# Preliminary evidence of prehistoric human activity by chemical analysis of sediments from Lapa Grande de Taquaraçu archaeological site using INAA

Diego R. G. Tudela[1] · Astolfo G. M. Araujo[2] · Sonia H. Tatumi[3] · Juan C. R. Mittani[3] · Casimiro S. Munita[1]

## Abstract

This paper analyzes the mass fractions of Ce, Co, Cr, Cs, Eu, Fe, Hf, K, La, Lu, Nd, Rb, Sc, Sm, Ta, Tb, Th, U, Yb and Zn determined from sixty sediment samples obtained from the Lapa Grande de Taquaraçu archaeological site using instrumental neutron activation analysis, INAA. The archaeological site is located in the municipality of Jaboticatubas, about 60 km from Belo Horizonte, Minas Gerais State, Brazil. The dataset was explored by means of cluster analysis, principal component analysis and discriminant analysis. The study showed three different compositional groups related to anthropogenic sediment, fluvial system and oxisols, respectively. The crystalline structure of each group were studied using X-ray diffraction, XRD.

**Keywords** Paleoindians · Sediment · INAA · Taquaraçu archaeological site · Geoarchaeology · Archaeometry

## Introduction

Soil chemistry can be applied in archaeological research as a technique to locate sites with few or no visible archaeological features [1, 2]. Moreover, the physical and chemical properties of sediments found in rockshelters and other archaeological sites may be altered by several factors, such as the human activities and by the environment nature. The degree of anthropogenic influence to identify the specific human activities may be studied by means of the analysis of minor and trace elements. Several authors used soil phosphorus analysis in the phosphate form, $PO_4$, as a chemical indicator of anthropogenic activity and to study the intensity of the occupation [3–8]. In order to achieve this, it is necessary to analyze multiple attributes of sediments formed by complex interactions [9]. In this paper were used the mass fractions of Ce, Co, Cr, Cs, Eu, Fe, Hf, K, La, Lu, Nd, Rb, Sc, Sm, Ta, Tb, Th, U, Yb and Zn determined by INAA and XRD to aid in understanding the formation processes of the Taquaraçu archaeological site.

## Study area

The archeological site of Lapa Grande de Taquaraçu is located inside a limestone rockshelter and dates from the Paleoindian period (Early Holocene). The rockshelter sits on the left margin of the Taquaraçu river, city of Jaboticatubas, Minas Gerais State, about 60 km from Belo Horizonte.

The karst, where the site is located, is part of the Serra de Santa Helena mountain range, which, in turn, along with the Jequitaí, Sete Lagoas, Lagoa do Jacaré, Serra da Saudade and Três Marias formations. They form part of the Bambuí group, which is a thick carbonate succession of the São Francisco basin, covering an extensive area of the São Francisco Craton, a Neoproterozoic geotectonic unit formed from the Brasiliana orogeny [10, 11] (Fig. 1).

The Bambuí Group is divided into two members, the upper Lagoa Santa Member shows chemical compositions reaching 96% of $CaCO_3$, 1.4% of $MgCO_3$, and 0.34% of $SiO_2$, whereas the lower Pedro Leopoldo Member is much reacher in silica, with only 67% of $CaCO_3$, 3% of $MgCO_3$, and 23% of $SiO_2$ [12, 13]. The rockshelter was formed into the Pedro Leopoldo Member. The sheltered area is

✉ Casimiro S. Munita
camunita@ipen.br; camunita@yahoo.com

1  Nuclear and Energy Research Institute, IPEN/CNEN - SP, Av. Professor Lineu Prestes 2242, São Paulo, SP 05508-000, Brazil

2  Museum of Archaeology and Ethnology, University of Sao Paulo, Av. Professor Almeida Prado 1466, São Paulo, SP 05508-070, Brazil

3  Federal University of Sao Paulo, Rua Dr. Carvalho de Mendonça 144, Santos, SP 11070-102, Brazil
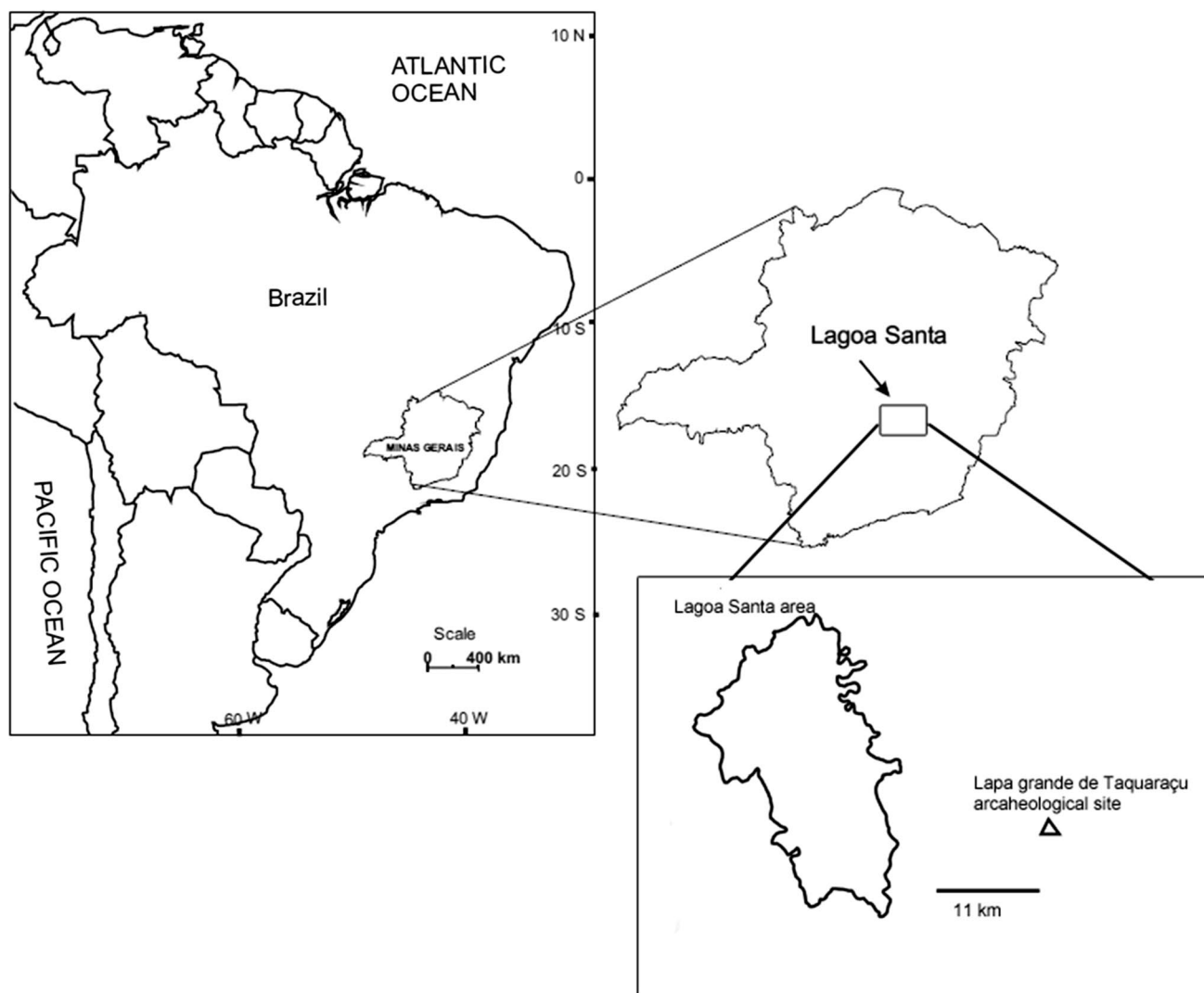
**Fig. 1** Location of the Lagoa Santa area and the Lapa Grande de Taquaraçu archaeological site. Map modified from Araujo, Neves and Kipnis [12]

approximately 30 m long by 9 m in wide at its widest point. The entrance faces west and its floor is elevated 7 m above the river that runs alongside it (Figs. 2 and 3). The radiocarbon ages obtained place the onset of the Paleoindian occupation at 11,500 cal BP ($9960 \pm 60$ RCYBP) [12]. The rockshelter was occupied in a continuous manner until 9000 cal BP ($8080 \pm 40$ RCYBP), Table 1.

Carbonate rocks (limestone, marble and dolomite), of which the main mineral is calcite (and/or dolomite), which dissociates in $Ca^{2+}$ and/or $Mg^{2+}$ and $CO_3^{2-}$ ions through water action, are among the most favorable for karstification.

A main corrosive rock agent is rainwater, initially acidified with atmospheric $CO_2$, which undergoes a significant enrichment in carbonic acid when it passes through the soil due to the $CO_2$ produced by the respiration of plant roots and the decomposition of organic in the soil. Carbonic acid is



**Fig. 2** Overview of Lapa Grande de Taquaraçu rockshelter. Photo by Astolfo G.M. Araujo
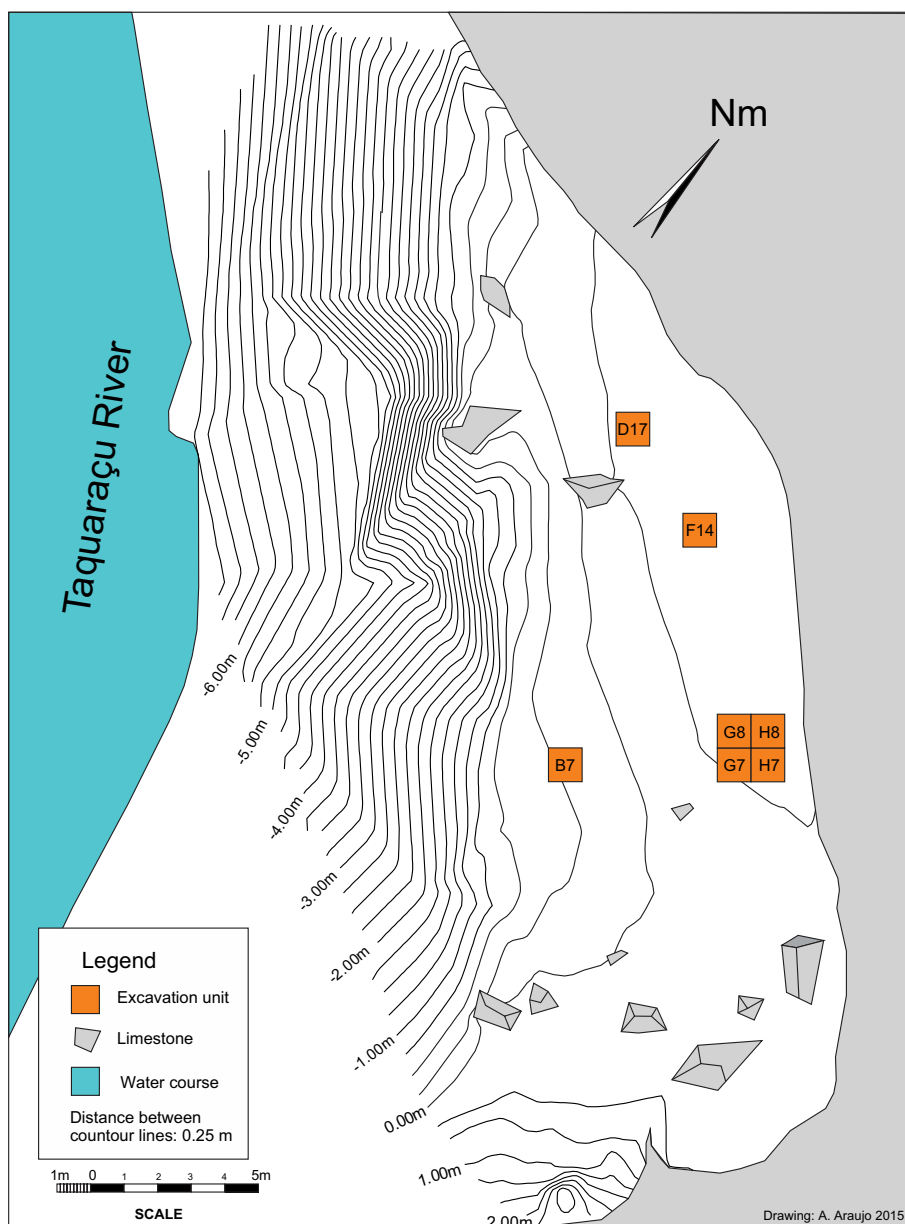
almost entirely consumed in the first few meters of groundwater seepage in the rocky batch, and in the deeper parts only a small part of this acid remains to dissolve the rock [15].

## Experimental

### Site stratigraphy and sampling

The four excavated units named G7, H7, H8 and D17 (Fig. 4) show a well-preserved stratigraphy, which makes it possible to detect the existence of an archaeological deposit approximately 80 cm thick, with a high density of lithic and faunal remains and a few ceramic fragments on the surface. In general, the archaeological deposit is gray and powdery in appearance, lying over a layer of limestone blocks with a reddish, archaeologically- sterile matrix that was part of the original cave floor (Fig. 4).

Previous studies have shown that the gray and powdery sediments are mainly anthropogenic, related to the human accumulation of wood ash [16–19]. However, the actual contribution of other sources of sediment remained ambiguous. The proximity of the river suggested a possible riverine input, and in spite of the geometry of the walls, colluvial input could not be entirely ruled out, especially due to the presence of thin, bright red layers of sediment that seemed burned soil (Fig. 5).

**Table 1** Calibrated dates from the Lapa Grande de Taquaraçu using calibration curve CalPal2007_HULU [14]

| Sample number | Level | Radiocarbon age, years BP | Calibrated age, years BP |
|---|---|---|---|
| TQ 421 | | 1 1160 ± 60 | 1087 ± 78 |
| TQ 417 | | 1 8080 ± 40 | 9028 ± 41 |
| TQ 402 | E profile | 8230 ± 50 | 9202 ± 82 |
| TQ 430 | | 3 8310 ± 40 | 9344 ± 59 |
| TQ 441 | | 4 8730 ± 40 | 9703 ± 82 |
| TQ 404 | E profile | 8730 ± 50 | 9720 ± 102 |
| TQ 454 | | 5 8910 ± 40 | 10,052 ± 96 |
| TQ 459 | | 6 9040 ± 40 | 10,218 ± 17 |
| TQ 297 | | 7 9540 ± 90 | 10,896 ± 164 |
| TQ 268 | | 8 9550 ± 60 | 10,911 ± 139 |
| TQ 295 | | 6 9620 ± 40 | 10,979 ± 131 |
| TQ 536 | | 9 8990 ± 60 | 11,477 ± 133 |
| TQ 544 | | 10 9900 ± 60 | 11,349 ± 89 |



**Fig. 4** Top view of excavated unit showing a clear difference in the color of the sediments. The bottom shows the limestone blocks immersed on reddish sediment (original cave floor) whereas the walls and surface consist of a grayish sediment, derived from anthropogenic activities. Photo by Astolfo G.M. Araujo
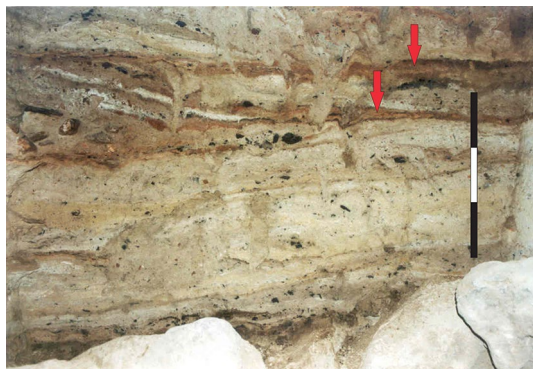


**Fig. 5** Aspect of the stratigraphy on north wall of excavation unit H8, where it is possible to see thin, reddish layers of soil (arrows) intermingled with gray ash layers. Bar scale showing 30 cm. Photo by Astolfo G.M. Araujo

The excavation was done by means of a mixed strategy, in 10 cm excavation levels for general recording, sketches and photographs, coupled with the individual observation, excavation and sieving of discrete layers, or facies [16, 20]. Each facies was subject to the collecting of 3-l bulk sediment samples. A small fraction (approximately 120 mg) of 60 of these sediment bulk samples was analyzed in this study.

The 60 sediment samples were prepared by manually grinding them with an agate mortar and pestle until a fine enough granulometry was obtained. The particle size required had to pass through a 100 mesh sieve. These materials dried in an oven for 24 h, at 105 °C were stored in a desiccator.

## INAA methodology

Constituent Elements in Coal Fly Ash (NIST-SRM-1633b) used as a standard and Trace Elements in Soil, RM-IAEA-Soil 7, used to check the results throughout the analysis. These materials dried in an oven for 2 h at 105 °C were stored in a desiccator until weighing [21].

About 120 mg of sediments samples, one standard, Constituent Elements in Coal Fly Ash, NIST-SRM 1633b, and the reference material, Trace Elements in Soil, IAEA-Soil-7, weighed in polyethylene bags wrapped in aluminum foil were irradiated in the research reactor, IEA-R1, at IPEN-CNEN/SP, Brazil, with a thermal neutron flux of about $5 \times 10^{12}$ cm$^{-2}$ s$^{-1}$ for 8 h.

Two measurement series were carried out using a Ge (hyperpure) detector, model GX 2519 from Canberra, with resolution of 1.9 keV at the 1332.49 keV gamma peak of $^{60}$Co and an S-100 MCA from Canberra with 8192 channels. K, La, Lu, Na, Nd, Sm, U, and Yb were measured after 7 days cooling time and Ce, Co, Cr, Cs, Eu, Fe, Hf, Rb, Sc, Ta, Th and Zn after 25–30 days. Gamma ray spectra analysis and the determination of the concentrations were carried out using the Genie-2000 Neutron Activation Analysis Processing Procedure from Canberra. Details of the analytical procedure have been shown elsewhere [21].

## X-ray diffraction

All diffraction measurements were performed on a RIGAKU diffractometer, Model Miniflex II using a step of 0.05° ($\theta$) with 1 s reading for each step. The measurements were initiated at 5° ($\theta$) and ended at 80° ($\theta$). The diffractograms were analyzed using the Crystallographica Search-Match program, version 1.11.

## Statistical treatment of the dataset

By INAA it is possible to determine several chemical elements and the interpretation of the data is complex,

especially when a large number of samples are analyzed. Therefore, multivariate statistical methods are generally used for data handling and assessment. In multivariate statistical techniques, each sample can be considered a point in multidimensional space, whose number of dimensions is equivalent to the number of elements measured. Groups of samples of similar composition form a cloud of points surrounding its centroids. In this paper, the dataset was studied using cluster analysis, principal component analysis and discriminant analysis.

## Results

The analytical quality control of the method were tested using the concentration of Ce, Co, Cr, Cs, Eu, Fe, Hf, K, La, Lu, Na, Nd, Rb, Sc, Sm, Ta, Tb, Th, U, Yb and Zn using a dataset of 19 independent determinations of the reference material IAEA-Soil 7, Trace Elements in Soil. The homogeneity of the set of results for each element was studied using Dixon's Q test for statistical treatment of rejection of deviant values at the 95% confidence level [22]. In the dataset, no outliers were found. The observed results were compared to the certified values. The RSD for most elements (Ce, Co, Eu, Fe, Hf, Na, Rb, Sc, Sm, Th, U, Yb, and Zn) were around 5% or less which are comparable with that found in the literature [23]. For the elements determined with a precision of 10% or more (Cs, K, La, Ta) ours results are also in agreement with those found in the literature [23]. The determination of Zn is not reliable as a consequence of strong γ-ray interference by $^{46}$Sc (1120.3 keV) and $^{182}$Ta (1121.41 keV). The interference by the $^{235}$U fission in the determination of La, Ce, and Nd was negligible because the U concentration did not exceed 5 ppm and the rare earth elements were not extremely low [24]. In this work, we considered only elements determined with a RSD better than 10%. This precision is considered by several authors as suitable for the choice of chemical elements using multivariate statistical methods [25]. Although Rb, Tb, Sm, K and Ta, showed a RSD less than 10%, they were not included in the dataset due to the natural heterogeneity.

Therefore, elements Ce, Co, Cs, Eu, Fe, Hf, La, Na, Sc, Th, U and Yb were used in the sediment data analyses. None of these elements contained missing values.

In geochemistry, concentration data is often assumed to follow a lognormal distribution after being log10 transformed [26, 27].

In addition, the dataset of the sediment was submitted to an outlying test using the Mahalanobis distance [28]. Outliers can have a considerable influence on multivariate statistical methods because they can disturb homogeneous groups. For each of n samples and p variables, the

Mahalanobis distance (Di) was taken from the sample to the centroid, as calculated by the expression [28]:

$$D_i = \sqrt{(x_i - \bar{x})' S^{-1} (x_i - \bar{x})} \tag{1}$$

where $'$ represents the transpose matrix; $S = \sum_{i=1}^{n} (x_i - \bar{x})' (x_i - \bar{x})$ is the variance–covariance sampling matrix. Also, $(x_i - \bar{x})$ represents the difference between the column vector of concentrations measured for one sample and the vector of concentration means for the group being analyzed [29, 30]. Each one of these values is compared with the critical value, cv, which can be calculated using the lambda Wilks criteria [29, 30], calculated as follows:

$$\frac{p(n-1)^2 F_{p,n-p-1;\alpha/n}}{n(n-p-1+pF_{p,n-p-1,\alpha/n})} \tag{2}$$

where $p$ is the number of variables; $n$ is the number of samples and; $F$ is the F test called "Fisher distribution" ($F = s_1^2/s_2^2$ where $s_1^2$ and $s_2^2$ are the sample variances), with $p$ degrees of freedom at a significance level of $\alpha/n$, $\alpha = 0.05$.
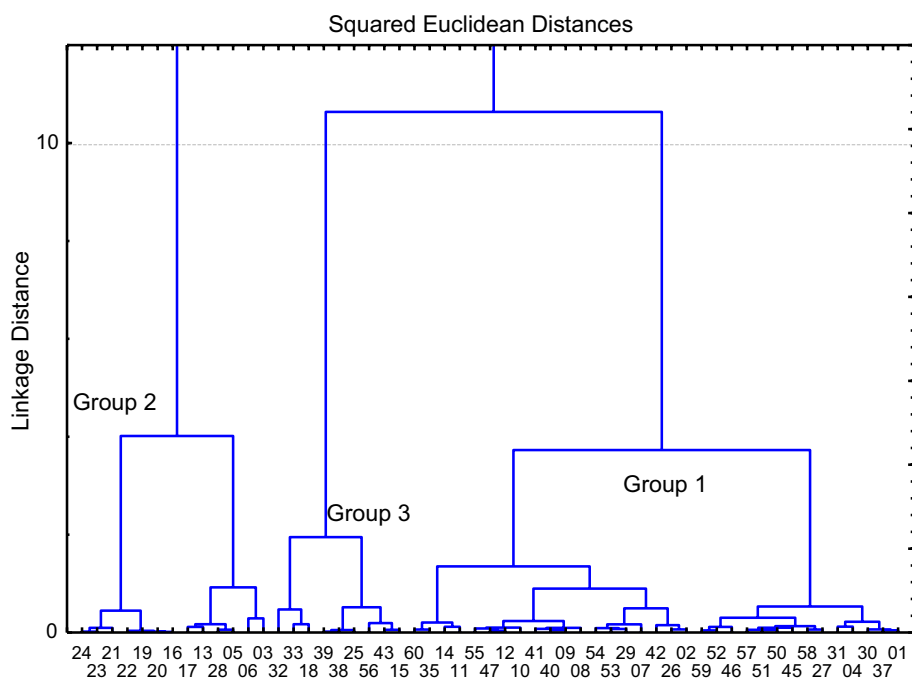
When the value found by expression (1) is larger than the critical value produced by expression (2), the sample is considered to be an outlier. Thus, the Mahalanobis distance values for each sample were calculated and compared to the critical value. In accordance with the Mahalanobis distance rule, five outlier samples were found. Due to the small number of samples, it is difficult to know whether it is an archaeological meaningful group. We believe that the outlier samples could be related to very specific processes that may have occurred either during burial or due to specific human activities, and this notion is strengthened by the fact that lower or higher concentrations of Ce, Na, and Th were found as compared with other samples. The outliers were subsequently removed from the dataset because multivariate statistical analysis are highly sensitive to outliers [31].

The 55 remaining samples were then analyzed using cluster analysis, principal component analysis and discriminant analysis.

Cluster analysis is usually used for the initial inspection of results because it is a fast and efficient way of evaluating relationships between data. The cluster technique groups similar objects according to their characteristics and analysis is based on the similarity/dissimilarity matrix where the distance between all pairs of samples is calculated using one of the several distances [32]. In this work, was used the Ward method and the squared Euclidean distance because it exaggerates differences in the data and produces a clearly identifiable cluster dendrogram.

Figure 6 shows the dendrogram of the cluster analysis, which in turn shows the presence of two principal groups

**Fig. 6** Dendrogram showing the results of hierarchichal cluster analysis of the 55 specimens



the three principal components and Cr into two principal components. In a way, all the elements selected for the characterization of sediments using multivariate statistical analysis at the Taquaraçu site are well correlated. Another fact that also merits attention is that the elements with higher weighting coefficients are the elements of the first transition series (Co, Cr, Fe and Sc) and the rare earth elements (Ce, Eu, La, Yb and Lu), which is consistent with the literature in other studies [34].

where, apparently, the largest group is divided into two subgroups (1 and 3).

Initially, the data obtained by cluster was worked by assuming the existence of only two groups, where it was supposed that these two groups corresponded to sedimentation of geological and anthropogenic origin. However, after using other statistical methods, the presence of a third group became apparent. To confirm the existence of the third group, the results were submitted to a principal component analysis, PCA.

PCA involves a transformation of data based on eigenvalues to determine the magnitude and direction of maximum variance in the distribution of the dataset in hyperspace [33]. The first principal component is a linear combination of the original variables and is aimed at maximum variance. The second component is estimated to be aimed at remaining maximum variance (2nd most important), the third component carries the third most important amount of transformation and so on.

Table 2 shows the weighting coefficients for each chemical element and the variances explained by the three first components. Also included are the loadings of the principal component for each element. In the first component, the explained variance was 90.27%, in the second, 3.85% and 2.46% in the third. Therefore, the variance explained by these three components was 96.58%, while the first two accounted for 94.12% of the total variance of the database.

According to Table 2, the elements that showed higher loading on the first component were Ce, Co, Cr, Cs, Eu, Fe, Lu, Th and Yb; Hf and Th in the second component and Na and U in the third component. Uranium is distributed in

**Table 2** Results of the variance–covariance matrix for the first three principal components

| Element | PC1 | PC2 | PC3 |
|---|---|---|---|
| Ce | 0.24589 | 0.10292 | − 0.2221 |
| Co | 0.37202 | − 0.3868 | − 0.0667 |
| Cr | 0.291011 | 0.01106 | 0.28186 |
| Cs | 0.31187 | 0.12484 | 0.11539 |
| Eu | 0.21976 | − 0.254 | − 0.3125 |
| Fe | 0.33139 | 0.03643 | − 0.0291 |
| Hf | 0.17734 | 0.45635 | 0.15125 |
| La | 0.13092 | − 0.0074 | − 0.3655 |
| Lu | 0.30123 | − 0.0067 | − 0.1139 |
| Na | 0.16077 | − 0.5819 | 0.54366 |
| Sc | 0.31815 | 0.02042 | − 0.064 |
| Th | 0.2468 | 0.30243 | − 0.1067 |
| U | 0.20697 | 0.34136 | 0.49828 |
| Yb | 0.30156 | − 0.053 | − 0.1618 |
| Var. explained (%) | 90.27 | 3.85 | 2.46 |

Figure 7 shows the graph of principal component 1 versus principal component 2 in the 55 sediment samples. As can be seen, the results confirm the existence of three sample groups. Group 1 is formed by 35 samples, Group 2 by 11 and Group 3 by nine samples. The scattering of the samples in Group 3 may be due to the fact that some of the samples have a certain probability of belonging to another group. Bivariate graphs were drawn up to confirm this possibility. Figure 8 shows the Co versus Th (a) and Sc versus Yb (b) biplot. In these figures, the existence of three distinct compositional groups can clearly be seen.
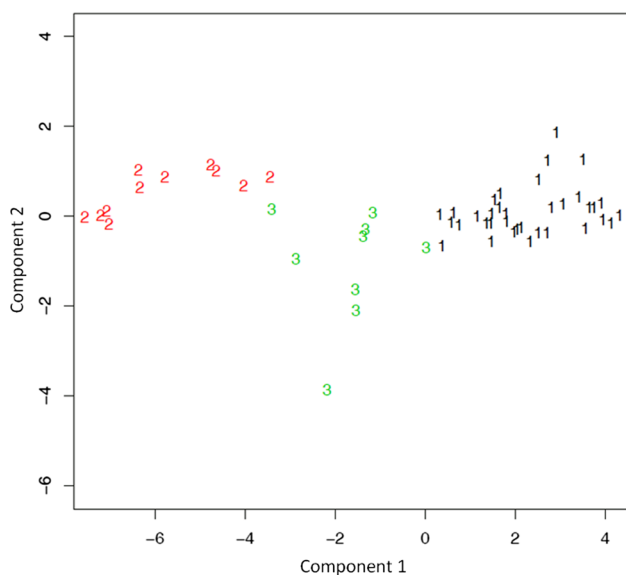
Although it was possible to separate the sediment samples of the Taquaraçu site into three compositional groups differentiated by bivariate graph using the combination of almost all the 14 elements that are included in the database, the greater difference between the groups (largest distance) involved the elements Co versus Th and Sc versus Yb. Nevertheless, such elements are part of the chemical groups of the periodical table that are important in archeometric studies from the geochemical standpoint, transition metals (Sc and Co) and inner transition metals (Yb and Th), where Yb is part of the heavy rare earths.

In order to confirm the existence of the three groups, the results were studied using discriminant analysis. Discriminant analysis is another technique used to reduce sample dimensionality. Discriminant analysis differs from PCA in the way that it extracts a new set of variables that maximize the difference of the total variance of the dataset [35]. The purpose of this technique is to find the functions of the original variables (discriminant function) that explain the differences between the samples.

Figure 9 shows discriminant function 1 versus discriminant function 2 in which the three different chemical composition groups in the sediment of the Lapa Grande de Taquaraçu archaeological site were identified.

The INAA method enables several elements to be determined simultaneously and many of which may not be very informative. Therefore, it is necessary to identify the most relevant subset of elements and to remove the variables with the least amount of relevant information, while preserving multivariate data structure and minimizing the loss of essential information [36]. The Procrustes analysis was used for this paper in conjunction with a stopping rule to select the variables that are in some way suitable for discrimination [36–38]. The concept of Procrustes uses the measured distance, $M^2$, between two $k$-dimensional configurations



**Fig. 7** Principal component 1 versus principal component 2 of the sediments samples. Where the numbers 1, 2 and 3 are the Groups 1, 2 and 3, respectively
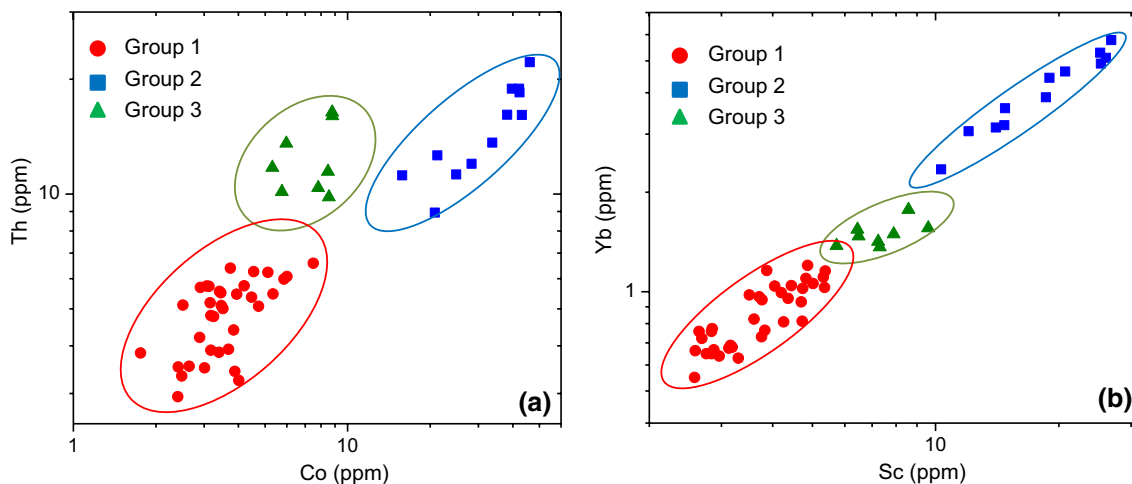


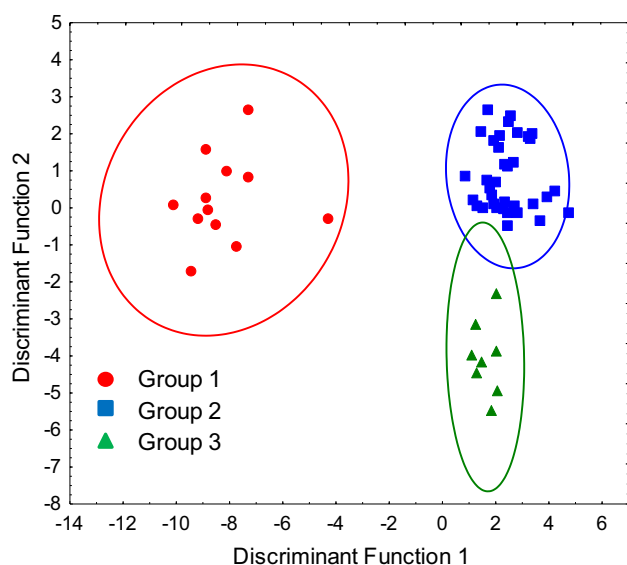**Fig. 8** Log-log scatterplot of Co versus Th (**a**) and Sc versus Yb (**b**). The ellipses represent a confidence level of 95%

**Fig. 9** Discriminant function 1 versus discriminant function 2. The ellipses represent a confidence level of 95%
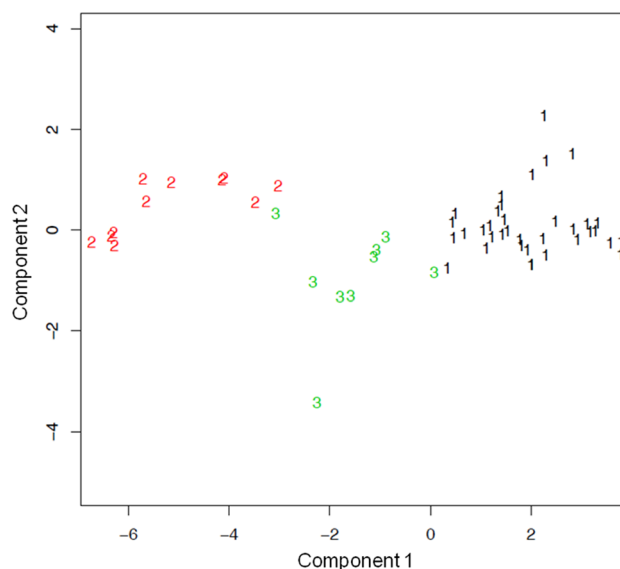


**Fig. 10** Principal component 1 versus principal component 2 using the variables Eu, Cr, Th, U, Co, Yb, Cs, Na, La, Sc and Hf. The numbers 1, 2 and 3 are the Groups 1, 2 and 3, respectively

of PCA. The value of $M^2$ is determined for each variable extracted and the resulting value indicates the effect in the configuration and identifies the variable that has the lowest effect when eliminated [38]. The extraction of variables continues up to point where the sum of residues ($M^2$) is higher than the critical value (cv) [37, 38]. Table 3 shows the results of the selection procedure, including the sequence of elimination.

In Table 3, the first element to be eliminated was Fe, where $M^2$ is 1.25 and the cv is 19.1. When each variable is eliminated, the associated configuration is changed up to the point that it becomes unacceptable, which occurs when $M^2$ becomes greater than the critical value [38]. This value was reached with Eu, whose value of $M^2$ was 21.1 and that of the critical value, 16.3. The procedure suggested that the Fe, Lu and Ce variables can be eliminated and the elements Eu, Cr, Th, U, Co, Yb, Cs, Na, La, Sc and Hf should be considered in the dataset.

In order to confirm this assumption, the results of 55 samples were subjected to a new principal component analysis using the eleven variables selected in the Procrustes method. Figure 10 shows the graph of principal component 1 versus principal component 2, which shows the existence of

three different compositional groups. When Figs. 7 and 10 are compared, it becomes apparent that using the variables selected by Procrustes, the complete data structure is represented perfectly without a loss of information. This result shows that the sediment of the Lapa Grande de Taquaraçu site can be characterized by the analysis of Eu, Cr, Th, U, Co, Yb, Cs, Na, La, Sc and Hf.

The results of the chemical composition show that the site is formed by three different groups. According to previous studies [12] the site has an archaeological stratum approximately 1 m deep, underneath which it appears the sedimentation composed of what would be the original floor of the shelter. Table 4 lists information pertaining to the samples according to the group.

Group 1 consists of samples from the excavation of four excavation units, consisting of 15 samples from excavation unit G7, one from H7, 15 from H8, and four from D17. Of the 12 samples that form Group 2, 11 are from excavation unit G7 and one is from H7. In Group 3, five samples derived from excavation unit H8, one from G7 and two of soil collected outside the rockshelter. As shown in Table 4, the separation of the sediment samples of Groups 1 and 2 occurred mostly according to the depth.

The elementary concentration values of the samples as a whole and by group were analyzed to study the behavior of the analyzed elements. The means of elemental concentrations are separated according to groups (Table 5).

The mean of the concentrations of the 55 samples showed a large scatter when the whole data set is considered, column 2. On the other hand, when the mean of the concentrations of each sample group are taken into

**Table 3** Variable selection with stopping rule, $n = 55$

|       | Fe   | Lu   | Ce   | Eu   | Cr, Th, U, Co, Yb, Cs, Na, La, Sc e Hf |
|-------|------|------|------|------|----------------------------------------|
| $M^2$ | 1.25 | 4.97 | 11.4 | 21.1 |                                        |
| cv    | 19.1 | 18.1 | 17.2 | 16.3 |                                        |

**Table 4** Distribution of samples according to the group, excavation unit and depth

| Group | Number of samples | Excavation unit | Depth (cm) |
|---|---|---|---|
| 1 | 15 | G7 | 10–90 |
| | 1 | H7 | 80 |
| | 15 | H8 | 10–80 |
| | 4 | D17 | 10–70 |
| 2 | 11 | G7 | 90–140 |
| | 1 | H7 | 80 |
| 3 | 5 | H8 | 20–50 |
| | 1 | G7 | 50 |
| | 2 | Outside soil | – |

consideration (columns 3, 4 and 5), the scatter is reduced. Table 5 shows that the concentration is higher for the elements of Group 2 as compared with the other groups, with the exception of Hf, whose concentration is higher in Group 3.

Table 5 also shows that samples of Group 1 presents lower levels of elemental concentrations of each element.

In order to interpret the data, it was take into consideration not only the results from the 55 samples, but also the five outliers, since they could also help convey important information about the site and its environs. Three of these outlier samples were excavated inside the rockshelter and two outside (one from the base of the limestone outcrop and the other from the modern river terrace). The concentrations of each element analyzed were significantly different between on and off site contexts.

To study the possible differences in crystalline structure between the three different compositional groups formed by the sediments of the Taquaraçu site, XRD

measurements were taken. From a total of 8 samples, three belong to Group 1, three to Group 2 and two to Group 3.

The samples in Group 1 contained calcite as their major constituent while the main constituent of Group 2 and 3 was quartz with no, or at least no observed calcite. Although Ca was not determined by INAA, its presence is inferred from the lithology of the rock on which the Lapa Grande de Taquaraçu site is formed, since this consists of limestone of high chemical purity.

Therefore, based on the XRD spectra, we can assume that Ca is present in samples of Group 1 in the form of calcite and this crystalline structure is much superior to the quartz one.

## Discussion

The chemical and macroscopic characteristics (such as the gray color and pulverulent aspect) of the samples pertaining to Group 1 converge to interpret this group as representative of anthropogenic sediments. The XRD results showed a large amount of calcite in sediments of Group 1, value that could be accounted for by the dissolution of the limestone from which the cave is formed. However, wood ash is predominantly composed of calcite due to decomposition of the calcium oxalate present in timber [39]. At the site of Lapa das Boleiras [16], situated near Lapa Grande de Taquaraçu, the presence of Ca was higher in the ashes, which were recognized as archaeological, and decreased for samples that were not ash, similar to what occurs in the current dataset. Thus, the possibility that the calcite present in Group 1 may have originated from the burning of wood, and as such, is of anthropic origin, is very strong.

Group 2, on the other hand, is most probably related to sediments coming from the riverine system. XRD

**Table 5** Mean and standard deviation of the elemental concentrations for each group in mg/kg ($n = 55$)

| Element | Complete set | Group1 | Group2 | Group3 |
|---|---|---|---|---|
| Ce | 45.83 ± 29.23 | 27.7 ± 6.06 | 92.02 ± 22.61 | 55.87 ± 13.9 |
| Co | 11.76 ± 13.17 | 4.85 ± 1.37 | 34.79 ± 10 | 7.44 ± 1.5 |
| Cr | 33.50 ± 26.96 | 18.26 ± 5.06 | 77.58 ± 23.83 | 34.06 ± 9.74 |
| Cs | 4.28 ± 3.65 | 2.2 ± 0.61 | 10.20 ± 3.33 | 4.51 ± 1.37 |
| Eu | 0.79 ± 0.51 | 0.51 ± 0.1 | 1.66 ± 0.36 | 0.73 ± 0.22 |
| Fe | 24,116 ± 22,152 | 11,651 ± 3200 | 60,658 ± 20,221 | 23,838 ± 5230 |
| Hf | 3.29 ± 1.74 | 2.24 ± 0.53 | 5.01 ± 1.05 | 5.29 ± 2.25 |
| La | 29.86 ± 10.82 | 23.44 ± 3.67 | 45.67 ± 9.0 | 34.02 ± 7.71 |
| Lu | 0.26 ± 0.22 | 0.14 ± 0.03 | 0.63 ± 0.17 | 0.26 ± 0.03 |
| Na | 1171 ± 638 | 893 ± 350 | 2143 ± 501 | 929 ± 282 |
| Sc | 7.63 ± 6.77 | 3.79 ± 0.9 | 18.97 ± 5.86 | 7.41 ± 1.26 |
| Th | 8.39 ± 5.26 | 4.94 ± 1.1 | 15.75 ± 4.32 | 12.44 ± 2.61 |
| U | 1.60 ± 0.9 | 1.09 ± 0.38 | 2.79 ± 0.78 | 2.01 ± 0.8 |
| Yb | 1.66 ± 1.42 | 0.86 ± 0.2 | 4.11 ± 1.07 | 1.50 ± 0.13 |

analysis showed quartz and mica in two samples. These two samples are related to river flooding episodes (small silt "mud-cracks"). The third group was collected at the bottom of the stratigraphic sequence, in the middle of the limestone blocks, and accordingly showed some calcite in its composition, but the overall composition of the elements determined by INAA shows it to be closer to Group 2. Finally, when the outlier sample from the modern river terrace was included in the analysis, it also clustered with the samples of Group 2. To summarise, Group 2 is related to the fluvial system, and its presence inside the rockshelter is mainly due to natural factors, because: (1) the bottom of the rockshelter was part of the ancient fluvial system, and (2) events of river flooding reached the site during human occupation.

Group 3 has a chemical signature that can be related to soils (oxisols) that are ubiquitous in the region. This group encompasses two off-site samples (soil collected outside the rockshelter) as well as samples from inside the rockshelter. Some of the samples analyzed by INAA that were collected inside the excavation were interpreted in the field as "burned earth". Although it is not possible to rule out that some soil could have reach the rockshelter by natural means, the local topography and the geometry of the entrance rule out a significant soil contribution by natural factors suggesting that the material was brought in by people. These thin layers of bright red soil, forming fairly continuous horizons, were also recognized at Lapa do Santo rockshelter and interpreted by other authors [40] as being related to natural deposition. Our data suggests otherwise.

As can be seen in the Table 5, the elements that contribute most to the distinction between the samples inside the rockshelter, (anthropogenic sediment and natural sediments), and outside the rockshelter, are Ce, Co, Cr, Fe, Hf, La, Na, Sc, and Th. These results are in agreement when the five outliers that were removed and the dataset was subjected to cluster analysis, discriminant analysis, principal component analysis and bivariate plots. In this study it was possible to distinguish clearly between natural and anthropogenic sediment.

Another factor that should be mentioned, going back to Table 5, is that the sediments which supposedly correspond to the anthropogenic group (Group 1) have lower concentrations of each element. This data contrasts with those from other studies of anthropogenic sediments because, generally, human activities generate a large volume of waste which causes increases in the elemental concentrations of soil, for example, from craft activities and food preparation [41–43]. Sheltered sites such Lapa Grande de Taquaraçu, are less affected by weathering and are therefore expected to be less eroded and to have increased soil enrichment due to chemical compounds [44, 45]. However, this was not what was detected at the Lapa Grande de Taquaraçu site, where the group that corresponds to the anthropogenic stratum shows the lowest concentration of elements between the three compositional groups.

Unless the sediment is severely leached, disturbed or altered, the residue that is incorporated by and that modifies the elementary concentration of the sediment can be detected and identified through chemical analysis [9, 41, 45–48]. Thus, the low elemental concentrations in Group 1 may be explained by the fact that ash, being formed by the combustion of wood, will tend to show lesser concentrations of elements than the original soil. On the other hand, Group 2, being related to the fluvial system, shows the highest values since the water carries a large amount of minerals, both in solution and suspension. Group 3, related to the oxisols, shows intermediate values; on the one hand, the soil is richer in elements than the trees that grow over it, and on the other hand such soils are very well-developed and stable, maintaining only the minerals that are more resistant to chemical attack, and therefore more depleted than the fluvial sediments.

## Conclusions

Sixty sediment samples from four excavation units excavated from Lapa Grande de Taquaraçu rockshelter were analyzed. The samples were studied using instrumental neutron activation analysis (INAA) in which the Ce, Co, Cr, Cs, Eu, Fe, Hf, La, Lu, Na, Sc, Th, U and Yb concentrations were determined. The results of the elementary analysis were submitted to cluster, principal component, and discriminant analysis. The study showed the existence of three different compositional groups, one of which is related to anthropogenic sediments. The depth of the samples had a significant role in the formation of groups, since the upper part of the stratigraphy is mainly anthropogenic, and the lower part is geogenic. A subset of variables was also determined from the data matrix using Procrustes analysis. The results showed that only eleven variables (Eu, Cr, Th, U, Co, Yb, Cs, Na, La, Sc and Hf) are suitable for discrimination purposes. Some sediment samples from each group were analyzed by XRD and the results showed that there are differences in the crystalline phases of the samples of each group. The samples in Group 1 contained calcite as their major constituent, which may be due to the burning of wood. In the other two groups, the main species found was quartz.

Our data and results show that the analysis of a large amount of bulk samples, subject to statistical treatment, are mandatory in terms of grasping the overall aspects of site formation. More detailed studies, with focus on specific, punctual aspects of the site, should follow only after this fundamental step is accomplished.

# References

1. Konrad VA, Bonnichsen R, Clay V (1983) Soil chemical identification of ten thousand years of prehistoric human activity areas at the Munsungun Lake Throroughfare, Maine. J Archaeol Sci 10:13–28

2. Cavanagh WG, Hirst S, Litton CD (1988) Soil phosphate, site boundaries, and change point analysis. J Field Archaeol 15:67–83

3. Arrhenius O (1929) Die phosphoatfrage. Z Pflanzenenernahrung Dungung Bodenkunde 10:185–194

4. Craddock PT, Gurney D, Pryor F, Hughs M (1986) The application of phosphate analysis to the location and interpretation of archaeological sites. Archaeol J 142:361–376

5. Holliday VT, Gartner WG (2007) Methods of soil P analysis in archaeology. J Archaeol Sci 34(2):301–333

6. Wells EC, Terry RF, Parnell JJ, Hardin PJ, Jackson MW, Houston SD (2000) Chemical analyses of ancient anthrosols in residential areas at Piedras Negras, Guatemala. J Archaeol Sci 27:449–462

7. Kämpf N, Woods WI, Sombroek W, Kern DC, Cunha TJ (2003) Classification of amazonian dark earth sand other ancient anthropic soils. In: Lehmann J et al (eds) Amazonian dark earths: origin, properties, management. Springer, Dordrecht, pp 77–102

8. Costa ADR, Silva Junior ML, Kern DC, Ruivo MDLP, Marichal R (2017) Forms of soil organic phosphorus at black earth sites in the Eastern Amazon. Rev Ciênc Agron 48(1):1–12

9. Middleton WD (2004) Identifying chemical activity residues on prehistoric house floors: a methodology and rationale for multi-elemental characterization of a mild acid extract of anthropogenic sediments. Archaeometry 46(1):47–65

10. Dardenne MA (1981). Os grupos Paranoá e Bambuí na faixa dobrada Brasiliana. Anais do Simpósio sobre o Cráton de São Francisco e suas Faixas Marginais, novembro 2–7, Salvador, Bahia, pp 140–157

11. Alkmin FF, Martins AMN (2001) A bacia intracratônica do São Francisco. Arcabouço estrutural e cenários evolutivos. In Sociedade Brasileira de Geologia, Núcleo de Minas Gerais (Ed.), Bacia do São Francisco: Geologia e Recursos Naturais. Belo Horizonte, Minas Gerais, pp 9–30

12. Flores RA, Sousa JCM, Araujo AGM, Ceccantini G (2016) Before Lagoa Santa: micro-remain and technological analysis in a lithic artifact from the Itaperica industry. J Lithic Stud 3(1):6–29

13. Piló LB (1998) Morfologia cárstica e materiais constituintes: dinâmica e evolução da depressão poligonal macacos-baú-carste de Lagoa Santa, MG. Doctoral dissertation, FFLCH—University of São Paulo

14. Danzeglocke U, Jöris O, Weninger B (2007) CalPal online (Version 1.5). University of Cologne, Germany. http://www.calpal-online.de. Accessed date August 2015

15. Teixeira W, Toledo MC, Fairchild TR, Taioli F (2008) Decifrando a Terra. Companhia Nacional (Ed.), 2nd ed. São Paulo

16. Araujo AG, Feathers JK, Arroyo-Kalin M, Tizuka MM (2008) Lapa das boleiras rock shelter: stratigraphy and formation processes at a paleoamerican site in Central Brazil. J Archaeol Sci 35(12):3186–3202

17. Araujo GM, Neves WA, Kipnis R (2012) Lagoa Santa Revisited: an overview of the chronology, subsistence, and material culture of paleoindian sites in eastern central Brazil. Lat Am Antiq 23:533–550

18. Araujo AGM, Piló LB (2017) Towards the development of a tropical geoarchaeology: Lagoa Santa as an emblematic case study. In: Da-Gloria P et al (eds) Archaeological and paleontological research in Lagoa Santa. Springer, Berlin, pp 373–391

19. Silva LV, Prous AP (2014) O papel dos resíduos de combustão na formação dos estratos sedimentares na Lapa do Niactor: o elemento antrópico como agente dominante na sedimentação em abrigos. Arq Museu Hist Nat UFMG Belo Horiz 23(2):105–139

20. Stein JK, Kornbacher KD, Tyler JL (1992) British camp shell midden stratigraphy. In: Stein JK (ed) Deciphering a shell midden. Academic Press, Cambridge, pp 95–134

21. Hazenfratz R, Munita CS, Neves EG (2018) Neural networks (SOM) applied to INAA data of chemical elements in archaeological ceramics from central Amazon. Sci Technol Archaeol Res 3(2):334–340

22. Rorabacher DB (1991) Statistical treatment for rejection of deviant values: critical values of Dixon's "Q" parameter and related subrange ratios at the 95% confidence level. Anal Chem 63:139–146

23. Kuleff I, Djingova R (1990) Activation analysis in archaeology. In: Alfassi ZB (ed) Activation analysis, vol II. CRC Press, Boca Raton, pp 428–477

24. Glascock MD (1992) Characterization of ceramics at MURR by NAA and multivariate statistics. In: Neff H (ed) Chemical characterization of ceramics paste in archaeology, monographs in world archaeology. Prehistory Press, New York, pp 11–26

25. Bishop RL, Canouts V, Crown PL, Atley SP (1990) Sensitivity, precision and accuracy: their roles in ceramic compositional data bases. Am Antiq 55(2):537–546

26. Arenhs LH (1954) The lognormal distribution of the elements. Geochim Cosmochim Acta 5(2):49–73

27. Baxter MJ, Freestone IC (2006) Log-ratio compositional data analysis in archaeometry. Archaeometry 48(3):511–531

28. Grubbs FE (1950) Sample criteria for testing outlying observations. Ann Math Stat 20:27–58

29. Penny KI (1996) Appropriate critical values when testing for a single multivariate outlier by using the Mahalanobis distance. J R Stat Soc C 45(1):73–81

30. Oliveira PMS, Munita CS (2003) Influência do valor critico na detecção de valores discrepantes em arqueometria. In Anais da 48° Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria, Lavras, MG, Brasil, 07 a 11 de Julho

31. Joossens K, Croux C (2004) Empirical comparison of the classification performance of robust linear and quadratic analysis. In: Huber M, Pison G, Struyf A, Van Aelst S (eds) Theory and application of recent robust methods. Verlag Basel, Birkhauser, pp 131–140

32. Sayre EV, Dodson RW (1975) Brookhaven procedures for statistical analyses of multivariate archaeometric data. Brookhaven National Laboratory Report BNL-23128, New York

33. Glascock MD, Neff H (2003) Neutron activation analysis and provenance research in archaeology. Meas Sci Technol 14:1516–1526

34. MacDonald BL, Hancock RGV, Cannon A, Pidruczny A (2011) Geochemical characterization of ochre from central coastal British Columbia, Canada. J Arhcaeol Sci 38:3620–3630

35. Hair JF Jr, Anderson RE, Tatham RL, Black WC (2005) Análise multivariada de dados, 5a edn. Bookman, Porto Alegre, p 593

36. Krzanowski WJ (1987) Selection of variables to preserve multivariate data structure, using principal components. Appl Stat 36(1):22–33

37. Krzanowski WJ (1996) A stopping rule for structure-preserving variable selection. Stat Comput 6:51–56

38. Munita CS, Barroso P, Oliveira PMS (2013) Variable selection study using procrustes analysis. Open J Archaeom 1:31–35

39. Canti MG (2003) Aspects of the chemical and microscopic characteristics of plant ashes found in archaeological soils. CATENA 54(3):339–361

40. Villagran XS, Strauss A, Miller C, Ligouis B, Oliveira R (2017) Buried in ashes: site formation processes at Lapa do Santo rockshelter, east-central Brazil. J Archaeol Sci 77:10–34

41. Middleton WD, Price TD (1996) Identification of activity areas by multi-element characterization of sediments from modern and

archaeological house floors using inductively coupled plasma-atomic emission spectroscopy. J Archaeol Sci 23:673–687

42. Entwistle JA (2000) The geoarchaeological significance and spatial variability of a range of physical and chemical soil properties from a former habitation site, Isle of Skye. J Archaeol Sci 27:287–303

43. King SM (2008) The spatial organization of food sharing in early postclassic households: an application of soil chemistry in ancient Oxaca, Mexico. J Archaeol Sci 35:1224–1239

44. Barba LA (1996) La química en el estudio de áreas de actividad. In: Manzanilla Linda (ed) Análisis de unidades habitacionales mesoamericanas y sus Áreas de actividad. DF, México, pp 21–39

45. Barba LA (2007) Chemical residues in lime-plastered archaeological floors. Geoarchaeology 22(4):439–452

46. Barba LA (1990) El análisis químico de pisos de unidades habitacionales para determinar sus áreas de actividad. In: Sigiura Y, Serra MC (eds) Etnoarqueología: coloquio Bosch-Gimpera. México, DF, pp 177–200

47. Barba LA, Ortiz A, Link KF, López Lujan L, Lazos L (1996) The chemical analysis of residues in floors and the reconstruction of ritual activities at the Templo Mayor, México. In Orna MV (ed.) Archaeological chemistry: organic, inorganic and biochemical analysis. ACS symposium series, vol 625, pp 139–156

48. Vyncke K, Degryse P, Vassilieva E, Waelkens M (2011) Identifying domestic functional areas. Chemical analysis of floor sediments at the classical-Hellenistic settlement at Düzen Tepe (SW Turkey). J Archaeol Sci 38:2274–2292