

ARQUEOESTATÍSTICA APLICADA AO ESTUDO COMPOSICIONAL DE CERÂMICAS ARQUEOLÓGICAS

J.O. SANTOS¹, C.S. MUNITA², M.E.G. VALÉRIO³, C. VERGNE⁴

ABSTRACT

Chemical studies of the ceramics are based on the assumption that ceramic produced from a specific clay will show a similar chemical composition and that its composition will be distinguished from that of ceramics produced from a different clay. It is common to determine the chemical composition of ceramic using techniques such as neutron activation analysis, inductively coupled plasma spectroscopy, x-ray fluorescence analysis. In this work 74 ceramics fragments and 1 clay from three archaeological sites (São José 26, Saco da Onça 24 and Curitiba: 24) located in the Brazilian Northeast were analyzed using INAA to determine the concentration of 24 chemical elements. The results were interpreted by mean of Archaeostatistics techniques, such as Linear Discriminant Analysis and principal components, which was utilized to facilitate identification of compositional groups. The Mahalanobis distance was applied for detecting outliers and Wilks's lambda provided the critical values. It was obtained that samples from São José and Curitiba are constitute of ceramic pastes different, while samples from Saco da Onça Site overlap the pottery samples from São José and Curitiba Sites.

Palavras chave:

Arqueoestatística, análise por ativação com nêutrons, arqueometria.

¹ Centro Federal de Educação Tecnológica de Sergipe, CEFET/SE-UNED LAGARTO. CEP 49.055-260, Sergipe, SE, Brasil. * E-mail: josantos@ipen.br

² Instituto de Pesquisas Energéticas e Nucleares, IPEN-CNEN/SP. C.P. 11049 CEP 05422-970, São Paulo, SP, Brasil

³ Departamentento de Física, UFS, Av Marechal Rondon s/n, São Cristóvão, SE, Brasil

⁴ Museu de Arqueologia de Xingó, MAX/UFS, Av Marechal Rondon s/n, São

Cristóvão, SE, Brasil

INTRODUÇÃO

Como resultado da alta resistência em relação ao meio circundante, os vestígios de origem cerâmica são os mais comumente encontrados em escavações arqueológicas, e indicam interações sócio-culturais e econômicas dos povos.

As principais abordagens para caracterização e classificação das cerâmicas arqueológicas têm sido realizadas, quase que exclusivamente, por meio de seus perfis técnicos - morfológicos (decoração, cor, forma, função, etc.) (Beaudry, 1991). Entretanto, nos últimos anos tem sido sugerida a complementação destas abordagens com métodos que permitam uma classificação cerâmica mais objetiva (Borone et al., 2002).

Diferentemente dos atributos macroscópicos da cerâmica, suas propriedades microscópicas, tal como composições químicas e mineralógicas, têm revelado informações a respeito das origens, níveis de intercâmbios, costumes, processos de migrações, entre outros fenômenos sociais.

Os estudos arqueológicos realizados com base nas propriedades físico-químicas dos vestígios encontrados têm constituído um ramo das ciências denominado Arqueometria. A Arqueometria tem se utilizado de diversas técnicas de caracterização físicas e químicas para obter o máximo de informações com relação aos materiais em estudo (composições químicas e mineralógicas, datações, níveis de desgastes, etc.). Em especial, as técnicas de caracterização química têm sido utilizadas em virtude das diversas questões que podem ser solucionadas (Munita et al., 2000). Historicamente, a espectroscopia de emissão atômica (OES) foi a técnica analítica pioneira no estudo composicional da cerâmica, sendo superada, posteriormente, pelas técnicas de espectroscopia de absorção atômica (AAS), análise por ativação com nêutrons instrumental (AANI), fluorescência de raios-X (XRF) e mais recentemente pela espectroscopia de massa com plasma indutivamente acoplado (ICP-MS) (Pillay, 2001).

Entre as técnicas utilizadas, a AANI é atualmente a técnica analítica mais bem sucedida nos estudos composicionais da pasta cerâmica (Bishop & Blackman, 2002). Por meio desta técnica é possível determinar, simultaneamente, mais de 30 elementos químicos (elementos maiores e traços) com alta precisão e exatidão. Além destas vantagens, por ser uma técnica instrumental, apresenta relativa facilidade na preparação das amostras o que resulta na redução dos erros experimentais.

A classificação das amostras analisadas é facilitada quando suas composições químicas são estatisticamente distinguíveis de acordo com seus perfis químicos. No estudo composicional da cerâmica, supõe-se que as características da pasta são determinadas pelas características químicas das argilas e temperos utilizados. Como o comportamento químico das rochas que deram origem aos materiais utilizados para fabricação da cerâmica é um reflexo dos processos geológicos, as presenças dos elementos são dependentes destes processos. Assim, a composição química elementar da cerâmica tem uma relação particular com o local do qual foi coletada a matéria-prima.

Durante os últimos anos a aplicação de técnicas multielementares para estudo composicional, tais como AANI, permitiu a geração de uma enorme quantidade de dados de tal forma que a sua interpretação por simples inspeção visual é complexa. Assim, para estudos multicomposicionais têm sido utilizadas técnicas estatísticas multivariadas com o propósito de simplificar a estrutura dos dados, ordená-los e agrupá-los, investigar a dependência entre as variáveis, realizar as previsões e executar os testes de hipóteses.

A formação de grupos estatisticamente significantes das amostras analisadas e a avaliação de amostras de origem desconhecidas para alocação da mesma em um dos grupos estabelecidos são objetivos básicos dos estudos composicionais da cerâmica. Para estes fins têm sido aplicadas técnicas multivariadas de agrupamento, análise fatorial, análise discriminante e análise por componentes principais (Mommensen et al., 1988). Embora estas técnicas tenham sido desenvolvidas há bastante tempo, somente, a partir dos anos 60 têm sido difundidas em estudos arqueológicos como resultado direto do desenvolvimento oriundo das Ciências da computação. Esta revolução proveniente da microinformática permitiu, sobretudo, a disponibilidade de pacotes computacionais que permitem ao usuário à aplicação das técnicas estatísticas multivariadas com relativa facilidade. Assim, na atualidade, as técnicas estatísticas estão presentes nas pesquisas arqueológicas com diversos objetivos, desde prospecção dos sítios até a reconstrução das culturas antigas. A aplicação destas técnicas estatísticas ao estudo de questões arqueológica tem constituído um novo campo de pesquisa denominado de arqueostatística (Fieller, 1993).

Todavia, a aplicação da arqueostatística requer uma base de dados bem comportada em relação aos valores discrepantes (“outliers”), aos valores perdidos (“missing values”), às condições de normalidade e

em relação às diferenças entre as magnitudes das variáveis medidas. De acordo com estas considerações, previamente à aplicação das técnicas multivariadas, é necessário um estudo sistemático dos “outliers”, das condições de normalização e padronização dos dados, visando diminuir os ruídos nas interpretações finais dos dados.

Com objetivo de apresentar à comunidade científica, sobretudo aos arqueólogos, algumas ferramentas estatísticas aplicadas ao estudo composicional de cerâmicas arqueológicas, neste trabalho são apresentadas técnicas que constituem a base da arqueostatística, e para exemplificar foi realizado um estudo composicional de cerâmicas arqueológicas provenientes de escavações realizadas no município de Canindé do São Francisco, localizado no estado de Sergipe. A análise química elementar das cerâmicas foi realizada por meio do método de análise por ativação com nêutrons instrumental.

ASPECTOS TEÓRICOS DA ANÁLISE POR ATIVAÇÃO COM NÊUTRONS INSTRUMENTAL

A análise por ativação foi descoberta em 1936 quando Hevesy e Levi verificaram que certos elementos tornavam-se radiativos após sua exposição em uma fonte de nêutrons (Hevesy & Levi, 1936). A partir dessa observação, perceberam-se rapidamente as potencialidades das reações nucleares para identificação, qualitativa e quantitativa, de elementos químicos, por meio da medida da radiatividade induzida por uma fonte de nêutrons ou partículas carregadas.

A sequência de eventos que ocorre durante as reações envolvidas na AANI é denominada captura radiativa de nêutrons (Figura 1). Quando um nêutron interage com o núcleo alvo por meio de uma colisão inelástica ou captura, há formação de um núcleo composto em um estado excitado. O núcleo composto é levado para estados energéticos estáveis, quase que instantaneamente, devido à emissão de um ou mais raios gamas, denominados raios gamas prontos. Na maioria dos casos, esta nova configuração resulta em um nuclídeo radiativo com emissão de raios gamas característicos, com taxa de decaimento governada pela meia-vida ($T_{1/2}$) própria do nuclídeo.

Figura 1 - *Representação dos fenômenos envolvidos da ativação de um núcleo.*

Dependendo da espécie radiativa em particular, a meia vida pode alcançar de frações de segundos ou anos (IAEA-TECDOC-564, 1990). A identificação e quantificação das concentrações elementares é realizada por meio de espectroscopia gama, que utiliza a energia e intensidade da radiação emitida para identificar o elemento e determinar sua concentração.

A determinação das concentrações elementares das amostras analisadas pode ser realizada por meio da medida direta das energias e intensidades da radiação emitida, ou por comparação com um material cuja composição química elementar é conhecida previamente (Tölgyessy & Kyr, 1989). No primeiro caso o método é denominado de método absoluto e no segundo caso é chamado de método relativo.

A aplicação do método absoluto requer uma medida de alta precisão da atividade da amostra em estudo. A atividade induzida na amostra pela absorção de nêutrons pelo núcleo pode ser determinada de acordo com a equação (Kuleff & Djingova, 1990).

$$(1)$$

onde N é o número de Avogadro, m a massa da amostra, θ a fração isotópica do elemento, λ a constante de decaimento, t_c o tempo de resfriamento, σ a seção de choque para captura radiativa, ϕ o fluxo de nêutrons e M a massa atômica do elemento a ser determinado.

Por meio da equação (1), pode-se determinar a massa (m) de um dado elemento presente na amostra. Neste caso, a sensibilidade e preci-

são da determinação irão depender da energia da radiação detectada e das condições da medida. No método absoluto é necessário determinar a atividade absoluta da amostra, sendo, assim, extremamente sensível à eficiência de contagem (geometria, absorção de fótons, espalhamentos, tempo morto, etc.) (Alamin & Spyrou, 1997).

Na prática o método absoluto tem sido utilizado com pouca frequência, uma vez que sua precisão é afetada significativamente por diversos fatores, tais como variação do fluxo e distribuição espectral da energia dos nêutrons incidentes. Tendo em vista que a precisão deste método é muito baixa, o método que tem sido utilizado mais comumente é o método relativo, visto que os fatores variáveis no método absoluto (fluxo, tempo de irradiação, seção de choque de absorção radiativa, eficiência de contagem) são praticamente desprezíveis.

O método relativo é baseado na comparação da atividade da amostra em estudo (A_x) com a atividade de um material cuja concentração elementar é conhecida, denominado de padrão (A_s). A amostra e padrão são irradiados (ativados) em condições bastante próximas e de forma simultânea. Neste método o padrão deve ser colocado o mais próximo possível da amostra para evitar erros devido à heterogeneidade do fluxo de nêutrons. Da mesma forma, as medidas das atividades devem ser realizada em condições idênticas para que a precisão do método não seja afetada (Balla et al., 2004).

O método relativo de análise por ativação é baseado na comparação da atividade da amostra em estudo (A), a qual é proporcional a área do pico correspondente, com a atividade de um material cuja concentração elementar é conhecida, denominado de padrão (A_s). A amostra e padrão são irradiados (ativados) em condições bastante próximas e de forma simultânea. Neste método o padrão deve ser colocado o mais próximo possível da amostra para evitar erros devido à heterogeneidade do fluxo de nêutrons. Da mesma forma, as medidas das atividades devem ser realizada em condições idênticas para que a precisão do método não seja afetada (Balla, et al., 2004).

Se o padrão contém uma quantidade conhecida de um dado elemento, cuja concentração é C_s , a concentração C_x do mesmo elemento presente na amostra é dada por

$$(2)$$

Pode-se observar a partir da Equação 2 que todos os parâmetros que podem influenciar nas atividades do material em estudo e padrão se cancelam, tornando a análise independente destes fatores.

Apesar do método relativo tornar a determinação das concentrações elementares mais precisas, deve ser ressaltado que os procedimentos envolvidos nas análises multielementares apresentam algumas dificuldades, tais como (Heydorn, 1990) preparação de um grande número de padrões, obtenção de padrões para determinados elementos, assegurar que as condições de irradiação e medida das amostras e padrões sejam idênticas, entre outras dificuldades.

Estas dificuldades podem ser contornadas, parcialmente, através do uso de padrões multielementares, sintéticos e padrões de materiais de referência. Padrões multielementares devem conter, de preferência, todos os elementos a serem determinados, em quantidades conhecidas.

ARGILAS

A principal fonte de matéria-prima para produção de artefatos cerâmicos é a argila, que é formada basicamente pelos argilominerais, podendo conter minerais que não são argilominerais (calcita, dolomita, quartzo, pirita, etc.), matéria orgânica e outras impurezas. Por ser constituída basicamente de argilominerais, a argila apresenta: estrutura cristalina definida; diâmetro inferior a 2 μm ; boa plasticidade quando umedecida; quando seca adquire alta resistência mecânica; possui alta capacidade de trocas catiônicas (Santos, 1975).

Durante o processo de queima da cerâmica os argilominerais apresentam transformações múltiplas em suas estruturas cristalinas. A caulinita, por exemplo, perde a água adsorvida nas superfícies dos cristais abaixo de 100 °C (desidratação da caulinita), continuando o aquecimento até 518 °C ocorre o processo de desidroxilação da matriz (perda de água interbasal), o que a transforma em metacaulinita (Frost & Vassallo, 1996; Okada et al., 1998; Murad & Wagner, 1991). Acima de 800 °C (com pico 970°C na curva de DTA) praticamente toda a água é perdida, resultando na nucleação do argilomineral denominado de mulita (Rice, 1987). Através do levantamento da composição mineralógica é possível estimar a faixa de temperatura que foi realizada a queima da cerâmica.

A seleção da argila para produção cerâmica requer uma avaliação de sua plasticidade para a formação do manufaturado. Em todos os processos de seleção das argilas é necessário encontrar um balanço entre o percentual de argila, que fornece a plasticidade, e o percentual de inclusões não-plásticas, os quais fornecem importantes propriedades mecânicas ao produto final.

A inclusões não-plásticas, denominadas de temperos, podem ser identificados por diversas técnicas de caracterização mineralógica. De acordo com a distribuição dos tamanhos dos grãos e suas formas pode ser inferido se estas inclusões foram adicionadas intencionalmente ou se são componentes intersticiais da própria matriz argilosa.

A investigação das composições químicas e mineralógicas dos componentes da cerâmica, argilas e temperos, desempenha um papel crucial para o entendimento das culturas que a utilizaram (Tite, 1999).

TÉCNICAS ESTATÍSTICAS

Estudo das Amostras Discrepantes (Outliers)

Os outliers são observações constantes em uma base de dados que diferem significativamente da maioria das observações, ou seja, são observações caracterizadas pela sua discrepância relativa às demais.

Estes valores discrepantes podem ser gerados por diferentes mecanismos, tais como: erros de medição, variabilidade populacional do atributo medido, erros humanos, instrumentos defeituosos, entre outros processos. A presença de amostras discrepantes pode conduzir a falsas estimativas e interpretações equivocadas. Independente de suas causas, o estudo dos outliers é realizado basicamente em três fases: detecção, testes para confirmação e destino final (Beckman & Cook, 1983).

Na literatura há poucos trabalhos sobre identificação de valores discrepantes em amostras que envolvem mais de uma variável. A maioria dos métodos propostos na atualidade é subjetiva e resume-se a métodos gráficos, tais como dendrograma obtidos por análise de conglomerado. Alguns autores propõem que a distância de Mahalanobis (D_i) é eficiente como método de detecção de amostras discrepantes em dados multivariados (Baxter, 1999a). Considerando uma base de dados com n amostras e p variáveis medidas, a distância de Mahalanobis é dada por

$$D_i = \sqrt{(X_i - \bar{X})^T S^{-1} (X_i - \bar{X})} \quad , \quad (3)$$

onde,

$$S = \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})^T / n - 1$$

X_i é o vetor de observação a i-ésima amostras

$\bar{X}$ é o vetor de média

S é a matriz de variância-covariância amostral

A identificação de valores discrepantes por meio de D_i é feita calculando-se D^2 para cada grupo de amostras e efetuando-se teste de hipóteses por meio da comparação deste valor com o valor crítico. Na literatura, tem sido sugerido que a determinação dos valores críticos da distância de Mahalanobis deve ser efetuada por meio da distribuição F, especialmente, para amostras de tamanho pequeno (Penny, 1987). Wilks sugeriu que o valor crítico para a distância de Mahalanobis é dado por

$$\frac{p(p-1)^2 F_{p;n-p-1;\alpha/n}}{n(n-p-1+pF_{p;n-p-1;\alpha/n})} \quad , \text{ caso } D^2 \text{ seja maior que este valor}$$

crítico a ocorrência é considerada outliers (Wilks, 1963). A variável denominada F na equação anterior é obtida a partir da distribuição de probabilidades denominada de distribuição F.

Neste trabalho a detecção de observações discordantes foi efetuada por meio da distância de Mahalanobis, sendo que o valor crítico será obtido por meio do critério de Wilks. A identificação destas observações foram importante para eliminar a possibilidade de erros nas interpretações.

Normalização e Padronização dos Dados

As inferências estatísticas consistem em generalizar informações obtidas a partir dos espaços amostrais para uma ou mais populações. Nos procedimentos inferenciais, nas estimativas e na obtenção dos intervalos de confiança admite-se, geralmente, que as médias amostrais são normalmente distribuídas, independente da forma da distribuição original da população.

A normalidade multivariada dos dados é requerida, em muitos casos, visto que muito dos métodos estatísticos multivariados, incluindo

análise de variância, análise fatorial com extração por função de verossimilhança, correlação canônica, entre outros métodos, assumem que os dados seguem este tipo de distribuição (Baxter, 1999b).

Segundo Mecklin & Mundfrom, a identificação da normalidade multivariada pode ser obtida por meio de gráficos e coeficientes de correlação (Mecklin & Mundfrom, 2004). Dentro deste ponto de vista o gráfico dos percentis (Q-Q plot) tem sido utilizado com bastante frequência. O procedimento gráfico consiste em construir o gráfico dos percentis amostrais em função dos percentis da distribuição normal padronizada. A normalidade é verificada por meio da avaliação da qualidade do ajuste da reta dos mínimos quadrados aos pontos do gráfico. Apesar da praticidade do método, o mesmo não constitui um teste formal para verificar desvio de normalidade multivariada. Para contornar esta limitação, diversos autores têm proposto métodos formais para verificação de normalidade por meio de estatísticas que utilizam o vetor de média, coeficiente de assimetria e curtose (Mardia, 1974; Mardia & Kent, 1991).

No estudo da composição química da cerâmica tem sido observado por diversos autores que os elementos analisados se distribuem log – normalmente. (Glascok, 1992) e colaboradores verificaram que o tratamento dos dados como uma distribuição log – normal é mais viável (Glascok, 1992), por duas razões básicas: primeiro, porque tem sido observado que diversas concentrações elementares são normalizadas quando tomado os seus valores logarítmicos. A segunda, é que a transformação dos dados compensa as diferenças de magnitudes das variáveis medidas.

Visando a padronização e normalização dos dados, neste trabalho a interpretação dos dados foi efetuada em uma escala logarítmica. Esta transformação viabilizou a aplicação das técnicas de agrupamentos das amostras.

Analise de Conglomerados

A análise de conglomerados (“Cluster Analysis”) é uma técnica estatística multivariada utilizada para produzir padrões de comportamento em bancos de dados, por meio da formação de grupos homogêneos de casos. O objetivo principal da técnica é agrupar objetos semelhantes de acordo com suas características.

Para formação dos grupos é necessário adotar critérios bem definidos. Um critério razoável para formação dos grupos é considerar a proximidade entre os pontos no espaço p -variado, visto que pontos que estão próximos representam regiões cujas amostras são similares. Logo, esta técnica exige a definição de um coeficiente de parença para indicar a proximidade entre as amostras.

Na literatura estatística são citadas duas medidas de parença: medidas de similaridades (quanto maiores os valores mais similares são os objetos) e medidas de dissimilaridades (quanto maior, menor a similaridade entre os objetos) (Johnson & Wichern, 1992). A partir da definição do critério de parença, é construída uma matriz de parença como ponto de partida do método de agrupamento.

Os elementos da matriz de similaridade ou dissimilaridades são determinados, geralmente, a partir das diversas medidas de distância (Massart & Kaufman, 1983). Entre as medidas de distância, a mais utilizada é a distância Euclidiana.

Após construção da matriz de parença, o passo seguinte é optar por um algoritmo de agrupamento. Há diversos algoritmos para formação dos grupos, os quais de forma geral podem ser classificados em métodos hierárquicos e de partição ou não-hierárquicos (Jolliffe et al., 1995).

Nos métodos hierárquicos os agrupamentos são formados a partir da matriz de parença de acordo com a seqüência: Inicialmente os n objetos formam n classes; Em seguida as observações mais semelhantes são agrupadas em uma mesma classe, resultando em $n-1$ classes; Este processo continua até que todas as amostras pertençam a uma única classe.

As regras de reconstrução da matriz de parença a cada formação de classe determinaram a homogeneidade dos grupos formados. Entre os métodos mais comuns para este fim estão: método do vizinho mais próximo, método do vizinho mais longe, método das médias das distâncias, métodos dos centróides e método de Ward.

Os métodos de agrupamentos não-hierárquicos buscam definir partições de n objetos em k grupos de acordo com algum critério estabelecido previamente. A prefixação dos critérios está associada à produção de medidas referentes à qualidade da partição produzida. Entre os métodos de partição os mais conhecidos são o método das k -médias e método dos k -medóides (Kaufman & Rousseeuw, 1990).

Uma vantagem dos métodos hierárquico em relação aos métodos não-hierárquico é que não requerem o conhecimento *a priori* do núme-

ro de grupos que serão formados. Todavia, as introduções dos métodos não-hierárquicos permitem uma verificação da significância da alocação de um dado objeto em um grupo. Assim, em muitos casos remenda-se a adoção de um método hierárquico para determinação do número inicial de grupos e posteriormente aplicar um dos métodos não-hierárquico.

Nas fases finais de aplicação da técnica de conglomerado os agrupamentos podem ser representados graficamente por meio dos dendrogramas. Os dendrogramas apresentam os elementos e os respectivos pontos de fusão ou divisão dos grupos formados em cada estágio. A inspeção visual dos dendrogramas permite a identificação dos grupos. Apesar de sua simplicidade a análise de conglomerado tem sido utilizada com bastante sucesso como ponto de partida para outras técnicas estatísticas mais refinadas e que produzam grupos mais homogêneos e de fácil identificação.

Análise de Componentes Principais

Uma questão importante na análise composicional da cerâmica está relacionada à representação dos dados. Como resultado direto do grande número de variáveis observadas para cada amostra, a representação dos dados é dificultada. Tendo em vista esta dificuldade, em muitos casos é necessário recorrer a técnicas estatísticas multivariadas no sentido de explicar a estrutura de dados em termos de poucas variáveis. Uma das técnicas mais utilizadas com este objetivo é a técnicas de componentes principais.

A análise de componentes principais é uma técnica que transforma linearmente um conjunto de p variáveis observadas em um conjunto com um número menor (k) de variáveis não correlacionadas e que explicam uma parcela substancial da estrutura de covariância dos dados (Jolliffe, 1989). As p variáveis transformadas ($Y_1, Y_2, \dots, Y_p$) a partir das variáveis originais são denominadas de componentes principais. As componentes principais estão ordenadas de forma que a primeira componente (Y_1) explique a maior parcela da variabilidade, (Y_2) a segunda maior parcela e assim sucessivamente. Apesar desta transformação gerar um número de componentes igual ao número de variáveis originais, na aplicação da técnica retém-se k componentes ($k < p$) para explicação de toda variabilidade dos dados. Assim, de uma forma geral a análise de componentes principais tem os objetivos de reduzir a dimensionalidade

dos dados, obter combinações interpretáveis das variáveis, descrever e entender a estrutura de correlação entre as variáveis observadas.

Algebricamente, as componentes principais são combinações lineares das variáveis originais. Geometricamente, as componentes principais representam as coordenadas dos pontos amostrais em um sistema de eixos ortogonais obtidos pela rotação do sistema de eixos originais em direção das máximas variabilidades (Figura 2).

Figura 2 - Representação das variáveis em termos de componentes principais (CP e CP)

É importante citar, que a transformação da matriz de dados em função das componentes principais não altera a variância total da estrutura de dados. De acordo com esta consideração a proporção da variância total devido ao i-ésimo componente principal é dada por

$$\frac{\lambda_i}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \quad i=1,2,\dots,p. \quad (4)$$

onde λ_i é o autovalor associado ao i-ésimo componente.

Em diversas situações a aplicação da análise por componentes principais permite que 70% ou mais da variância total seja explicada pelas primeiras k componentes. Sendo assim, estes primeiros k componentes podem substituir as p variáveis originais sem perda significativa de informações.

Em análise de componentes principais uma questão freqüente é saber o número de componentes devem ser retidos para explicar toda

estrutura de dados. Os aspectos que devem ser considerados são as quantidades de variâncias explicadas, os tamanhos relativos dos autovalores e a interpretação subjetiva das componentes. Alguns critérios tais como critério de Kaiser, screen plot, entre outros critérios podem auxiliar na escolha do número de componentes (Kaiser, 1958).

No estudo composicional da cerâmica a técnica de componentes principais é extremamente útil, visto que as modernas técnicas de análise fornecem um grande número de variáveis para as amostras, e na maioria dos casos estas estão correlacionadas. A composição de cada espécie original pode ser convertida em seus escores principais tornando-se mais facilmente interpretáveis. Vários pesquisadores descrevem que no estudo da cerâmica cerca de 70% ou mais da variância total dos dados é explicada em termos das três primeiras componentes principais. Desta forma, por meio das componentes principais é possível realizar os agrupamentos das amostras de uma forma mais simples em virtude da redução da dimensionalidade dos dados.

Análise Discriminante

A análise discriminante é uma técnica estatística multivariada utilizada com o objetivo de discriminar populações e/ou classificar objetos em populações previamente definidas. Os principais objetivos da técnica são encontrar funções das variáveis originais (funções discriminantes) que expliquem as diferenças entre as populações e que permitam alocar novos objetos em uma das populações envolvidas na análise.

Diferentemente da análise de agrupamento, a análise discriminante é uma técnica supervisionada, pois neste tipo de análise há necessidade do conhecimento *a priori* das populações às quais pertencem os objetos. Para aplicação da análise discriminante as g populações devem ser bem definidas. Estas características diferem da análise de agrupamento visto que nesta técnica não se conhecem *a priori* quais as populações envolvidas.

Para determinação das funções discriminantes é utilizado com muita frequência o método de Fisher (Poston & Marchette, 1998). Este método consiste em obter novos eixos a partir de combinações lineares das variáveis originais que permitam diferenciar significativamente as populações. Com este propósito, busca-se a combinação linear $y = 1^T X$ (onde X é o vetor de observação) que maximize as distâncias entre os

vetores de média de cada população (μ_i) e minimize as variâncias internas.

Considerando que as matrizes de covariâncias das populações ($\Sigma_1, \Sigma_2, \dots, \Sigma_g$) são iguais, ou seja $\Sigma_1 = \Sigma_2 = \dots = \Sigma_g = \Sigma$, a combinação linear para discriminação das populações segundo o método Fisher implica na maximização de $\frac{B^T \Sigma^{-1} B}{B^T B}$, onde $B = \sum_{i=1}^g (\mu_i - \bar{\mu})(\mu_i - \bar{\mu})^T$ e $\bar{\mu}$ é o vetor de média das médias.

Os coeficientes (β_i) da primeira função discriminante são os elementos do autovetor padronizado de $\Sigma^{-1} B$, associado ao maior autovalor dessa matriz. O vetor ($\tilde{\beta}$) que maximiza a razão referida no parágrafo anterior, sujeita à restrição $\tilde{\beta}^T \tilde{\beta} = 1$, é o vetor padronizado associado ao segundo maior autovalor de $\Sigma^{-1} B$, sendo que a combinação linear resultante é chamada de segunda função discriminante. Assim, a k-ésima função discriminante é $\tilde{\beta}_k^T X$, onde $\tilde{\beta}_k$ é o correspondente k-ésimo autovetor padronizado sujeito à condição $\tilde{\beta}_i^T \tilde{\beta}_k = 0$ com $i < k$. Como $\tilde{\beta}_i$ e $\tilde{\beta}_k$ são, em geral, desconhecidos estes são substituídos por suas estimativas na análise discriminante.

Nos estudos Arqueométricos a análise discriminante tem sido aplicada para estudos de proveniência das fontes de matérias-primas e agrupamentos de vestígios cerâmicos conforme suas similaridades. A grande vantagem desta técnica em relação às técnicas de conglomerados é que fornece grupos individuais mais homogêneos facilitando as observações dos grupos.

PROCEDIMENTO EXPERIMENTAL

Preparação das Amostras

Inicialmente, cerca de 2-3 mm das superfícies externas dos fragmentos cerâmicos foram extraídos com uma lima rotativa de carbeto de tungstênio, adaptada a uma furadeira, cuja velocidade de rotação é regulável através de potenciômetro. Após este procedimento, extraírem-se cerca de 500 mg de amostra, na forma de pó, através da perfuração

de orifícios transversais, com broca de carbeto de tungstênio, distribuídos uniformemente pelo fragmento. O pó obtido foi secado em estufa à temperatura de 105 °C por 24 h e armazenado em dessecador (Munita et al., 2004). O material de referência Coal Fly Ash (NIST – SRM – 1633b) foi utilizado como padrão e os materiais Brinck Clay (NIST – SRM-679) e IAEA – SOIL – 7 foram utilizados para verificar a qualidade analítica dos resultados. Estes materiais foram secados a 105°C por 2 horas e postos em dessecador até a pesagem.

Procedimentos Analíticos e Medidas

Análises das Cerâmicas

Cerca de 100 mg de amostras e padrões, na forma de pó, foram pesados em envelopes de polietileno, e submetidos à irradiação sob um fluxo de nêutrons da ordem de $10^{12} \text{ n.cm}^{-2}.\text{s}^{-1}$ durante 8 h no reator IEA-R1m do IPEN-CNEN/SP.

As amostras irradiadas foram submetidas à espectrometria gama após um tempo resfriamento de 7 dias para determinação das concentrações de As, Ba, K, La, Na, Nd, Sm e Yb. A segunda medida foi realizada após 30 dias para determinação das concentrações de Co, Cr, Cs, Eu, Fe, Hf, Rb, Sb, Sc, Ta, Tb, Th, U e Zn (Glascock & Neff, 2003).

As medidas da radiação gama das amostras foram realizadas no espectrômetro de raios gama constituído por um detector de Ge hiperpuro, modelo GX 2020, da Canberra, com resolução de 1,90 keV no pico de 1332 keV do ^{60}Co , acoplado a um analisador multicanal, constituído de 8192 canais. As determinações das concentrações elementares foram realizadas por meio do programas Genie -2000 Neutron Activation Processing Procedure da Canberra.

RESULTADOS E DISCUSSÕES

Controle de Qualidade Analítico

Um requisito básico para a caracterização composicional da cerâmica arqueológica é que a técnica analítica empregada apresente boa

precisão. Elementos que são determinados com baixa precisão podem interferir, significativamente, na interpretação dos resultados. Estas interferências reduzem, na maioria dos casos, os efeitos discriminantes da composição química da pasta cerâmica utilizada na manufatura das peças (Gouveia & Prudêncio, 2000). Para verificação da precisão e exatidão da metodologia a ser empregada neste trabalho foram realizadas 10 determinações para os materiais de referência Brick Clay (NIST – SRM-679) e IAEA – SOIL – 7.

As concentrações elementares dos materiais de referência foram estatisticamente comparadas com os valores certificados. Para os resultados preliminares apresentados nesta seção, diversos elementos apresentaram precisão abaixo de 5% para os materiais de referência analisados (Tabelas 1 e 2) concordando com valores obtidos por outros autores (Ni et al., 1995). Alguns elementos apresentaram CV maiores que 10% (Tabelas 1 e 2), valores comparáveis com outros trabalhos apresentados na literatura (Kuleff & Dingova, 1990).

O coeficiente de variação para os elementos Ba, K, Nd, Sm, Ta, Tb e U no material de referência foram maiores que de 10% no Brick Clay (NIST – SRM-679). Esses elementos, com exceção do Ba apresentaram o mesmo comportamento no IAEA – SOIL – 7. Para os elementos Ba, K, Sm e Nd este comportamento pode ter sido provocado pelo grande número de “missing values” presentes nas análises, que pode ser reduzido pela otimização do tempo de resfriamento. O elemento tântalo apresenta-se em concentrações muito baixas e possui uma interferência importante no pico de 1221,3 keV com o pico de 1120,5 keV do Sc, resultando na redução de sua precisão. Finalmente, as determinações do U e Tb são prejudicadas em virtude da baixa concentração nos materiais de referência.

Tabela 1 – *Resultados obtidos para o material de referência Brinck Clay. Associados aos valores determinados estão os intervalos de confiança.*(valores em ppm a não ser quando indicado)

Elem.	V. Determinados	V. referência	CV (%)	ER (%)
Ba	549,16 ± 204,99	432,2 ± 9,8	42,59	16,10
Ce	107,91 ± 2,22	(105)	2,97	4,77
Co	26,73 ± 0,43	(26)	2,30	0,10
Cr	108,54 ± 4,15	109,7 ± 4,9	5,52	0,50
Cs	9,94 ± 0,35	(9,6)	5,06	2,92
Eu	1,87 ± 0,06	(1,9)	4,12	10,65
Fe, %	9,16 ± 0,15	9,05 ± 0,21	2,31	1,44
Hf	4,46 ± 0,23	(4,6)	7,30	2,43
K, %	2,64 ± 0,67	2,433 ± 0,047	28,74	8,35
La	56,28 ± 2,03	(56,4)	5,21	0,21
Lu	0,59 ± 0,04	(0,615)	9,84	4,27
Na, %	0,15 ± 0,01	0,1304 ± 0,0038	12,42	13,11
Nd	36,02 ± 7,39	(46,7)	29,61	22,87
Rb	186,47 ± 11,20	(190)	8,67	14,46
Sc	23,33 ± 0,38	(22,5)	2,37	1,42
Sm	6,36 ± 2,49	(9,16)	56,36	30,55
Ta	1,34 ± 0,15	(1,24)	16,96	7,76
Tb	1,30 ± 0,24	(1,21)	27,14	7,54
Th	14,36 ± 0,31	(14)	3,14	0,45
U	3,03 ± 0,50	(2,41)	23,89	25,73
Yb	3,89 ± 0,18	(4,11)	6,82	5,41
Zn	123,58 ± 7,07	(150)	8,25	4,20
n = 10				

Tabela 2 – *Resultados obtidos para o material de referência IAEA – SOIL – 7. Associados aos valores determinados estão os intervalos de confiança.*
(valores em ppm a não ser quando indicado)

Elem.	V. Determinados	V.Referência	CV(%)	ER(%)
Ba	184,39 ± 4,00	(159 ± 32,5)	2,21	15,97
Ce	59,47 ± 2,38	61 ± 6,5	5,40	2,51
Co	9,08 ± 0,29	8,9 ± 0,85	4,30	1,98
Cr	70,75 ± 0,92	60 ± 12,5	1,75	0,92
Cs	5,76 ± 0,25	5,4 ± 0,75	5,91	6,60
Eu	1,04 ± 0,05	1,0 ± 0,2	7,10	4,25
Fe, %	2,67 ± 0,07	(2,57 ± 0,06)	3,65	3,93
Hf	5,07 ± 0,24	(5,1 ± 0,35)	6,32	0,54
K, %	1,19 ± 0,17	(1,21 ± 0,07)	19,46	1,76
La	31,28 ± 0,95	28 ± 1	4,10	11,73
Lu	0,38 ± 0,02	(0,3 ± 0,2)	6,64	27,08
Na, %	0,24 ± 0,01	(0,24 ± 0,01)	3,38	1,56
Nd	25,88 ± 3,86	30 ± 6	20,16	13,75
Rb	46,61 ± 6,20	51 ± 5	17,96	8,60
Sc	8,68 ± 0,22	8,3 ± 1,1	3,44	4,61
Sm	5,58 ± 1,53	5,1 ± 0,4	19,68	4,47
Ta	0,64 ± 0,09	0,8 ± 0,2	18,29	20,31
Tb	0,65 ± 0,19	0,6 ± 0,2	8,70	8,75
Th	8,26 ± 0,25	8,2 ± 1,1	4,17	0,76
U	2,48 ± 0,36	2,6 ± 0,6	19,68	4,47
Yb	2,33 ± 0,06	2,4 ± 0,4	3,50	2,76
Zn	103,89 ± 7,85	104 ± 6	10,20	0,10
n = 10				

Para avaliar a exatidão das análises, os desvios das concentrações obtidas em relação aos valores certificados estão mostrados nas Figuras 3 e 4. Pode –se verificar que os maiores desvios são atribuídos aos elementos K, Sm e Nd para Brick Clay (NIST – SRM-679), e Sm, Lu e La para IAEA – SOIL – 7.

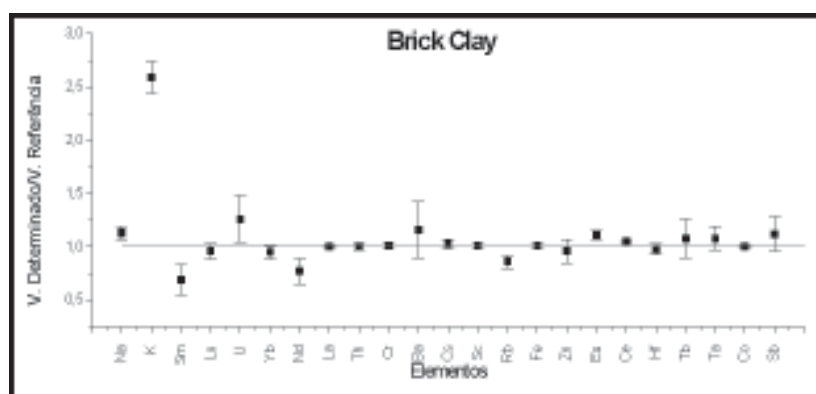


Figura 3 – Avaliação dos desvios das concentrações obtidas em relação aos valores certificados para Brinck Clay (NIST – SRM-679).

$$S = \sum_{i=1}^n (x_i - \bar{x}) (x_i - \bar{x})^T / n - 1$$

Figura 4 – Avaliação dos desvios das concentrações obtidas em relação aos valores certificados para IAEA – SOIL – 7.

Estudo da Normalidade dos Dados

Inicialmente construíram-se os histogramas das concentrações elementares para observar a distribuição dos valores obtidos em torno de um valor médio. Tem sido apontado por alguns autores que a distribuição dos dados para material geológico, geralmente, seguem uma distribuição lognormal (Ahrens, 1954).

Na Figura 5 é mostrado um exemplo da distribuição das concentrações para tântalo. Pode-se observar que os dados apresentam uma forte assimetria positiva, que é característico da distribuição lognormal, todavia quando são tomadas as distribuições logarítmicas das concentrações dos dados observa-se que a curva normal ajusta-se muito bem ao histograma. Este comportamento é desejável, primeiro porque os elementos traços aparecem mais normalmente distribuídos, em amostras geológicas, quando são tomados os logaritmos das concentrações. A segunda razão reside no fato que as transformações das concentrações na forma logarítmica compensa as diferenças de magnitudes entre os elementos maiores, tais como K e Fe, e os elementos traços, tais como terras raras, entre outros.

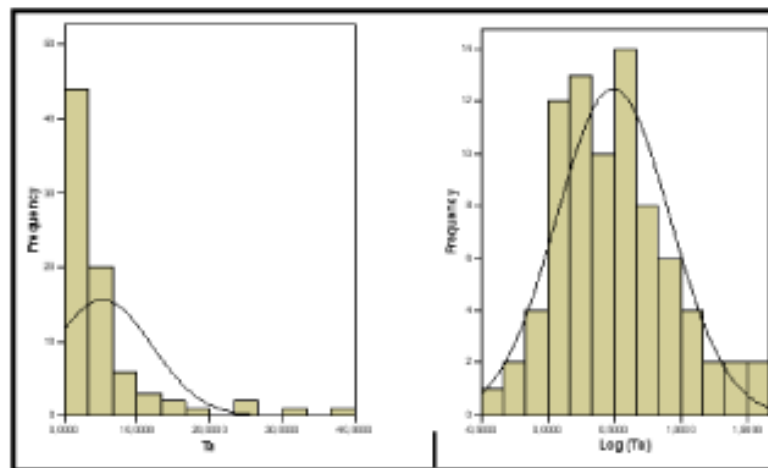


Figura 5 – Distribuição das concentrações para Ta .

Apesar das distribuições das concentrações elementares apresentarem uma distribuição lognormal univariada, um teste formal, o qual constitui uma extensão do teste de Shapiro – Wilks para o caso multivariado, mostrou que a distribuição conjunta dos dados não se ajusta a uma normal multivariada. Entretanto, este resultado não configura um problema crítico para interpretação dos dados, visto que é mostrado teoricamente que se as variáveis individuais forem normais, a conjunta não será, necessariamente, normal multivariada. Mas a recíproca é verdadeira, ou seja, se a distribuição dos dados é normal multivariada, as variáveis individuais terão uma distribuição gaussiana no espaço univariado. Além deste fato, a maioria das técnicas estatísticas para agrupamento é robusta ao tipo de distribuição dos dados.

Identificação das Amostras Discrepantes

Após transformação logarítmica, os dados foram submetidos a uma rotina com base no critério lambda Wilks, com significância de 5%, para identificação das amostras discrepantes.

Considerando a distância Mahalanobis e o lambda Wilks, 7 amostras do sítio São José e 4 amostras do sítio Saco da Onça foram consideradas discrepantes em relação aos seus respectivos grupos. As amostras discrepantes foram eliminadas da base de dados, obtendo-se uma matriz de dados livre de ruídos estatísticos.

Agrupamentos das Amostras dos Sítios São José, Curitiba e Saco da Onça

Os sítios analisados e interpretados neste trabalho, com objetivo de exemplificar a aplicação da arqueostatística, foram os sítios São José, Curitiba e Saco da Onça, localizados no Noroeste do estado de Sergipe, na divisa com o estado de Alagoas (Figura 6). As pesquisas arqueológicas nesta área foram iniciadas por volta da década de 1960 e intensificadas devido à construção do complexo Hidrelétrico de Xingó. Em virtude do alagamento da região para compor a represa da hidrelétrica, foi necessário implementar um projeto de salvamento arqueológico, visto que na região foi detectado um enorme volume de vestígios de grupos pré – históricos (Relatório – MAX, 1994).

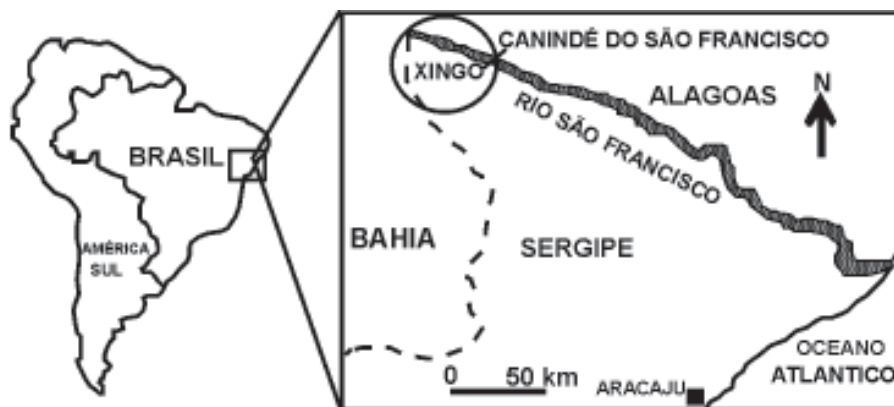


Figura 6: Mapa de localização da área estudada. A área de Xingó está localizada no Nordeste brasileiro, entre os estados de Sergipe, Alagoas e Sergipe.

O salvamento arqueológico efetuado na área que foi inundada pela barragem permitiu a identificação, sondagem e escavação de 28 sítios classificados como de acampamento; 11 classificados como habitação e 2 considerados como de habitação e enterramento (São José e Justino), que acabam sendo os mais importantes. A importância dos sítios Justino e São José é reconhecida em virtude das suas cronologias diferenciadas em relação a outras tradições ceramistas estabelecidas no Nordeste brasileiro, e em virtude da associação do material resgatado nestes sítios com todo um contexto ritualístico e funerário (Vergne, 1997).

Estudos baseados na tipologia cerâmica mostraram que apresentam o mesmo perfil técnico e morfológico. Segundo estes estudos, os grupos que habitaram esta área eram constituídos de caçadores-coletores e pescadores, apresentando uma forma de agricultura incipiente (Luna, 1997). A cerâmica era utilizada basicamente para fins de armazenamento e aquecimento de alimentos, bem como urnas fúnebres. De acordo com alguns autores, os grupos que habitaram a região caracterizam uma tradição ceramista independente, não filiada às tradições Tupi e Aratu, amplamente difundidas no Nordeste do Brasil (Martin, 1997).

A importância arqueológica dos sítios ceramistas que foram estudados neste trabalho deve-se às suas localizações, e volumes de material cerâmico encontrado.

Para avaliar as similaridades e dissimilaridades químicas entre as amostras provenientes dos sítios São José, Saco da Onça e Curituba,

foram determinadas as funções discriminantes a partir da matriz das concentrações padronizadas de acordo com uma transformação logarítmica. Na Figura 7 são mostrados os escores discriminantes para as amostras provenientes dos três sítios estudados neste trabalho, onde é possível observar uma clara distinção entre as composições químicas elementares das cerâmicas. Na Figura 7, fica evidente a dispersão das amostras dos sítios São José. De fato, pode ser visto na Tabela 3 que as amostras do sítio São José são caracterizadas por uma alta dispersão (%CV) em relação aos elementos Cr, Lu, Rb e Yb.

analisadas. As elipses representam um nível de confiança de 95%

Na Tabela 3, as amostras identificadas como pertencentes ao sítio Saco da Onça apresentaram maiores concentrações em relação a Th e menor concentração de Cr, todavia as amostras do sítio Curitiba apresentaram maiores concentrações em relação aos elementos Ce, Sc, Fe e Yb. As amostras do sítio São José apresentaram concentrações intermediárias em relação aos dois outros sítios. Do ponto de vista geológico e arqueológico, a maior dispersão da composição química do sítio São José pode ser um resultado da mistura da argila utilizada para produção da pasta cerâmica com antiplásticos.

Tabela 3. Estatística descritiva dos resultados por AANI para as amostras cerâmicas analisadas (resultados em $\mu\text{g.g}^{-1}$ a não ser quando indicado).

Elem. Saco da Onça (n = 21) São José (n = 18) Curitiba (n = 24)

	Média	Desvio	%CV	Média	Desvio	%CV	Média	Desvio	%CV
Ce	81.23	37.38	46.51	90.22	21.83	24.20	100.18	31.02	30.96
Cr	24.08	22.70	94.29	53.75	58.12	08.13	50.07	47.11	94.10
Cs	4.18	3.74	89.30	4.00	2.55	63.67	4.21	1.72	40.85
Eu	1.31	0.35	26.76	1.76	0.34	19.15	1.99	0.26	13.10
Fe, %	2.93	0.90	30.75	3.88	0.96	24.78	5.19	1.69	32.95
Hf	5.31	1.92	36.09	5.65	1.85	32.80	6.38	1.52	23.83
La	40.65	19.65	48.35	45.97	17.26	37.54	49.87	13.23	26.54
Lu	0.35	0.08	24.28	0.36	0.16	45.17	0.50	0.05	10.46
Na, %	1.00	0.43	43.02	1.02	0.27	26.65	1.20	0.36	29.79
Rb	90.33	62.15	68.81	77.74	49.14	63.21	62.89	24.91	53.07
Sc	8.58	2.97	34.63	11.32	4.11	36.30	17.06	6.00	35.14
Th	15.78	10.66	67.59	11.02	6.92	62.82	10.31	3.68	35.69
Yb	2.26	0.70	30.90	2.43	1.09	44.59	3.58	0.45	12.49

A despeito da proximidade geográfica entre os três sítios estudados, a análise discriminante mostrada na Figura 7, indicou que a pasta cerâmica proveniente dos sítios apresentaram diferenças significativas. Estes resultados podem ser consequência de alterações no uso dos espaços, organização da produção cerâmica e diferenciações sociais durante a ocupação da área, as quais levaram o ceramista pré-histórico a procurar diferentes matérias-primas para manufatura da cerâmica. Os estudos arqueológicos sugerem que a tipologia dos vasos cerâmicos provenientes dos sítios São José, Saco da Onça e Curitiba são similares, pode-se inferir que o ceramista de Xingó fazia sua opção, em relação ao uso de uma determinada argila, de acordo com a disponibilidade nas áreas próximas ao local que habitava. Ainda, durante a análise discriminante, obteve-se que a amostra argilosa coletada nas proximidades do sítio São José não pertence a qualquer um dos sítios estudados, portanto, pode-se inferir que este depósito argiloso não foi utilizado para obtenção dos produtos cerâmicos produzidos nestes três sítios.

CONCLUSÕES

As concentrações elementares apresentaram distribuição lognormal. A normalidade logarítmica é desejada para os estudos de proveniência, tendo em vista que esta normalização permite reduzir os efeitos das diferentes magnitudes das concentrações determinadas.

A qualidade analítica das concentrações multielementares determinadas por ativação com nêutrons foi verificada por meio das análises dos materiais de referência Brinck Clay (NIST – SRM-679) e IAEA – SOIL – 7. O valor obtido para cada elemento foi comparado com os valores certificados. Vários elementos apresentaram CV menor que 10%; entretanto, Ba, K, Nd, Sm, Ta, Tb e U apresentaram CV maior que 10% no material de referência Brinck Clay (NIST – SRM-679). Estes mesmo elementos, com exceção do Ba, apresentaram o mesmo comportamento para o material IAEA – SOIL – 7.

A identificação dos grupos composicionais de cerâmicas arqueológicas por meio da análise química elementar, por ativação neutrônica, foi realizada com êxito, apesar da proximidade geográfica entre os sítios estudados. Foi possível verificar que as amostras provenientes do sítio São José são caracterizadas por uma alta dispersão composicional, a qual pode ser um resultado da influência do antiplástico presente na pasta cerâmica. As diferenças composicionais entre as cerâmicas a partir dos três sítios podem ser entendidas em termos da influência cultural no preparo da pasta, alterações na organização e modo de produção e na disponibilidade da matéria-prima durante as sucessivas ocupações dos sítios. A interpretação arqueostatística dos dados por meio da análise discriminante permitiu, ainda, inferir que a argila coletada nas proximidades do sítio São José não foi utilizada no preparo das cerâmicas. Uma coleta sistematizada de argilas na região poderia revelar as fontes de matéria-prima utilizada na manufatura da cerâmica, todavia, na atualidade, os sítios estão submersos, em virtude da construção do reservatório que compõe a hidrelétrica de Xingó. Desta forma, os resultados obtidos neste trabalho contribuem para a formação de uma conjuntura que permita inferir, com fundamentação científica, a respeito da ocupação da região e da evolução tecnológica dos grupos que habitaram Xingó na pré-história.

REFERÊNCIAS

- AHRENS, L. H. *The lognormal distribution of the elements*. *Geochimica et cosmochimica acta*, 6 (2-3), 121 - 131, 1954.
- ALAMIN, M. B., SPYROU, N. M. *Semi-empirical determination of detector absolute efficiency in AANI of voluminous samples*. *Journal of Radioanalytical and Nuclear Chemistry*, 215 (2), 205 - 209, 1997.
- BALLA, M., MOLNAR, Z., KOROS, A. *Uncertainty budget and validation of NAA using reference materials*. *Journal of Radioanalytical and Nuclear Chemistry*, 259 (3), 395 - 400, 2004.
- BAXTER, M. J. *Detecting multivariate outliers in artefact compositional data*. *Archaeometry*, 41 (2), 321 - 338, 1999a.
- BAXTER, M. J. *On the multivariate normality of data arising from lead isotope fields*. *Journal of Archaeological Science*, 26, 117 - 124, 1999b.
- BEAUDRY, M.P. New world paste compositional investigations. In: BISHOP, R.L., LANGE, F.W. *The ceramic legacy of A.O. SHEPARD*. Univ. Press of Colorado, 224-256, 1991.
- BECKMAN, R. J., COOK, R. D. *Outlier*. *Technometrics*, 25 (2), 119 - 163, 1983.
- BISHOP, R. L., BLACKMAN, M. J. *Instrumental Neutron Activation Analysis of Archaeological Ceramics: Scale and Interpretation*. *Acc. Chem. Res.*, 35, 603 - 610, 2002.
- BORONE, G., IOPPOLO, S., MAJOLINO, D., MIGLIORDO, P., TIGANO, G. *A multidisciplinary investigation on archaeological excavation in Messina (Sicily), a comparison of pottery findings in the "The Strati of Messina area"*. *Journal of Cultural Heritage*, 3(2) 145-153, 2002.
- CHESF. Relatório de Operação de Resgate da UHE – Xingó, 1994.
- FIELLER, N. R. J. *Archaeostatistics: old statistics in ancient contexts*. *The Statistician*, 42 (3), 279 - 295, 1993.
- FROST, R. L., VASSALLO, A. M. *The dehydroxylation of the kaolinite clay minerals using infrared emission spectroscopy*. *Clays And Clay Minerals*, 44 (5), 635 - 651, 1996.
- GLASCOCK, M. D, NEFF, H. *Neutron activation analysis and provenance research in archaeology*. *Measurement Science & Technology*, 14 (9), 1516-1526, 2003.
- GLASCOCK, M.D. *Characterization of ceramics at MURR by NAA and Multivariate Statistics*. In: *Chemical Characterization of Ceramic Paste in*

Archaeology, Monographs in World Archaeology, H. NEFF, Prehistory Press, Section 1, 11, 1992.

GOUVEIA, M. A., PRUDENCIO, M. I. *New data on sixteen reference materials obtained by AANI*. Journal Of Radioanalytical And Nuclear Chemistry, 245 (1), 105 – 108, 2000.

HEVESY, G., LEVI, H. *Kgl. Danske Videnskab. Selskab. Math. Fys.*, 15 (5), 1936.

HEYDORN, K. *AANI - Application and limitation*. Fresenius Journal of Analytical Chemistry, 337 (5), 498 – 502, 1990.

INTERNATIONAL ATOMIC ENERGY AGENCY. *Practical Aspects of Operating a Neutron Activation Analysis Laboratory*, IAEA-TECDOC-564, Vienna, 1990.

JOHNSON, R. A., WICHERN, D. W. *Applied Multivariate Statistical Analysis* (3^a ed). Ed. Prentice Hall, New Jersey, pp 575 – 581 – 584, 1992

JOLLIFFE, I. T. *Rotation of ill-defined principal components*. Applied Statistics-Journal Of The Royal Statistical Society Series C, 38 (1), 139 – 147, 1989

JOLLIFFE, I. T., JONES, B., B. MORGAN, J. T. *Identifying influential observations in hierarchical cluster analysis*. Journal of Applied Statistics, 22 (1), 1995.

KAISER, H. F. *The varimax criterion for analytic rotation in factor analysis*. Psychometrika, 35, 401 – 415, 1958.

KAUFMAN, L., ROUSSEEUW, P. J. *Finding groups in data: An introduction to cluster analysis*. John Wiley & Sons, New York, 1990.

KULEFF, I., DINGOVA, R. *In: Activation Analysis, Vol.2*. Ed. CRC Press, Boca Raton, Florida, pag. 440, 1990.

KULEFF, I., DJINGOVA, R. *Activation analysis in archaeology*. In: Activation Analysis, edited by Z. Alfassi, CRC Press, vol. 2, pp 431, 1990.

LUNA, S., NASCIMENTO, A. *Os Grupos Ceramistas do Baixo São Francisco*. Cadernos de Arqueologia, doc.12, Universidade Federal e Sergipe, 1997.

MARDIA, K. V. *Applications of some measures of multivariate skewness and kurtosis for testing normality and robustness studies*. Sankhya B, 36, 115–128, (1974).

MARDIA, K. V., KENT, J. T. *Rao score tests for goodness of fit and independence*. Biometrika, 78, 355–363, 1991.

MASSART, D. L., KAUFMAN, L. *The interpretation of analytical chemical data by the use of cluster analysis*. John Wiley & Sons, New York, pp 17 – 37, 1983.

- MECKLIN, C. J., MUNDFROM, D. J. *An appraisal and bibliography of tests for multivariate normality*. International Statistical Review, 72 (1), 123 – 138, 2004.
- MOMMSEN, H., KREUSER, A., WEBER, J. *A method for grouping pottery by chemical composition*. Archaeometry, 30, 47-57, 1988.
- MUNITA, C. S., NASCIMENTO, A., SCHREIBER, S. B., LUNA, S., OLIVEIRA, P. M. S. *Chemical study of some ceramics from Brazilian Northeast*. Journal of Radioanalytical and Nuclear Chemistry, 259(2), 305 – 309, 2004.
- MUNITA, C. S., PAIVA, R. P., OLIVEIRA, P. M. S., MOMOSEA, E. F., PLA, R., MORENO, M., ANDONIE, O., FALABELLA, F., MUNOZ, L., KOHNENKAMP, I. *Intercomparison among three activation analysis laboratories in South América*. Journal Of Trace And Microprobe Techniques, 19 (2), 189 - 197, 2001.
- MUNITA, C.S., PAIVA, R.P., ALVES, M.A., OLIVEIRA, P.M.S., MOMOSE, E.F. *Contribution of neutron activation analysis to archaeological studies*. Journal Of Trace and Microprobe Technique, 18(3), 381 - 387, 2000.
- MURAD, E., WAGNER, U. *Mössbauer spectra of kaolinite, halloysite and firing products of kaolinite – new results and reappraisal of published work*. Neues Jahrbuch Für Mineralogie – Abhandlungen, 162, 281 – 309, 1991.
- NI, B. F., WANG, P. S., TIAN, W. Z. *AANI Of Iaea-331 (Spinach), Srm, For 40 Elements*. Journal of Radioanalytical and Nuclear Chemistry, 196 (2), 387 – 392, 1995.
- OKADA, K., SHIMAI, A., TAKEI, T., HAYASHI, S., YASUMORI, A., MACKENZIE, K. J. D. *Preparation of microporous silica from metakaolinite by selective leaching method*. Microporous And Mesoporous Materials, 21 (4-6), 289 - 296, 1998.
- PENNY, K. I. *Appropriate critical values when testing for a single multivariate outlier by using the Mahalanobis distance*. Applied Statistical, 35, 153 – 162, 1987.
- PILLAY, A. E. *Analysis of archaeological artefacts: PIXE, XRF or ICP-MS?* Journal of Radioanalytical and Nuclear Chemistry, 247(3), 593 – 595, 2001.
- POSTON, W. L., MARCHETTE, D. J. *Recursive dimensionality reduction using Fisher's Linear Discriminant*. Pattern Recognition, 31 (7), 881 – 888, 1998.
- RICE, P. *Pottery Analysis: a Sourcebook*. University of Chicago Press, Chicago, pp 90 – 93, 1987.

SANTOS, P.S. *Tecnologia de argilas*. Ed. Edgar Blücher, São Paulo, pág. 802, 1975.

TITE, M. S. *Pottery production, distribution, and consumption – The contribution of the physics sciences*. *Journal of Archaeological Method and Theory*, 6 (3), 181 – 233, 1999.

TÖLGYESSY, J., KYRS, M. *Radioanalytical Chemistry II*. Ed. John Wiley & Sons, New York, pág. 21 – 27, 1989.

WILKS, S. S. *Multivariate statistical outliers*. *Sankhya*, 25, 407 – 426, 1963.

AGRADECIMENTOS

Os autores deste trabalho agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, 155373/2006-4) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Processo 0056/014) pelo apoio financeiro. Agradecimentos especiais ao Museu Arqueológico de Xingó (MAX) e à Petróleo Brasileiro S.A (PETROBRAS) pelo apoio técnico-logístico e por permitir a análise da coleção.

