Archaeological and Anthropological Sciences (2024) 16:70
https://doi.org/10.1007/s12520-024-01965-y

RESEARCH

Comparison of log-ratio and $\log_{10}$ chemical elemental data analysis of Central Amazonian pottery and archaeological implications

Roberto Hazenfratz[1] · Guilherme Z. Mongeló[2,3] · Casimiro S. Munita[1] · Eduardo G. Neves[4]

Received: 8 June 2023 / Accepted: 9 March 2024 / Published online: 10 April 2024
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

The additive log-ratio (alr) transformation is recommended as one of the most robust data transformations for multivariate analysis of archaeometric compositional data. However, alr and other transformations are not mutually exclusive and can be combined to assess different aspects of an archaeometric data set, such as the addition of temper, post-depositional effects in pottery and associated archaeological implications. This study presents a comparative analysis of a multi-element data set of pottery from Lago Grande and Osvaldo archaeological sites in the Central Amazon, which are considered a microcosm of the region. The concentrations of nine chemical elements (La, Lu, Yb, Ce, Cr, Eu, Fe, Sc, and Th) measured by instrumental neutron activation analysis (INAA) were subjected to alr transformation, prior to chemical fingerprinting by cluster analysis (CA) and principal component analysis (PCA). The results were compared to a previous work using the $\log_{10}$ transformation. Multivariate analysis of variance (MANOVA) was employed to test for statistical differences between the chemical groups, and self-organizing maps (SOMs), a type of artificial neural network, were used for comparison due to their advantage of not depending on any specific data distribution assumption. In general, the results suggest the existence of socio-cultural interactions between Lago Grande and Osvaldo, which could have occurred through trade, exogamic marriage and territory sharing. In a broader perspective, the exchange networks corroborated by the results favor theories that minimize the role of ecological constraints in the emergence of social complexity and sedentary occupations in the Amazon region.

Keywords INAA · Compositional data · Log-ratio analysis · Self-organizing maps · Amazonian pottery · Central Amazon archaeology

Introduction

In archaeometric studies of pottery, it is commonly recommended to rescale multi-elemental concentration data prior to multivariate analysis to account for variations in concentration magnitudes among macro, micro and trace elements. It contributes to homogenizing the weights of the variables (Harbottle 1982) and improving the condition of normality, often assumed in many multivariate statistical techniques. In many archaeometric studies, the elemental concentrations are frequently logarithmically transformed based on the assumption that trace elements follow a lognormal distribution in many geochemical contexts (Glascock 1992; Beier and Mommsen 1994; Koch and Link 2002). However, formal tests should be conducted to confirm the data distribution before drawing conclusions.

The elemental concentrations measured in chemical characterization experiments of archaeological pottery are

✉ Roberto Hazenfratz
robertohm@alumni.usp.br

[1] Research Reactor Center, Nuclear and Energy Research Institute, Av. Prof. Lineu Prestes 2242, Cidade Universitária, São Paulo, SP 05508-000, Brasil

[2] Centro de Estudios Arqueológicos y Antropológicos (CEAA), Escuela Superior Politécnica del Litoral (ESPOL), Vía Perimetral 5, Guayaquil, Ecuador

[3] Escuela Superior Politécnica del Litoral, Facultad de Ciencias Sociales y Humanisticas, ESPOL Polytechnic University, P.O. Box 09015863, Campus Gustavo Galindo Km 30.5 Vía Perimetral, Guayaquil, Ecuador

[4] Museum of Archaeology and Ethnology, University of São Paulo, Av. Prof. Almeida Prado 1466, Cidade Universitária, São Paulo, SP 05508-900, Brasil

typically expressed in units of ng.g^{-1}, µg.g^{-1}, and mg.g^{-1}. Due to its nature, there is a linear restriction underlying such data regarding the constant value that must be obtained by adding the mass fractions (the sum must be 1). Statistical methods usually developed to be applied to data without such a restriction could, in principle, generate spurious effects in the correlation structure, which would mislead geochemical and archaeological patterns of interest (Pawlowsky-Glahn et al. 2007).

A nxp matrix is termed compositional when the sum of values in any row is always a constant. Furthermore, it is termed subcompositional if the elements characterizing each of its columns are a subset of the compositional data. In this manner, all matrices of relative elemental concentrations obtained in chemical characterization experiments (with n samples and p variables) of pottery are subcompositional, since not all naturally occurring chemical elements are measured. In this sense, the utilization of logarithmic concentration ratios was advocated as one of the most adequate and robust options for data transformation prior to the application of multivariate statistical methods because it considers the compositional character inherent to the elemental concentrations used routinely in archaeometry (Baxter and Freestone 2006; Buxeda i Garrigós 2018).

From the statistical point of view, Aitchison (1982, 1986) recommended using elemental ratios to handle problems due to the potential existence of a covariance structure that is not directly interpretable when calculated without considering the underlying linear restriction of compositional data, besides other difficulties in parametric modelling. In the archaeometric context, Buxeda i Garrigós (1999) reported that the additive log-ratio transformation (alr), a type of log-ratio analysis (LRA), is the most appropriate data transformation method. It deals with alterations and contaminations that can directly affect the concentration of a unique chemical element, and unknown indirect perturbations in the remaining elemental concentrations due to the compositional nature of such data. A review of the main statistical developments in the treatment of compositional data and exploratory tools may be found in Egozcue & Pawlowsky-Glahn (2019) and Greenacre (2021). Besides the importance of this issue, many applications in geochemistry and archaeometry have not considered the compositional nature of geochemical-related data (Buccianti and Grunsky 2014; Baxter 2016), although many works attempting compositional coherence in the archaeometric literature may also be found (e.g. Tsantini et al. 2016; Fantuzzi and Cau 2017; Hein and Kilikoglou 2017; Montana et al. 2018; Zhou et al. 2018; Cau Ontiveros et al. 2019; Greenacre 2019; Leroy et al. 2020; Luneau et al. 2020; Wood and Greenacre 2021).

Furthermore, using elemental ratios is advisable to prevent analytic errors between different batches of samples due to systematic instrumental variations among them (Baxter et al. 2005). The division between two variables that present the same systematic variation would cancel such an effect, avoiding possible misclassifications of samples in chemical compositional groups defined previously by multivariate analysis. Another advantage of using chemical elemental ratios in the context of archaeological pottery is related to the heterogeneous addition of temper during the ceramic production (Buxeda i Garrigós 1999), the possibility of diagenesis affecting the buried ceramics in the depositional context or even when there are considerable natural geochemical variations in the clay sources.

Neglecting the compositional nature of a data set can be the most critical when analyzing archaeological glasses, for which it is common to determine the concentrations of major elements. For pottery, the combined analyzes of major and minor elements can be important to assess the chemical fingerprinting and uses of raw materials in a more complete perspective (Munita et al. 2001; Civici 2007; Tite 2008). However, in many ceramic studies the analysis can be focused solely on trace elements (Martini et al. 2004; Speakman and Glascock 2007; Munita et al. 2019). In this case, LRA might be equivalent to the traditional analysis based on logarithms, making the compositional nature of data less critical (Baxter and Freestone 2006). Hence, it can be a good practice to test different data transformations in archaeometric multivariate statistical analysis to address these issues.

Although transforming data via LRA could be considered the most appropriate approach for compositional/subcompositional data, it has not been used widely in archaeometric studies. One of the reasons would be the generation of unsatisfactory practical results. It was reported (Baxter and Freestone 2006; Baxter 2016) that LRA might work inadequately, misleading the identification of archaeologically interpretable structures or identifying patterns less directly than other frequently used data transformation methods. As Baxter et al. (2005) pointed out, this behavior may be partially due to the overestimated importance of elements with high relative variations due to their absolute presence in relatively low concentrations, which may have little interpretative utility in archaeology, despite any geochemical importance. Moreover, absolute differences in elemental concentrations may still be archaeologically meaningful and obscured in LRA.

Since LRA can map data from a simplex space (a natural space for compositional data) to log-ratios in a real space, several methodologies based on such transformations, which remove the linear restriction of data, have been developed (Pawlowsky-Glahn et al. 2007). Those methodologies allow one to use standard multivariate statistical techniques to the transformed data without further restrictions, where it is possible to translate the inferences in terms of the original compositions.

The ceramic material from the Central Amazon have temper in many cases (Lima 2008; Keding et al. 2010). The well-known dilution effect in the elemental concentrations has the potential to hide other useful archaeological information, as the provenance of pottery (Neff et al. 1989). In this case, the alr transformation can handle it properly, as the addition of non-plastic materials may be mathematically represented by a perturbation vector (Buxeda i Garrigós 1999). However, given the still incipient knowledge of the geochemical variability of clay sources in the Central Amazon, tests are needed to assess possible differences between different data transformations for pottery shards from this region. In Central Amazon some chemical elements can be highly susceptible to diagenetic effects, compared to other "well-behaved" depositional contexts. The region is prone to seasonal rainfall regime, and its hydrological dynamics produces high variations in the level of the Negro and Solimões rivers, with the flooding of areas of interest that brings sediments from Andean and sub-Andean regions (e.g. Martinelli et al. 1993). These effects combined can make the geochemical patterns of Central Amazonian pottery fuzzier.

Within this context, this study comprises the analysis of a data set of chemical elemental concentrations of pottery shards from two large and representative archaeological sites in Central Amazon, Brazil, which was already analyzed with $\log_{10}$ data transformation (Hazenfratz et al. 2016, 2017). Our main objectives can be summarized as follows:

1. Compare the $\log_{10}$ transformation to the LRA regarding the identification of archaeological patterns, evaluating whether diagenesis, tempering with non-plastic materials and other culturally driven modifications in pottery shards are dominating clustering.
2. After compensating for tempering, evaluate whether the use of different raw materials or recipes can be discerned for the purpose of determining provenance.
3. Establish more robust chemical groups for interpretation by complementing CA + PCA with other methods which do not rely on any prior assumption about data distribution.
4. Assess whether different modes of exchange networks can be distinguished: trade, exogamic marriages and/or territorial integration within Central Amazon.

The pottery pieces were obtained during excavations under the Central Amazon Project (PAC, acronym in Portuguese), coordinated by Neves (Neves 2018), from the University of São Paulo, in collaboration with the Federal University of Amazon.

This paper is organized as follows. Section 2 summarizes the main rationale regarding questions of Amazonian archaeology implied in this work. The methodological aspects are presented in Section 3. Section 4 comprises the main results of the multivariate statistical analysis for LRA-transformed data and comparison with the $\log_{10}$ transformation. Section 5 discusses the results comprising the comparison of data transformations and some archaeological implications. Finally, a summary of the main points and a conclusion are given in Section 6.

Archaeological framework

Lago Grande and Osvaldo archaeological sites

The Central Amazon region, an area of 900 km², is primarily located at the confluence of the Solimões and Negro Rivers, which archaeologists have systematically studied since the 1980s. This area is critical for understanding the archaeological past of the Amazon since, according to the initial publications of Lathrap (1970), it would have been an important center of South American cultural and technological development. To test such hypotheses, Eduardo Neves, Michael Heckenberger and James Petersen started in 1995 the Central Amazon Project, an archaeological investigation program that operated for 20 years in the area and whose archaeological data presented here are part of this effort.

The archaeological sites of Central Amazonia have some notable characteristics: they are composed of thick layers of Terra Preta, dark-colored anthropic soil resulting from the intense interaction between sedentary communities and the environment. The dimensions of the sites in the region are also notable, many exceeding 10 hectares. Finally, most archaeological sites are multicomponent, i.e., composed of successive and different cultural occupations (Neves 2010, 2012).

These cultural occupations are defined by archaeologists mainly from ceramic material's morphological and technological analysis, a vestige of the most abundant material culture in most Amazonian sites. It is assumed that different cultural groups produce, in turn, elements of material culture with specific characteristics, and archaeologists call these ceramic groups *Cultural Phases*.

In the Central Amazon region, four Cultural Phases predominate (Hilbert 1968; Neves 2012): Açutuba Phase (300 BC – 360 AD), Manacapuru Phase (550–650 AD), Paredão Phase (700–1200 AD), and Guarita phase (950–1450 AD). In large sites of Central Amazonia, these phases appear sequentially in the stratigraphy, with rare evidence of their overlapping in the same site. An exception is precisely the relationship between the Osvaldo and Lago Grande sites (Fig. 1).

Both sites are on plateaus on the banks adjacent to large lakes (Lago do Limão and Lago Grande), lacustrine formations characteristic of the Iranduba peninsula, which are 10 km apart in a straight line. Channels and streams interconnect these water sources during the rainy season.

The Lago Grande site is a large-scale habitation site identified in 1999. It was the target of three different excavation campaigns, whose main objectives were the delimitation and definition of the cultural layers of occupation of the site. In addition to the thick layer of Terra Preta, many pottery fragments dispersed over the surface characterize the area. The archaeological fieldwork identified eight earthwork structures, like mounds type, arranged in a circular format around a central square demarcated by a defensive ditch. The excavations corroborated the initial hypothesis that the artificial mounds were built to establish housing units (Donatti 2003; Mongeló 2011).

The Lago Grande site is one of the rare cases in the region to have been occupied by only one archaeological culture, during the 9th and 10th centuries, the Paredão culture. This cultural phase is marked mainly by its distinctive ceramics, which have considerable technological refinement. The vessels have a vitreous appearance after being burned, and the walls are generally thin. Plastic or painted decoration is not very abundant, drawings with fine incisions in red color predominate, and urns are commonly decorated with simple anthropo-zoomorphic appliqués.

The Osvaldo site lies on the shores of Lago do Limão, a body of water with 15 other sites identified (Moraes 2006), and is considered the largest (30 hectares) and the oldest there. It is located on a plateau adjacent to the lake and is marked by a thick layer of Terra Preta. However, no housing mounds were identified, and the predominant ceramic material belongs to the Manacapuru phase.

Archaeological work at the Osvaldo site has demonstrated, through 19 dated samples (Lima 2008), that it was uninterruptedly occupied between 630 AD and 990 AD, predominantly by groups that produced Manacapuru-type ceramics. This ceramic material (Hilbert 1968) is characterized by the use of cauixí (a freshwater sponge which is the main temper in the Incised Rim Tradition) as a non-plastic

Fig. 1 Location of Lago Grande (3° 14′ 21″ S, 60° 11′ 43″ W) and Osvaldo (3° 11’ 15’’ S, 60° 20’ 41’’ W) archaeological sites. Sources: Portocarrero (2006) and Lima (2008)

additive and a wide variety of decorations, with incisions, dots, red englobe and zoomorphic appendages (Lima 2008).

Although no mounds were identified, the excavations showed that the archaeological materials were circularly arranged in well-marked concentrations at the site (Portocarrero 2006). It represents strong evidence that the concentrations of pottery shards are indicators of housing units and that the indigenous people who occupied this site did so in a circular village.

The issue under consideration, which has been discussed for some years (Mongeló 2011; Neves 2012; Hazenfratz et al. 2012), starts from the commonalities that share the archaeological characteristics of the two sites within the framework of a notable difference: the cultural affiliation of the ceramic material. The ceramics of the Manacapuru and Paredão phases are technologically very different and are found in the region at different ages, with the Manacapuru Phase being older than the Paredão.

However, when we compare the two sites, we observe that they share two centuries of human occupation (9th to 11th centuries), and both are villages with housing units arranged in a circular format, with a well-signposted central square. Mongeló (2011) studied the exogenous archaeological material in each of the sites, i.e., ceramic fragments of the Paredão phase from the Osvaldo site and fragments of the Manacapuru phase from the Lago Grande site. Furthermore, Hazenfratz et al. (2012, 2016) sought to find archaeometric evidence of exchange networks between these two populations.

This article deals precisely with the analysis some of these fragments of non-preponderant cultural affiliation in the two sites to establish increasingly concrete points relating to the two populations that occupied the region.

Previous results from $\log_{10}$ analysis

Previous work by Hazenfratz et al. (2016) comprised the use of instrumental neutron activation analysis to determine the concentration of 24 chemical elements in pottery shards from Lago Grande and Osvaldo archaeological sites. The multidimensional data set was analyzed by a hierarchical clustering method, Average Linkage, compared to Ward and K-means methods in Hazenfratz-Marks (2014); and principal component analysis for refining chemical groups of pottery. Fifteen elements were selected according to the analytical quality parameters: As, K, La, Lu, Na, Yb, Ce, Co, Cr, Cs, Eu, Fe, Hf, Sc and Th. After analysis of potential interferences in the analytical procedure, quantity of missing values and diagenetic effects for the post-depositional context in the Central Amazon, the final set of variables chosen for analysis was comprised by the nine chemical elements La, Lu, Yb, Ce, Cr, Eu, Fe, Sc and Th. Two chemical groups of pottery shards were defined for each site. Their pairwise superposition in the three first principal components indicated geochemical similarities which could be archaeologically meaningful. The results were correlated to complex and multimodal exchange networks potentially present in the Central Amazon, driven by three mechanisms: trade, exogamic marriage and territorial integration in the region, documented also in the archaeological literature (e.g. Moraes 2006; Neves and Petersen 2006; Lima 2008; Mongeló 2011).

Materials and methods

Instrumental neutron activation analysis (INAA)

The sample of pottery shards (Fig. 2) for this research comprised 187 fragments selected from excavation units with a high density of archaeological vestiges from Lago Grande and Osvaldo archaeological sites in Central Amazon (Brazil), including all the stratigraphic levels identified on both sites. Additionally, 13 clay samples from the surrounding region were included in the analysis. The INAA experimental methodology for ceramic sampling appeared in previous manuscripts (Portocarrero 2006; Lima 2008; Mongeló 2011; Hazenfratz et al. 2016). In this paper, we focus on the comparison of two different data transformations and their archaeological implications.

Statistical techniques

The data set comprising 24 elemental concentrations of Lago Grande and Osvaldo ceramic material was measured by INAA and it was first examined for its analytical data quality and identification of outliers according to the procedures described elsewhere (Hazenfratz-Marks 2014; Hazenfratz et al. 2016). The criteria for selecting the optimal set of chemical elements were the acceptance in the quality control analysis, a minimum value for missing values (10%), and the inexistence of interferences due to the INAA experimental procedure. The fourteen chemical elements selected for chemical fingerprinting were then As, K, La, Lu, Na, Yb, Ce, Cr, Cs, Eu, Fe, Hf, Sc, and Th.

This work focuses on a scenario extracted from Hazenfratz-Marks (2014) in which nine chemical elements were used (La, Lu, Yb, Ce, Cr, Eu, Fe, Sc, and Th) with alr data transformation for chemical fingerprinting after the identification of possible diagenetic effects in the concentrations of Na, K, As, Cs and Hf. Such elements were extracted due to the spurious effects identified statistically (by preliminary PCA and calculations of

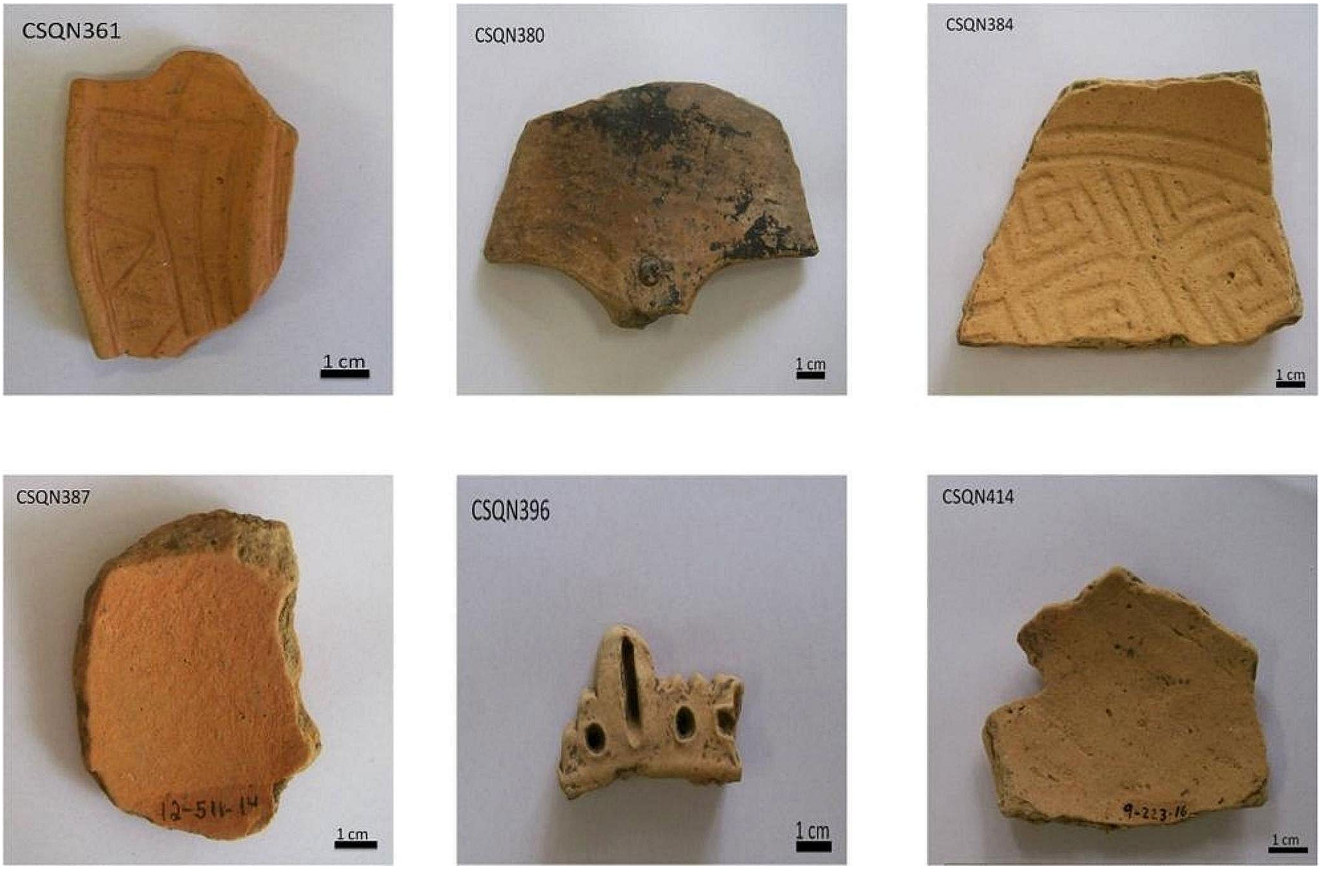


Fig. 2 Examples of pottery shards from Lago Grande (CSQN361, 380, 384, 387) and Osvaldo (CSQN396, 414) analyzed in this work

the variation matrix) and based on evidence in the literature (Gaudette et al. 1966; Buxeda i Garrigós et al. 2001, 2002; Bishop 2003; Mommsen 2004; Schwedt et al. 2004; Schwedt and Mommsen 2007). The maintenance of those five elements in the analysis would obscure cultural patterns of interest due to their concentration anomalies and high variance. This scenario is compared to the one where the same chemical elements were used but with the $\log_{10}$ transformation of data (Hazenfratz et al. 2016).

For the compositional data transformation, it was employed the alr transformation of asymmetric nature (Buxeda i Garrigós 1999)

$$\boldsymbol{x} \in S^d \rightarrow \boldsymbol{y} = log\left(\frac{\boldsymbol{x_{-D}}}{x_D}\right) \in R^d \qquad (1)$$

where S^d is the d-dimensional simplex space ($d = D - 1$) and x_D is the element chosen as the denominator. The alr transformation would emphasize the presence of disturbances in the elemental concentrations, making them more easily identifiable by the variances in the variation matrix. Furthermore, it is possible to choose the adequate element, from the statistical point of view, to be used as the denominator in the alr transformation by calculating the same matrix. Although some authors do not recommend using the Euclidian metric in cluster analysis with alr data transformation due to possible distortions on preserving distances (Pawlowsky-Glahn et al. 2007), it is recommended by other authors for its suitability for archaeological ceramics and glasses compositional data (Buxeda i Garrigós 1999; Greenacre 2021).

The selection of the chemical element to be used as the denominator in the alr transformation was made by calculating the variation matrix of the subcompositional data, choosing the element which imposes the slightest variations in the original covariance matrix.

After data transformation, the first multivariate pattern recognition technique employed was the method of hierarchical clustering by Average Linkage with Squared-Mean Euclidean Distance (SMED), as also implemented in the MURRAP software (Glascock 1992). This metric allows the possibility of using samples whose data are incomplete. As a cut criterion in the dendrogram, it was established that the minimum quantity of chemical groups should at least differentiate between the clay samples and pottery shards (see chemical profiles in Hazenfratz-Marks 2014). Cuts at lower linkage distances should only be considered if any archaeological correlation justified the division into more chemical groups.

After the cluster analysis (CA) and a preview of the results from principal component analysis (PCA), the membership to each group was then refined by a procedure based on the Mahalanobis distance to calculate the membership probability of each sample to the defined chemical groups, considering the correlations among variables (Glascock 1992).

The statistical differences between the chemical groups were tested by a multivariate analysis of variance (MANOVA) based on Hotelling´s T^2 test, a multivariate generalization of the Student´s t-statistic test (Meyers et al., 2006). This test is appropriate when the dependent variables (in this case, the elemental concentrations) have any degree of intercorrelation. It tests the null hypothesis in the following general form (Warner 2013)

$$H_0 = \begin{bmatrix} \mu_{11} \\ \mu_{21} \\ . \\ . \\ . \\ \mu_{p1} \end{bmatrix} = \begin{bmatrix} \mu_{12} \\ \mu_{22} \\ . \\ . \\ . \\ \mu_{p2} \end{bmatrix} = \cdots = \begin{bmatrix} \mu_{1k} \\ \mu_{2k} \\ . \\ . \\ . \\ \mu_{pk} \end{bmatrix} \tag{2}$$

where μ represents the mean of a dependent variable (elemental concentration), p is the number of dependent variables and k is the number of groups being tested, which will depend on the clustering and principal component analysis of pottery. This test assumes multivariate normality, homoscedasticity and is sensitive to outliers.

Finally, an artificial neural network method was used for comparison with the multivariate patterns identified by CA and PCA of the alr-transformed data. The main applications of artificial neural networks (ANN) in archaeometric studies regards the comparison or even substitution of geochemical data classification by algorithms which rely on hypotheses about data normality and homoscedasticity (Pagnotta et al. 2018; López and Ramil 2018; Barone et al. 2019). One main advantage is that ANNs in general do not depend on any specific data distribution assumption (Bell and Croson 1998). They also may be a powerful tool for multivariate modeling of systems that do not present linear correlation between variables, as well as to visualize high-dimensional data. One special kind of neural network of interest in archaeometric studies is the Self-Organizing Map (SOM).

The SOMs were developed by Teuvo Kohonen (Kohonen 2013) and preserve the topological features of the original multivariate space. They are frequently applied with a non-supervised learning strategy that does not require a predefined set of samples from a known structure. The maps consist of nodes (neurons), each possessing a weight vector with the same dimension as the data set. The data inputs representing different samples are assigned to the neuron unit with the most similar weight vector (BMU – best matching unit) based on a chosen metric. The BMU and its neighbors have their weights adjusted according to the input vector using the following updating equation (Lopez-Molinero et al. 2000)

$$\boldsymbol{W}_v(t+1) = \boldsymbol{W}_v(t) + \theta(v,t,d)\,\alpha(t)\left[\boldsymbol{D}(i) - \boldsymbol{W}_v(t)\right] \tag{3}$$

where t is the epoch (step) of the algorithm, $\boldsymbol{D}(i)$ is the input vector, $\theta(v,t,d)$ is the neighborhood function, d is the distance from the neighboring neuron to the BMU, and $\alpha(t)$ is the monotonically decreasing learning coefficient.

The neuronal geometry can be hexagonal or rectangular, influencing the number of neighbors. A training epoch finishes when all input vectors have been presented once to the neural network to update the weight vectors. At the end, the output nodes (neurons) may be associated to the groups present in the original multidimensional data space. In the batch version of the SOM algorithm, all input vectors are presented simultaneously to the network.

Hazenfratz et al. (2017) previously applied this technique on the $\log_{10}$-transformed data set. Kohonen and CP-ANN packages (implemented in Matlab 7.6.0) were used to compute the SOMs (Ballabio et al. 2009; Ballabio and Vasighi 2012). The non-supervised algorithm was adopted to prevent interference from previous classifications of samples by CA and PCA during the learning procedure. The parameters selected for the SOMs are presented in Table 1 (Lopez-Molinero et al. 2000; Tanevska et al. 2007; Toyota et al. 2009; Ballabio and Vasighi 2012). Martín-Fernández et al. (2019) also described a methodological framework to perform SOM analysis for compositional data.

Table 1 Setup of the self-organizing maps (SOMs) for chemical elemental data of pottery shards from Lago Grande and Osvaldo archaeological sites

Parameter	Type / Value
Model	Non-supervised Kohonen map
Network geometry	Hexagonal
Number of neurons	7×7
Number of epochs	500
Neighborhood function	Gaussian
Training algorithm	Batch
Initialization of neuronal weights	Linear (from data eigenvectors)
Data transformation	alr
Initial learning rate	0.5
Final learning rate	0.01

Table 2 Variation matrix for chemical elemental concentrations in pottery shards from Lago Grande and Osvaldo archaeological sites

	Sc	Cr	Fe	La	Ce	Eu	Yb	Lu	Th
Na	0.623	0.694	0.622	0.705	0.694	0.600	0.699	0.690	0.806
K	0.200	0.234	0.218	0.279	0.290	0.291	0.267	0.278	0.283
Sc	0	0.011	0.027	0.052	0.051	0.058	0.050	0.058	0.034
Cr	0.011	0	0.027	0.048	0.053	0.067	0.041	0.048	0.023
Fe	0.027	0.027	0	0.082	0.087	0.086	0.068	0.074	0.057
As	0.179	0.164	0.182	0.210	0.237	0.264	0.193	0.195	0.179
Cs	0.219	0.210	0.239	0.249	0.254	0.283	0.245	0.256	0.215
La	0.052	0.048	0.082	0	0.042	0.035	0.024	0.031	0.038
Ce	0.051	0.053	0.087	0.042	0	0.045	0.067	0.076	0.057
Eu	0.058	0.067	0.086	0.035	0.045	0	0.056	0.065	0.073
Yb	0.050	0.041	0.068	0.024	0.067	0.056	0	0.015	0.041
Lu	0.058	0.048	0.074	0.031	0.076	0.065	0.015	0	0.043
Hf	0.180	0.138	0.195	0.132	0.174	0.199	0.110	0.099	0.093
Th	0.034	0.023	0.057	0.038	0.057	0.073	0.041	0.043	0
$\mathrm{tr}\,(\Sigma)$	1.743	1.758	1.963	1.926	2.128	2.122	1.876	1.928	1.943
$\mathrm{Var}/\mathrm{tr}\,(\Sigma)$	0.859	0.852	0.763	0.778	0.704	0.706	0.799	0.777	0.771
Var	1.498								

Results

Alr data transformation

Initially, the variation matrix was computed for the nine chemical elements selected to determine the element with the lowest variability to be used as the denominator for the alr transformation. This matrix was calculated using MURRAP 8.6 (with routines for statistical analysis in archaeometry) and the R software with missing values substitution by minimizing the Mahalanobis distance. Essentially the same results were obtained (Table 2).

Each column in Table 2 represents the diagonal of the data covariance matrix when an element (in this case, the element naming the column) is selected as the denominator in the alr transformation (Eq. 1). The $\mathrm{tr}\,(\Sigma)$ term refers to the trace of the covariance matrix when each element is used; the term $\mathrm{Var}/\mathrm{tr}\,(\Sigma)$ refers to the fraction of the trace of each covariance matrix explained by a measure of total variation in the covariance structure of the original compositional data (Var). Note that $(1 - \mathrm{Var}/\mathrm{tr}\,(\Sigma))$ represents a variability measure imposed on the original covariance matrix due to the selection of each element in the alr transformation. In this regard, higher values of $\mathrm{Var}/\mathrm{tr}\,(\Sigma)$ are obtained for the most "stable" elements, which impose the slightest variations in the original covariance structure when selected as the denominator for alr data transformation.

Therefore, by analyzing Table 2, cerium is the chemical element that imposes the highest variability when chosen as a reference in the alr transformation. Scandium, the selected denominator of the alr transformation for the data, is the element that imposes the lowest variability.

Table 3 Descriptive parameters of Average Linkage clustering of pottery shards from Lago Grande (LG) and Osvaldo (O) for $\log_{10}$ and alr transformations of chemical elemental data

Parameter	$\log_{10}$	alr
LG groups	2	2
O groups	2	2
LG groups size (number of samples)	68 – 36	54 – 50
O groups size (number of samples)	38 – 17	25 – 23
Total number of samples	159	152

Clustering

Table 3 shows the distribution of samples in each compositional group comprising the minimum cut in the dendrogram from Average Linkage clustering after removing obvious outliers (Hazenfratz et al. 2016). The initial sample size was 187. Besides analytical issues, the outliers identified could be related to diagenetic effects in highly perturbed depositional contexts as the ones found in the Central Amazon, or even chemical groups misrepresented in our data set.

Clustering refinement and PCA

After the clustering analysis, the membership of each pottery shard to the chemical groups was refined by a procedure based on the Mahalanobis distance and principal component analysis, according to the procedure described in Glascock (1992). However, this method needed more iterations for alr (3) than for $\log_{10}$ (1) transformation, indicating that the Average Linkage clustering procedure generated a better initial approximation of data partition for the latter. Similar results were observed with the Ward clustering method.

Figs. 3 and 4 show the comparative plots for the three first principal components after the refinement of classification of samples into chemical groups.

They show that the general patterns for identifying two chemical groups and their pairwise superposition for each site (A-A and B-B) were not altered by changes in the data transformation. Furthermore, both analyzes show a reasonable group separation, with some degree of superposition (A-B for each site), which is expected in chemical analyzes of pottery from nearby archaeological sites located in geochemically similar environments (e.g. Bishop 2003; Baria, 2015). More specifically, the alr transformation generated chemical groups that are more overlapping than the groups of the $\log_{10}$ analysis (comparison among groups A and B), with a relative worsening in their chemical resolution. Furthermore, the chemical group A of Osvaldo is contained within group A of Lago Grande in the alr scenario, indicating lower geochemical variability for this group in Osvaldo.

Table 4 shows the characteristics of each group in terms of the number of samples and similarities in group composition. The pottery shards from Lago Grande were the ones that showed the most remarkable similarities in group membership between the $\log_{10}$ and alr transformations. The similarity coefficients are the percentage of the samples with the same classification into chemical groups for both transformations, with respect to the total number of samples in the $\log_{10}$ transformation.

Table 4 Characterization of the chemical groups of reference established for Lago Grande and Osvaldo archaeological sites

Site	A-log	A-alr	Similarity (%)	B-log	B-alr	Similarity (%)
LG	61	57	72	28	33	61
OSV	37	20	46	15	26	60

MANOVA

A MANOVA was conducted to test for statistical differences between the chemical groups for both transformations, represented by their mean elemental concentration vectors. The results are organized in Table 5, where the last column informs the conclusion concerning the rejection of the null hypothesis of equality between the mean vectors. Due to the fact that most of the elements are at trace level, the variables considered in the test, which are actually proportions (compositional data), were approximated to variables with absolute values (Woronow & Love 1990). Besides the direct test here, indirect evidence for the statistical significance of the chemical groups is also discussed in Sect. 5.

Table 5 shows statistically significant differences between groups A and B for each archaeological site for both data transformations. When analogous groups defined in different data transformation analyzes are compared, the null hypothesis is rejected only for the groups B of Osvaldo (alr vs. $\log_{10}$). Hence, it is not possible to consider groups A and B from Lago Grande and groups A from Osvaldo as different chemical groups when the data transformation is changed from $\log_{10}$ to alr for the test conditions.

Self-organizing maps

Shenton-Bowman and Shapiro-Wilk tests for multivariate normality and Box´s M test for the homoscedasticity were performed with in the chemical groups defined here (Hazenfratz-Marks 2014). Both underlying assumptions could be rejected in most cases. Since we are testing compositional data, the impact of such deviations in the results must be carefully analyzed prior to any conclusion.

In this regard, the SOM for the alr transformation is presented in Fig. 5 for both archaeological sites combined. The map converged to a configuration where the pottery shards of group g1 (A in PCA) of Lago Grande and group g3 (A in PCA) of Osvaldo presented a pairwise superposition and were allocated in a common region in the map (white hexagons). This region is distinct from the region where the pottery shards from groups g2 (B in PCA) of Lago Grande and g4 (B in PCA) from Osvaldo were allocated (blue hexagons), and it presents the same pairwise superposition observed for groups g1 and g3. The three neurons highlighted with ellipses represent a transition region with different mixing of samples. However, they represent the smallest portion of the map.

The general patterns identified in the SOM for pottery shards from Lago Grande and Osvaldo agree with the results from CA and PCA, which identified two chemical groups for each archaeological site, with pairwise superposition (A-A and B-B) and a region of mixture between different groups (A-B for both sites), as illustrated in Figs. 3 and 4.

Typological correlation in chemical groups

The results of the multivariate analyzes were explored to find possible archaeological correlations regarding provenance and typological analyzes carried out by archaeologists in Central Amazon. Figures 6 and 7 show the histograms for the Manacapuru and Paredão pottery shards distribution in terms of the number frequency of pottery shards in each chemical group. It is important to emphasize here that the differentiation between Manacapuru and Paredão ceramics is sometimes complex, due to homogenizing elements in ceramics from the Incised Rim Tradition, as the use of cauixí as the preferred temper and specific decorative motifs. In this sense, imprecisions in the classifications are expected (Portocarrero 2006).

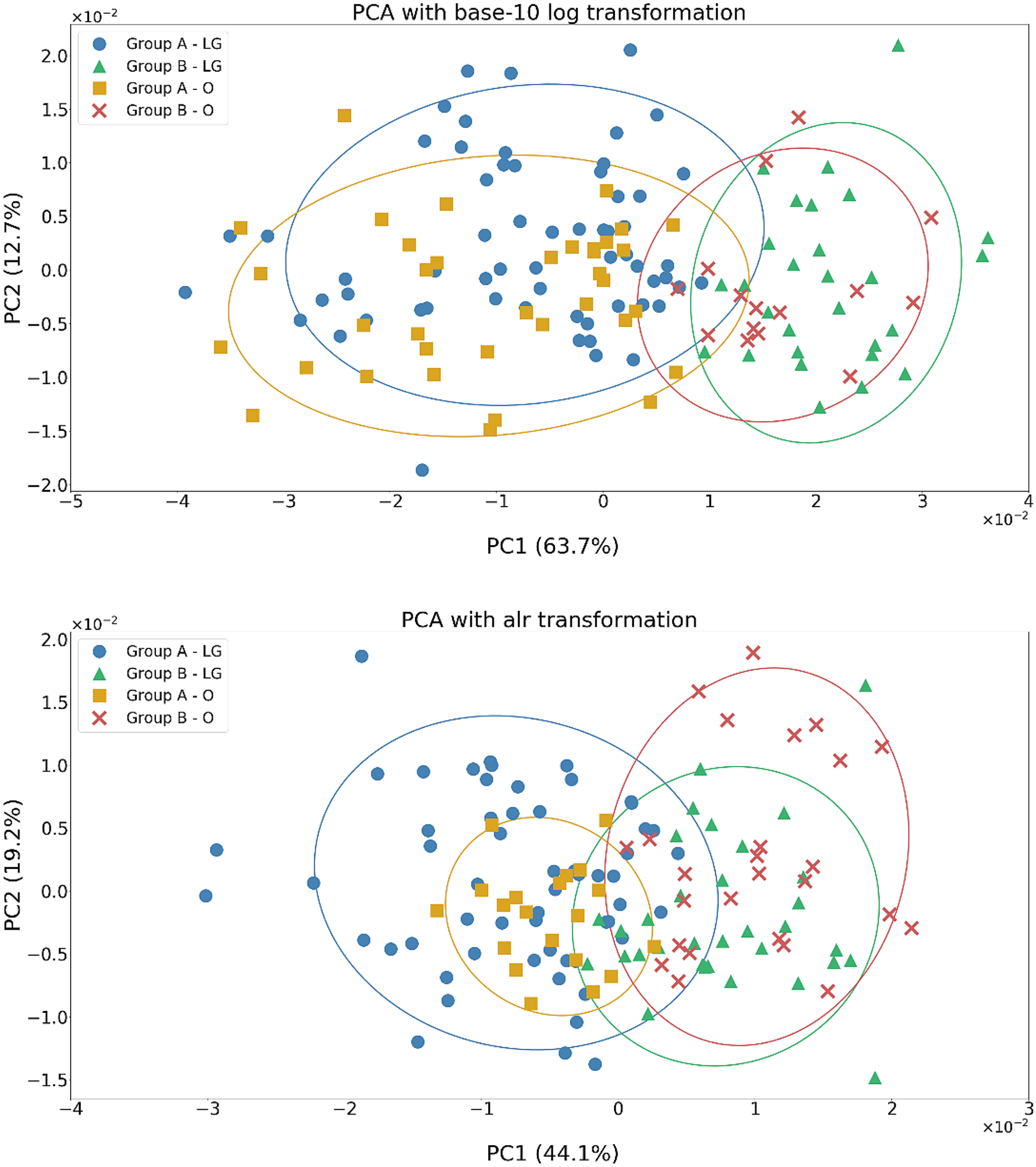
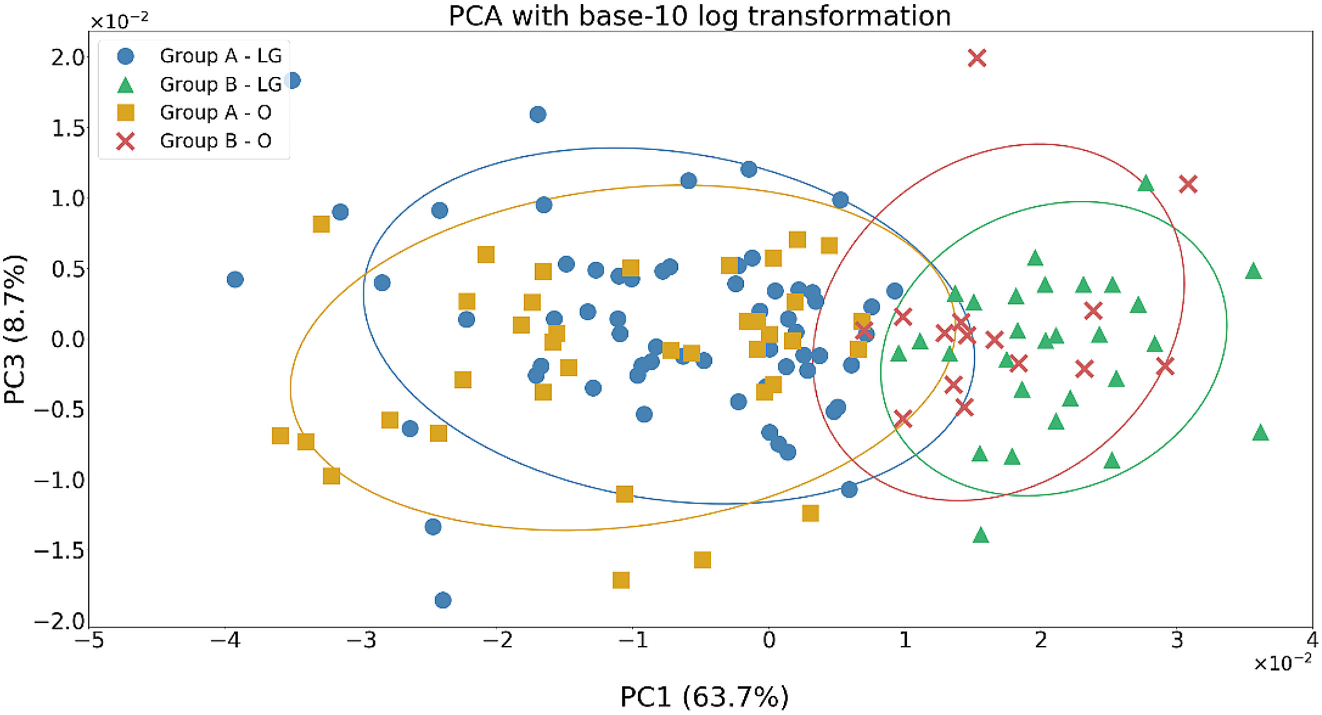


Fig. 3 Comparative principal component analysis (PC1 x PC2) for chemical elemental data of pottery from Lago Grande and Osvaldo archaeological sites. Top: $\log_{10}$ analysis. Bottom: alr analysis. The ellipses represent the region for the 95% confidence level

Figure 6 shows that Manacapuru ceramics concentrate in groups A of each site, regardless of the data transformation, and more pronouncedly in Lago Grande. This finding indicates that the ceramic recipes (regarding different combination of clays and temper) represented by such groups would be used most in Osvaldo, which is considered a Manacapuru site. Regarding Paredão ceramics (Fig. 7), similar patterns were not observed, with a higher



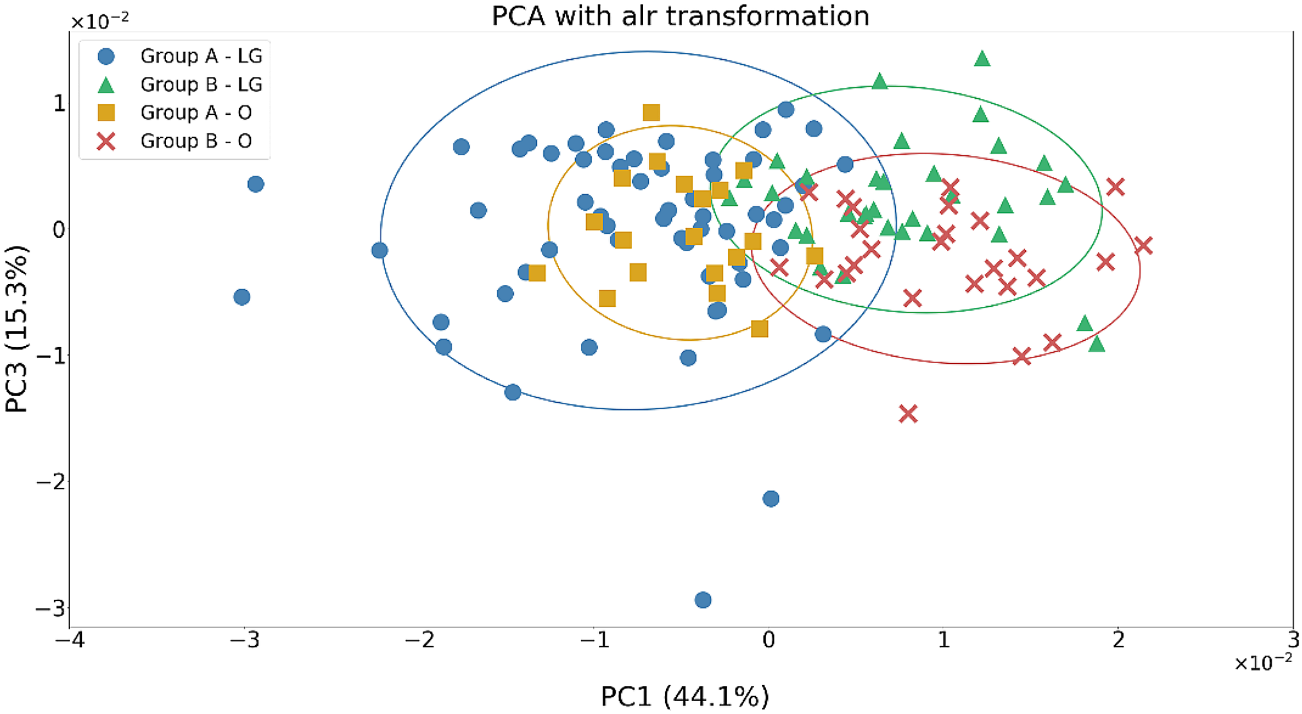


Fig. 4 Comparative principal component analysis (PC1 x PC3) for chemical elemental data of pottery from Lago Grande and Osvaldo archaeological sites. Top: $\log_{10}$ analysis. Bottom: alr analysis. The ellipses represent the region for the 95% confidence level

Table 5 Multivariate analysis of variance (MANOVA) for the chemical groups of pottery shards determined by hierarchical clustering and PCA-refined methods for the archaeological sites of Lago Grande (LG) and Osvaldo (O). Significance level: $\alpha = 0.05$

Test	Group	n_1, n_2	Site	Transformation	p	Conclusion
1	$A_{LG}^{log}, B_{LG}^{log}$	61, 28	LG	$\log_{10}$	3.9×10^{-19}	$A_{LG}^{log} \neq B_{LG}^{log}$
2	A_{O}^{log}, B_{O}^{log}	37, 15	O	$\log_{10}$	6.5×10^{-8}	$A_{O}^{log} \neq B_{O}^{log}$
3	$A_{LG}^{alr}, B_{LG}^{alr}$	57, 33	LG	alr	1.1×10^{-13}	$A_{LG}^{alr} \neq B_{LG}^{alr}$
4	A_{O}^{alr}, B_{O}^{alr}	20, 26	O	alr	1.3×10^{-7}	$A_{O}^{alr} \neq B_{O}^{alr}$
5	$A_{LG}^{log}, A_{LG}^{alr}$	61, 57	LG	$\log_{10}$ vs. alr*	0.44	$A_{LG}^{log} = A_{LG}^{alr}$
6	$B_{LG}^{log}, B_{LG}^{alr}$	28, 33	LG	$\log_{10}$ vs. alr*	0.38	$B_{LG}^{log} = B_{LG}^{alr}$
7	A_{O}^{log}, A_{O}^{alr}	37, 20	O	$\log_{10}$ vs. alr*	0.08	$A_{O}^{log} = A_{O}^{alr}$
8	B_{O}^{log}, B_{O}^{alr}	37, 26	O	$\log_{10}$ vs. alr*	0.02	$B_{O}^{log} \neq B_{O}^{alr}$

*Results for $\log_{10}$ transformation were divided by scandium concentrations for comparison with alr groups

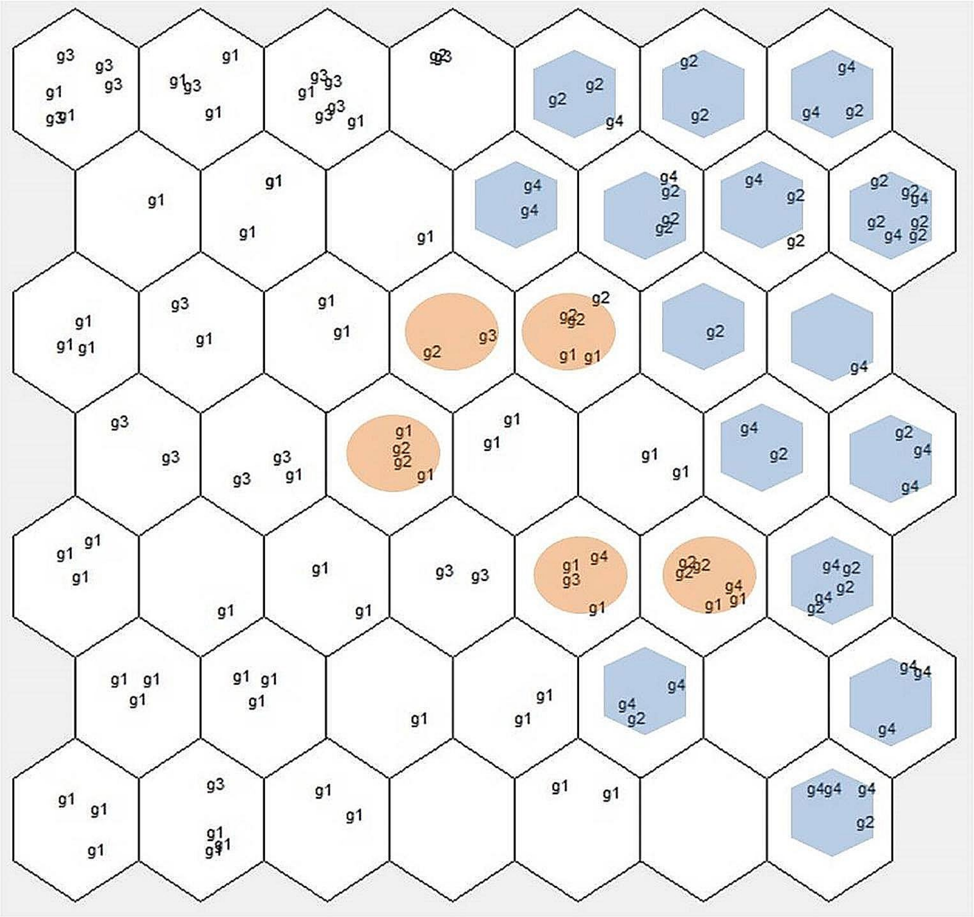


Fig. 5 Self-organizing map (SOM) for the chemical elemental data of Lago Grande and Osvaldo archaeological sites. Nomenclature: g1 = group A of Lago Grande; g2 = group B of Lago Grande; g3 = group A of Osvaldo; g4 = group B of Osvaldo

concentration in groups A with $\log_{10}$ transformation and groups B with the alr transformation. Thus, the effect of data transformation was more evident in the Paredão than in the Manacapuru samples.

Comparison with clay samples

After further confirmation of the chemical groups obtained by SOMs, we performed a comparison of the multi-elemental patterns between clays and pottery shards, Fig. 8 shows the projection onto the first two principal components of ten clay samples from Osvaldo (eight within the site and two near it) and three clay samples near Lago Grande.

In Fig. 8, two clay samples have PCA scores similar to groups B of Lago Grande and Osvaldo. They were extracted near Osvaldo, but it cannot be concluded univocally that the potters used such a clay source in the past. Moreover, we infer that a hypothetic clay mixture from Lago Grande

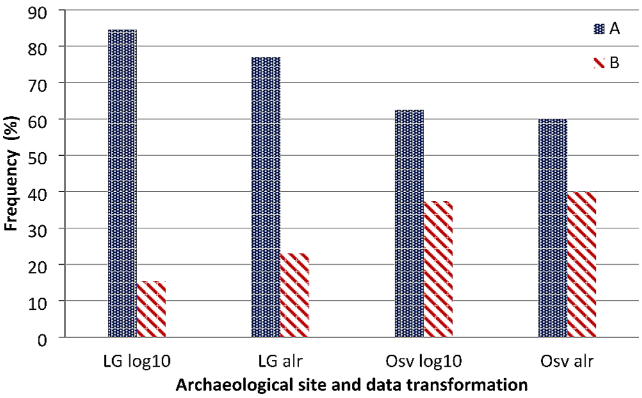


Fig. 6 Histogram for the number frequency of Manacapuru pottery shards in each chemical group for Lago Grande (LG) and Osvaldo (Osv) archaeological sites for different data transformations. The $\log_{10}$ results are further explored in Hazenfratz et al. (2016)

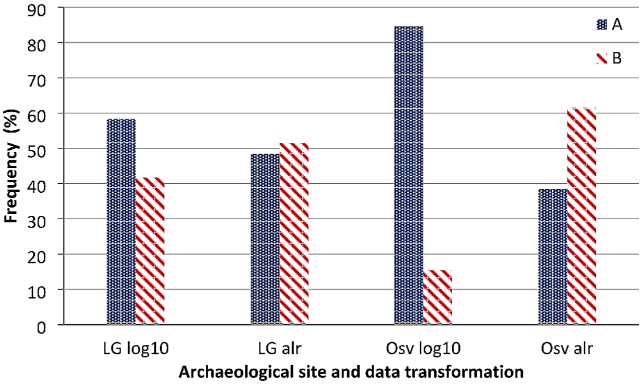


Fig. 7 Histogram for the number frequency of Paredão pottery shards in each chemical group for Lago Grande (LG) and Osvaldo (Osv) archaeological sites for different data transformations. The $\log_{10}$ results are further explored in Hazenfratz et al. (2016)

(to the left of the groups) and Osvaldo (to the right of the groups) could generate a hybrid ceramic paste with a similar composition to one of the pottery shards.

Discussion

Comparison of $\log_{10}$ and alr data transformations

The alr data transformation may help minimize the dilution effect caused by the addition of non-plastic materials, enabling the comparison of the chemical composition of pottery concerning the use of different clay sources. Here, the general patterns of the comparative PCA in Figs. 3 and 4 agree with the conjecture of Aitchison et al. (2002), which reported that the LRA of subcompositional data that consist basically of trace elements could be equivalent to the logarithmic transformation, ignoring dilution effects. Neither did the change in the data transformation generate more chemical groups with a clear archaeological correlation, nor did the two chemical groups identified in both analyzes (for each site) show a clear correlation with the origin, ceramic phase, typological classification, or stratigraphic level. It was not an expected result, given the considerably perturbed post-depositional context of Central Amazon, and considering the insertion of temper (cauixí) in the pottery shards. Furthermore, the insertion of crushed pottery, as identified by Portocarrero (2006), could add non-linear perturbations to the data.

Although the general multivariate structure from the statistical point of view is similar for both data transformations (Table 5), the classification into groups changes (see Table 4; Figs. 6 and 7). It indicates that temper addition to pottery influences some of the samples, justifying the comparison of data transformations performed in this work.

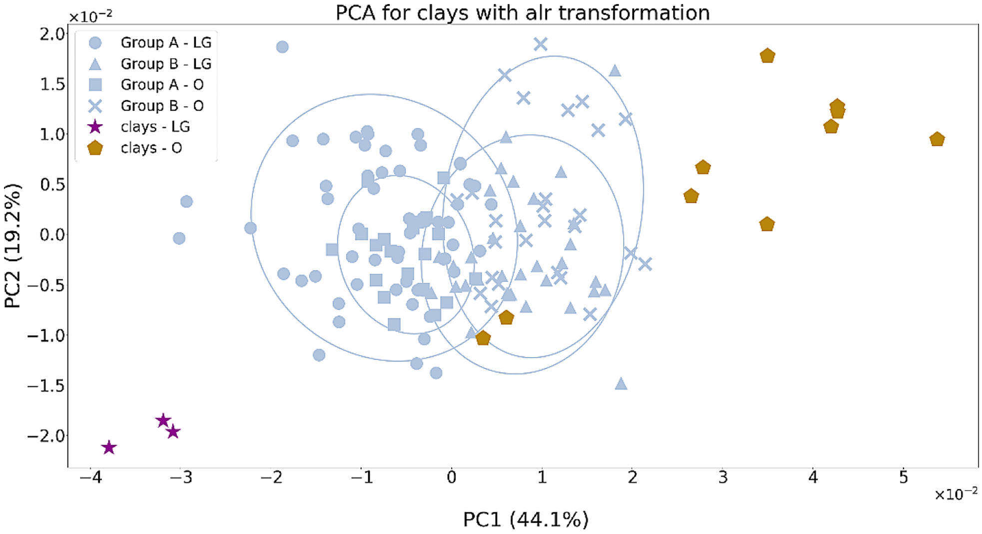


Fig. 8 Scores of the first and second principal components for the alr-transformed data of clay samples projected against the two chemical groups defined for Lago Grande (LG) and Osvaldo (O). The ellipses represent the region for the 95% of confidence

In principle, more variables than the nine chemical elements chosen here could be used to assess further differences between the data transformations. However, Baxter and Freestone (2006) warned about the need to remove variables that do not carry relevant information in the multivariate structure, and which have high variance, concluding that PCA based on LRA is more sensitive to the insertion of such variables than PCA based on standardized data. It was observed a similar trend for our samples, where chemical elements altered by diagenetic effects impacted more the alr than the $\log_{10}$ analysis. For example, in a previously analyzed scenario where the elements Hf, Cs, and As were included, it was observed that the dominant chemical elemental ratios Hf/Sc, Cs/Sc, and As/Sc were related to the elements with the highest variances, presenting coefficients of variation equal to 70%, 37%, and 44% respectively, for the combined data set of Lago Grande and Osvaldo ($n = 172$ samples). These elements influenced the results of PCA in the alr analysis, resulting in higher weighting coefficients and differences between the classification of the samples into chemical groups. In this sense, we may infer that chemical elements prone to diagenetic effects (Na, K, Hf, Cs and As identified in Hazenfratz et al. 2016) would influence the alr analysis more critically.

We can use the coefficients of variation as a metric for an indirect validation of the chemical groups defined in this work. Hazenfratz et al. (2016) used the coefficients for the $\log_{10}$ analysis, so here we focused on the alr analysis. Moreover, Bishop (2003) observed that the procedures to confirm the statistical significance of chemical groups of pottery could be carried out internally or by comparison among different groups. His observations of several archaeometric studies of pottery performed worldwide indicated that well-characterized chemical groups had univariate dispersion of elements between 10 and 15%, depending on the precision of measurements. Previously, Harbottle (1982) suggested an interval between 15 and 20%. In this work, the coefficients of variation for the chemical elements in the groups defined for Lago Grande and Osvaldo (alr transformation) agree with such intervals, according to the following:

- Group A (Lago Grande): 9–18% ($\log_{10}$: 11–20%, 22% for Fe).
- Group B (Lago Grande): 7-15% ($\log_{10}$: 11–15%, 17% for Ce).
- Group A (Osvaldo): 6-14% ($\log_{10}$: 11–20%).
- Group B (Osvaldo): 8-14% ($\log_{10}$: 10–14%, 23% for Ce).

This indirect comparison can be considered another indication that the chemical groups defined are statistically meaningful and with lower dispersion than in the $\log_{10}$ analysis. For a direct test, see MANOVA in Table 5.

The element scandium was chosen as the denominator in the alr transformation in line with the criterion presented in Sect. 4 (Table 2), but it is also a relevant element from the geochemical point of view. Using a conservative lithogenic element (not influenced by diagenetic effects like mobilization and leaching) as a reference, may assist in the compensation of anthropogenic effects that cause data variability, such as the addition of non-plastic materials, and improve the chemical fingerprinting to identify clay sources. Dias and Prudêncio (2007) reported that Si, Al, Li, Sc, Fe, Mn, organic carbon, and carbonates are good chemical factors of geochemical data normalization to compensate for concentration anomalies of natural origin. Furthermore, Li and Sc are two chemical elements that are structurally combined in clay minerals and micas and good indicators of such minerals. Finally, from an experimental perspective, scandium was the element that presented the second lowest relative concentration (12% - sponge/ceramics) in the temper cauixí among the elements used in this work, being only surpassed by chromium (10% - sponge/ceramics) (Hazenfratz-Marks 2014). In this regard, the element Sc is adequate as the alr transformation reference according to the statistical criterion of lowest variability and the geochemical suitability.

Regarding a possible correlation between the chemical groups of pottery with a geographical provenance, the *abundance criterion* is frequently employed (e.g. Harbottle 1982; Buxeda i Garrigós et al. 2001). According to this principle, the group with more members would be characteristic of the archaeological site where the pottery shards were collected. In contrast, smaller groups would refer to exogenous ceramics and correlate with cultural exchanges, such as trade. However, in this work, it is impossible to apply such a criterion because the sampling of both archaeological sites was not random, but partially biased to increase the number of samples of non-dominant ceramic phases in each site to improve their chemical characterization statistically.

Furthermore, in Hazenfratz et al. (2016), the superposition of the chemical groups from Lago Grande and Osvaldo defined in the $\log_{10}$ PCA indicates that pottery from both sites may be related. In this paper, after revisiting such data set with the alr transformation, the same general patterns of pairwise superposition were identified (Figs. 3 and 4). From the analysis of the distribution of pottery in the groups (Figs. 6 and 7), no clear association of different ceramic phases with distinct chemical groups can be uniquely determined, as it is possible to find Paredão and Manacapuru pottery in both groups for each archaeological site.

From the nature of the alr transformation, we can infer that the ceramic recipes used in both sites were geochemically similar in many cases, even after discounting for the effect of dilution by temper addition. However, Manacapuru pottery seems to concentrate in groups A for both sites (Fig. 6) regardless of the data transformation used, while Paredão pottery change dominance in the groups for different data transformation (Fig. 7). These results indicate that the Paredão pottery is more affected by the addition of non-plastic materials than Manacapuru´s, indicating a possible technological change since the Paredão ceramic phase is more recent (Lima 2008). However, expanding the sampling of Osvaldo pottery is advisable for better statistics (see Table 3).

Similarly to the $\log_{10}$ analysis (Hazenfratz et al. 2017), the departure from multivariate normality across all chemical groups did not significantly alter the overall multivariate patterns identified for the alr-transformed data. While the rejection of multivariate normality could potentially impact the analysis, particularly when groups closely cluster in the multivariate space, as observed in this study (Figs. 3 and 4), it is noteworthy that the same general multivariate patterns were consistently identified in the SOMs and CA+PCA (compare Figs. 3, 4 and 8). Therefore, we can infer that the deviations from multivariate lognormality were not critical for our dataset.

Archaeological implications

Generally, the similarities observed between the $\log_{10}$ and alr transformations (Figs. 3 and 4) regarding the pairwise superposition of the chemical groups from both archaeological sites are suggestive of cultural interactions, which could happen through commercial exchange, exogamic marriage and/or territory sharing between the former inhabitants of Lago Grande and Osvaldo (e.g. Moraes 2006; Lima 2008; Mongeló 2011).

Trade would be a means of establishing social and political integration in the region. Exogamic marriage, identified by ethnographic research in recent times (Moraes 2006), would imply women carrying gifts and expertise as pottery-makers from their original village, developing their skills in a hybrid fashion over time in terms of techniques and choice of raw materials. The identification of hybrid pottery in Lago Grande would corroborate such a mechanism (Lima 2008). Finally, territorial integration would imply that the ancient inhabitants of Lago Grande, Osvaldo, and other contemporary Paredão and Manacapuru sites used common sources of clays and other materials. In favor of this hypothesis, note that the mixture of clay sources near both sites plotted in Fig. 8 could generate the average chemical composition of pottery groups.

There is evidence in the ethnographic literature that those three mechanisms could have occurred simultaneously (Lima 2008; Mongeló 2011). In this sense, the multivariate geochemical patterns identified in this work with different data transformations would reflect complex and multimodal means of social and cultural interaction and integration in Central Amazon, as it is not possible to distinguish between different mechanisms solely based on clear correlations between chemical groups and types of ceramics in Figs. 6 and 7, for example.

Narrow water channels (igarapés) connecting the lakes around Lago Grande and Osvaldo could have facilitated the contact and exchanges between the two populations during the flooding periods. Archaeological research suggests that exchange networks would be a condition for long-term occupations, strengthened by evidence of pre-Columbian forms of agriculture in interfluvial Amazonian uplands. It contrasts with the short-term occupations documented in the ethnographic literature and strengthens theories that minimize the role of ecological constraints in the emergence of social complexity and sedentary occupations in the region (e.g. Machado 2006; Neves and Petersen 2006; Silva 2009).

Regarding the use of raw materials in pottery-making, despite the compensation of temper addition by the alr data transformation, there are other features which make it impossible to correlate the chemical groups to different clay sources univocally. Portocarrero (2006) and Lima (2008) studied the first and most important aspect, regarding the use of different types of clay in the chromatic analyzes of pottery from Lago Grande and Osvaldo. Furthermore, for the Incised Rim Tradition pottery, Lima (2008) concluded that specific forms of ceramic objects were produced with different recipes (types of clay, temper, and post-treatments) for different uses according to their characteristics.

As Bishop (2003) has already pointed out, a second aspect regarding pottery-making in societies where the ceramic production is not centralized is that different producers may have unrestricted access to different clay sources and other raw materials, imposing a high variability in the multivariate chemical patterns of pottery. In such cases, the chemical groups would actually represent clusters of smaller groups associated with different ceramic recipes. Even in a scenario of centralized ceramic production, it is possible to find subgroups of similar objects located in larger groups.

Tests with a deeper partitioning in cluster analysis defining more groups for each archaeological site (up to seven for Lago Grande and four for Osvaldo) were performed. However, we did not find any obvious archaeological correlations that would justify the formation of more groups in our data set. In the future, more samples and the analysis of other Central Amazonian archaeological sites will be necessary to assess the possibility of working with a higher geochemical resolution to define groups in PCA. Thus, improving the analytical precision for more chemical elements than those considered here would be desirable, although defining more chemical groups would only be useful if it brings additional archaeological information, which is the final objective in archaeometric studies.

The addition of temper did not cause significant non-linear disturbances in the multivariate geochemical structure for the general definition of groups (for our chemical resolution). Such perturbations could change the multivariate patterns in a more complex and heterogeneous way than just diluting the concentration of chemical elements by the same factor. As indicated previously, former INAA of cauixí showed no significant amounts of trace elements (Hazenfratz-Marks 2014), confirming the scenario where it would serve only as a means of dilution of the elemental concentrations in the ceramic paste. However, differences in the group memberships were observed for analogous groups from both sites when data transformation was changed (see Table 4). It might reflect the fact that some ceramic pieces were more affected by the addition of temper, due to the referred variability in the ceramic paste recipes.

Regarding the correlation of the ceramic phases and temper, Figs. 6 and 7 indicate that the chemical profiles of Paredão ceramics were more affected by the addition of non-plastic materials than the profiles of Manacapuru ones, being even more pronounced in Osvaldo than in Lago Grande. If we assume a context of trade exchanges, a first assumption is that it would be related to variation in the Paredão ceramic recipes intended for trade with Manacapuru sites compared to ceramics produced for local use in Lago Grande. This observation would provide further evidence in support of Lima's (2008) conclusion that specific forms of ceramic objects were produced with different recipes to fulfill different uses and performance requirements.

Another assumption for the results in Fig. 7 is that the higher degree of variation in the group classification of Paredão ceramics could be related to non-exclusive commercial exchanges of Osvaldo with Lago Grande, but also with other Paredão sites, which could have access to different clay sources and/or might have used different recipes. This assumption is based on studies that infer a possible interconnection among different sites in the pre-colonial period (Donatti 2003; Machado 2006; Moraes 2006; Portocarrero 2006; Neves 2008; Heckenberger and Neves 2009).

Chronologically, the Paredão ceramic phase is more recent than Manacapuru and coincident with a demographic peak in Central Amazon (Lima 2008; Neves 2008). Hence, the results could also reflect a technological change in the production of such ceramics regarding the addition of temper and choices of clay sources. The increased variability for Paredão classification into the chemical groups when data transformation is changed might indicate a greater variation in the selection of raw materials and/or addition of temper by potters from Lago Grande. This variability could be attributed to the need for producing more utilitarian pottery during the demographic peak, combined with the seasonal availability of clays and other raw materials during the flood and drought periods, which are typical throughout the year in the Central Amazon region (Viers et al. 2005; Moraes 2006; Neves 2006; Heckenberger and Neves 2009).

Portocarrero (2006) verified that around 42% of Manacapuru pottery contained ground pottery shards. Adding such a non-plastic material to the ceramic paste reduces the geochemical resolution in the elemental analysis of pottery, because, unlike the effects of cauixí addition, such ground shards do not only dilute the concentration of chemical elements but introduce non-linear perturbations in the multielemental data set, affecting the separation of groups and the precision of archaeological interpretations (e.g. Neff et al. 1989). Considering such a scenario, a systematic study regarding the chemical fingerprinting of clay sources in the confluence region of Negro and Solimões rivers could assess the expected natural geochemical variation and separate it from the anthropically-induced ones. However, until the present date, the authors identified no systematic study comprising the determination of trace elements in clay sources of the region.

Regarding the stratigraphic levels of pottery, no evident correlation was found with the chemical groups defined in any scenario of data transformation. However, three excavation units were associated with discarding piles (Donatti 2003), so stratigraphic perturbations are expected.

The provenance postulate cannot be strictly applied for the pottery shards in this study (Weigand et al. 1977; Wilson and Pollard 2001) since a critical assumption regarding the absence of mixture among raw materials cannot be held, as discussed before, and due to the insufficient characterization of clay sources in the confluence region of Negro and Solimões rivers (see Fig. 8). Furthermore, the region's complex flood and drought dynamics would accelerate diagenetic effects in pottery, reducing the geochemical resolution for identifying multivariate patterns of archaeological interest.

Finally, the analysis of chemical elemental data of pottery from Lago Grande and Osvaldo archaeological sites, comprising two data transformations, suggests that the archaeological conjecture of a regional exchange network for Central Amazon is highly probable by the previous arguments. Furthermore, Mongeló (2011) reported at least 15 Manacapuru and 14 Paredão sites, which might have been contemporary to Lago Grande and Osvaldo during a specific period, complicating the interpretation of the geochemical patterns as direct correlates of exclusive interactions between both sites.

Comparison with other log-ratio transformations

To compare the results of the alr and $\log_{10}$ analyzes with other log-ratio transformations for compositional data, the centered log-ratio (clr) and the isometric log-ratio transformation were also explored. The results are not presented here because the alr transformation is the most appropriate for this study (Buxeda i Garrigós 1999). The results from clustering and PCA with clr and isometric log-ratio showed no correlation with alr or $\log_{10}$ transformations. Furthermore, they weakened previously observed correlations between chemical groups and ceramic phases, having more in common with themselves than with the other two transformations. The interested reader should refer to the supplementary material comprising cluster and PCA for each archaeological site with these data transformations.

The possible distortions in cluster analysis due to these transformations would be more critical for the samples farthest from the chemical group centroids. However, this problem can be minimized since the cluster classification is refined by a more robust method using the Mahalanobis distance (operating in the space of principal components) and the F distribution, thus considering the correlation structure among the chemical elements (Glascock 1992).

Summary and conclusions

Buxeda i Garrigós (1999) reported that the additive log-ratio transformation (alr) is the most robust data transformation method to deal with alterations and contaminations in archaeometric compositional data. The advantages of using elemental ratios would be related to the compensation of dilution effects from the heterogeneous addition of temper in pottery, diagenetic effects in the depositional context, or natural geochemical variations in the clay sources.

Within this context, this study analyzed a data set of chemical elemental concentrations of pottery shards from two large and representative archaeological sites in Central Amazon, Lago Grande and Osvaldo. The aim was to compare the $\log_{10}$ transformation to the LRA concerning its effect on the archaeological interpretation, assessing whether the addition of non-plastic materials in the ceramic production had a critical effect on Central Amazonian pottery analysis.

This study focused on a scenario extracted from Hazenfratz-Marks (2014), in which nine chemical elements measured by INAA were used (La, Lu, Yb, Ce, Cr, Eu, Fe, Sc, and Th) with alr data transformation for chemical fingerprinting by CA and PCA. The selection of scandium as the denominator in the alr transformation was made by calculating the variation matrix of the subcompositional data, choosing the element which imposed the lowest variations into the original covariance matrix. MANOVA was employed to test for statistical differences between the chemical groups, and SOMs, a type of artificial neural network, were used for comparison with CA and PCA, since they do not depend on any specific hypothesis about the data distribution and the covariance structure.

Figures 3 and 4 showed that the general patterns regarding the definition of two chemical groups and their pairwise superposition were not altered by changing the data transformation, although the classification of some samples into chemical groups changed. Furthermore, both analyzes showed a reasonable group separation, with some degree of superposition, according to what is expected for geochemically similar environments (Bishop 2003). It was verified by MANOVA that it is impossible to consider groups A and B from Lago Grande and A from Osvaldo as different chemical groups when the data transformation was changed from $\log_{10}$ to alr. Due to the nature of the alr transformation, we can then infer that the ceramic recipes identified in both sites are geochemically similar in many cases, even after discounting the effect of dilution by the temper addition.

The SOM for the alr transformation yielded comparable general patterns to those identified in CA and PCA, indicating that deviations from multivariate lognormality were not critical for the PCA results. The indirect comparison of the statistical significance based on recommended intervals for coefficients of variations by Harbottle (1982) and Bishop (2003) indicated that the chemical groups defined are statistically meaningful.

The results of the multivariate analyzes were explored to find possible archaeological correlations with provenance and typological analyzes. Neither the change in the data transformation generated more chemical groups with a clear archaeological correlation, nor the two groups identified in both analyzes (for each site) showed a clear correlation with the origin, ceramic phase, typological classification, or stratigraphic level.

It should be noted that the inclusion of other minor and major chemical elements in the analysis could lead to more striking differences when changing the data transformation. In this work, two major elements measured, Na and K, were not used due to abnormal enrichment in the ceramic samples due to diagenetic effects. In other scenario, where Hf, Cs and As were used (Hazenfratz-Marks 2014), striking differences were observed in the size of chemical groups and membership of pottery shards to them. The analysis of contamination and diagenetic effects were performed also with the aid of the alr transformation. In this work, the identified geochemical patterns (although not strictly the pottery classification into groups) were similar for the $\log_{10}$ and alr analysis. However, it could not be expected a priori for our data set, due to potentially perturbed chemical profiles by the use of cauixí temper and crushed shards in the ceramic production, with the latter inserting non-linear perturbations with the potential to affect the discriminatory power of the analyzes (e.g. Neff et al. 1989).

Differences in the classification of samples were observed for the chemical groups from both sites when the data transformation was changed, and Paredão pottery were the most affected by temper addition, implying more variability in the ceramic paste recipes, which could be intentional when considering the hypothesis of exchange networks in the region.

Archaeologically, in the context of commercial exchange, a first assumption was related to variation in the Paredão ceramic recipes intended for trade when compared to ceramics of local use in Lago Grande, in agreement with Lima´s (2008) conjecture that specific forms of ceramic objects were produced with different clay recipes. Another assumption was that the higher degree of variation in the patterns of Paredão ceramics (mainly for Osvaldo), when the data transformation was changed, could be related to non-exclusive commercial exchanges of Osvaldo with Lago Grande, but also with other Paredão sites in the past. The Paredão ceramic phase is more recent than Manacapuru and coincident with a demographic peak in Central Amazon. Thus, the last hypothesis is that the higher chemical variability for the Paredão samples might indicate more variation in the selection of raw materials due to the necessity of producing more utilitarian pottery, compensating the seasonality of clay sources and other materials over the flood and drought periods (Viers et al. 2005; Moraes 2006; Neves 2006; Heckenberger and Neves 2009).

Still regarding the use of raw materials in pottery-making, the use of different types of clay identified in the chromatic analyzes of pottery shards of Lago Grande and Osvaldo by Portocarrero (2006) and Lima (2008) makes it impossible to correlate univocally the chemical groups to different clay sources, even after discounting the temper effect by the alr transformation. Furthermore, it was verified by Portocarrero (2006) that many Manacapuru samples contained ground pottery shards, which might reduce the geochemical resolution of an elemental analysis by introducing non-linear perturbations into the data.

The pairwise superposition of the chemical groups from both sites (Figs. 3 and 4) regardless of the data transformation corroborates the existence of cultural interactions between Lago Grande and Osvaldo in the past, which could have occurred through commercial exchange, exogamic marriage and/or territory sharing. There is archaeological evidence that those three mechanisms could have occurred simultaneously (Lima 2008; Mongeló 2011) and the results of this work support this conjecture. In this sense, the multivariate geochemical patterns identified here and in previous works would reflect complex and multimodal means of socio-cultural interaction and integration in Central Amazon. Moreover, the existence of other contemporary Manacapuru and Paredão sites makes difficult the interpretation of the geochemical patterns as direct correlates of exclusive interaction between Lago Grande and Osvaldo. In a broader perspective, the exchange networks corroborated by the results favor theories that minimize the role of ecological constraints in the emergence of social complexity and sedentary occupations in the Amazon (Machado 2006; Neves and Petersen 2006; Silva 2009).

Concerning archaeometric compositional data analysis, it is not possible to consider that the use of multivariate statistical methods with log-ratio transformation is better *a priori*. It is advisable to use different data transformations and statistical tools to approach different research questions for a broader perspective. It could extract the greatest potential for reliable archaeological interpretation of the multivariate data structure. It is far from our intention to determine which data transformation would be the most appropriate for chemical analyzes of pottery shards. The $\log_{10}$, alr and other data transformations not considered in this work are not mutually exclusive and can be combined in archaeometric studies as needed.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12520-024-01965-y.

Acknowledgements This research study was conducted with the support of the São Paulo Research Foundation—FAPESP—Brazil (Process Number: 2010/07659-0), and partially with the support of the Conselho Nacional de Desenvolvimento Científico e Tecnológico—CNPq—Brazil (Process Number: 134116/2009-7). We would like to acknowledge the Archaeometry Laboratory at the University of Missouri Research Reactor and Dr. Michael Glascock, for making the MURRAP freeware software available. The authors also acknowledge the valuable discussions and suggestions provided by anonymous reviewers, which have significantly contributed to enhancing the quality of this text.

Author contributions RH worked on experimental procedures, data acquisition, analysis and interpretation. GZL contributed with the selection of samples, archaeological introduction and interpretation. CSM contributed with the selection of samples, analytical methods and general discussions. EGN contributed with the selection of samples, archaeological introduction and interpretation. All authors reviewed the manuscript.

Funding This research study was conducted with the support from São Paulo Research Foundation—FAPESP—Brazil (Process Number: 2010/07659-0), and partially with the support of Conselho Nacional de Desenvolvimento Científico e Tecnológico—CNPq—Brazil (Process Number: 134116/2009-7).

Data availability A spreadsheet with the raw data was included as supplementary material.

Code availability Not applicable.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

Aitchison J (1982) The Statistical Analysis of Compositional Data. Journal of the Royal Statistical Society: Series B (Methodological) 44:. https://doi.org/10.1111/j.2517-6161.1982.tb01195.x

Aitchison J (1986) The Statistical Analysis of Compositional Data

Aitchison J, Barceló-Vidal C, Pawlowsky-Glahn V (2002) Some comments on compositional data analysis in archaeometry, in particular the fallacies in Tangri and Wright's dismissal of logratio analysis. Archaeometry 44. https://doi.org/10.1111/1475-4754.t01-1-00061

Ballabio D, Vasighi M (2012) A MATLAB toolbox for Self Organizing maps and supervised neural network learning strategies. Chemometr Intell Lab Syst 118:24–32. https://doi.org/10.1016/j.chemolab.2012.07.005

Ballabio D, Consonni V, Todeschini R (2009) The Kohonen and CP-ANN toolbox: a collection of MATLAB modules for Self Organizing maps and Counterpropagation Artificial neural networks. Chemometr Intell Lab Syst 98. https://doi.org/10.1016/j.chemolab.2009.05.007

Baria R, Cano NF, Silva-Carrera BN, et al (2015) Archaeometric studies of ceramics from the São Paulo II archaeological site. J Radioanal Nucl Chem 306:721–727. https://doi.org/10.1007/s10967-015-4183-5

Barone G, Mazzoleni P, Spagnolo GV, Raneri S (2019) Artificial neural network for the provenance study of archaeological ceramics using clay sediment database. J Cult Herit 38. https://doi.org/10.1016/j.culher.2019.02.004

Baxter M (2016) Multivariate analysis of Archaeometric Data an introduction. academia.eduSign in

Baxter MJ, Freestone IC (2006) Log-ratio compositional data analysis in Archaeometry. Archaeometry 48

Baxter MJ, Beardah CC, Freestone IC (2005) Compositional analysis of archaeological glasses. CoDaWork'05 Compositional Data Analysis Workshop

Beier T, Mommsen H (1994) Modified mahalanobis filters for grouping pottery by chemical composition. Archaeometry 36. https://doi.org/10.1111/j.1475-4754.1994.tb00971.x

Bell S, Croson C (1998) Artificial neural networks as a tool for archaeological data analysis. https://doi.org/10.1111/j.1475-4754.1998.tb00829.x. Archaeometry 40:

Bishop RL (2003) Instrumental neutron activation analysis of archaeological ceramics: progress and challenges

Buccianti A, Grunsky E (2014) Compositional data analysis in geochemistry: are we sure to see what really occurs during natural processes? J Geochem Explor 141. https://doi.org/10.1016/j.gexplo.2014.03.022

Buxeda i Garrigós J (1999) Alteration and contamination of archaeological ceramics: the perturbation problem. J Archaeol Sci 26. https://doi.org/10.1006/jasc.1998.0390

Buxeda i Garrigós J (2018) Compositional data analysis. The Encyclopedia of Archaeological sciences. John Wiley & Sons, Inc., Hoboken, NJ, USA

Buxeda i Garrigós J, Kilikoglou V, Day PM (2001) Chemical and mineralogical alteration of ceramics from a late bronze age kiln at Kommos, Crete: the effect on the formation of a reference group. Archaeometry 43. https://doi.org/10.1111/1475-4754.00021

Buxeda i Garrigós J, Mommsen H, Tsolakidou A (2002) Alterations of na, K and rb concentrations in mycenaean pottery and a proposed explanation using X-ray diffraction. Archaeometry 44. https://doi.org/10.1111/1475-4754.t01-1-00052

Cau Ontiveros MÁ, Fantuzzi L, Tsantini E et al (2019) Archaeometric characterization of common and cooking wares from the late antique city of Valentia (Valencia, Spain). Archaeol Anthropol Sci 11. https://doi.org/10.1007/s12520-019-00802-x

Civici N (2007) Analysis of illyrian terracotta figurines of Aphrodite and other ceramic objects using EDXRF spectrometry. X-Ray Spectrom 36:92–98. https://doi.org/10.1002/xrs.945

Dias MI, Prudêncio MI (2007) Neutron activation analysis of archaeological materials: An overview of the ITN NAA Laboratory, Portugal. Archaeometry 49

Donatti PB (2003) A ocupação pré-colonial Da área do Lago Grande, Iranduba, AM. Dissertação (Mestrado em Arqueologia). Universidade de São Paulo

Egozcue JJ, Pawlowsky-Glahn V (2019) Compositional data: the sample space and its structure. Test. https://doi.org/10.1007/s11749-019-00670-6

Fantuzzi L, Cau MA (2017) Investigating the provenance of the baetican amphorae Dressel 23: New archaeometric evidence from late roman consumption centres. Mediterranean Archaeol Archaeometry 17. https://doi.org/10.5281/zenodo.258085

Gaudette HE, Grim RE, Metzger CF (1966) Illite: a model based on the sorption behavior of Cesium. Am Mineral 51:1649–1656

Glascock MD (1992) Characterization of Archaeological Ceramics at MURR by Neutron Activation Analysis and Multivariate statistics. In: Neff H (ed) Chemical characterization of ceramic pastes in Archaeology. Pre-History, Madison, pp 11–26

Greenacre M (2019) Variable selection in Compositional Data Analysis using Pairwise Logratios. Math Geosci 51. https://doi.org/10.1007/s11004-018-9754-x

Greenacre M (2021) Compositional data analysis. Annu Rev Stat Appl 8. https://doi.org/10.1146/annurev-statistics-042720-124436

Harbottle G (1982) Chemical characterization in Archaeology. In: Contexts for Prehistoric Exchange

Hazenfratz R, Mongeló G, Munita CS, Neves EG (2012) Comparison of INAA elemental composition data between Lago Grande and Osvaldo archaeological sites in the central Amazon: a first perspective. J Radioanal Nucl Chem 291:43–48. https://doi.org/10.1007/s10967-011-1251-3

Hazenfratz R, Munita CS, Glascock MD, Neves EG (2016) Study of exchange networks between two Amazon archaeological sites by INAA. J Radioanal Nucl Chem 309. https://doi.org/10.1007/s10967-016-4758-9

Hazenfratz R, Munita CS, Neves EG (2017) Neural networks (SOM) Applied to INAA Data of Chemical Elements in Archaeological Ceramics from Central Amazon. Sci Technol Archaeol Res 3. https://doi.org/10.1080/20548923.2018.1470218

Hazenfratz-Marks R (2014) Nêutrons, radiação e arqueologia: estudo de caso multianalítico de cerâmicas da tradição borda incisa na Amazônia Central. Universidade de São Paulo

Heckenberger M, Neves EG (2009) Amazonian archaeology. Annu Rev Anthropol 38. https://doi.org/10.1146/annurev-anthro-091908-164310

Hein A, Kilikoglou V (2017) Compositional variability of archaeological ceramics in the eastern Mediterranean and implications for the design of provenance studies. J Archaeol Sci Rep 16. https://doi.org/10.1016/j.jasrep.2017.03.020

Hilbert PP (1968) Archäologische Untersuchungen am mittleren Amazonas: Beiträge Zur Vorgeschichte Des südamerikanischen Tieflandes. Mit 164 Abbildungen und 11 Karten. vom Verfasser), Reimer

Keding R, Jensen M, Yue Y (2010) Characterization of the mesoporous amorphous silica in the fresh water sponge Cauxi. In: Ceramic Transactions

Koch GS, Link RF (2002) Statistical analysis of geological data. Courier Corporation

Kohonen T (2013) Essentials of the self-organizing map. Neural Netw 37:52–65. https://doi.org/10.1016/j.neunet.2012.09.018

Lathrap DW, Lathrap DW (1970) The upper amazon. Thames & Hudson Southampton

Leroy S, Bauvais S, Delqué-Količ E et al (2020) First experimental reconstruction of an angkorian iron furnace (13th–14th centuries CE): Archaeological and archaeometric implications. J Archaeol Sci Rep 34. https://doi.org/10.1016/j.jasrep.2020.102592

Lima HP (2008) História das Caretas: a Tradição Borda Incisa na Amazônia Central. Tese (Doutorado em Arqueologia). Universidade de São Paulo

López AJ, Ramil A (2018) Neural networks. The Encyclopedia of Archaeological sciences. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp 1–4

Lopez-Molinero A, Castro A, Pino J et al (2000) Classification of ancient roman glazed ceramics using the neural network of Self-Organizing maps. Fresenius J Anal Chem 367. https://doi.org/10.1007/s002160000433

Luneau E, Martínez Ferreras V, Abdykanova A et al (2020) The first combined archaeological and archaeometric analyses on bronze age pottery from Kyrgyzstan (Uch Kurbu site). J Archaeol Sci Rep 31. https://doi.org/10.1016/j.jasrep.2020.102302

Machado JS (2006) Dos Artefatos às aldeias: Os vestígios arqueológicos no entendimento das formas de organização social Da Amazônia. Revista De Antropologia 49. https://doi.org/10.1590/s0034-77012006000200009

Martín-Fernández JA, Engle MA, Ruppert LF, Olea RA (2019) Advances in self-organizing maps for their application to compositional data. Stoch Env Res Risk Assess 33:817–826. https://doi.org/10.1007/s00477-019-01659-1

Martinelli LA, Victoria RL, Dematte JLI et al (1993) Chemical and mineralogical composition of Amazon River floodplain sediments, Brazil. Appl Geochem 8:391–402. https://doi.org/10.1016/0883-2927(93)90007-4

Martini M, Milazzo M, Piacentini M (2004) Physics methods in archaeometry. IOS

Meyers LS, Gamst G, Guarino AJ (2006) Applied multivariate research: Design and interpretation.

Mommsen H (2004) Short note: provenancing of pottery - the need for an integrated approach? https://doi.org/10.1111/j.1475-4754.2004.00156.x. Archaeometry 46:

Mongeló GZ (2011) Processos De interação entre os sítios Lago Grande E Oswaldo (AM) baseados no material cerâmico. Revista do Museu de Arqueologia e Etnologia, pp 109–114

Montana G, Randazzo L, Tsantini E, Fourmont M (2018) Ceramic production at Selinunte (Sicily) during the 4th and 3rd century BCE: New archaeometric data through the analysis of kiln wastes. J Archaeol Sci Rep 22. https://doi.org/10.1016/j.jasrep.2018.09.022

Moraes CP (2006) Arqueologia na Amazônia Central Vista De Uma Perspectiva Da Região do Lago do Limão. Dissertação (Mestrado em Arqueologia). Universidade de São Paulo

Munita CS, Paiva RP, Alves MA et al (2001) Major and trace element characterization of prehistoric ceramic from Rezende archaeological site. J Radioanal Nucl Chem 248:93–96. https://doi.org/10.1023/A:1010682209370

Munita CS, Glascock MD, Hazenfratz R (2019) Neutron activation analysis. An Overview

Neff H, Bishop RL, Sayre EV (1989) More observations on the problem of tempering in compositional studies of archaeological ceramics. J Archaeol Sci 16. https://doi.org/10.1016/0305-4403(89)90056-3

Neves EG (2006) Arqueologia Da Amazônia. Jorge Zahar, Rio de Janeiro

Neves EG (2008) Ecology, ceramic chronology and distribution, long-term history, and political change in the amazonian floodplain. In: The Handbook of South American Archaeology

Neves EG (2010) A arqueologia da Amazônia Central E as classificações na Arqueologia amazônica. Arqueologia amazônica 2:54–72

Neves EG (2012) Sob os tempos do equinócio: oito mil anos de história na Amazônia Central (6.500 AC—1.500 DC). Universidade de São Paulo, São Paulo, Brazil

Neves EG (2018) The heart of lightness. Engaging Archaeology. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp 79–86

Neves EG, Petersen JB (2006) 9. Political Economy and Pre-columbian Landscape transformations in Central Amazonia. In: Time and Complexity in Historical Ecology

Pagnotta S, Lezzerini M, Botto A et al (2018) Evaluation of potsherds features using hyperspectral maps generated by µ-LIBS scanner. Estudos De Conservação E Restauro 55–63. https://doi.org/10.34618/ecr.9.18872

Pawlowsky-Glahn V, Egozcue JJ, Tolosana Delgado R (2007) Lecture notes on compositional data analysis

Portocarrero RC (2006) A variabilidade espacial no sítio Osvaldo. Estudo De um assentamento da tradição barrancóide na Amazônia central. Dissertação (Mestrado em Arqueologia), Universidade de São Paulo

Schwedt A, Mommsen H (2007) On the influence of drying and firing of clay on the formation of trace element concentration profiles within pottery. https://doi.org/10.1111/j.1475-4754.2007.00316.x. Archaeometry 49:

Schwedt A, Mommsen H, Zacharias N (2004) Post-depositional elemental alterations in pottery: Neutron activation analyses of surface and core samples. Archaeometry 46:85–101

Silva FA (2009) A etnoarqueologia na Amazônia: contribuições e perspectivas. Bol do Museu Paraense Emílio Goeldi Ciências Humanas 4. https://doi.org/10.1590/s1981-81222009000100004

Speakman RJ, Glascock MD (2007) Acknowledging fifty years of neutron activation analysis in archaeology. Archaeometry 49

Tanevska V, Kuzmanovski I, Grupče O (2007) Provenance determination of Vinica terra cotta icons using self-organising maps. Ann Chim 97. https://doi.org/10.1002/adic.200790036

Tite MS (2008) Ceramic production, provenance and use—a review. Archaeometry 50:216–231. https://doi.org/10.1111/j.1475-4754.2008.00391.x

Toyota RG, Munita CS, Boscariolli C et al (2009) Neural Network Applied to Elemental Archaeological Marajoara Ceramic Compositions. In: Anais de International Nuclear Atlantic Conference (INAC 2009)

Tsantini E, Martínez Ferreras V, Ariño Gil E et al (2016) Pottery production in the Buddhist communities in Central Asia: the Kushan-Sassanian pottery workshop of Kara Tepe (Termez, Uzbekistan). https://doi.org/10.1111/arcm.12161. Archaeometry 58:

Viers J, Barroux G, Pinelli M et al (2005) The influence of the amazonian floodplain ecosystems on the trace element dynamics of the Amazon River mainstem (Brazil). Sci Total Environ 339. https://doi.org/10.1016/j.scitotenv.2004.07.034

Warner RM (2013) From Bivariate through Multivariate techniques. From Bivariate through Multivariate techniques, Applied Statistics

Weigand PC, Harbottle G, Sayre EV (1977) Turquoise sources and source analysis: mesoamerica and the southwestern U.S.A. In: Exchange Systems in Prehistory

Wilson L, Pollard AM (2001) The provenance hypothesis. Handbook of Archaeological Sciences

Wood JR, Greenacre M (2021) Making the most of expert knowledge to analyse archaeological data: a case study on parthian and sasanian glazed pottery. Archaeol Anthropol Sci 13. https://doi.org/10.1007/s12520-021-01341-0

Woronow A, Love KM (1990) Quantifying and testing differences among means of compositional data suites. Math Geol 22. https://doi.org/10.1007/BF00890666

Zhou S, Zhou K, Wang J et al (2018) Application of cluster analysis to geochemical compositional data for identifying ore-related geochemical anomalies. Front Earth Sci 12:491–505. https://doi.org/10.1007/s11707-017-0682-8

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.