 40, n. 3, p. 14-23, 2018.
- BEZERRA, F. A. Análise Fatorial. In: CORRAR, L. J.; PAULO, E.; DIAS FILHO, J.M. (ed.). **Análise Multivariada**. São Paulo: Atlas, 2014, p. 73-130.
- CALLEGARI-JACQUES, S. M. **Bioestatística: princípios e aplicações**. Porto Alegre: Artemed, 2003. 255p.

- CARVALHO, I. G. Chemical deposits associated to meta volcano sedimentary sequences of the central portion of the São Francisco craton. The state of Bahia, Brazil: a review. **Revista Brasileira de Geociências**, v. 2, n. 30, p. 279-284, 2000.
- CETESB. **Informações Toxicológicas**. 2018. Available at: <https://cetesb.sp.gov.br/laboratorios/servicos/informacoes-toxicologicas/#1530563846759-63efbc36-4e83> Access: 09 Apr. 2020.
- CONAMA (Brasil). 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] União**: seção 1, Brasília, DF, n. 249, p. 81-84, 30 dez. 2009.
- CUNHA, F. G. da; VIGLIO, E. P.; ANJOS, J. A. S. A.; LOUREIRO, T. B. **Estudos geoquímicos no município de Boquira - Estado da Bahia**. Salvador: CPRM, 2016. 37p.
- FINLAYSON-PITTS, B. J.; PITTS, J. N. P. **Chemistry of the Upper and Lower Atmosphere**. San Diego: Academic Press, 2000. 696p.
- GARCIA, P. M. P. **Análise comparativa de dados geológicos, litogeoquímicos e geofísicos das formações ferríferas do Complexo Boquira e supergrupo espinhaço na região de Boquira, BA**. 2011, 165f. Monografia (Graduação em Geologia) - Instituto de Geociências, Universidade Federal da Bahia, Salvador, 2011.
- GOMES, M. C. R.; ANJOS, J. A. S. A.; DALTRO, R. R. Multivariate statistical analysis applied to the evaluation of groundwater quality in the central-southern portion of the state of Bahia – Brazil. **Revista Ambiente & Água**, v. 15, n. 1, 2020. <https://dx.doi.org/10.4136/ambi-agua.2408>
- JÄRUP, L. Hazards of heavy metal contamination. **British Medical Bulletin**, v. 68, n. 1, p. 167–182, 2013. <https://doi.org/10.1093/bmb/ldg032>
- JOHNSON, R. A.; WICHERN, D. W. **Applied Multivariate Statistical Analysis**. 6. ed. Upper Saddle River: Pearson/Prentice Hall, 2007. 794p.
- LANDIM, P. M. B. **Análise estatística de dados geológicos multivariados**. São Paulo: Oficina de textos, 2011. 208 p.
- MACHADO, S. L.; PORTELLA, R. B; CESANA, E.; RABELO, T. S.; LOPES, D. O. M. Estudo da influência na contaminação do solo por metais pesados derivada das emissões atmosféricas de uma metalúrgica desativada no município de Santo Amaro – BA. In: SIMPÓSIO ÍTALO-BRASILEIRO DE ENGENHARIA SANITÁRIA E AMBIENTAL, 10. 2010. **Artigos[...]** São Paulo: ABES, 2010.
- PEREIRA, J. L. G.; FORTES, J. D. N.; MARTINS, E. M. Poluição do ar por material particulado em área intraurbana no Rio de Janeiro: aspectos metodológicos. **Revista Eletrônica de Engenharia Civil**, v. 10, n. 3, p. 53–67, 2015. <https://doi.org/10.5216/reec.v10i3.32901>
- QUITERIO, S. L.; SILVA, C. R. S; VAITSMAN, D. S.; MARTINHON, P. T; MOREIRA, M. F. R; ARAÚJO, U. C.; MATTOS, R. C. C. C.; SANTOS, L. S. C. Uso da poeira e do ar como indicadores de contaminação ambiental em áreas circunvizinhas a uma fonte de emissão estacionária de chumbo. **Cadernos de Saúde Pública**, v. 17, n. 3, p. 501-508, 2001.

- SILVA, A. P. S.; ASMUS, C. I. F.; PAVIH, J. L. P.; LACERDA, J. C. V.; SALES, L. B. F.; RESENDE, M. T.; SARAIVA, R. D. S.; CARMA, T. F. M. **Estudo de avaliação de risco à saúde humana em localidades atingidas pelo rompimento da barragem do Fundão – MG**. Relatório técnico final. São Paulo: Ambios Engenharia e Processos, 2019. 369p.
- VIANNA, N. A.; GONÇALVES, D.; BRANDÃO, F.; BARROS, R. P.; AMADO FILHO, G. M.; MEIRE, R. O.; TORRES, J. P. M.; MALM, O.; D'OLIVEIRA JÚNIOR, A.; ANDRADE, L. R. Assessment of heavy metals in the particulate matter of two Brazilian metropolitan areas by using *Tillandsia usneoides* as atmospheric biomonitor. **Environmental Science and Pollution Research**, v. 18, n. 3, p. 416-427, 2011. <https://doi.org/10.1007/s11356-010-0387-y>