# Compositional Analysis in Archaeology

# Oxford Handbooks Online

## Abstract and Keywords

Compositional analysis in archaeology involves the analysis and interpretation of chemical fingerprints obtained from archaeological materials. The primary objective of compositional analysis is to identify groups of related artifacts and/or raw materials that provide insight into archaeological questions. The evidence has application to questions about human behavior, provenance, technology, and artifact authentication. Because compositional data are inherently multivariate, a familiarity with the methods from multivariate statistics is essential. This article presents an overview of the history of compositional anlaysis and the analytical methods employed. The main multivariate procedures used for data interpretation and presentation of results are also described. Finally, before selecting an analytical method and laboratory to perform the analysis on archaeological specimens, the article recommends obtaining answers to a list of questions.

Keywords: chemical analysis, chemometrics, multivariate statistics, provenance, classification

Compositional analysis of archaeological materials involves the application of methods from the physical sciences to generate chemical fingerprints for raw materials and artifacts. Unlike dating methods that provide archaeologists with information on *when* artifacts were used by humans, compositional analysis identifies *groups* of like objects. The information is used to locate the sources of raw materials, discover production sites, and research the technologies used to manufacture artifacts. Archaeologists use the evidence obtained from these investigations to study questions about past human behavior. Other applications include artifact authentication and improving techniques for conservation and preservation of artifacts.

# Compositional Analysis in Archaeology

## History of Compositional Analysis

Reviews of the early literature on archaeological chemistry by Caley (1951), Harbottle (1982), and Pollard and Heron (1996) describe the works of several famous scientists who pioneered the earliest compositional studies of archaeological materials. Among the most notable individuals were Martin Heinrich Klaproth (1743–1817), Sir Humphry Davy (1778–1829), Michael Faraday (1791–1867), Karl Christian Göbel (1794–1851), Jan Erazim Vocel (1803–1871), Augustin Alexis Damour (1808–1902), Friedrich Wilhelm Rathgen (1862–1942), and Theodore William Richards (1868–1928).

Klaproth was a distinguished German chemist, often acknowledged as the father of analytical chemistry for establishing a systematic methodology for performing analytical chemistry. Klaproth conducted the first known compositional analysis of ancient coins from Greece and Rome and discovered they were made from alloys of copper. He also estimated their composition. A report of his work was read before an audience of the Academie Royale des Sciences et Belles-Lettres de Prusse in Berlin on July 9, 1795, and later published in a volume of memoirs by the Royal Academy. Three years later, Klaproth presented a second paper to the Royal Academy describing his investigation of three samples of Roman glass in a mosaic recovered from the ruins of Emperor Tiberius on the island of Capri. The samples of glass were highly colored and opaque: one was red, one was green, and the third was blue. He concluded the different colors were due to the presence of copper and iron. He also produced a nearly complete analysis of the major constituents.

Sir Humphry Davy was an English chemist best known for discovering the alkali and alkaline earth metals. His major contribution to archaeological chemistry was the characterization of pigment samples from Rome and Pompeii (Davy 1815). One of Davy's laboratory assistants was Michael Faraday. Although Faraday is more widely recognized for his work in the field of electromagnetism, he also studied glazed Roman pottery and discovered the glaze was lead based.

An article by Göbel (1842), a German pharmacist and chemist, reported on his study of a collection of brass artifacts excavated from the Russian Baltic provinces. He compared the composition of the brass artifacts to prehistoric brass artifacts of known Greek and Roman origins. His discovery that the Baltic artifacts were similar to those from the Roman Empire led Göbel to conclude the artifacts were acquired by trade instead of having been made from local materials.

The Austrian scholar Vocel is considered the forefather of Austrian archaeology as a science. His contribution to archaeological chemistry was made in the mid-nineteenth

century when he recommended that the correlations between the compositions of raw materials and artifacts might be used to identify the sources of raw materials.

The French mineralogist Damour (1864) proposed that the sources of stone axes might be identified by determining the density and composition of different rock types, including jade and obsidian. He also encouraged archaeologists to collaborate with specialists from other disciplines such as chemistry, geology, and zoology, which he predicted would increase the amount of knowledge gained.

During the latter portion of the nineteenth century, compositional analyses of archaeological metals were conducted by John Percy, a metallurgist at the Royal School of Mines in London, who contributed to several major works on the early production and use of metals. Among the most significant of these were his analyses of metal artifacts recovered from the sites of Nineveh and Babylon by Austen Henry Layard (1853) and from Mycenae by Heinrich Schliemann (1878).

The German chemist Rathgen was hired by the Royal Museums of Berlin to be its first chemist. Later, he became the first director of its chemical laboratory. He is considered to be the father of modern archaeological conservation for his work in developing procedures for the removal of corrosion from antiquities and methods for artifact preservation (Rathgen 1898).

Until the end of the nineteenth century, most quantitative analyses of artifacts were performed on metallic alloys using gravimetric methods. It was a logical method because all determinations could be made with great precision by direct weighing of the solids produced after dissolution, precipitation, filtration, and evaporation. The main disadvantage was that measurements were limited to one element at a time.

One of the first major investigations of ceramic artifacts was a gravimetric study of Athenian pottery from the Boston Museum of Fine Arts conducted by T. W. Richards (1895), the first American to receive the Nobel Prize in Chemistry. Richards and his graduate students at Harvard University found the gravimetric method to be extremely laborious and not practical for application to large numbers of samples.

During the 1920s and 1930s, the development of instrumental methods of analysis, such as optical emission spectroscopy (OES), made it possible to determine multiple trace elements with greater sensitivity and to analyze more samples faster. These advances opened the door to systematic studies of ancient metals by dedicated laboratories in Britain, Germany, and the United States. Within a couple of decades, thousands of copper-based artifacts from Europe were analyzed (Otto and Witter 1952; Junghans, Sangmeister, and Schröder 1960; Caley 1964a, 1964b).

# Compositional Analysis in Archaeology

From the 1950s until the present, the number of scientific techniques grew significantly and their application to a wider variety of archaeological materials has increased. One of the major breakthroughs occurred in the mid-1950s when neutron activation analysis (NAA) was first employed by Sayre and Dodson (1957) to study pottery. With NAA, sample dissolution was no longer required, and more elements could be measured with greater precision and accuracy. Another advantage of NAA was that many of the procedures could be automated. In the 1960s and 1970s, dedicated NAA laboratories were established at national laboratories in the United States (Brookhaven National Laboratory and Lawrence Berkeley National Laboratory), in museums (British Museum and Smithsonian Institution), and at major universities (University of Michigan, University of Manchester, Hebrew University, and University of Toronto). Almost all of the first-generation NAA laboratories have since been replaced by a second generation of laboratories located at the University of Missouri, Oregon State University, University of Budapest, and so forth.

In addition to NAA, other analytical techniques were developed for bulk compositional analysis, including X-ray fluorescence (XRF), particle-induced X-ray emission (PIXE), electron probe microanalysis (EPMA), inductively coupled plasma-mass spectrometry (ICP-MS), and laser ablation ICP-MS (LA-ICP-MS). Most of the analytical methods are capable of producing high-precision data, but they often require that a portion of the artifact be destroyed to prepare the sample for analysis. However, recent developments with portable XRF spectrometers make it possible to perform rapid, nondestructive analysis at archaeological sites and museums for a limited suite of elements when lower precision is acceptable.

Most of the chemical analytical methods mentioned earlier are used to determine the compositions of inorganic archaeological materials (i.e., stone tools, pottery, metals, and glass). Because the organic archaeological materials (i.e., human remains, food residues, textiles, plants, etc.) are rarely preserved, measurements of bulk compositions of organic materials are unusual. Techniques that measure isotope ratios in organic artifacts have proven more productive.

## Strategies for Interpreting Compositional Data

The science of extracting information from data acquired by chemical analysis is known as chemometrics. Although the methods used in chemometrics originated in data-analytic disciplines such as applied mathematics, statistics, and computer science, many have

direct application to problems in archaeology being studied by way of chemical analysis of artifacts.

This discussion will focus on the interpretation of data from inorganic artifacts, which can be categorized as either *natural* or *chemically altered*. Natural materials such as obsidian, basalt, flint, gemstones, and limestone were used to create stone tools, weapons, jewelry, statuary, and so forth. Although these natural artifacts were modified physically, their chemical compositions were not significantly modified. On the other hand, chemically altered artifacts were created from raw materials by processes such as mixing, firing, and smelting to produce pottery, bricks, man-made glass, metals, and so forth. Matching chemically altered artifacts to raw materials is more difficult (Neff 2000; Pernicka 2014).

Studies of the technologies used to produce chemically altered artifacts rely mainly on data for the major elements present in the artifacts and raw materials (concentrations >1%). On the other hand, studies to establish the sources of both natural and chemically altered artifacts tend to rely heavily on the concentrations of the minor and trace elements whose presence in most artifacts is effectively "accidental." The latter studies are commonly referred to as *provenance* investigations, and their main objective is to study the movement of artifacts from their raw material source or place of manufacture to the archaeological site where they were recovered.

The basic proposition underlying provenance determination was understood by the early 1970s (e.g., Perlman and Asaro 1969; Sayre 1975; Bieber et al. 1976). However, it was Weigand et al. (1977) who first stated explicitly that linking artifacts to sources through compositional analysis depends on satisfying the postulate "the differences in chemical composition between different natural sources [must] exceed, in some recognizable way, the differences observed within a given source." (Although Weigand's use of the term *source* might seem to imply only the location from which raw materials originated, in the case of chemically altered artifacts, *source* usually refers to the workshop or production site where the artifacts were made.)

The strategies used for provenance determination are different for certain types of artifacts, as explained by Neff (2000). As shown in Table 1, when investigating natural artifacts, the raw material sources should be analyzed first to establish statistically valid compositional groups against which the chemical fingerprints of artifacts can be compared. For chemically altered artifacts, the artifacts should be analyzed first to establish statistically valid compositional groups against which the chemical fingerprints of kiln wasters, unfired sherds, and raw materials can be compared. Finally, as more chemically altered artifacts are analyzed, the compositional groups should continue to be refined.

# Compositional Analysis in Archaeology

| Table 1 Contrasting Strategies for Provenance Determination of Inorganic Artifacts | | |
|---|---|---|
| **Procedures** | **Natural** | **Chemically Altered** |
| Select samples for analysis | Obsidian, basalt, limestone, gems, etc. produced by physical modification of raw materials but maintaining a direct chemical relationship between artifacts and raw materials. | Pottery, bricks, coins, glass, etc. produced by mixing, firing, and smelting of raw materials resulting in an indirect chemical relationship between artifacts and raw materials. |
| Determine compositions | Collect raw materials from known locations and analyze. | Collect artifacts from archaeological sites and analyze. |
| Create compositional groups | Create statistically valid compositional groups from raw materials. | Create statistically valid compositional groups from artifacts. As more artifacts are analyzed, groups should be further refined. |
| Assign compositional groups to geographic locations | Link compositional groups to locations of raw materials. | Use abundance of artifacts from different cultural contexts to infer production site locations. |
| Compare artifacts to compositional groups | Analyze and compare artifacts to raw material groups. Eliminate improbable sources in order to identify the most likely source for each artifact. | Analyze and compare kiln wasters, unfired sherds, and raw materials to artifact groups. Search for correlations between geographic locations of artifacts and production sites. |

# Compositional Analysis in Archaeology

## Preparation of Compositional Data for Interpretation

Aitchison (1986) defines compositional data as the parts of some whole that convey relative information. For geological and archaeological materials, a single sample is characterized by a series of positive concentrations whose sum total must equal unity (i.e., 1 or 100%). Due to this constant sum constraint, the concentrations are not free to vary independent of one another. If the concentration of one element increases, the concentrations of one or more other elements must decrease. Therefore, an artificial correlation between the elements exists, making compositional data intrinsically multivariate. According to Aitchison (1986), the only real information in compositional data is given by the ratios between components.

The volume of compositional data generated can be substantial (e.g., >30 elements each for hundreds of samples), usually arranged in the form of a matrix of *m* samples (rows) by *n* elements (columns). As a result, data management, statistical analysis, and interpretation can be very challenging. Large multivariate datasets necessitate the use of multivariate analysis (MVA) programs to discover patterns and extract knowledge from the data. All MVA programs include data visualization routines capable of generating plots, tables, and charts that assist in the recognition of patterns, trends, and anomalies not otherwise apparent from numerical summaries. Visualization routines also aid the process of data summarization and simplify explaining the interpretation of a complex dataset to the audience.

A range of MVA procedures for interpreting compositional data have been adapted from the discipline of statistics (Bieber et al. 1976; Glascock 1992; Baxter 1994; Beier and Mommsen 1994; Neff 2000; Buxeda i Garrigos and Kilikoglou 2003). No single procedure (or combination of procedures) has been accepted as the best method for interpreting compositional data in the field of archaeology. Every dataset has its own unique characteristics. A procedure that works well for one dataset may be inappropriate for another. For a best practice, several methods should be employed and the results compared to find the best overall solution.

## Data Screening

Before investigating a compositional dataset with a series of MVA routines, the data should be inspected to ensure it is reliable for testing. The *data screening* process is intended to detect possible errors and abnormal values that may interfere with the data evaluation. Some of the possible errors are sample misidentification, data entry errors,

# Compositional Analysis in Archaeology

contamination, and so forth. When inconsistent values are found, they should be investigated for the possible need to correct, reanalyze, or reject completely.

Screening of data also includes the investigation of missing values (i.e., nondetects). These may occur for a number of reasons: values were below the detection limit, data were lost, or data were otherwise unavailable. Because missing values cannot be used directly by MVA routines, a number of different methods are available to supply replacement values (Arteaga and Ferrer 2002). When a particular element has a large number of missing values, it may be necessary to eliminate the element from the MVA.

## Frequency Distribution for Compositional Data

When the data for a single variable are arranged to display the number of instances in which the measured variable takes each of its values, it becomes a frequency distribution. This provides a picture of how individual data are distributed while also revealing extreme values (i.e., outliers). A frequency distribution is considered *normal* if it can be characterized by a bell-shaped curve symmetrically centered about its mean. If the number of data points is very large, the bell-shaped curve is referred to as a Gaussian distribution, which takes the functional form

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

Here, $\mu$ is the mean of the distribution, $\sigma$ is its standard deviation, and $\sigma^2$ is the variance.

Geologists and statisticians have long debated over whether chemical data for archaeological and geological materials are distributed normally or lognormally (Ahrens 1954; Vistelius 1960). According to a study by Zhang et al. (2005), almost all elements show neither a normal nor lognormal data distribution. However, a majority of archaeological researchers have found it convenient to use a logarithm transform on geochemical data because it acts as a "quasi-standardization" (Sayre 1975; Neff 2000) that compensates for the differences in absolute magnitude between the major elements, such as Al, K, and Fe present at percent levels, and the trace elements, such as the rare earth elements (REEs) present at parts per million and below. Thus, the transformation places the elements onto more or less the same scale.

As the dispersion in a distribution of concentrations for a particular element increases, it becomes more difficult to use the element to distinguish between groups. A parameter that quantifies the dispersion is the measured variance,
$$\sigma_m^2$$
, in the mean concentration. The measured variance represents the sum of the natural variance,

$$\sigma_n^2$$
, and variance due to errors in sampling and analysis,
$$\sigma_a^2$$
, as defined by the expression

$$\sigma_m^2 = \sigma_n^2 + \sigma_a^2$$

where $\sigma_m$ is equal to the standard deviation of the measured mean. Ideally, the analytical variance should be minimized in such a way that a majority of the variance is due to the natural variance of the material. Replicate sampling and analysis of a highly homogeneous material can be used to produce a spread of values assessing overall measurement precision.

## Correlations and Covariance

The existence of correlations between many of the elements present in geological materials is well known. Some examples include the following: (1) Ca is usually correlated with Sr; (2) Sc is associated with siderophile elements like Fe, Cr, and Co; (3) Hf and Zr are correlated in mantle rocks; and (4) the rare-earth elements exhibit strong correlations with one another.

The correlation coefficient, $r_{jk}$, between two elements $j$ and $k$ is defined by the ratio

$$r_{jk} = \frac{\sigma_{jk}}{\sigma_j \sigma_k}$$

where $\sigma_j$ and $\sigma_k$ are the respective sample standard deviations and $\sigma_{jk}$ is the covariance between them. Correlation coefficients range in value from +1 to –1. A correlation of +1 indicates a perfect 1:1 relationship between the two variables; a correlation of –1 indicates two variables that are inversely related. The sample covariance can be calculated from the expression

$$\sigma_{jk} = \frac{\sum_{i=1}^{m}[C_{ij} - A_j][C_{ik} - A_k]}{m - 1}$$

where $C_{ij}$ and $C_{ik}$ denote the respective concentrations of the $j$th and $k$th elements in sample $i$. The mean element concentrations are given by $A_j$ and $A_k$, and $m$ represents the total number of samples.

# Compositional Analysis in Archaeology

## Searching for Structure in Multivariate Compositional Data

MVA procedures are used to identify compositional groups within a multivariate dataset. Most MVA procedures are based on measures of similarity or difference (variability) between objects. Objects that are similar will be grouped together, while objects that are dissimilar will be placed in different compositional groups. These groups can be thought of as "centers of mass" or "clusters" in the multivariate space of variables described by data.

The main purpose for examining multivariate compositional data with MVA routines is to identify groups from which new knowledge can be obtained. Learning from multivariate data is often described as either *supervised learning* or *unsupervised learning*. In supervised learning, the model or structure of the data is already known; some variables may have more weight than others, but the structure is not modified during the learning process. In unsupervised learning, the structure is unknown, all variables are treated equally, and patterns within the data are sought. However, there is no certainty that the patterns identified from unsupervised learning are correct. When the sources are already known, the approach to assigning sources for obsidian artifacts involves supervised learning through the use of discrimination or classification methods. When the sources are unknown, as is the case with most investigations of archaeological ceramics, the MVA procedures involve unsupervised learning with procedures that identify patterns of compositional groups.

## Cluster Analysis

One of the most basic unsupervised learning procedures applied to multivariate compositional data is hierarchical cluster analysis (HCA). HCA is a procedure that groups samples according to their similarity to one another. HCA works by calculating the Euclidean distances between pairs of samples. The HCA procedure employs a "bottoms-up" approach toward building a hierarchy of clusters. Using this approach, every sample is assumed to be the lowest member of a cluster. As the distances between pairs of samples are compared, those with the smallest distances between are merged together as one moves upward in the hierarchy. Samples separated by large distances are placed in separate clusters.

# Compositional Analysis in Archaeology
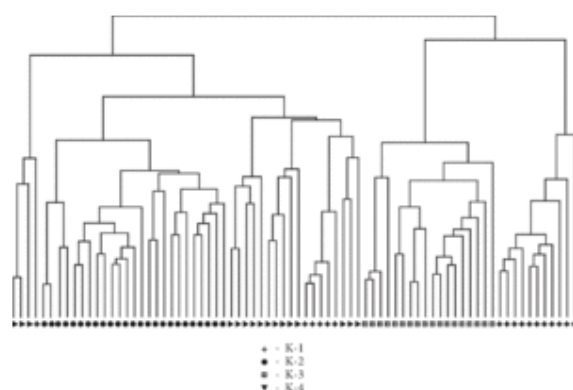


*Click to view larger*

*Figure 1* Dendrogram generated from applying cluster analysis to compositional data to 80 samples of Bronze Age pottery from a site in Turkey. Four compositional groups were identified.

HCA results are typically presented in the form a *dendrogram* indicating the relationships between samples in the dataset, as shown in Figure 1 for pottery from a site in Turkey. Although HCA is an efficient tool for displaying the visual relationships between samples, its major weakness is the failure to take into account the correlations between elements. This can be a serious problem when dealing with a highly correlated data.

HCA is based on a dissimilarity matrix in which the distances between all pairs of specimens are calculated using one of several possible distance measures (Sayre 1975). The most popular measure is the squared-mean Euclidean distance (SMED) by which one calculates a distance between specimens *j* and *k* according to the equation

$$d^2{}_{jk} = \frac{1}{n}\sum_{i=1}^{n}[C_{ij} - C_{ik}]^2.$$

The scaling factor, *n*, corresponds to the number of elements actually determined. It removes the possible problem of missing values, since the measure averages over only those elements for which data are included. For an assemblage containing *m* samples, there are $m(m-1)/2$ pairs of samples, and a $d^2$ value will be calculated for each pair. At this stage, it is advisable to use the information on all elements having few missing values to use the maximum amount of information possible when generating clusters.

Visual inspection of the dendrogram can be used to identify potential groups. However, the inclination for HCA to force data into spherical groups can lead to erroneous assignments of samples due to the tendency for compositional groups to be elongated due to interelement correlation. Groups created from cluster analysis should be considered tentative until validated later by a more robust evaluation that takes the effects of correlation into account.

# Compositional Analysis in Archaeology

A clustering method known as *k-means clustering* is popular among some experts who study compositional data. The *k*-means method differs from HCA by partitioning the *n*-sample dataset into *k* clusters, where each sample is assigned to a cluster with the nearest mean. A major challenge with *k*-means clustering is that the number of clusters must be estimated in advance.

## Total Variation Method

The total variation procedure is another unsupervised learning approach to compositional data first described by Aitchison (1986) and subsequently employed to examine archaeometric data by Buxeda i Garrigos and Kilikoglou (2003). It evaluates the degree of variability in a dataset from the measured variance,

$$\sigma_m^2$$

. Assuming the errors due to sampling and analysis are minimal compared to the natural variance, the degree of variability can be useful for predicting whether compositional data is monogenic or polygenic.

Using Aitchison's procedure, the variation matrix is constructed from all possible variances of the log-ratios for all *n* elemental concentrations. For example, the covariances are defined by

$$\sigma_{ij,kl} = \mathrm{cov}\left\{\log\left(\frac{x_i}{x_j}\right), \log\left(\frac{x_j}{x_k}\right)\right\} \rightarrow \text{for i , j, k, l } = \ 1, \dots \text{n}.$$

All diagonal elements of the variation matrix are zero, and the off-diagonal elements of the variances are defined as

$$\tau_{ij} = \sigma_{ii,jj} = \mathrm{var}\left\{\log\left(\frac{x_i}{x_j}\right)\right\} \rightarrow for\ i, j = 1,\ \dots n.$$

As a result, the covariances can be calculated from the log-ratio variances using the equation

$$\sigma_{ij,kl} = (\tau_{il} + \tau_{jk} - \tau_{ij} - \tau_{kl})/2$$

and the total variation *vt* is given by

$$vt = \frac{\sum_{i=1}^{i=n}\sum_{j=1}^{j=n}\tau_{ij}}{2n}$$

A large value for *vt* indicates greater variability and suggests the dataset is polygenic (i.e., subgroups), whereas a small value for *vt* indicates less variability and suggests a possible monogenic dataset. Aitchison (1992) also demonstrated that the total variation is directly related to the Euclidean distances between all samples in the dataset.

In addition to total variation parameter, Buxeda i Garrigos and Kilikoglou (2003) show that the variation matrix of the entire dataset provides other useful information. For

example, sources of variability within the dataset can be revealed by examining the log-ratio variances, $\tau_{.t}$, for each element as a divisor. It provides an estimate regarding each element's contribution to the total variation. The information from total variation and log-ratio variances can also be used to isolate samples suffering from various processes such as alternation and contamination.

## Principal Components Analysis

Principal components analysis (PCA) is a variable reduction technique used to simplify the structure of compositional data. The aim is to explain as much of the variance as possible using the minimum number of variables. Although the PCA is not explicitly a group-formation technique, it can make important contributions toward finding groups and understanding the differences between groups.

PCA performs an orthogonal transformation on the data to convert potentially correlated variables into a set of linearly uncorrelated variables called the principal components (PCs). The goal is to transform the original multivariate data into a new representative dataset. However, PCA is only useful if the data are correlated in some way. For uncorrelated data, PCA offers no advantages.

The PCA transformation applies eigenvector methods on the variance-covariance matrix to determine the directions and magnitudes of maximum variance in the dataset. The first PC is a linear combination of the original variables that maximizes the amount of variance explained. The first PC is oriented in the direction of maximum variance in the dataset. The second PC is calculated to lie along the direction of maximum remaining variance, with the additional constraint that it must be orthogonal to the first PC. The third PC lies in the direction of maximum remaining variance after the variance attributed to the second PC has been removed and is orthogonal to the earlier PCs. The procedure concludes when the number of PCs is equal to the number of original variables. One of the requirements of PCA is that the number of samples in the dataset undergoing transformation must be equal to or greater than the number of variables.

Assuming the compositional data matrix has $m$ samples and $n$ elements, the first step in PCA involves centering the data on the mean values for each element. This insures that the transformed data will also be centered in the same location, but it will have no effect on the spatial relationships between samples or the variances between elements. The first PC is described by the linear combination of element concentrations according to the equation

$$PC_1 = a_{11}C_1 + a_{12}C_2 + \dots + a_{1n}C_n$$

where the coefficients $a_{11}$, $a_{12}$, ... $a_{1n}$ of the eigenvector represent the *weighting factors* for each element, which represent the correlation coefficients between the elements and the PCs. By definition, the weighting factors are constrained so that the sum of their squares must equal 1.

$$a^2_{11} + a^2_{12} + \ ... \ + a^2_{1n} = 1$$

The second PC is calculated in the same way as the first PC, with the condition that it must be uncorrelated (i.e., perpendicular to) and it must account for the maximum amount of remaining variance. This process is continued until the total number of PCs is equal to the number of elements, *n*. Thus, $PC_2$ and $PC_n$ can be written as:

$$PC_2 = a_{21}C_1 + a_{22}C_2 + \ ... \ + a_{2n}C_n$$
$$PC_n = a_{n1}C_1 + a_{n2}C_2 + \ ... \ + a_{nn}C_n$$

.

The new reference axes produced by PCA represent a new coordinate system, which usually offers an improved perspective for viewing the dataset. The position of every specimen in element concentration space can be converted to its principal score in the new PCA space.
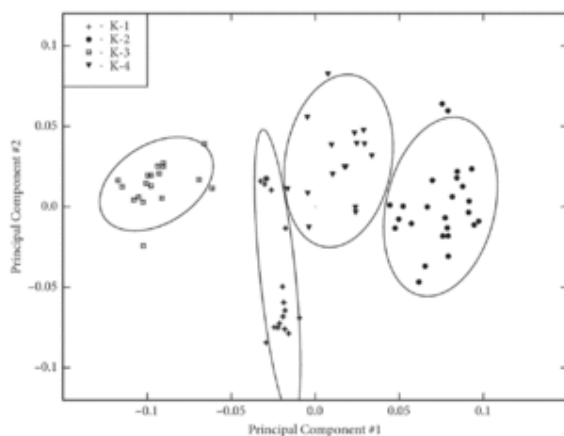


*Click to view larger*

*Figure 2* Scatterplot of principal component #1 and principal component #2 for Bronze Age pottery from a site in Turkey. Ninety percent confidence ellipses surround the four compositional groups.

Ideally, a greater portion of the structure for the compositional dataset under examination will be explained by the lowest PCs. The results of PCA can be studied by inspecting in a table of weighting factors or by viewing one or more scatterplots in two or three dimensions. Table 2 lists the weighting factors for the same Turkish pottery dataset illustrated in Figure 1. The weighting factors contributing most strongly to each PC are highlighted in bold. The scatterplot in Figure 2 shows the scores on PC#1 and PC#2 for the same data. The total number of unique two-dimensional scatterplots possible for

inspection is $n(n-1)/2$; however, the scatterplots of higher ordered PC scores gradually explain lesser amounts of the variation. Although different criteria have been proposed for determining how many PCs should be investigated and how many should be ignored, one of the most common criteria is to include all PCs until the total percent of variance explained reaches an acceptable level. For a majority of archaeological work, this is about 90% of the variance.

# Compositional Analysis in Archaeology

| Table 2 Weighting Factors from Principal Components Analysis for Bronze Age Pottery from Site in Turkey along with Percent Variance Explained by Each Principal Component | | | | | |
|---|---|---|---|---|---|
| **Element** | **PC1** | **PC2** | **PC3** | **PC4** | **PC5** |
| As | –0.134 | –0.232 | **–0.803** | **0.445** | 0.063 |
| Sr | –0.081 | 0.077 | –0.223 | **–0.394** | **–0.683** |
| Mn | –0.323 | **0.612** | –0.065 | 0.053 | **0.347** |
| Ca | –0.088 | 0.171 | **–0.426** | **–0.455** | 0.130 |
| Cr | **–0.610** | –0.050 | 0.197 | 0.067 | 0.023 |
| Cs | –0.141 | –0.104 | 0.059 | 0.264 | –0.362 |
| Co | **–0.354** | 0.286 | 0.009 | 0.062 | –0.063 |
| Th | 0.287 | 0.168 | –0.009 | 0.180 | –0.034 |
| Na | 0.056 | 0.329 | 0.023 | –0.079 | –0.100 |
| Eu | –0.052 | 0.191 | –0.003 | 0.119 | –0.269 |
| Ba | 0.043 | 0.088 | –0.240 | –0.192 | –0.035 |
| V | –0.237 | –0.055 | 0.001 | 0.147 | –0.157 |
| Nd | 0.123 | 0.200 | –0.047 | 0.156 | –0.107 |
| Sm | 0.110 | 0.207 | –0.021 | 0.143 | –0.090 |
| La | 0.125 | 0.175 | –0.021 | 0.152 | –0.113 |
| K | 0.213 | 0.160 | –0.093 | –0.016 | 0.024 |
| Fe | –0.136 | 0.020 | 0.045 | 0.124 | –0.179 |

| | | | | | |
|---|---|---|---|---|---|
| Ta | 0.123 | 0.016 | 0.031 | 0.201 | –0.104 |
| Ce | 0.119 | 0.190 | –0.019 | 0.104 | –0.059 |
| Rb | 0.204 | 0.127 | 0.002 | 0.078 | 0.014 |
| Dy | 0.061 | 0.169 | 0.002 | 0.149 | –0.039 |
| Hf | 0.050 | –0.030 | –0.014 | 0.115 | 0.179 |
| Yb | 0.049 | 0.136 | 0.010 | 0.158 | –0.005 |
| Sc | –0.099 | 0.041 | 0.061 | 0.123 | –0.101 |
| Lu | 0.039 | 0.104 | 0.008 | 0.158 | –0.047 |
| Al | 0.045 | 0.068 | 0.009 | 0.078 | –0.148 |
| Variance | 54.0% | 16.7% | 10.6% | 5.5% | 3.2% |
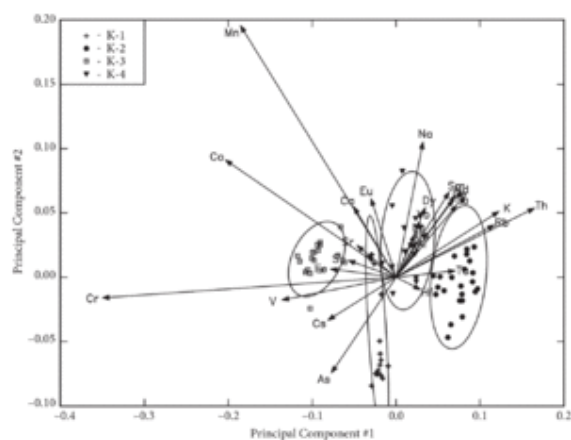


*Click to view larger*

*Figure 3* Biplot showing results from RQ-mode principal components analysis. Vectors indicate the direction and magnitude of weighting factors for each element. Ninety percent confidence ellipses surround the four compositional groups.

# Compositional Analysis in Archaeology
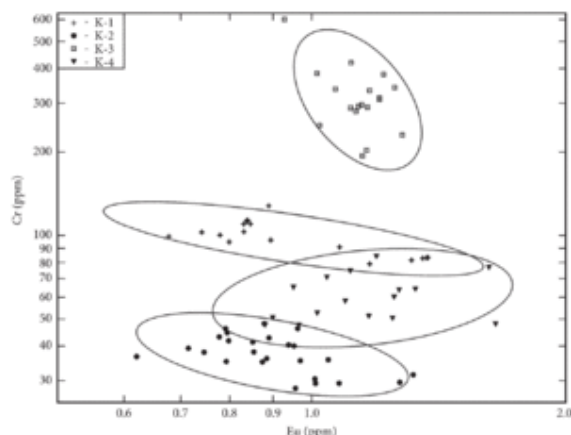


*Click to view larger*

*Figure 4* Scatterplot of Eu and Cr concentrations for the Bronze Age pottery from Turkey plotted a logarithmic scale. Ninety percent confidence ellipses surround the four compositional groups.

PCA can also be used to examine the basis for differences between groups. In R-mode analysis, the weight on each PC can be used to display the scores for samples in the new PCA space. In Q-mode analysis, the factor scores for the variables (elements) can be inspected as vectors. As described by Neff (1994), the MVA method that performs both simultaneously is known as RQ-mode PCA. When the RQ-mode PCA technique is used, it is possible to display both samples and element vectors simultaneously on a single plot known as a *biplot*. A biplot for the Turkish pottery is shown in Figure 3. The directions and lengths of the element vectors can be easily interpreted in terms of explaining which elements are responsible for differentiating compositional groups from one another and indicating the degree of correlation between elements. Examination of the element vectors can also be a useful tool for identifying which element pairs have the most potential for use in scatterplots. In other words, scatterplots should show that the differences between compositional groups are real and they are not a consequence of the PCA procedure. Figure 4 shows a scatterplot of Cr versus Eu for the Turkish pottery example.

## Mahalanobis Distance

Compositional groups are characterized by the location of their centroids and the unique correlations between the elements. The existence of correlations between the elements in geological and archaeological materials mandates the use of Mahalanobis distance (MD) measures to properly handle compositional data (Sayre 1975; Bishop and Neff 1989).

The MD is defined as the squared Euclidean distance between a sample and a group centroid, divided by the group variance in the direction of the sample (Sayre 1975). It is

# Compositional Analysis in Archaeology

equivalent to measuring the number of standard deviations between a sample and the group mean along each principal component axis. Mathematically, the MD between specimen $k$ and the centroid of group $A$ can be written as

$$D^2_{kA} = \sum_{i=1}^{n}\sum_{j=1}^{n}[C_{ik} - A_i]I_{ij}[C_{jk} - A_j]$$

$A_i$ and $A_j$ are the mean concentrations of elements $i$ and $j$ in the group, and $I_{ij}$ is the $ij$th element of the inverse of the variance-covariance matrix. The MD is both unitless and scale invariant, and it accounts for all correlations between pairs of elements as derived from the off-diagonal terms of the variance-covariance matrix, which the simple Euclidean distance does not. If all axes were rescaled such that they have unit variance, the MD distance is equal to the Euclidean distance. Calculation of the probability that a particular sample belongs to a group is based not only on its proximity to the group centroid in Euclidean terms but also on the rate at which the density of samples decreases away from the centroid in the direction of the sample of interest.

The significance of differences between two groups of specimens can be tested by Hotelling's $T^2$ statistics (the multivariate equivalent of the student's $t$) calculated from $\mathrm{D}^2$ according to the equation

$$T^2 = \frac{D^2}{\frac{1}{m_1} + \frac{1}{m_2}}$$

where $m_1$ and $m_2$ are the numbers of samples in the two groups. Hotelling's $T^2$ statistic is equivalent to the MD for individual data points. Therefore, the probabilities of membership are easily calculated after transforming the $T^2$ statistic into the related $F$ value by

$$F = \frac{[m_1 + m_2 - v - 1]}{[m_1 + m_2 - 2]n}T^2$$

where $n$ is the number of elements.

Variations in the procedures using MD distances have been employed by some laboratories to overcome some of the limitations of cluster analysis and PCA. A modified MD filter method combined with chi-squared statistics (Beier and Mommsen 1994) establishes groups by considering correlations, measurement uncertainties, and dilution effects. They argue that their procedure produces improved separations between compositional groups and reduces the effects of data shifts due to possible dilution and weathering effects.

MD distance calculations can also be used to replace missing values (Sayre 1975). When the number of samples with missing values is modest, it is possible to calculate a replacement value for each sample relative to its presumed compositional group based on minimizing the effect on the MD distance with and without the replacement value.

# Compositional Analysis in Archaeology

## Group Validation

The initial groups created from the examination of scatterplots or cluster analysis should be validated by using the MD to calculate membership probabilities for individual samples. Calculation of the MD from a sample to a group requires that number of samples in the group exceeds the number of variables (elements or PCs) by at least one. To avoid bias, individual samples should not be compared to a compositional group in which they have are already a member. The solution to this problem is to use a *jackknifing* procedure by which membership probability for each sample is calculated by temporarily excluding the sample from the group to which it is being compared.

For most accurate probability calculations, theoretical studies have shown that sample-to-element ratio, *m/n*, should be at least 3—5, and the larger the better (Foley 1972). Unfortunately, the *m/n* problem affects most compositional analysis projects involved in archaeology because most archaeological sites lack the number of artifacts or more likely because archaeologists lack the financial resources necessary to analyze the number of artifacts needed to achieve the recommended sample-to-element ratio. The most common method for circumventing the small sample-to-element ratio is to base the MD distance measures on a reduced number of principal components rather than use the original element concentrations.

## Classification and Discriminant Analysis

Classification and discriminant analysis are supervised learning techniques related to the MVA methods used to create and validate compositional groups. As new archaeological samples are analyzed, the same procedures used to validate group membership are applicable to classifying (or assigning) the new samples to the existing compositional groups. After membership to an existing group is confirmed, it may be necessary to reevaluate the entire dataset. If the new samples do not belong to existing groups, they may be outliers or representatives of yet-to-be-identified groups.

Canonical discriminant analysis (CDA) is a procedure for dimension reduction similar to PCA. CDA contrasts with PCA by extracting a new set of variables that maximize the differences between two or more groups instead of maximizing the total variance. CDA relies on the assumption that the pooled variance-covariance matrix is an accurate representation of the total variance and covariance (Davis 1986). By definition, CDA also requires all of the samples to belong to one of the known groups.

The CDA procedure constructs a series of canonical discriminant functions (one fewer than the number of groups), which for each group maximizes the likelihood for specimens to belong to their assigned group and minimizes the likelihood of belonging to all other groups. The main requirements for CDA are as follows: (1) two or more groups; (2) at least two samples per group; and (3) the number of variables must be at least two fewer than the number of samples. The individual discriminant functions are linear combinations of the original data that successively describe decreasing amounts of the separation between the groups. Two-dimensional scatterplots of the discriminant functions are used to illustrate the success in separating groups by CDA.

## Selecting an Analytical Method and Laboratory

The compositional analysis methods mentioned earlier have specific advantages and disadvantages for application to analysis of different types of archaeological samples. Because most archaeologists do not operate a laboratory for compositional analysis, they necessarily seek collaborations with professionals from analytical laboratories. Selecting the appropriate technique and choosing a competent laboratory to perform compositional analysis are extremely important. Before engaging a laboratory to perform an analysis on archaeological material, the archaeologist should seek answers to the following questions:

**1.** Which techniques are appropriate for the specific sample material?
**2.** Will the elements measured have the potential to answer the archaeological questions?
**3.** Is the sensitivity adequate for the elements of interest?
**4.** Are the accuracy and precision for the measured elements satisfactory?
**5.** Is the analysis destructive or nondestructive?
**6.** How much sample material is required?
**7.** Does the analytical laboratory have experience with the archaeological material?
**8.** Is there an existing database available for comparative purposes at the selected laboratory?
**9.** Can the data from the laboratory be reliably compared to data from other laboratories?
**10.** How much will the analysis cost?

All laboratories performing analysis of archaeological samples should have a history of analyzing standard and quality control samples. Standards and quality controls are samples of material from a matrix similar to the type of artifacts to be measured.

# Compositional Analysis in Archaeology

Standards are usually reference materials that have been certified by analysis at a number of laboratories such that a set of consensus values exist describing the "best" values for the concentration of elements in the standard. Standards are generally available for purchase from organizations such as the National Institute of Standards and Technology (NIST) and the United States Geological Survey (USGS) in the United States. Most other countries have similar organizations that offer standard reference materials for chemical analysis measurements. Standards are used to develop calibration curves for the analytical procedure that will be applied to the archaeological samples. Quality control samples should be included with each batch of standards and artifacts undergoing analysis to provide an independent check on data quality. Quality control of data is essential to detecting and correcting systematic errors to ensure the long-term variability of measured data is minimized.

Before submitting samples to an analytical laboratory for analysis, advice on the project design would be advantageous and may avoid collection of useless data. Instructions on how to collect, prepare, and ship samples should be provided. In exchange for collaboration, some laboratories will offer assistance with MVA and data interpretation. These factors should be considered before undertaking a large and expensive program of analysis.

## References

Ahrens, L. H. (1954). "The lognormal distribution of the elements (A fundamental law of geochemistry and its subsidiary)." *Geochimica Cosmochimica Acta* 5: 49–74.

Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. Chapman and Hall: London and New York.

Aitchison, J. (1992). "On criteria for measures of compositional difference." *Mathematical Geology* 24: 365–379.

Arteaga, F., and Ferrer, A. (2002). "Dealing with missing data in MSPC: Several methods, different interpretations, some examples." *Journal of Chemometrics* 16: 408–418.

Baxter, M. J. (1994). *Exploratory Multivariate Analysis in Archaeology*. Edinburgh: Edinburgh University Press.

Beier, T., and Mommsen, H. (1994). "Modified Mahalanobis filters for grouping pottery by chemical composition." *Archaeometry* 36(2): 287–306.

# Compositional Analysis in Archaeology

Bieber, A. M, Jr., Brooks, D. W., Harbottle, G., and Sayre, E. V. (1976). "Application of multivariate techniques to analytical data on Aegean ceramics." *Archaeometry* 18(1): 59–74.

Bishop, R. L., and Neff, H. (1989). "Compositional data analysis in archaeology." In *Archaeological Chemistry IV*, edited by R. O. Allen, pp. 57–86. Washington, D.C.: American Chemical Society.

Buxeda i Garrigos, J., and Kilikoglou, V. (2003). "Total variation as a measure of variability in chemical datasets." In *Patterns and Process: A Festschrift in Honor of Dr. Edward V. Sayre*, edited by L. Van Zelst, pp. 185–198. Washington, D.C.: Smithsonian.

Caley, E. R. (1951). "Early history and literature of archaeological chemistry." *Journal of Chemical Education* 28(2): 64–66.

Caley, E. R. (1964a). *Analysis of Ancient Metals*. International Series of Monographs on Analytical Chemistry, vol. 19. Oxford: Pergamon Press.

Caley, E. R. (1964b). *Orichalcum and Related Ancient Alloys: Origin, Composition, and Manufacture with Special Reference to the Coinage of the Roman Empire*. New York: American Numismatic Society.

Damour, A. (1864). "Sur la composition des haches en pierre trouvees dans les monuments celtiques et chez les tribus sauvages." *Comptes Rendus* 61: 313–311, 357–368.

Davis, J. C. (1986). *Statistics and Data Analysis in Geology*. New York: John Wiley and Sons.

Davy, H. (1815). "On the colours used in painting by the ancients." *Philosophical Transactions of the Royal Society of London* 105: 97–124.

Foley, D. H. (1972). "Considerations of sample and feature, size." *IEEE Transactions in Information Theory* IT-18(5): 618–626.

Glascock, M. D. (1992). "Characterization of archaeological ceramics at MURR by neutron activation analysis and multivariate statistics." In *Chemical Characterization of Ceramic Pastes in Archaeology*, edited by H. Neff, pp. 11–26. Madison, Wis.: Prehistory Press.

Göbel, C. C. T. F. (1842). *Über den Einfluss der Chemie auf die Ermittelung der Völker der Vorzeit: Oder Resultate det chemischen Untersuchung metallishcher Alterthümer*. Erlangen, Germany: Enke.

# Compositional Analysis in Archaeology

Harbottle, G. (1982). "Chemical characterization in archaeology." In *Contexts for Prehistoric Exchange*, edited by J. E. Ericson and T. K. Earle, pp. 13–51. New York: Academic Press.

Junghans, S., Sangmeister, E., and Schröder, M. (1960). *Metallanalyses kupferzeitlicher und frühbrozenseitlicher Bodenfunde aus Europa*. Berlin: Gebr. Mann Verlag.

Layard, A. H. (1853). *Discoveries in the Ruins of Nineveh and Babylon*. London: John Murray.

Neff, H. (1994). "RQ-mode principal components analysis of ceramic compositional data. "*Archaeometry* 36(1): 115–130.

Neff, H. (2000). "Neutron activation analysis for provenance determination in archaeology." In *Modern Analytical Methods in Art and Archaeology*, edited by E. Ciliberto and G. Spoto, pp. 81–134. New York: John Wiley & Sons.

Otto, H., and W. Witter, (1952). *Handbuch der ältesten vorgeschichtlichen Metallugrie in Mittleeuropa*. Leipzig, Germany: Barth Verlag.

Perlman, I., and Asaro, F. (1969). "Pottery analysis by neutron activation." *Archaeometry* 11: 21–52.

Pernicka, E. (2014). "Provenance determination of archaeological metal objects." In *Archaeometallurgy in Global Perspective*, edited by B. W. Roberts and C. P. Thornton, pp. 239–268. New York: Springer.

Pollard, A. M., and Heron, C. (1996). *Archaeological Chemistry*. Cambridge, England: Royal Society of Chemistry.

Rathgen, F. (1898). *Die Konservierung von Altertumsfunde*. Berlin: W. Spemann.

Richards, T. W. (1895). "The composition of Athenian pottery." *Journal of the American Chemical Society* 17: 152–154.

Sayre, E. V., and Dodson, R. W. (1957). "Neutron activation study of Mediterranean potsherds." *American Journal of Archaeology* 61: 35–41.

Sayre, E. V. (1975). "Brookhaven procedures for statistical analysis of multivariate archaeometric data." Brookhaven National Laboratory report, BNL-21693. Upton, N.Y.: Broohaven National Laboratory.

Schliemann, H. (1878). *Mycenae*. (Reissued 2010; Cambridge, England: Cambridge University Press).

Vistelius, A. B. (1960). "The skew frequency distributions and the fundamental law of the geochemical processes." *Journal Geology* 68: 1–22.

Weigand, P., Harbottle, G., and Sayre, E. V. (1977). "Turquoise sources and source analysis: Mesoamerica and the southwestern USA." In *Exchange Systems in Prehistory*, edited by T. K. Earle, and J. E. Ericson, pp. 15–32. New York: Academic Press.

Zhang, C. S., Manheim, F. T., Hinde, J., and Grossman, J. N. (2005). "Statistical characterization of a large geochemical database and the effect of sample size." *Applied Geochemistry* 20: 1857–1874.

**Michael D. Glascock**

University of Missouri