Chapter 2

A Systematic Approach to Obsidian Source Characterization

MICHAEL D. GLASCOCK, GEOFFREY E. BRASWELL, AND ROBERT H. COBEAN

ABSTRACT

A systematic approach to sample collection, chemical analysis, and statistical evaluation of obsidian sources is recommended before significant numbers of artifacts should be analyzed. Multivariate statistical methods based on Mahalanobis D^2 enable robust rejection of erroneous source assignments and lead to sourcing of artifacts with a very high degree of confidence. These multivariate procedures can also assist in identifying and evaluating abbreviated analytical methods that are more rapid and less expensive. Test cases based on obsidian data from sources in Mesoamerica are presented.

INTRODUCTION

Archaeologists seeking to reconstruct past human behavior and interaction often examine the materials that early peoples made and used. Lithic

---

MICHAEL D. GLASCOCK • Research Reactor Facility, University of Missouri, Columbia, Missouri. **GEOFFREY E. BRASWELL** • Department of Anthropology, State University of New York at Buffalo, Buffalo, New York. **ROBERT H. COBEAN** • Instituto Nacional de Antropología e Historia, Mexico City, Mexico.

Archaeological Obsidian Studies, edited by Shackley.
Plenum Press, New York, 1998.

materials are undoubtedly the most widely studied of all archaeological materials because: (1) they are found practically everywhere that prehistoric peoples lived or traveled; (2) they are well preserved over time; and (3) they occupied an important role in the material culture of most prehistoric peoples. Although nearly every type of stone was used by one prehistoric culture or another, the lithic material providing archaeologists with the most definitive evidence concerning contact between different cultures is the volcanic glass obsidian.

Obsidian, with a composition similar to rhyolite, is formed when highly viscous molten lava cools rapidly such that the process of crystallization is precluded. The glass is usually black or gray in color (although other colors are possible) and is sometimes banded or streaky in appearance. As a glass, obsidian has physical properties that readily yield sharp-edged, conchoidal-shaped fractures. Its glassy consistency, predictable fracturing qualities, and sharp working edge contributed to obsidian being a highly sought-after material. Obsidian generally makes better edges than most other types of tool-making stones. In regions where obsidian was available, it was often the most commonly used material for making sharp-edged implements, weapons, and jewelry. Obsidian raw materials and finished products were valuable commodities that were often exchanged over long distances from their sources.

There are two types of obsidian sources: primary and secondary. Primary sources consist of lava flows or pyroclastic bomblets surrounding volcanic cones. Secondary sources (or deposits) are the result of the erosional activities of glaciers, streams, gravity or other geologic processes that involve the transport of primary material. Due to the abundance of raw material, most prehistoric quarries were located near the primary sources.

Obsidian artifacts have been used to examine resource procurement patterns, manufacturing processes, distribution networks, and to develop use-site chronologies (Elam 1993). In addition, hydration-rim measurements of obsidian artifacts can be used to investigate chronological changes in socioeconomic conditions (Freter 1992).

Clark (1981) identifies several attributes of obsidian that give it a unique role in archaeology:

1. *The number of sources is limited.* Because nearly all pre-Tertiary obsidians have lost their glassy properties, obsidian is generally restricted to the tectonic regions that experienced vulcanism since the beginning of the Tertiary period. In addition, not all lava flows are sufficiently rich in silica (e.g., Hawaii) for glass formation to take place. The most important obsidian bearing regions are the Mediterranean, eastern Africa, central Mexico, South America, western North America, Japan, and New Zealand.
2. *Obsidian objects are found in a variety of locations and contexts.* Obsidian artifacts are found in a much wider distribution area than that of

their sources. Obsidian artifacts are found in many villages, rural locations, elite areas, and ceremonial sites that are often hundreds of kilometers removed from their originating sources.

3. *Vast amounts of debitage are found near the ancient quarries.* The large amounts of waste obsidian found near sources help archaeologists locate prehistoric quarries and make it possible to estimate the quantities of raw obsidian excavated and exchanged.
4. *Artifacts retain many of their physical features from fabrication.* Scars, platform angles, and other distinguishing marks on artifacts help archaeologists to investigate the production techniques, the skills of prehistoric craftsmen, and the probable functions of the tools.
5. *In most archaeological contexts, obsidian artifacts are indestructible.* Thus, archaeologists have confidence that artifacts they are investigating closely resemble the tools or trinkets that prehistoric peoples were using before they were either cached, lost, or discarded.
6. *The fragile nature of obsidian means that it had a high replacement rate.* Examining the quantities of artifacts at use sites and the rate of consumption provides information through which prehistoric economic conditions can be estimated.
7. *The chemical composition of obsidian at any particular source or flow is, with few exceptions, homogeneous and different sources or flows are compositionally different from each other.* The fortunate circumstance that obsidian, by virtue of having been molten during the period of volcanic activity, is well mixed material means that sources are typically homogeneous. The homogeneous character of most obsidian sources and the fact that different sources are customarily dissimilar in composition is a fundamental attribute of obsidian that enables successful differentiation between sources.
8. *A new obsidian surface will begin absorbing water gradually with time.* The hydration process offers archaeologists a technique, through microscopic examination of obsidian thin sections, by which artifacts can be dated and compared to one another to learn more about the chronology of an archaeological site. The only qualification to this is that because obsidian from chemically dissimilar sources and subsources hydrates at different rates, each artifact must be traced to its source before its date of manufacture can be determined.

In order to take advantage of these and other attributes of obsidian for archaeological examination, archaeologists have explored a variety of methods to characterize obsidian for sourcing purposes. For archaeological purposes, adequate characterization entails identifying a property or set of properties that are the same for obsidian specimens from the same source and different between specimens from different sources. In other words, archaeologists seek to observe or measure attributes of obsidian that are both homogeneous within sources and heterogeneous between them.

SOURCING METHODS

Archaeologists have long sought methods for obsidian sourcing that are rapid, reliable, non-destructive, and low-cost. As a result, a variety of methods have been explored to identify characteristics of obsidian that can be employed to associate artifacts with sources.

Visual appearance is very likely one of the first properties of obsidian to have been investigated as a method for sourcing (Fuller 1927). It is possible to describe obsidian in terms of several visual traits, including: color in transmitted light, color in reflected light, refractive index, success of fracture, opacity, internal structure, and luster. However, appearance alone is not a reliable guide by which the source of obsidian can be determined. Because assessment of the visual attributes of individual obsidian specimens is a consequence of one's subjective judgement, appearance is often incapable of being employed reliably by different individuals as a method for obsidian differentiation. Assorted colors (black, gray, clear, reddish-black, reddish-brown) and various degrees of banding or mottling often occur within a single obsidian deposit. In spite of these inadequacies, many modern-day archaeologists still employ visual methods with marginal to moderate success (Bettinger *et al.* 1984; Clark 1988; Moholy-Nagy and Nelson 1990; Jackson and Love 1991). In general, only when supported by other types of analyses can meaningful statements be made about obsidian materials on the basis of appearance.

Several other previously investigated methods for obsidian sourcing include: (1) density measurements (Reeves and Armitage 1973); (2) thermoluminescence (Huntley and Bailey 1978); (3) Mössbauer spectroscopy (Longworth and Warren 1979); (4) fission-track analysis (Duranni *et al.* 1971); (5) measurement of magnetic properties (McDougall *et al.* 1983); and (6) measurement of natural radioactivity (Leach *et al.* 1983). Although some of these methods identified systematic differences between sources, the overlap between different sources is such that their reliability for differentiation purposes is low, and none show much promise.

Compositional Analysis

The chemical composition of most volcanic obsidian ranges from about 70–75% SiO_2, 10–15% Al_2O_3, 3–5% Na_2O, 2–5% K_2O and 1–5% total Fe_2O_3 + FeO. Peralkaline obsidians are typically higher in Fe composition than rhyolitic obsidians. In addition, the intrinsic water content of obsidian ranges from 0.1–0.5% and water content increases to about 3.5% by weight as rhyolitic obsidian gradually transforms into a less useful form of the glass known as perlite. Most of the remaining elements in obsidian are present in concentrations of much less than 1% and are therefore referred to as trace elements.

The success of chemical sourcing is due to the fact that these trace element compositions may differ by one or two orders of magnitude between sources, while within-source variation is usually much smaller. In a few cases, within-source variation in elemental composition is considerable (e.g., Bowman *et al.* 1973). Nevertheless, the correlation in variation between certain elements is so extraordinary that the ability to assign artifacts to the source is just as certain as if the flow were homogeneous. Other sources have been discovered where a collection of individually homogeneous and discrete flows have been reported (e.g., Hughes 1988). Obsidian artifacts from these sources can be assigned to specific subsources.

Over the past three decades, many analytical techniques have been used to chemically source obsidian artifacts. These methods include: (1) optical emission spectroscopy (OES), used to examine obsidian from the Mediterranean (Cann and Renfrew 1964); (2) atomic absorption spectroscopy (AAS), used to characterize obsidian from Alaska (Wheeler and Clark 1977); (3) particle-induced X-ray emission spectroscopy (PIXE), used to study obsidian from New Zealand (Duerden et al. 1984); (4) inductively-coupled plasma emission spectroscopy (ICPES), used to analyze obsidian from New Mexico (Stevenson and McCurry 1990); (5) X-ray fluorescence spectroscopy (XRF), used to characterize obsidian from northern California (Hughes 1982); and (6) neutron activation analysis (NAA), used to study obsidian from the highlands of Guatemala (Asaro *et al.* 1978).

In order to be successful, a technique for chemical fingerprinting must be quantitative, capable of simultaneously measuring several elements, sensitive to the elements of interest, independent of sample matrix, and independent of artifact size and shape. In addition, the choice of a method for analysis may depend upon its availability, cost, speed, accuracy, ability to differentiate between sources, existence of comparative data, etc. XRF and NAA have proven to be highly cost effective and, therefore, are the methods most frequently used to source artifacts.

In XRF, a beam of X-rays irradiates the specimen causing displacement of atomic electrons from the inner energy levels. As electrons from outer (higher) levels repopulate these vacant inner (lower) levels, energy is emitted in the form of fluorescent X-rays. Because the energy levels are unique for each element, these X- rays have characteristic energies that permit identification of the element. By measuring the intensities of emitted X-rays, one can determine the quantities of elements present in the specimen. In obsidian, the elemental compositions of Na, K, Ti, Mn, Fe, Rb, Sr, Y, Zr, Nb, and Ba are frequently determined by XRF.

The principles of NAA differ from XRF in that samples are irradiated by thermal neutrons (usually from a nuclear reactor). During irradiation, a few neutrons are captured by the nuclei of atoms in the specimen. This process, called activation, causes some of the nuclei to become unstable. During and after neutron irradiation, these unstable nuclei emit gamma rays with unique energies at rates defined by the characteristic half-lives of the radioactive nuclei. Identification of the radioactive nucleus is possible by measuring the gamma-ray energies. Determination of their intensities permits quantitative analysis of the elements in the sample. Use of NAA on obsidian specimens by a combination of two or three irradiation, decay, and measurement strategies enables determination of the elements Ba, Ce, Cl, Co, Cs, Dy, Eu, Fe, Hf, K, La, Lu, Mn, Na, Nd, Rb, Sb, Sc, Sm, Sr, Ta, Tb, Th, U, Yb, Zn, and Zr. If necessary, the sensitivities of different subsets of these elements can be enhanced by changing the irradiation, decay, and measurement schemes.

In general, XRF is more available, more rapid, and less expensive than NAA. NAA offers a far greater number of elements, more sensitivity, superior precision, and greater accuracy than XRF. In XRF, the artifact can be studied without destruction. On the other hand, NAA can be performed on extremely small samples (~10 mg), meaning that only minor damage to valuable artifacts may be required.

PREVIOUS WEAKNESSES IN INTERPRETING COMPOSITIONAL DATA

A review of previous obsidian research reveals that, in some studies, only a fraction of the information potentially available was recovered. In other instances, poorly planned studies came to conclusions that were later refuted. Most problems involved the unsatisfactory characterization of sources (i.e., failure to locate all possible sources, collection and analysis of too few specimens from each source, poor descriptions of source areas, analysis of samples from modern road cuts rather than prehistoric quarries, analysis of too-few elements, and failure to identify the most critical elements useful for discriminating between particular regions or individual sources).

Another fallacy is the anticipation by a few archaeologists that the same elements successful in differentiating the sources in one region will succeed in others. Unfortunately, one cannot know in advance which elements will be best or which analytical technique will be most efficient for a particular region until after the sources have been thoroughly characterized. Other problems pointed out by Hughes (1984) are due to the uncritical acceptance of the results of artifact sourcing studies analyzed by inadequate statistical procedures. These include various analytical or statistical shortcuts with unproven reliability.

Although many thousands of obsidian-source and artifact analyses have been made throughout the world, no single databank exists that makes available all of the compositional and descriptive data. Instead, individual research groups have concentrated on sources in particular regions, sometimes in complete isolation from other groups working in the same or neighboring regions. The use of different analytical calibrations by different analytical laboratories makes data exchange problematic. As a result, exchanges of data and source samples between laboratories are a rare rather than routine occurrence.

One example illustrating several of these problems is the NAA work of Pires-Ferriera (1973) which dealt with a large collection of obsidian artifacts found in the Valley of Oaxaca. In her study, Pires-Ferriera employed a procedure that measured two short-lived elements (i.e., Mn and Na). Although her work was mostly successful, Pires-Ferriera assigned a large fraction of artifacts to the source at Altotonga which more recent data calls into question. From more recent work reported by Cobean *et al.* (1991), we now realize that Altotonga obsidian overlaps perfectly with the nearby Zaragoza source on the basis of Mn and Na. However, both sources are easily differentiated by a number of other elements measured in the Cobean *et al.* (1991) study. Recent research by Elam (1993), reporting six elements from an analysis of ~500 obsidian artifacts from the Valley of Oaxaca, identifies Zaragoza as one of the most important source supplying obsidian to Oaxaca. Not a single artifact in this recent study was traced to Altotonga. Pires-Ferriera also suggested that several unassigned artifacts came from an unknown Oaxacan source that she was unable to locate. The existence of this hypothetical source is disputed by more comprehensive data (Glascock et al. 1988) indicating that the unassigned artifacts very likely came from sources in the state of Hidalgo that Pires-Ferriera did not adequately characterize. Thus, on the basis of minimal data, questionable source assignments for obsidian artifacts from Oaxaca were accepted for nearly twenty years.

A study of obsidian artifacts from San Lorenzo Tenochtitlan, Veracruz by Cobean *et al.* (1971) also reported the similarity of the Otumba, Mexico source to sources in Guatemala (over 1,000 km to the south) for elements commonly measured by XRF. Later, Charlton et al. (1978), using a small suite of elements, found chemical similarity for the obsidian source at Paredon (~50 km from Otumba) with the Otumba source. A larger suite of elements (Cobean et al. 1991) now enables identification of inter- and intra-regional differences between these and other sources. As a result, it is quite probable that the artifacts identified as the Unknown A group in Cobean's earlier study were actually made of Paredon obsidian.

Other examples of studies with inadequate source information involve work in the Andean region of South America. There, hundreds of artifacts have been analyzed and results suggest the existence of 25 or more sources (Avila-Salinas 1975; Burger and Asaro 1977). Thus far, less than ten sources have been spatially located and chemically characterized in the entire Andes region. Thus, a great deal of work still remains before substantial reconstruction of the obsidian exchange networks in South America will be possible.

As Hughes (1986) points out, some researchers (e.g., Sappington 1981, 1984) have failed to consider the significance of particular elements and, instead, assigned equal weight to all measured elements. Unfortunately, this practice may result in an increased number of misclassifications. Hughes suggests that a more critical approach to variable selection in multivariate applications can reduce the number of errors. Finally, Leach and Manly (1982) argue that the power of any sourcing algorithm to reject wrong answers should be demonstrated.

A MORE SYSTEMATIC APPROACH

In order to make reliable source attributions, researchers must: (1) locate each source precisely and obtain a representative range of source specimens from all sources; (2) demonstrate that flake-quality obsidian can be obtained from the source; (3) verify that the source was accessible to prehistoric peoples and it was not exposed by more recent (historical period) mining or construction activities; (4) locate prehistoric quarries or identify materials at an archaeological site that originated from the primary source area or a secondary deposit; (5) analyze the source specimens thoroughly to determine the number of different chemical fingerprints represented; (6) correlate the source fingerprints to specific source locations; and (7) show that artifacts assigned to a particular source cannot be mistaken for those from another source. A systematic program of sampling, chemical analysis, and statistical evaluation is required to accomplish these goals.

A MORE SYSTEMATIC APPROACH

In order to make reliable source attributions, researchers must: (1) locate each source precisely and obtain a representative range of source specimens from all sources; (2) demonstrate that flake-quality obsidian can be obtained from the source; (3) verify that the source was accessible to prehistoric peoples and it was not exposed by more recent (historical period) mining or construction activities; (4) locate prehistoric quarries or identify materials at an archaeological site that originated from the primary source area or a secondary deposit; (5) analyze the source specimens thoroughly to determine the number of different chemical fingerprints represented; (6) correlate the source fingerprints to specific source locations; and (7) show that artifacts assigned to a particular source cannot be mistaken for those from another source. A systematic program of sampling, chemical analysis, and statistical evaluation is required to accomplish these goals.

Sampling and Chemical Analysis

Before the benefits of artifact sourcing can be fully realized, it is necessary

The number of source specimens collected and analyzed should be sufficient to enable determination of the range of compositional variation in the source. In general, we recommend that at least a dozen source samples be analyzed, although for large and complex sources more than 100 specimens may be desirable. If possible, specimens should be taken from the actual locations of prehistoric quarries rather than from road cuts or other sites that are the consequence of modern human activity. Additional specimens should be collected beyond those slated for analysis in order to maintain samples for future reference, exchange with other laboratories, and conduct possible hydration-rate experiments. If the amount of compositional variation in particular sources proves to be greater than anticipated or if artifact analyses indicate the existence of undiscovered sources, additional field surveys and collections should performed (Shackley 1988).

A comprehensive data bank or source catalog should be compiled to store all of the descriptive, physical, and chemical information gathered. Previous efforts by Skinner (1983) and Hughes (1986) were an initial step in this direction. More recently, members of the International Association for Obsidian Studies (IAOS) have started collecting information for a computerized database that would organize all known descriptive and chemical information on obsidian sources. Informational categories urged for each source include: (1) geographic location and name(s) of source; (2) type of source (either primary or secondary); (3) geologic setting of the source area; (4) evidence of quarrying or chipping activity and the quality of the obsidian for tool making; (5) visual characteristics (both macroscopic and microscopic) of the glass; (6) refractive index of the glass; (7) age of source and method of determination (e.g., K-Ar, fission-track); (8) source specimens were collected from flows, bomblets, or nodules; (9) names of persons who are knowledgeable of the source area; (10) evidence from use sites of artifacts that came from the source area; (11) types of chemical analyses performed; (12) a tabulation of all compositional data (not only averages); (13) results from hydration rate measurement investigations; (14) a listing of relevant publications.

Source specimens should be analyzed comprehensively by all available analytical techniques (preferably both NAA and XRF) and in sufficient numbers to establish the concentration ranges, means, and standard deviations for each source. Use of well-known standards such as the National Institute of Standards and Technology (NIST) SRM-278 Obsidian Rock and SRM-1633a Fly Ash are recommended for calibration (Graham *et al.* 1982). In addition, routine quality-control samples should be analyzed and rigorously monitored between analytical runs in order to certify that the data are as accurate and precise as reasonably possible (Glascock and Anderson 1993). These steps will insure that data from separate batches and different laboratories are comparable and bias is minimized.

Statistical Evaluation of Obsidian Source and Artifact Data

Most researchers recognize the wisdom of collecting many source specimens and analyzing for several elements to achieve a statistically significant differentiation between two sources. Unfortunately, as Ward (1974) and, more recently, Hughes (1986) have pointed out, it is difficult to visualize more than two or three elements at once during statistical analysis. For example, low-dimensional plots (bivariate scattergrams and ternary plots) of the measured elements are the usual way most researchers display and examine their data. When considering a small number of sources, this technique may be satisfactory. However, when the number of sources in a geographic region is quite large, such as in central Mexico (~20 sources) or in the California-Oregon-Nevada region (100+ sources), low-dimensional plots of raw elemental concentrations are frequently inadequate. This can be particularly problematic in XRF where the number of high-precision, element-concentration pairs is small. As a result, the failure to consider multivariate methods of data analysis and the unsystematic use of source and artifact data has led to a number of erroneous source assignments.

One of the most notable examples of mistaken source assignments concerning failure to use multivariate methods concerns the Chatham Island artifacts. Using XRF data and simple univariate methods, Leach (1973) mistakenly concluded the artifacts to be of Mayor Island origin, located approximately 600 km to the west of the Chatham Islands. Subsequent use of more comprehensive NAA data and multivariate methods found the Chatham Island artifacts to be from Easter Island located almost 7,000 km to the east (Leach and Warren 1981).

Some researchers (Perlman and Asaro 1969; Wilson 1978) have argued against consideration of correlated elements because little extra discrimination between sources is achieved by analyzing a number of correlated elements. However, the concentrations of various elements in obsidian and other geological materials are far from independent from one another. Hence, the practice of not considering correlated elements results in a loss of potentially useful information. For example, Bowman *et al.* (1973) demonstrated that in spite of the existence of extreme correlations in compositional data from Borax Lake obsidian their data interpretations are just as definitive as if the correlations had not existed. Multivariate statistical procedures described below permit fully accounting for interelement correlation without sacrificing the useful information from all measured elements.

MULTIVARIATE ANALYSIS

The main objectives of statistical analysis when applied to compositional data are data exploration, hypothesis generation, hypothesis verification, and data summarization. Ward (1974), Sayre (1975), Bieber *et al.* (1976), Leach and Manly (1982), Bishop and Neff (1989) and others have pioneered the use of the multivariate analytical techniques in order to achieve these objectives in archaeometric studies. Before describing the statistical procedures employed by the Archaeometry Laboratory at the Missouri University Research Reactor (MURR), a brief review of the terminology and techniques employed in multivariate analysis is presented. More detailed explanations are available in Sayre (1975), Bieber *et al.* (1976), Davis (1986), and Bishop and Neff (1989), and Neff (1994).

Sample Distributions and Outlier Specimens

Previous authors (e.g., Picon *et al.* 1975; Bieber 1976) have differed over whether concentration data for geological materials are distributed normally or lognormally. Fortunately, as Ahrens (1954) has noted, if the dispersion (or spread) in the data is small compared to the mean concentration, the lognormal distribution approximates a normal distribution. The assumption of lognormality is also convenient because log transformations compensate for the large differences in magnitude between the high-concentration elements (such as K, Fe, and Na) and the trace elements (e.g., rare earths or REEs). Thus, log transformation creates a convenient *quasi-standardization* (Sayre 1975) or rescaling of the original data that facilitates the application of various multivariate methods.

Specimens that chemically differ from all others and have a disruptive influence on the shapes of sample distributions are commonly known as outliers. During the application of multivariate methods, preliminary screening and group refinement procedures are used to eliminate statistical bias caused by outlying specimens (Bishop and Neff 1989).

Euclidean Hyperspace, Variance-Covariance, and Correlation

Multivariate statistical techniques are based on the idea that each specimen can be considered a point in multidimensional space (hyperspace), with the number of dimensions equivalent to the number of elements measured. In such a hyperspace, groups of specimens with similar compositions form a cloud of points surrounding a mathematically calculable "center of mass." An individual compositional group is characterized by the location of its centroids and by the unique hyperellipsoidal shape of the group distribution (Sayre 1975). The shape is due to the correlations between different elements. Distinct source groups are represented by two or more clouds that do not overlap.

As the dispersion in concentrations for an element increases, it becomes more difficult to use the element to differentiate between source groups because they will have an increasing tendency to overlap with one another. A parameter that quantifies dispersion in compositional data is the measured variance, σ_m^2, in the mean element concentration. The measured variance represents the sum of the natural variance in group composition, σ_n^2, and the variance due to sampling and analytical errors, σ_a^2, as defined by expression (1).

$$\sigma_m^2 = \sigma_n^2 + \sigma_a^2 \tag{1}$$

And it is equal to the square of the standard deviation, σ_m, of the group's mean concentration. In analyses of obsidian source data by NAA, we have found that about half of the elements are measurable with standard deviations of 2–5%.

Just as the sample variance is a measure of the spread of concentration values for an element around a central mean, the sample covariance is a measure of the mutual variability of a pair of elements around their common mean. The sample covariance between any two elements j and k can be calculated from expression (2).

$$\sigma_{jk} = \frac{\sum_{i=1}^{n}\left(C_{ij} - A_j\right)\left(C_{ik} - A_k\right)}{n-1} \tag{2}$$

where C_{ij} and C_{ik} denote the respective concentrations of the jth and kth elements in specimen I. The mean element concentrations are given by A_j and A_k, and n is the total number of specimens under consideration. Although the covariance is somewhat like a correlation coefficient, it has not been standardized for the magnitudes of the elements. Therefore, the covariance can take on any value and is not limited to the range from −1 to +1.

For computational purposes, the covariances between elements are arranged in the form of a square matrix with one row and one column for each element. The intersection of each row and column contains the covariance for the corresponding pair of elements. The diagonal entries contain the variances, since the covariance for an element against itself is the same as the variance. Appropriately, the matrix is called the variance-covariance matrix.

Finally, the correlation coefficient, r_{jk}, between any two elements j and k is defined by the ratio in expression (3).

$$r_{jk} = \frac{\sigma_{jk}}{\sigma_j \cdot \sigma_k} \tag{3}$$

where σ_j and σ_k are the respective sample standard deviations, and σ_{jk} is the sample covariance between the elements. Correlation coefficients range in value from +1 to −1, with +1 indicating a perfect relationship, 0 indicating no relationship, and -1 indicating an inverse relationship between elemental compositions. For obsidian it is not uncommon to find correlation coefficients of 0.9 or higher for certain elements (e.g., REEs).

Bivariate Analysis

As stated above, two-dimensional scattergrams are frequently used to make visual distinctions between obsidian source-groups. In addition, highly-correlated element pairs and outliers can sometimes be identified through inspection of bivariate plots. Confidence ellipses (or probability intervals surrounding source groups) are frequently calculated and drawn at a constant Mahalanobis distance from group centroids. In regions where there are relatively few possible sources, bivariate plots often permit unambiguous assignment of artifacts to sources. Bivariate plots can be used to graph original concentration data or linear combinations of the data such as principal components or discriminant functions.

Cluster Analysis

Cluster analysis is often used in the initial inspection of obsidian-source data because it is a rapid and efficient technique for evaluating relationships between large numbers of samples, between which distance measures have been calculated. Some investigators resort to cluster analysis routinely (Hurtado de Mendoza and Jester 1978; Blackman 1984). Cluster analysis allocates samples into distinct groups. The results are commonly presented as dendrograms showing the order and levels of clustering, as well as the distances between individual samples.

Cluster analysis 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; Harbottle 1976). One of the most frequently employed distance measures is the squared-mean Euclidean distance (SMED), defined between specimens *j* and *k* according to equation (4).

$$d_{jk}^2 = \frac{1}{n}\sum_{i=1}^{n}\left[C_{ij} - C_{ik}\right]^2 \tag{4}$$

where the scaling factor, *n*, corresponds to the number of elements actually measured.

Most clustering procedures employ Euclidean distances between individual samples in the calculation of similarity or difference. In doing so, they assume that the elements under consideration are uncorrelated (hyperspherical). Consequently, dendrograms may accurately represent the differences between within-group members, but greatly distort the differences between clusters of samples (Sneath and Sokal 1973; Sayre 1975; Harbottle 1982). Thus, the effectiveness of dendrograms as a means of faithfully illustrating differences between groups is questionable. Dendrograms can be used as an initial step in the identification of groups, but they are inadequate as a quantitative measure of the ability to differentiate between groups. Fortunately, as illustrated below, there are other procedures that enable a more reliable determination of both within- and between-group differences.

Principal Components Analysis

One of the more widely utilized techniques for exploring underlying relationships in multivariate data is factor analysis. In particular, the factor analytical approach known as principal components analysis (PCA) is employed when the main objective is to derive a set of uncorrelated linear combinations of the original variables that retain the original information. If the original variables are significantly correlated, a reduced number of principal components may be adequate to describe the significant features of the original dataset.

In PCA, one performs a transformation of the data using eigenvector methods to determine the direction and magnitude of maximum variance in hyperspace (Davis 1986). Computation of the eigenvectors is derived from either the variance-covariance or correlation matrix of the original variables. The first principal component (PC) is oriented in the direction of maximum variance in the data. The second PC lies in the direction of the maximum remaining variance, with the additional constraint that it must be orthogonal to the first PC. The third PC is orthogonal to the first two PCS, again representing the direction of maximum remaining variance. The procedure continues until the number of PCS is equal to the number of original dimensions. All resulting PCS are then linear combinations of the original axes. The loading factors (coefficients) used to construct the PCS can be used to identify attributes important to the individual components.

The PCA technique can be used to describe the entire distribution of elemental concentrations in a compositional dataset on a different basis. The positions of samples in the original element concentration space are converted into principal scores calculated in terms of the linear combinations of the original data and are projected onto the new principal component axes. For correlated data, a larger portion of the data structure (often >90% of the total population variance) can be inspected from two- or three-dimensional plots using the first few PCS than is possible with most bivariate plots of the original elements. Thus, PCA eliminates repetitive data and finds a new basis which facilitates dimensionality reduction. If PCA provides a satisfactory differentiation between source groups, it is often possible to assign artifacts by projecting their principal component scores against the principal axes.

The PCA approach is widely employed in archaeometric studies of ceramics (Neff *et al.* 1988) and manmade glasses (Baxter 1991) but has only rarely seen application in obsidian studies (Stark *et al.* 1992). However, PCA can be especially useful to understanding the chemical basis for group separation and in identifying original elements that by themselves may be satisfactory to describe the most significant compositional differences between particular obsidian source groups.

In their discussions of taxonomic methods, Cattell (1952) and Sneath and Sokal (1973) point out that multivariate factor analysis can be used in two ways. Associating variables over all specimens is called a *Q-mode* technique. On the other hand, associating specimens over all variables is called an *R-mode* technique. The emphasis of most PCAs has been the use of R-mode techniques with displays of variable scores plotted in R-space to reveal the contributions of individual variables. However, applications involving Q-mode techniques on compositional data are becoming more common (Neff *et al.* 1988).

Recent articles by Baxter (1992) and Neff (1994) demonstrate a simultaneous RQ-mode PCA as a means of examining relationships between specimens and between elements simultaneously. This technique of presenting R-mode and Q-mode plots together is called *biplotting* (Gabriel 1971). The simultaneous use of both Q-mode and R-mode plots allows the individual contributions of all elements to group separation to be displayed.

Discriminant Function Analysis

In multivariate analysis, the term discriminant analysis is frequently used to describe two different processes: *discrimination* and *classification*. The process of discrimination involves identifying a mathematical transformation of the original variables that best reveals the differences between known groups of observations. On the other hand, the process of classification involves the categorization of a number of observations into known groups.

Discriminant function analysis (DFA) is a multivariate technique differing from PCA in that it is used to summarize relationships between groups. DFA is based on the assumption that the pooled variance-covariance matrix is an accurate representation of the total variance and covariance of the data (Davis 1986). When employing DFA, one assumes that all specimens in the dataset necessarily belong to one of the known source groups.

Two computational methods can be used to derive discriminant functions: *concurrent* or *stepwise*. The concurrent method finds a set of linear combinations of the original variables that best summarizes between-group differences by considering all variables simultaneously, regardless of their individual discriminating powers. The concurrent method computes a series of discriminant functions (one less than the number of source groups under consideration) that minimize the ratio of the difference between group multivariate means to the multivariate variance within each group. As a result, the procedure maximizes the total score of specimens belonging to their assigned group and minimizes the total score of belonging to all other groups. The resulting discriminant functions are linear combinations of the original data each successively describing lesser amounts of separation between compositional groups.

In the stepwise method, the variables are entered into the discriminant function one at a time on the basis of their discriminating power. The stepwise method begins by identifying the single best discriminating variable. The initial variable is then paired with all others in order to identify the next best discriminating variable. This process is successively repeated for all remaining variables. A reduced set of most discriminating variables, is often as good at differentiating as the entire suite of variables. In their studies of obsidian from western North America and the Maya lowlands, Nelson (1985) and Hughes (1986) used stepwise discriminant analysis procedures available in the SPSS statistical package.

Bivariate plots of discriminant functions are useful for visually displaying group separation. In the same manner as PCA, one can employ discriminant functions to calculate the discriminant scores for artifacts. By comparing artifact scores with scores for source specimen groups, decisions regarding source assignments for the artifacts can be made.

Mahalanobis Distances and Probability Calculations

Discriminant analysis for classification purposes and related techniques is based on the standardized-squared distance or *Mahalanobis distance*. The Mahalanobis distance (D^2) statistic is defined as the squared Euclidean distance between a group centroid and an individual specimen divided by the group standard deviation in that direction. Mathematically, the D^2 from specimen k to the centroid of cluster A is given by expression (5).

$$D^2_{kA} = \sum_{i=1}^{n}\sum_{j=1}^{n}\left[C_{ik} - A_i\right]\cdot I_{ij}\cdot\left[C_{jk} - A_j\right] \tag{5}$$

where A_i and A_j are the mean concentrations of elements I and j in the cluster and I_{ij} is the *ij*th element of the inverse of the variance-covariance matrix. The Mahalanobis distances from a specimen to each of several alternative group centroids can be used to estimate the relative probabilities of membership in each group (Bishop and Neff 1989). The D^2 statistic can also be used to measure the separation distance between pairs of multivariate means in units of pooled variance (Davis 1986).

As explained by Klecka (1980) and Hughes (1986), the classification aspect of discriminant analysis is based on a number of statistical assumptions:

(1) there are two or more groups; (2) there are at least two observations in each group; (3) the number of discriminating variables must be at least two less than the total number of specimens; (4) the discriminating variables are measured at the interval level; (5) the discriminating variables must not be linear combinations of other discriminating variables; (6) each group is drawn from a sample population with a multivariate normal distribution on the discriminating variables; and (7) the variance-covariance matrices for each group must be approximately equal. Of these assumptions, the normality of multivariate distributions and homogeneity of group variance-covariance matrices are the most important and difficult to satisfy in practice (Sneath and Sokal 1973).

The assumption that each group is drawn from a multivariate normal distribution is critical for tests of significance. Significance tests are computed by comparing a statistic calculated for a sample relative to a theoretical probability distribution for that statistic. If the sample population does not satisfy this requirement, the distribution for the sample statistic will be different from the theoretical distribution resulting in some inaccuracy. As long as the deviation is minor, Lachenbruch (1975) has shown that the discriminant analysis is not particularly sensitive. The assumption of multivariate normality is also important for classification based on relative probability of group membership. Probabilities are calculated from a chi-square distribution which is appropriate only if the variables are normally distributed. If the normality assumption is violated, probabilities for membership in some groups may be overstated while the probabilities for other groups may be underestimated. According to Ahrens (1954), the evidence from basic statistics (i.e., means, ranges, etc.) seems to indicate that most geochemical data distributions are approximately normal (or lognormal) and the multivariate normality assumption is rarely violated.

The more serious assumption concerns the homogeneity of group variance-covariance matrices. If the matrices are not equal, distortions in the classification equations will occur such that specimens have a greater likelihood of being allocated into groups with greater dispersion as measured by the determinant of the variance-covariance matrix (Klecka 1980). Leach and Manly (1982) employed a power transformation to alleviate this problem in obsidian data from the New Zealand and Pacific region. Fortunately, as Klecka (1980) points out, for classification purposes violation of the homogeneity assumption is only important when comparing probabilities that are nearly equal in magnitude. As a result, cautious consideration of similar probability values is mandatory when making classifications.

The D^2 statistic, unlike the simple Euclidean measure, incorporates information about the correlations between pairs of elements as derived by the off-diagonal terms of the variance-covariance matrix. Thus, it permits calculation of the probability that a particular specimen belongs to a group based not only on the proximity of the specimen to the group centroid, but also on the rate at which the density of specimen data points decreases away from the group centroid in the direction of the specimen of interest.

Calculation of probabilities using D^2 requires that the number of specimens defining the source groups exceed the number of elements under consideration by at least one. Probability calculations with small-sized groups suffer from a problem that Harbottle (1976) calls *stretchability* because of the tendency of individual specimens to stretch the envelope so as to include that specimen. Excluding the specimen from the group before calculating its probability of membership helps to eliminate this problem. Nevertheless, it is recommended that each compositional group consist of three to five times as many specimens as the number of elements under consideration (Harbottle 1976).

The significance of differences between two groups of specimens can be tested by Hotelling's T^2 statistic (the multivariate equivalent of the Student's t) according to expression (6).

$$T^2 = \frac{D^2}{\dfrac{1}{n_1} + \dfrac{1}{n_2}} \tag{6}$$

where n_1 and n_2 are the number of specimens in each of two groups. Hotelling's T^2 statistic is essentially equivalent to D^2 for individual data points. If ν is the number of elements considered in the analysis, the T^2 statistic can be transformed into an F ratio by expression (7).

$$F = \frac{(n_1 + n_2 - \nu - 1)}{(n_1 + n_2 - 2)\nu} T^2 \tag{7}$$

This is easily transformed into a probability percentage using tables of the F distribution.

Finally, as Sayre (1975) recommends, Mahalanobis distance calculations are quite useful for handling missing data. When large numbers of specimens are analyzed, there is almost a certainty that a few concentration measurements will be missed for some specimens. This occurs most often when the group mean concentration for an element is near the detection limit. For example, the elements Sr and Ba are near or below the NAA detection limit for a number of obsidian sources. Rather than completely eliminate specimens with occasional missing values from consideration, it is possible to substitute by computing a value that minimizes the Mahalanobis distance for that specimen from the group centroid. As a result, the substitute concentration values for these few specimens will have little effect on the statistical characteristics of the group and enables them to be included in all group calculations.

EXAMPLES FROM THE MURR MESOAMERICAN OBSIDIAN DATABASE

To demonstrate the approach advocated here, we present examples from our experience in building a database for obsidian source materials from Mesoamerica. Braswell and Cobean were actively involved in acquiring the source specimens and Glascock primarily handled the data analysis and statistical interpretation. Subsequently, artifact analyses have been performed in collaboration with more than two dozen different colleagues interested in Mesoamerica. As a result of the source characterization work, nearly 4,000 source specimens and artifacts from Mesoamerica have been analyzed and sourced in the MURR Archaeometry Laboratory.

Obsidian Sources in Central Mexico

One of the most ambitious obsidian source characterization studies carried out to date was initiated in 1980 by Cobean with support from the National Science Foundation and the University of Missouri. In collaboration with other archaeologists from Mexico's Instituto Nacional de Antropologia e Historia (INAH), Cobean collected more than 800 obsidian source specimens (weighing 710 kg) from the volcanic sources located throughout central Mexico shown in Figure 2.1.

The main purpose of Cobean's project was to establish a reference sample repository and geochemical data bank for Mexican obsidian. In addition to collecting samples for analysis at MURR and for future reference, portions of the source specimens were sent to Fred Nelson (Brigham Young University) for XRF analysis and Jonathon Ericson (California-Irvine) for hydration rate measurements. A major goal was to increase the confidence with which obsidian artifacts from archaeological sites could be associated with specific source areas. A final goal of the project was to employ a detailed statistical analysis to identify the most discriminating elements between regions, sources, and, if possible, different flows or outcrops from the same source. Identification of highly discriminating elements might facilitate the development of more rapid and less expensive analytical procedures for artifacts that are as reliable as the comprehensive analysis applied to characterize source specimens.

Several of the obsidian sources from which Cobean collected samples extend over vast areas greater than 300 km^2. An intensive field survey mapped the obsidian flows, located prehistoric quarrying areas, and gathered other essential information. Partial reports of this field work appear in Stocker and Cobean (1984) and Cobean *et al.* (1991).

In following with the thorough design of the sampling strategy, a comprehensive analytical program using NAA evolved in the MURR laboratory. Sample preparation techniques (Glascock *et al.* 1988), standards for analytical

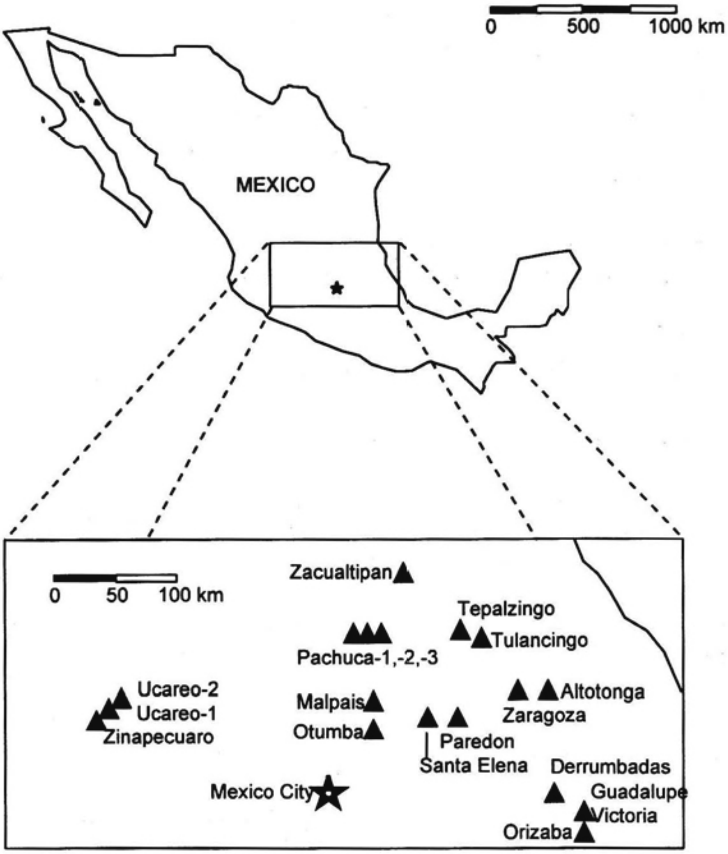

Figure 2.1. Locations of obsidian sources in central Mexico mentioned in this study.

calibration (Graham *et al.* 1982), and quality control procedures (Glascock and Anderson 1993) were developed in order to provide the most complete and accurate analyses possible. In addition, database management and statistical methods originally developed for compositional studies of ceramic materials are being employed (Bishop and Neff 1989; Neff 1990).

MURR's analytical program for obsidian routinely measures 27 elements in every source specimen by employing two irradiations and three measurements. A five-second irradiation in a neutron flux of 8×10^{13} n/cm^2/s is followed by a 25-minute decay and 12-minute count to determine the short-lived elements Ba, Cl, Dy, K, Mn, and Na. To measure the medium- and long-lived elements, samples are irradiated for 70 hours in neutron flux of 5×10^{13} n/cm^2/s. The medium-lived elements (Ba, La, Lu, Nd, Sm, U, and Yb) are determined by a 2,000-second count after 7–8 days of decay. The long-lived elements (Ce, Co, Cs, Eu, Fe, Hf, Rb, Sb, Sc, Sr, Ta, Tb, Th, Zn, and Zr) are determined by a 10,000-second count after 4–5 weeks of decay.

A total of 575 of the original 800 Mexican specimens have been analyzed. The most intensive analyses were made on specimens from the source regions near Pachuca in the state of Hidalgo, Ucareo-Zinapecuaro in the state of Michoacan, and the Pico de Orizaba-Guadalupe Victoria source area straddling the border between the states of Veracruz and Puebla.

Obsidian Sources in Guatemala

Obsidian sources in the highlands of Guatemala were sampled by Braswell who collected 137 samples from the seven areas shown in Figure 2.2.

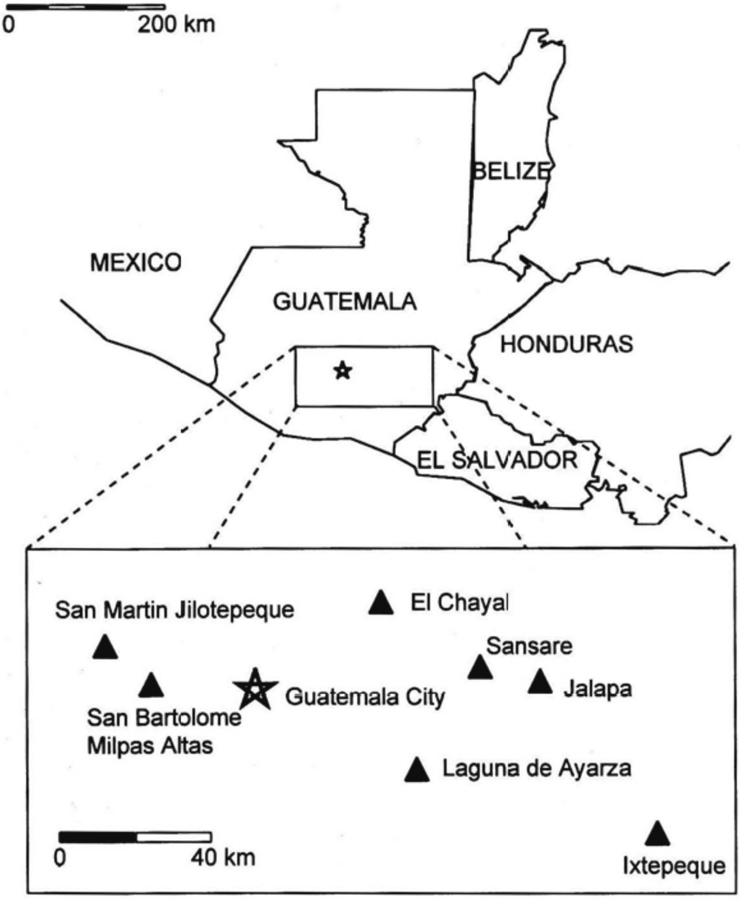

Figure 2.2. Locations of obsidian sources in Guatemala mentioned in this study.

In addition, the longitude and latitude coordinates for each specimen were recorded. Special emphasis was placed on obtaining specimens from the complex series of obsidian outcrops and river cobble areas surrounding the town of San Martin Jilotepeque, Chimaltenango (an area also known as Rio Pixcaya to many archaeologists). More than 70 specimens from this source area have been analyzed to date. A specific goal of Braswell's project was to determine whether obsidian quarries that were near one another could be adequately differentiated such that study of procurement trends and differences between archaeological sites, regions, and time periods might be possible.

In addition to providing samples to build an obsidian database for the Guatemalan sources, Braswell searched for a pair of sources suggested by previous research on artifacts from the neighboring La Entrada region of Honduras (Glascock and Aoyama 1990). Although Braswell's Guatemalan search failed to locate the sources supplying La Entrada, a new obsidian source was located near the village of Sansare. Details concerning the Sansare source are reported in Braswell and Glascock (1992).

Statistical Analysis of Source Data

At MURR, we employ a series of computer-based procedures, written in both FORTRAN and GAUSS languages, that facilitate the application of multivariate methods on personal computers. The GAUSS language routines written by Neff (1990) replace an earlier series of Brookhaven National Laboratory (BNL) programs by Sayre (1975) and Bieber *et al.* (1976). Our repertoire of statistical capabilities also includes the PC-based program SYSTAT.

After the obsidian specimens are analyzed, the concentration data for each is entered as a single record in a Master dBASE file with other analytical and descriptive information for that specimen. From the Master file, it is a simple matter to extract specimens as either individuals or groups for a particular statistical analysis. First, individual specimens that appear to be outliers rather than reliable core-group members are eliminated. If the number of missing values for an element is quite large, the element may be omitted from further consideration. If the number of missing values is small, then a GAUSS procedure based on minimizing the Mahalanobis distances from group centroid is used to calculate substitute values. In our mesoamerican database, Sr is below our detection limit of ~25 ppm in about half of the samples we analyze, thus causing us to eliminate it from the list of elements we consider. Dependent upon the particular source, Ba is occasionally below our detection limit of ~15 ppm and a missing value substitution is necessary. In either instance, one can sometimes make a rapid evaluation of candidate sources for many artifacts based on the presence or absence of Sr and Ba.

Before comparing artifacts to sources, it is necessary to demonstrate the ability of the source database to answer source-related questions. For example, can the source data be used to differentiate between geographic regions, between the sources within a region, or between outcrops within a source? How many different source profiles are represented in the mesoamerican database? Which elements or combinations of elements are best for differentiating between sources? Can a limited but satisfactory suite of elements be identified that can be measured more rapidly and at lower cost that works for a majority of artifacts? What is our degree of confidence in making source assignments?

In order to answer these questions, the GAUSS language routines are designed to establish chemical profiles for a collection of sources describing a region, individual sources within a region, and different outcrops within a source. Examples from MURR's mesoamerican database are presented here for illustration.

Differentiation between Regions

Are the obsidian sources in different geographic subregions of Mesoamerica more different than sources within each subregion?

In order to answer this question, the mesoamerican database was subdivided into nine geographic subregions with one or more obsidian source located in each. Based on the geographic proximity of sources to one another, the regions created were: Zaragoza (consisting of the Zaragoza and Altotonga sources); Orizaba (Pico de Orizaba, Guadalupe Victoria, and Derrumbadas sources); Paredon (Paredon and Santa Elena sources); Otumba (Otumba and Malpais sources); Tulancingo (Tulancingo and Tepalzingo sources); Pachuca (all Pachuca subsources); Zacualtipan (Zacualtipan source); Ucareo (Ucareo and Zinapecuaro sources); and the Guatemalan region (consisting of all sources in the highlands of Guatemala). A total of 712 source specimens were extracted from the Master dBASE file in order to create the tentative regional groupings.

On the basis of their poor analytical precision, the elements Cl, Dy, K, Nd, and U were omitted. Due to its large number of missing values, the element Sr was also omitted. Concentrations for the remaining 21 elements were transformed into their base-10 logarithms to compensate for the differences in magnitude between major elements such as Fe and Na and trace elements like the REEs. In order to identify the dimensions of greatest variance, the variance-covariance matrix of the nine-region dataset was input to the GAUSS routine designed for Q-mode PCA calculation.

Table 2.1 lists the calculated eigenvalues, percent variance described by each principal component, and the cumulative variance from the PCA. Table 2.2 presents the eigenvectors (coefficients) for the first six linear equations ranked from largest to smallest. As shown in Table 2.1, the first four PCS describe over 96% of the variance-covariance structure of the dataset. It often

Table 2.1. Eigenvalues and Percentage of Variance Explained by Principal Components Calculated from the Variance-Covariance Matrix of Concentration Data in Log Base-10 ppm for Nine Sub-regions in Mesoamerica

Principal Component	Eigenvalue	% Variance	Cumulative % Variance
1	1.4934	65.58	65.58
2	0.3069	13.48	79.06
3	0.2722	11.95	91.01
4	0.1235	5.42	96.43
5	0.0253	1.11	97.54
6	0.0153	0.67	98.22
7	0.0118	0.52	98.73
8	0.0084	0.37	99.10
9	0.0055	0.24	99.34
10	0.0045	0.20	99.54
11	0.0034	0.15	99.69
12	0.0022	0.10	99.78
13	0.0017	0.08	99.86
14	0.0011	0.05	99.91
15	0.0007	0.03	99.94
16	0.0005	0.02	99.96
17	0.0004	0.02	99.97
18	0.0003	0.01	99.99
19	0.0002	0.01	99.99
20	0.0001	0.00	100.00
21	0.0000	0.00	100.00

turns out that a transformation of the dataset using the first few PCS may delineate the different source regions better than the original variables. For example, Figure 2.3 presents a Q-mode plot of scores for the source specimens in each region plotted on PC01 versus PC02. Except for the Otumba and Guatemalan regions which overlap slightly, the source regions are well separated on these dimensions. Many of the regions in Figure 2.3 consist of multiple sources and with the exception of the Guatemalan and Zaragoza regions, anticipated internal structure is not readily apparent.

Table 2.2. Eigenvectors for the First Six Principal Components Best Describing Nine Sub-regions in Mesoamerica

Element	PC01	PC02	PC03	PC04	PC05	PC06
Ba	-0.4402	0.6812	-0.0656	0.2856	0.2466	0.0617
La	0.1407	0.1954	0.2881	0.0277	-0.4028	0.0243
Lu	0.2910	0.0262	0.0590	0.0344	0.1797	-0.1913
Sm	0.2032	0.0837	0.1350	0.1396	-0.0842	0.1250
Yb	0.2968	0.0261	0.0655	0.0529	0.1825	-0.1264
Ce	0.1678	0.1697	0.2548	0.0426	-0.3257	0.0493
Co	-0.2266	0.1971	0.5872	-0.6060	0.1012	-0.2226
Cs	-0.0030	-0.1136	0.2293	0.2119	0.4822	0.2740
Eu	0.2028	0.4857	-0.2195	-0.2281	0.1870	0.2221
Fe	0.1403	0.1490	0.0712	-0.0353	-0.0240	0.0775
Hf	0.2996	0.1003	-0.0390	-0.0158	0.1017	0.0770
Rb	0.0672	-0.0760	0.1046	-0.0293	0.2991	0.0913
Sb	0.0101	0.0709	0.3689	0.4985	0.0676	-0.5310
Sc	0.0532	-0.1286	0.0617	-0.3099	0.1850	0.0559
Ta	0.2350	-0.0849	0.0227	0.0032	0.1845	-0.1539
Tb	0.2719	0.0430	0.0795	0.1284	0.0774	0.0683
Th	0.1023	-0.0770	0.2307	-0.0314	0.2959	0.1850
Zn	0.2559	0.1166	-0.1054	0.0778	0.0183	-0.0472
Zr	0.3441	0.2652	0.0194	-0.1090	-0.1634	0.0723
Mn	0.1017	0.1185	-0.3773	-0.2141	0.1629	-0.6121
Na	0.0306	0.0185	-0.0704	0.0069	-0.0143	-0.0319

Differentiation between Outcrops in Complex Source Areas

Do the obsidian outcrops that cover large geographic areas at San Martin Jilotepeque, Sierra de Pachuca, and Ucareo-Zinapecuaro have multiple compositional fingerprints? Which elements best illustrate these differences? How many fingerprints are present in the mesoamerican database?

Obsidian from the San Martin Jilotepeque (SMJ) source area was obviously important to the earliest peoples of southern Mesoamerica as it appears in Formative period contexts at many archaeological sites. The source area is quite large and raw obsidian is found in numerous surface outcrops shown in Figure 2.4. The outcrops at Pachay, Choatalum, and the site labelled Group 6 have the greatest archaeological evidence for utilization during prehistoric times.

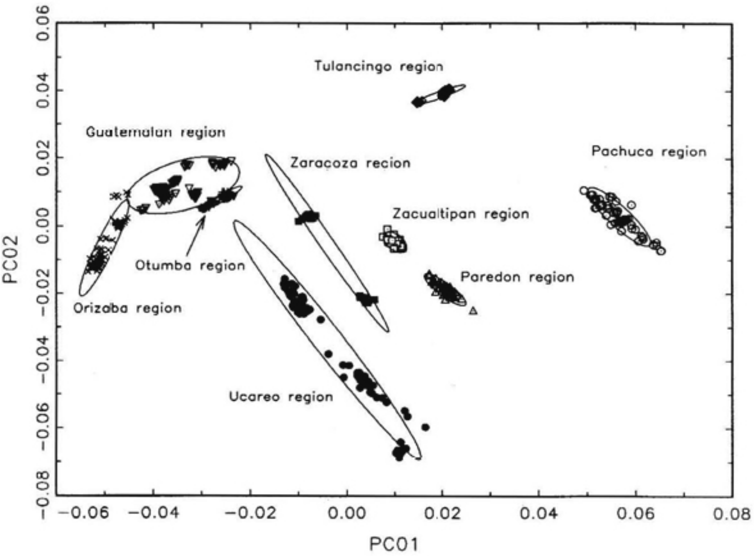

Figure 2.3. Q-mode bivariate plot of specimen scores on first and second principal components illustrating differentiation between obsidian source regions in Mexico and Guatemala. Probability ellipses at the 95% confidence level are included to indicate the different source regions.

The compositional data for seventy SMJ source samples were extracted from the database and submitted to cluster analysis (Figure 2.5). Examination of the clustering dendrogram suggests existence of six compositional subgroups in the SMJ source area. This observation is readily confirmed by the bivariate plot of Cs versus Sc presented in Figure 2.6. A comparison with the outcrop locations shown in Figure 2.4 indicates the geographical distribution of fingerprints. In particular, the SMJ-1 subgroup fingerprint appears in five different areas (labelled Group 1A through Group 1E) and in three river cobble samples representing a secondary deposit. The archaeological evidence to date indicates that the vast majority of artifacts sourced to SMJ came from quarries associated with the SMJ-1 fingerprint. Thus far, there are no confirmed assignments of artifacts to any of the other SMJ subsources.

The Sierra de Pachuca source located ten kilometers from the city of Pachuca, Hidalgo is one of the largest and most important in central Mexico. Most Pachuca obsidian is readily characterized by a dark-green color. Pachuca obsidian is of such high quality that it was one of the most widely utilized sources in all of Mesoamerica, with Pachuca artifacts being found several thousands of kilometers from the source area (Neff *et al.* 1993). Some variation in color was noted in specimens collected by Cobean from the source area and it was hypothesized that more than one compositional fingerprint exists. To examine

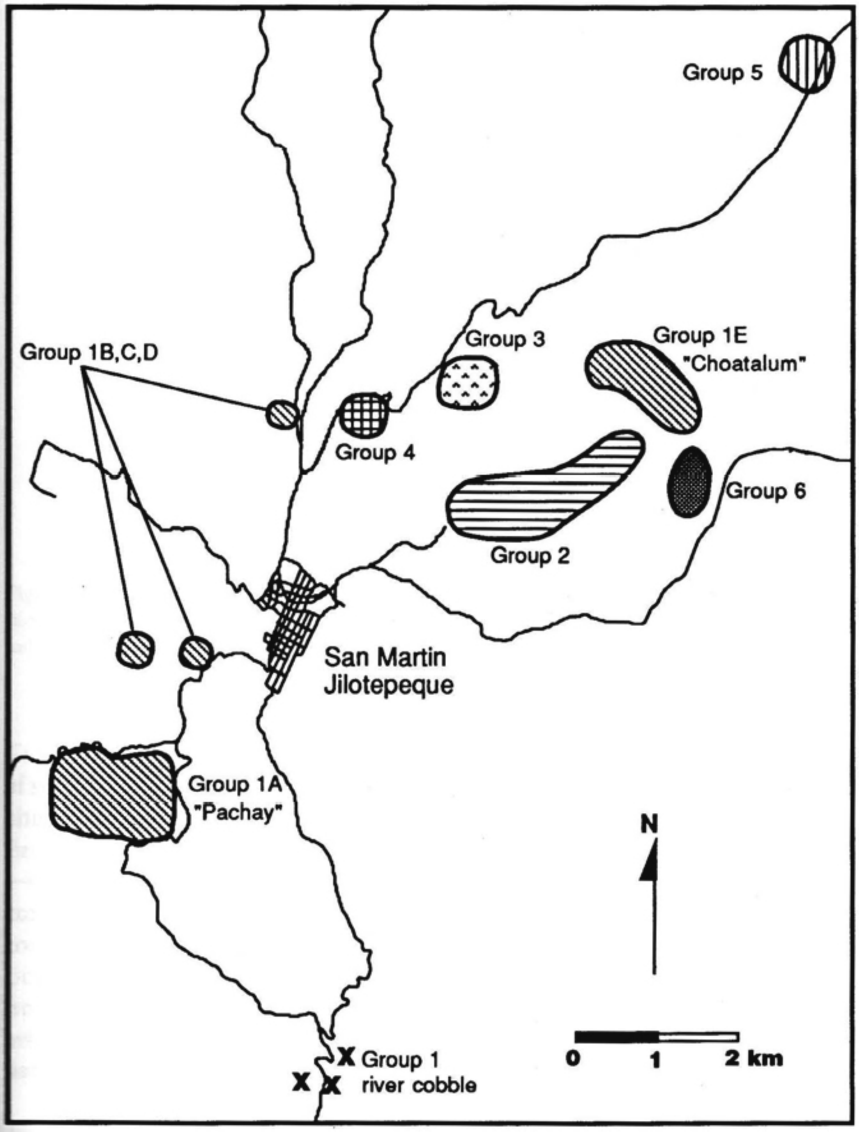

Figure 2.4. Locations of obsidian outcrops in the vicinity of San Martin Jilotepeque, Guatemala.

this hypothesis, the data for 167 specimens from the Pachuca source area was extracted from the database. A cluster analysis dendrogram indicted three possible subgroups (labelled Pachuca-1, Pachuca-2, and Pachuca-3 in Figure 2.1). Additional confirmation of these groups is shown in the bivariate plot of

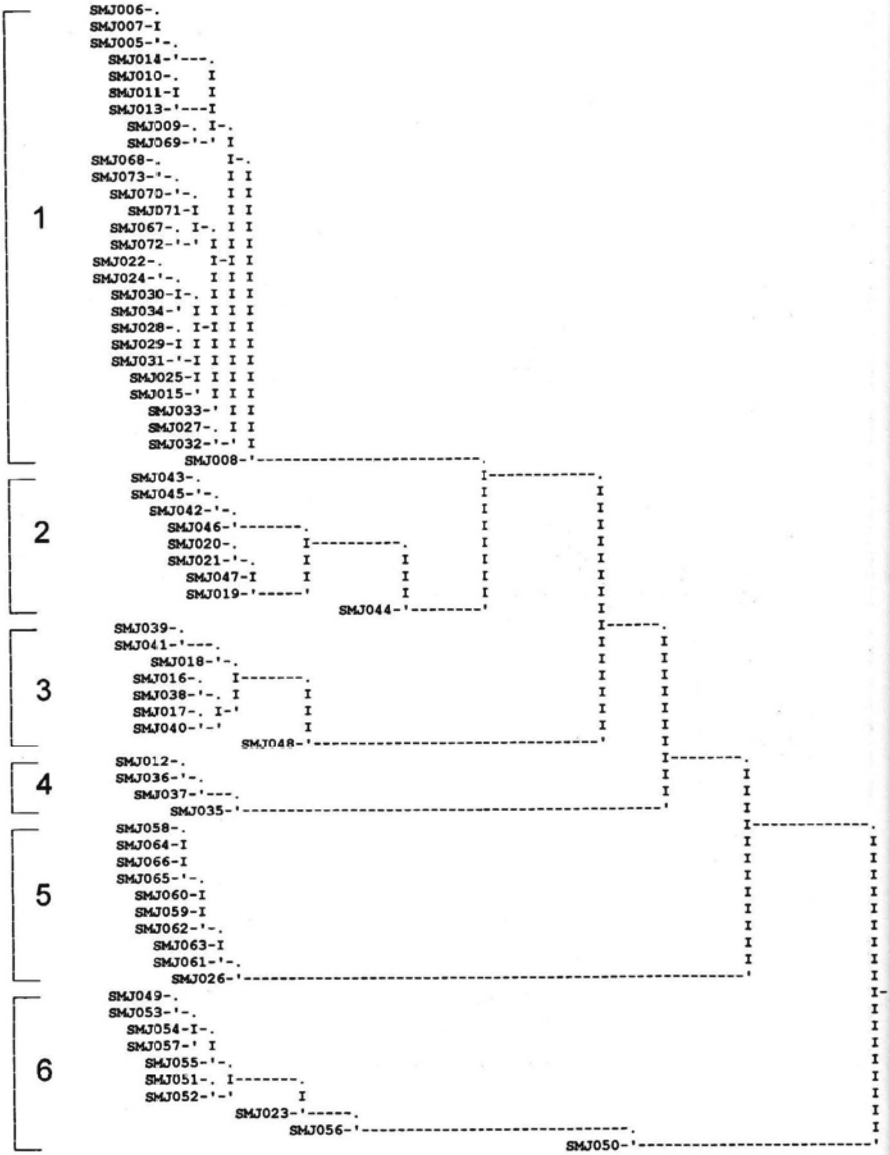

Figure 2.5. Dendrogram showing cluster analysis results for obsidian source specimens from San Martin Jilotepeque based. Squared-mean euclidean distance calculations were based on log base-10 transformations of the 21 element concentrations. Six tentative subgroups are identified.

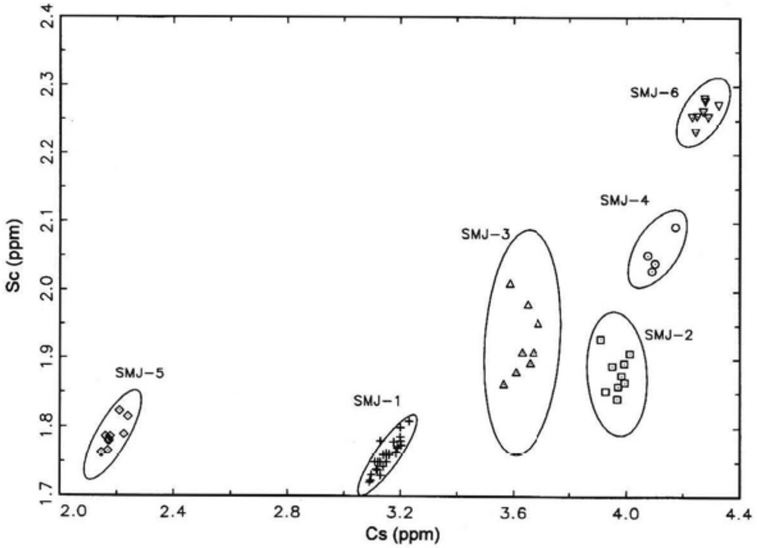

Figure 2.6. Bivariate plot of Cs versus Sc for the compositional subgroups of the San Martin Jilotepeque, Chimaltenango obsidian source. Probability ellipses at the 95% confidence level indicate the locations of the subgroups.

Cs versus Eu for Pachuca obsidian shown in Figure 2.7. Subsequent comparison of specimens with Cobean's sample collection records has been used to relate our chemical subgroups to geographic areas. All dark-green source material and Pachuca artifacts sourced to date have been linked to the Pachuca-1 fingerprint.

A similar cluster analysis and bivariate plot inspection performed on compositional data for source specimens from the Ucareo-Zinapecuaro region confirm the existence of three compositional fingerprints (labelled Ucareo-1, Ucareo-2, and Zinapecuaro in Figure 2.1). A bivariate plot of source specimens from the Ucareo-Zinapecuaro region is shown in Figure 2.8. The archaeological evidence supports the dominance of the Ucareo-1 subsource fingerprint relative to both the Ucareo-2 and Zinapecuaro subsources.

At first glance, the ability to source artifacts to specific outcrops or quarries within a multi-fingerprint source would seem to have little importance with respect to studies of long-distance trade. However, if raw material extraction is focussed on a particular type of obsidian that is distributed widely (i.e, certain colors or fracturing qualities) subsource identification may be extremely important. For example, highly specific subsourcing has the potential of supplying data on the organization of extraction activities and perhaps distribution of the material (i.e., mining claims by individuals or groups). More importantly, subsource identification is absolutely essential for obsidian-hydration studies.

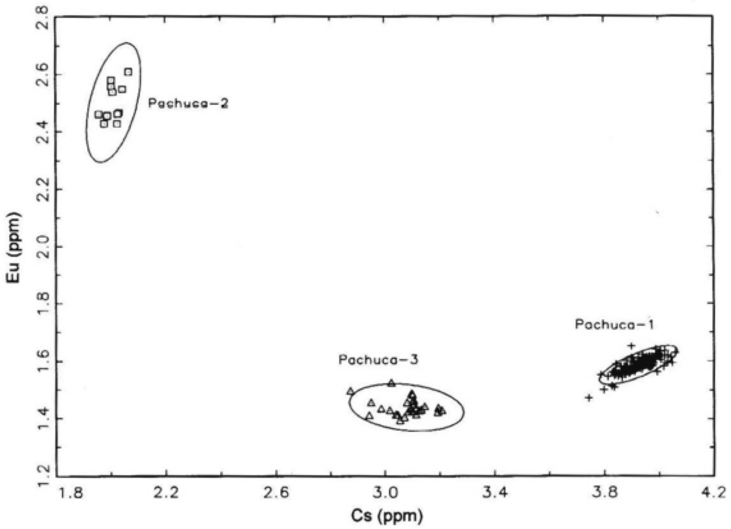

Figure 2.7. Bivariate plot of Cs versus Eu for the compositional subgroups of the Sierra de Pachuca, Hidalgo obsidian source. Probability ellipses at the 95% confidence level indicate the locations of the subgroups.

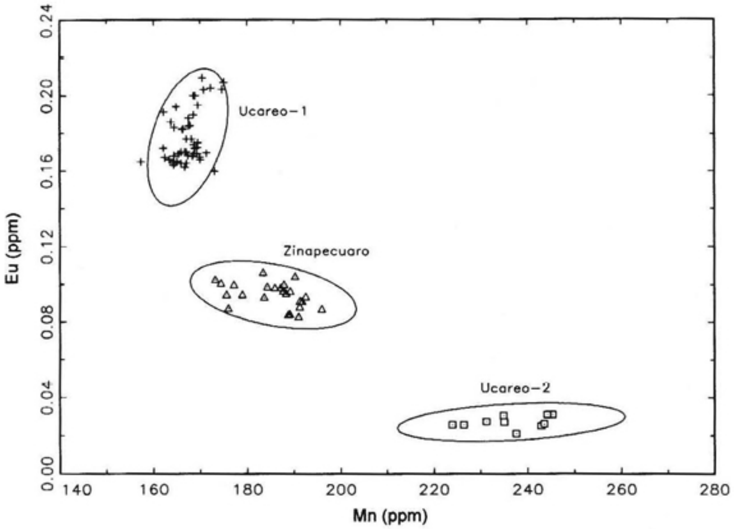

Figure 2.8. Bivariate plot of Mn versus Eu for compositional subgroups of the Ucareo-Zinapecuaro, Michoacan obsidian source. Probability ellipses at the 95% confidence level indicate the locations of the subgroups.

Ericson (1981) has shown that chemical differences in specimens from multiple-fingerprint sources in California have a significant effect on hydration rates. In each of the regions where multiple fingerprints are known, a more careful analysis is necessary because misidentification of chemical subgroup for artifacts also being dated may lead to erroneous dates. In addition, if certain subsources were exploited at different times, then archaeologists may be able to use their data to examine trends in obsidian procurement and exploitation.

Through application of the above mentioned PCA, cluster analysis, and bivariate plotting techniques, thirty different fingerprints were discovered in the regions of Mesoamerica under investigation. Eighteen sources are located in central and eastern Mexico and twelve sources are located in the highlands of Guatemala. The next step is to identify procedures through which the individual sources can be differentiated from one another.

Distinguishing between Individual Sources in Mexico

Can we formulate a reliable procedure for differentiating between all sources in central Mexico? What are its limitations? Does the procedure reject specimens from the Guatemalan sources?

Since we previously demonstrated that most of the Mexican sources are easily distinguished from the Guatemalan sources, the problem of differentiating between individual sources is separated into two portions (i.e., Mexico and Guatemala). Data for the eighteen source groups involving 575 source specimens from central Mexico was extracted from the Master dBASE file to form individual source groups. Six outlier specimens from five different sources were eliminated on the basis of a preliminary screening of elemental means, ranges, and standard deviations.

To identify the dimensions of greatest variance and elements likely to be important in differentiating between sources, an RQ-mode PCA was performed on the remaining 569 specimens. The resulting Q-mode and R-mode plots for PC01 versus PC02 are shown in Figures 2.9 and 2.10, respectively. Table 2.3 lists the individual eigenvalues, percent variance, and cumulative variance explaining the PCA results for these data. With exception for the three subsources in the Pachuca region, all of the central Mexican sources are separated at the 95% confidence level on the basis of the first two PCS. Although not shown here, further inspection of bivariate plots using other PCS finds the confidence ellipses for the Pachuca sources are not separable until we include component PC05.

Inspection of the R-mode plot in Figure 2.10 finds that Ba, Co, Zr, Hf, and the REEs are contributing most to separation between individual sources on the first two principal components. On the other hand, Na and Rb are contributing least to this view of the data. A bivariate plot of two high- precision elements (Cs versus Hf) shown in Figure 2.11 offers the greatest visual differentiation between confidence ellipses for the sources in central Mexico.

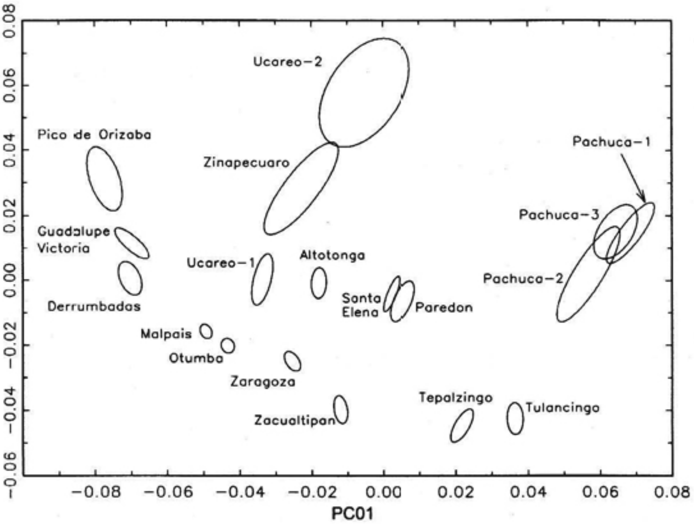

Figure 2.9. Q-mode bivariate plot on the first two principal components showing 95% confidence ellipses for obsidian sources in central Mexico. To avoid data congestion, the scores for individual specimens are not shown.

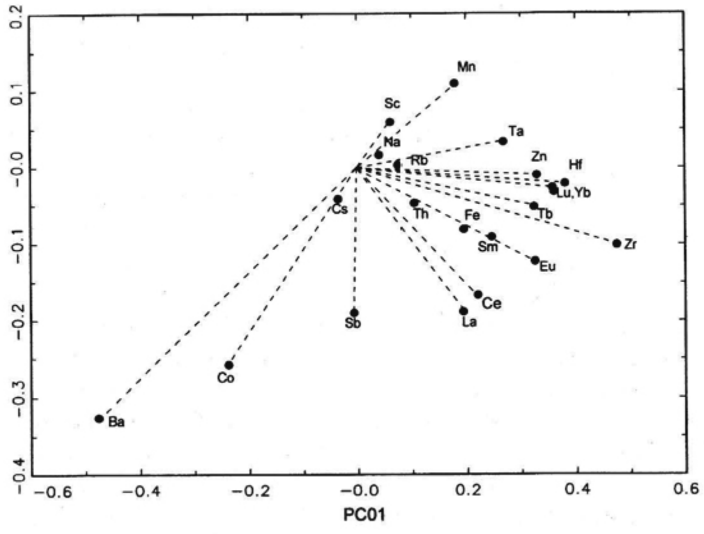

Figure 2.10. R-mode bivariate plot on the first two principal components showing scaled-loading factors for elements in central Mexican obsidian sources.

Table 2.3. Eigenvalues and Percentage of Variance Explained by Principal Components Calculated from the Variance-Covariance Matrix of Concentration Data in Log Base-10 ppm for the Eighteen Sources in Central Mexico

Principal Component	Eigenvalue	% Variance	Cumulative % Variance
1	1.5461	64.82	64.82
2	0.3392	14.22	79.04
3	0.3010	12.62	91.66
4	0.1235	5.18	96.84
5	0.0265	1.11	97.95
6	0.0173	0.72	98.67
7	0.0097	0.41	99.08
8	0.0073	0.30	99.38
9	0.0049	0.21	99.59
10	0.0029	0.12	99.71
11	0.0023	0.10	99.81
12	0.0018	0.07	99.88
13	0.0011	0.05	99.93
14	0.0005	0.02	99.95
15	0.0004	0.02	99.96
16	0.0003	0.01	99.98
17	0.0002	0.01	99.99
18	0.0001	0.01	99.99
19	0.0001	0.00	100.00
20	0.0000	0.00	100.00
21	0.0000	0.00	100.00

Two of the source groups (Santa Elena and Derrumbadas) have only five analyzed specimens each. Therefore, in order to consider all of the central Mexican sources simultaneously, probability calculations using Mahalanobis distances are limited to a maximum of three variables. Approximately 92% of the variance in the dataset listed in Table 2.3 is subsumed by the first three PCS. Mahalanobis distances and F-value probabilities based on the first three PCS were calculated to generate the posterior classification matrix presented in Table 2.4. The classification table is organized with the obsidian source *from* listed vertically and the source *assigned to* listed horizontally. As shown, the Santa Elena source

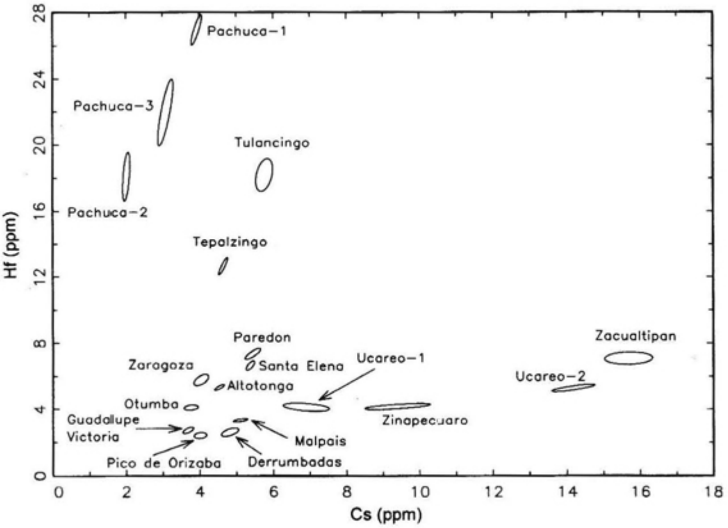

Figure 2.11. Bivariate plot of Cs versus Hf showing differentiation of the eighteen central Mexican compositional groups. Probability ellipses at the 95% confidence level indicate the locations of the subgroups.

and two of the Pachuca subsources are the only ones causing difficulty. Santa Elena has so few specimens that the Mahalanobis distance calculation extends its probability hyperellipsoid to overlap with fringe samples from several sources. This, of course, supports the earlier arguments that sources be thoroughly characterized through analysis of a sufficient number of source samples. Until more Santa Elena specimens are analyzed, we must critically examine all artifacts assigned to Santa Elena. Fortunately, there is no archaeological evidence that either Santa Elena or Derrumbadas sources were of any importance.

As mentioned earlier, the erroneous Pachuca classifications are caused by the relative similarity of these subsources to one another on the first few PCS. In order to eliminate errors in Pachuca assignments we can revert to inspection of bivariate plots of Cs versus Hf or Cs versus Eu both of which easily differentiate the Pachuca subsources. On the other hand, a procedure for robustly differentiating all sources in central Mexico (after omission of Santa Elena and Derrumbadas sources) can be developed by repeating our posterior classification procedure using the first five PCS which describe about 98% of the total variance. Table 2.5 presents the posterior classification results for central Mexican sources when using the first five PCS. As shown, the classification is completely successful as there are no misplaced specimens. Additional support for the robustness of this procedure is found in Table 2.6 which presents the average

Table 2.4. Posterior Classification Matrix Based on the First Three PCs for Obsidian Source Specimens from Central Mexican Sources—Jackknife Method

Source	n	AV	ZP	PV	GP	DP	PP	LH	OM	MH	TH	PH	ZH	SH1	SH2	SH3	UM1	UM2	ZM
AV	14	14																	
ZP	31		29					2											
PV	58			58															
GP	23				23														
DP	5					5													
PP	38						35	3											
LH	5							5											
OM	47								47										
MH	19									19									
TH	40										40								
PH	10							1				9							
ZH	30												30						
SH1	129													125		4			
SH2	11														11				
SH3	27													1	1	25			
UM1	48																48		
UM2	10																	10	
ZM	24																		24

Source abbreviations: AV = Altotonga, Ver.; ZP = Zaragoza, Pue.; PV = Pico de Orizaba, Ver.; GP = Guadalupe Victoria, Pue.; DP = Derrumbadas, Pue.; PP = Paredon, Pue.; LH = Santa Elena, Hgo.; OM = Otumba, Mex.; MH = Malpais, Hgo.; TH = Tulancingo, Hgo.; PH = Tepalzingo, Hgo.; ZH = Zacualtipan, Hgo.; SH1, SH2, & SH3 = Sierra de Pachuca, Hgo. subsources; UM1 & UM2 = Ucareo, Mich. subsources; and ZM = Zinapecuaro, Mich.

Table 2.5. Posterior Classification Matrix Based on the First Five PCs for Obsidian Source Specimens from Central Mexican Sources—Jackknife Method

Source	n	AV	ZP	PV	GP	PP	OM	MH	TH	PH	ZH	SH1	SH2	SH3	UM1	UM2	ZM
AV	14	14															
ZP	31		31														
PV	58			58													
GP	23				23												
PP	38					38											
OM	47						47										
MH	19							19									
TH	40								40								
PH	10									10							
ZH	30										30						
SH1	129											129					
SH2	11												11				
SH3	27													27			
UM1	48														48		
UM2	10															10	
ZM	24																24

Table 2.6. Average Probabilities for Group Membership Based on the First Five PCs for Obsidian Sources Specimens from Central Mexican Sources—Jackknife Method

Source	AV	ZP	PV	GP	PP	OM	MH	TH	PH	ZH	SH1	SH2	SH3	UM1	UM2	ZM
AV	54.292	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000
ZP	0.000	50.291	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000
PV	0.000	0.000	55.665	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GP	0.000	0.000	0.000	50.201	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
PP	0.000	0.000	0.000	0.000	53.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000
OM	0.000	0.000	0.000	0.000	0.000	47.523	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MH	0.000	0.000	0.000	0.000	0.000	0.000	48.699	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
TH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	49.109	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
PH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	49.710	0.000	0.000	0.000	0.000	0.000	0.001	0.000
ZH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	50.717	0.000	0.000	0.000	0.000	0.002	0.000
SH1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	60.663	0.000	0.004	0.000	0.001	0.000
SH2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	49.159	0.000	0.000	0.001	0.000
SH3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	49.373	0.000	0.003	0.000
UM1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	54.608	0.000	0.000
UM2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	50.353	0.000
ZM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	51.523

probabilities for membership in each source group. (Ideally, the average probability of membership in one's own group will be ~50%.) The results show that when using five PCS, even in the worst case, the average probability of being assigned to a wrong source group in central Mexico is less than 4 in 100,000.

As a test of our ability to use the latter procedure to reject specimens from sources in other regions, the principal component scores for Guatemalan source specimens were calculated. All of the Guatemalan source specimens were then projected against the Mexican source groups and membership probabilities calculated from Mahalanobis distances. Again, the procedure's ability to reject wrong specimens were found to be an outstanding success. Two Sansare specimens had probabilities of about 1 in 100,000 of belonging to the Ucareo-2 subsource while all other Guatemalan source specimens were much less probable of being associated with any of the eighteen central Mexican sources.

Distinguishing between Individual Sources in Guatemala

Can we formulate a reliable procedure for differentiating between all sources in Guatemala? What are its limitations? Does the procedure reject specimens from Mexican sources?

The process used to examine data for the obsidian sources from Guatemala is analogous to that applied to the Mexican sources. Data for the twelve Guatemalan source groups was extracted to yield a total of 137 specimens. Six outlier specimens from five different sources were eliminated on the basis of a preliminary screening of elemental means, ranges, and standard deviations. An RQ-mode PCA was performed on the remaining 131 source specimens.

Table 2.7 lists the resulting eigenvalues, percent variance, and cumulative variance described by successive components. Figures 2.12 and 2.13, respectively, show the Q-mode and R-mode plots on components PC01 and PC02. Confidence ellipses for several of the SMJ subsources and the San Bartolome Milpas Altas source overlap one another in Figure 2.12. However, plots of components PC03 and PC04 (not shown here) are successful in separating these overlapping groups. In order of importance, we note that the elements Co, Cs, Sb, and Mn make the largest contributions to the variance described by the first two components in the Guatemalan source data.. Contrarily, the REEs provide the least amount of source differentiation information. A bivariate plot of Cs versus Co shown in Figure 2.14 successfully separates the confidence ellipses for all twelve source groups.

Since the SMJ-4 subsource is characterized by only four specimens, we are limited to use of the first two PCS (subsuming approximately 76% of the variance) when calculating Mahalanobis distances and F-value probabilities for all twelve sources. As shown in the Table 2.8, limiting our classification to the first two PCS results in a large number of misclassifications, particularly into

Table 2.7. Eigenvalues and Percentage of Variance Explained by Principal Components Calculated from the Variance-Covariance Matrix of Concentration Data in Log Base-10 ppm for the Twelve Sources in Guatemala

Principal Component	Eigenvalue	% Variance	Cumulative % Variance
1	0.1351	55.55	55.55
2	0.0509	20.92	76.48
3	0.0416	17.09	93.57
4	0.0045	1.84	95.41
5	0.0032	1.31	96.72
6	0.0026	1.16	97.78
7	0.0021	0.87	98.64
8	0.0011	0.47	99.12
9	0.0005	0.23	99.34
10	0.0004	0.18	99.52
11	0.0004	0.16	99.68
12	0.0002	0.09	99.77
13	0.0002	0.07	99.84
14	0.0001	0.04	99.88
15	0.0001	0.04	99.91
16	0.0001	0.03	99.94
17	0.0000	0.02	99.96
18	0.0000	0.02	99.98
19	0.0000	0.01	99.99
20	0.0000	0.01	100.00
21	0.0000	0.01	100.00

the source at Jalapa and the two San Martin Jilotepeque subsources with the fewest specimens.

If the SMJ-4 subsource is omitted and three PCS (describing 94% of the variance) are used the results are much more satisfactory. Tables 2.9 and 2.10 show the resulting classification matrix and average probabilities, respectively.

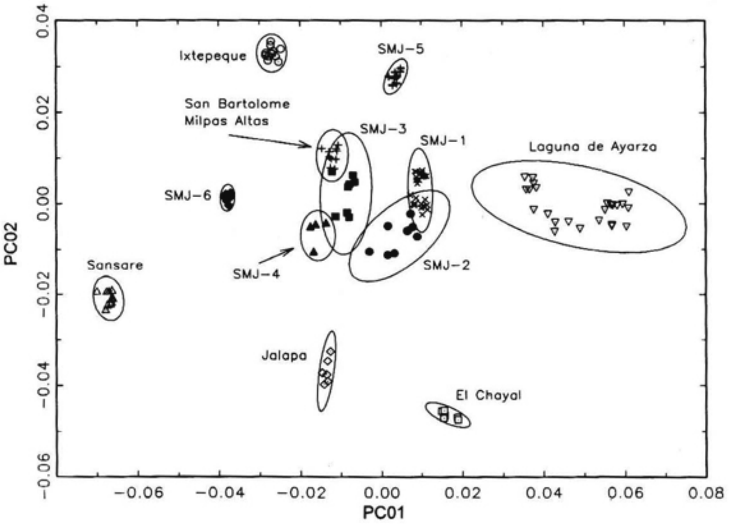

Figure 2.12. Q-mode bivariate plot on the first two principal components showing 95% confidence ellipses for obsidian sources in Guatemala and Otumba regions.

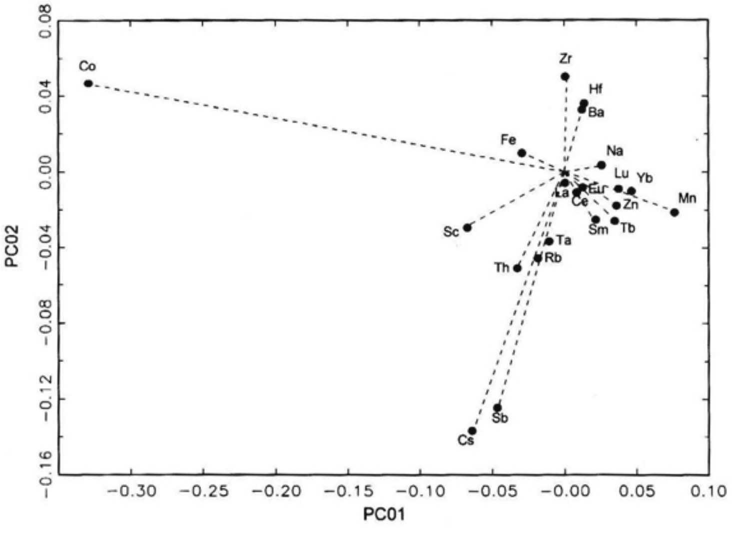

Figure 2.13. R-mode bivariate plot on the first two principal components showing scaled-loading factors for elements in Guatemalan and Otumba obsidian sources.

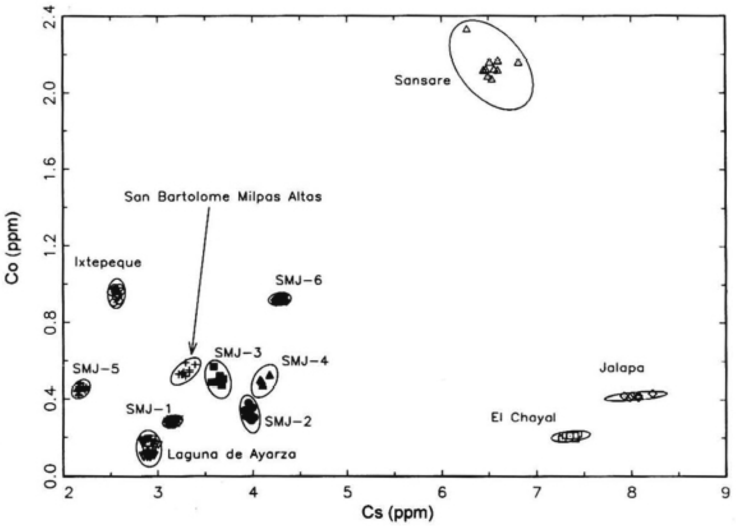

Figure 2.14. Bivariate plot of Cs versus Co showing differentiation of the twelve Guatemalan compositional groups. Probability ellipses at the 95% confidence level indicate the locations of the subgroups.

Table 2.8. Posterior Classification Matrix Based on the First Two PCs for Obsidian Source Specimens from Guatemalan Sources—Jackknife Method

Source	n	IG	CG	JG	SSG	SBM	LAG	SMJ1	SMJ2	SMJ3	SMJ4	SMJ5	SMJ6
IG	12	12											
C7	7		7										
JG	6			6									
SSG	10				10								
SBM	10					8				1	1		
LAG	24						24						
SMJ1	28							23	5				
SMJ2	9								9				
SMJ3	8					1				6	1		
SMJ4	4										4		
SMJ5	9											9	
SMJ6	8												8

Source abbreviations: IG = Ixtepeque; CG = El Chayal; JG = Jalapa; SSG = Sansare; SBM = San Bartolome Milpas Altas; LAG = Laguna de Ayarza; SMJ1, SMJ2, SMJ3, SMJ4, SMJ5, & SMJ6 = San Martin Jilotepeque subsources.

Table 2.9. Posterior Classification Matrix Based on the First Three PCs for Obsidian Source Specimens from Guatemalan Sources—Jackknife Method

Source	n	IG	CG	JG	SSG	SBM	LAG	SMJ1	SMJ2	SMJ3	SMJ5	SMJ6
IG	12	12										
CG	7		7									
JG	6			6								
SSG	10				10							
SBM	10					10						
LAG	24						24					
SMJ1	28							28				
SMJ2	9								9			
SMJ3	8									8		
SMJ5	9										9	
SMJ6	8											8

None of the source specimens are assigned to a wrong source and, with the exception of two SMJ entries in Table 2.10, the average membership probabilities are less than 1 in 100, and the vast majority are far less than 1 in 1,000. The smaller sizes of these Guatemalan source groups explains why these probabilities are much greater than those we calculated for the central Mexican sources. Even so, with probabilities of misclassification on the order of 1% or less, we are generally satisfied.

To test the Guatemalan sourcing procedure's ability to reject specimens from Mexican sources, principal component scores were calculated using the three PC basis. The Mexican source specimens were then projected against the Guatemalan sources and membership probabilities calculated from the Maha-

Table 2.10. Average Probabilities for Group Membership Based on the First Three PCs for Obsidian Sources Specimens from Guatemalan Sources—Jackknife Method

Source	IG	CG	JG	SSG	SBM	LAG	SMJ1	SMJ2	SMJ3	SMJ5	SMJ6
IG	51.149	0.000	0.018	0.000	0.001	0.000	0.000	0.000	0.015	0.000	0.000
CG	0.000	49.819	0.011	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000
JG	0.000	0.057	52.040	0.000	0.000	0.000	0.000	0.003	0.100	0.000	0.000
SSG	0.000	0.001	0.002	52.246	0.000	0.000	0.000	0.001	0.001	0.000	0.000
SBM	0.000	0.000	0.254	0.000	49.860	0.000	0.000	0.042	1.828	0.000	0.000
LAG	0.000	0.000	0.003	0.000	0.000	45.909	0.000	0.001	0.000	0.000	0.000
SMJ1	0.000	0.000	0.028	0.000	0.001	0.000	46.314	4.446	0.397	0.000	0.000
SMJ2	0.000	0.000	0.059	0.000	0.001	0.000	0.730	48.687	0.965	0.000	0.000
SMJ3	0.000	0.000	0.586	0.000	0.299	0.000	0.000	0.835	49.891	0.000	0.000
SMJ5	0.000	0.000	0.120	0.000	0.000	0.000	0.000	0.014	0.012	50.750	0.000
SMJ6	0.000	0.000	0.014	0.000	0.000	0.000	0.000	0.003	0.000	0.000	48.879

lanobis distances. Again, the method is quite successful with specimens from all but two sources having membership probabilities less than 1 in 1,000. The sources at Otumba and Malpais have a few specimens with probabilities of approximately 1 in 100 of being placed into the two Guatemalan source groups with the fewest members (Jalapa and SMJ-3). Therefore, we recommend use of a probability cutoff limit of 1% when using this procedure which would virtually eliminate all possibility of misclassification.

Examining the Efficacy of an Abbreviated-NAA Procedure

Can we develop a sourcing method that relies on the elements available by an abbreviated-NAA procedure? What are its limitations? As noted earlier, Pires-Ferriera (1973) used the short-lived elements Mn and Na to source artifacts with success for most but not all sources in Mexico. Later, Stross *et al.* (1983) employed a similar procedure on obsidian from Guatemala with five elements being measured (Ba, Dy, K, Mn, and Na). Our short irradiation procedure which measures Ba, Cl, Dy, K, Mn, and Na is similar to the method employed by Stross. The procedure is described in detail elsewhere (Glascock *et al.* 1994). Although it is anticipated the classification error rate will increase, the abbreviated method offers the advantages of reduced costs and rapid analysis. Therefore, we should test its applicability by performing a critical examination of the data in our database. In particular, the examination should identify those sources for which the method may be unsatisfactory.

For obsidian from Mexico, the variability of Ba, Dy, Mn, and Na concentrations between different sources suggests that these four elements might be employed in an abbreviated procedure. To examine this possibility, data for the 569 Mexican source samples was extracted and log base-10 transformations were made. A posterior classification matrix using these four elements was generated as shown in Table 2.11. The number of misclassifications is small and all occur for specimens within related regions (i.e., Orizaba specimens into Guadalupe Victoria, Otumba specimens into Malpais, Pachuca specimens from one subsource into one another, and Ucareo specimens into Zinapecuaro). More importantly, the average probabilities for these misclassification ranged up to a maximum of 5%. Therefore, if we establish a cutoff of 5%, or more conservatively 10%, for use of the abbreviated-NAA method, misclassifications in central Mexico should be extremely rare. For example, Figures 2.15 and 2.16 show bivariate plots of Mn versus Na and Mn versus Dy for almost 700 artifacts projected against the central Mexican sources to which they were sourced. None of the obsidian sources not shown on these figures were linked to any of these artifacts. Specimens with probabilities below our cutoff or with significant probabilities in two or more Mexican sources can be resolved by submitting them to the full-NAA and employing the more robust classification procedures described earlier.

Table 2.11. Posterior Classification Matrix Based on Log Base-10 ppm Concentrations for Ba, Dy, Mn, and Na for Obsidian Specimens from Central Mexican Sources—Jackknife Method

Source	n	AV	ZP	PV	GP	PP	OM	MH	TH	PH	ZH	SH1	SH2	SH3	UM1	UM2	ZM
AV	14	14															
ZP	31		31														
PV	58			56	2												
GP	23				23												
PP	38					38											
OM	47						46	1									
MH	19							19									
TH	40								40								
PH	10									10							
ZH	30										30						
SH1	129											128		1			
SH2	11												11				
SH3	27											1		26			
UM1	48														46		2
UM2	10															10	
ZM	24																24

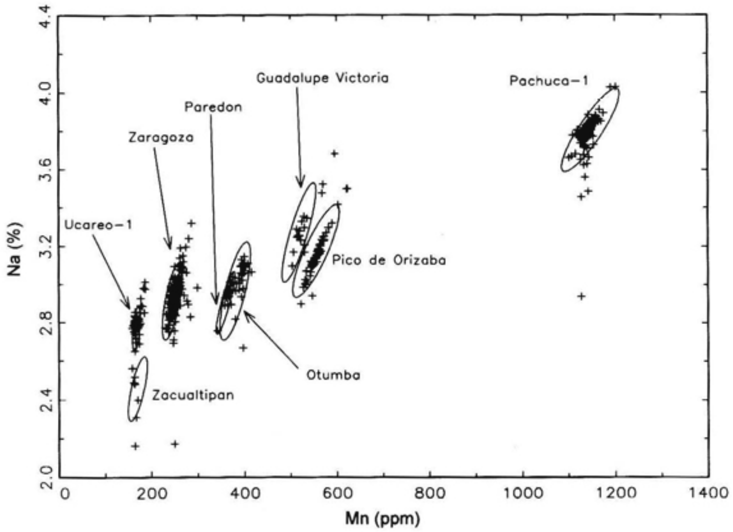

Figure 2.15. Bivariate plot of Mn versus Na showing artifacts projected against 95% confidence ellipses for the eight important sources in central Mexico. To avoid data congestion, sources to which artifacts were not assigned are not shown.

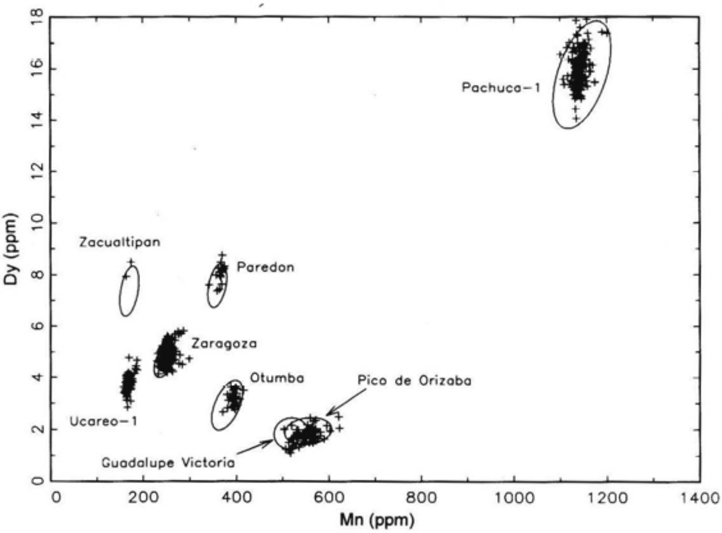

Figure 2.16. Bivariate plot of Mn versus Dy showing artifacts projected against 95% confidence ellipses for the eight important sources in central Mexico. Sources to which artifacts were not assigned are not shown.

Table 2.12. Posterior Classification Matrix Based on Log Base-10 ppm Concentrations for Ba, Dy, Mn, and Na for Obsidian Specimens from Guatamalan Sources—Jackknife Method

Source	n	IG	CG	JG	SSG	SBM	LAG	SMJ1	SMJ2	SMJ3	SMJ5	SMJ6
IG	12	12										
CG	7		7									
JG	6			6								
SSG	10				10							
SBM	10					10						
LAG	24			1			23					
SMJ1	28							26	2			
SMJ2	9		1						8			
SMJ3	8								1	7		
SMJ5	9										9	
SMJ6	8											8

A test of the same four elements in the Guatemalan sources results in the classification matrix shown in Table 2.12. Misclassifications occur primarily into the sources established by small numbers of specimens (i.e., El Chayal and Jalapa) or from one San Martin Jilotepeque subsource into another. The probabilities for misclassification ranged up to 7% in the worst cases involving the El Chayal and Jalapa sources and 15% between the San Martin Jilotepeque subsources. As we have recommended for the Mexican sources, a conservative probability cutoff of 10% should result in about 90% of all artifacts being properly classified on the basis of data from an abbreviated analysis using concentrations for Ba, Dy, Mn, and Na. Figure 2.17 shows a bivariate plot of Mn versus Na for 500 artifacts from sites in Belize and Guatemala projected against the three main Guatemalan sources to which they were sourced. As

Source	n	IG	CG	JG	SSG	SBM	LAG	SMJ1	SMJ2	SMJ3	SMJ5	SMJ6
SMJ5	9										9	
SMJ6	8											8

A test of the same four elements in the Guatemalan sources results in the classification matrix shown in Table 2.12. Misclassifications occur primarily into the sources established by small numbers of specimens (i.e., El Chayal and Jalapa) or from one San Martin Jilotepeque subsource into another. The probabilities for misclassification ranged up to 7% in the worst cases involving the El Chayal and Jalapa sources and 15% between the San Martin Jilotepeque subsources. As we have recommended for the Mexican sources, a conservative probability cutoff of 10% should result in about 90% of all artifacts being properly classified on the basis of data from an abbreviated analysis using concentrations for Ba, Dy, Mn, and Na. Figure 2.17 shows a bivariate plot of Mn versus Na for 500 artifacts from sites in Belize and Guatemala projected against the three main Guatemalan sources to which they were sourced. As

expected, the number of artifacts requiring the full-NAA is quite small.

Finally, our work has shown that only about 10% of all artifacts in Mexico and Guatemala are likely to require the full-NAA. Thus, the abbre-

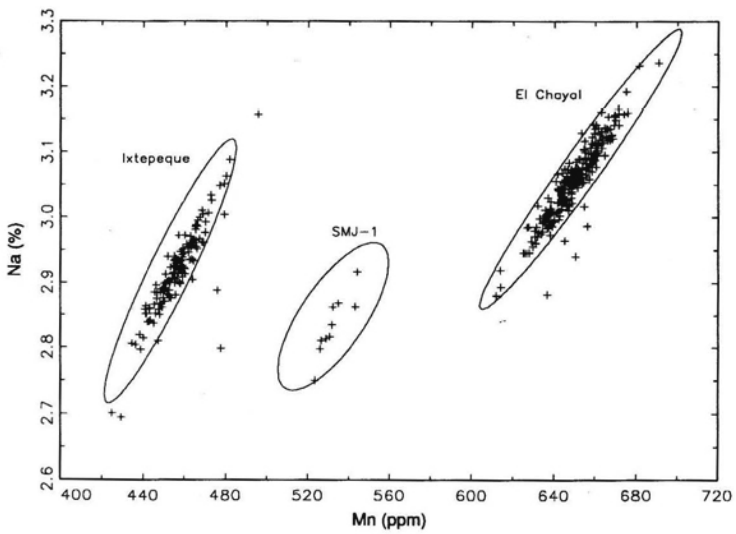

Figure 2.17. Bivariate plot of Mn versus Na showing artifacts projected against 95% confidence ellipses for the three important sources in Guatemala. Sources to which artifacts were not assigned are not shown.

elements (correlated and uncorrelated) at high precision and using the data to develop a powerful sourcing procedure. Once a satisfactory sourcing procedure is identified, applications to artifacts on a step-by-step basis are possible: 1) to identify the specific region from which a specimen came by rejecting all other regions, 2) to identify the individual sources by rejecting all other sources, and 3) if possible, to identify the specific subsource (quarry or outcrop) within the source by rejecting all other subsources. The multivariate methods described above offer a highly rigorous way to perform these operations.

ACKNOWLEDGMENTS

We acknowledge the support of three National Science Foundation grants (BNS 79-15409 to Cobean, DBS 91-02016 to Glascock, and DBS 93-01152 to Braswell) and a Fulbright scholarship to Braswell that supported various aspects of this work. In no particular order, E.V. Sayre, G. Harbottle, A.M. Bieber, Jr., R.L. Bishop, and H. Neff are gratefully acknowledged for their significant contributions in the advocacy of applying multivariate methods to compositional studies over the past 25 years. The assistance of many other colleagues in Mexico, Guatemala, and the United States who provided artifacts

for analysis are much appreciated. Finally, the encouragement and advice from H. Neff who patiently reviewed this manuscript are much appreciated.

REFERENCES

Ahrens, L.H. 1954 The lognormal distribution of the elements. *Geochimica et Cosmochimica Acta* 5: 49–73.

Asaro, F., Michel, H.V., Sidrys, R., and Stross, F. 1978 High-precision chemical characterization of major obsidian sources in Guatemala. *American Antiquity* 43: 436–443.

Avila-Salinas, W. 1975 Elementos trazas de algunas obsidianas Bolivianas, Centro de Investigaciones Arqueologicas, publicacion No. 4, La Paz, Bolivia, 16 pp.

Baxter, M.J. 1991 Principal component and correspondence analysis of glass compositions: an empirical study. *Archaeometry* 33: 29–41.

Baxter, M.J. 1992 Archaeological uses of the biplot--a neglected technique? *In* Lock, G. and Moffett, J., eds., *Computer applications and quantitative methods in archaeology, 1991*. British Archaeological Report International Series, S577, p. 141–148, Tempus Raparatum, Oxford.

Bettinger, R.L., Delacores, M.G., and Jackson, R.J. 1984 Visual sourcing of central eastern California obsidians. *In* R.E. Hughes, ed., Obsidian Studies in the Great Basin. *Contributions of the University of California Archaeological Research Facility* 45: 63–78.

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: 59–74.

Bishop, R.L. and Neff, H. 1989 Compositional data analysis in archaeology. *In* R.O. Allen, ed., *Archaeological Chemistry--IV.* American Chemical Society: Washington: 57–86.

Blackman, M.J. 1984 Provenance studies of Middle Eastern obsidian from sites in highland Iran. *In* Lambert, J.B., ed., *Archaeological Chemistry--III* Washington, American Chemical Society: 19–50.

Bowman, H.R., Asaro, F., and Perlman, I. 1973 On the uniformity of composition in obsidians and evidence for magmatic mixing. *Journal of Geology* 81: 312–327.

Braswell, G.E. and Glascock, M.D. 1992 A new obsidian source in the highlands of Guatemala. *Ancient Mesoamerica* 3: 47–49.

Burger, R.L., and Asaro, F. 1977 Trace element analysis of obsidian artifacts from the Andes: new perspectives on pre-hispanic economic interaction in Peru and Bolivia. University of California Lawrence Berkeley Laboratory preprint 6343: Berkeley, California.

Cann, J.R. and Renfrew, C. 1964 The characterization of obsidian and its application to the Mediterranean region. *Proceedings of the Prehistoric Society* 30: 111–133.

Cattell, R.B. 1952 *Factor Analysis*. Harper: New York, 462 pp.

Charlton, T.H., Grove, D.C., and Hopke, P.K. 1978 The Paredn, Mexico obsidian source and Early Formative exchange. *Science* 201: 807–809.

Clark, J.E. 1981 Multi-faceted approach to the study of Mesoamerican obsidian trade: an example from early Chiapas. Paper presented at the 46th annual meeting of the Society for American Archaeology in San Diego, California.

Clark, J.E. 1988 The lithic artifacts of La Libertad, Chiapas, Mexico: an economic perspective. *In Papers of the New World Archaeological Foundation*, No. 52., Provo, Utah.

Cobean, R.H., Coe, M.D., Perry, E.A. Jr., Turekian, K.K., and Kharkar, D.P. 1971 Obsidian trade at San Lorenzo Tenochtitlan, Mexico. *Science* 174: 141–146.

Cobean, R.H., Vogt, J.R., Glascock, M.D., and Stocker, T.R. 1991 High precision trace element characterization of major Mesoamerican obsidian sources and further analyses of artifacts from San Lorenzo Tenochtitlan, Mexico. *Latin American Antiquity* 2: 69–91.

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

Duerden, P., Bird, J.R., Clayton, E., Cohen, D.D., and Leach, B.F. 1984 Provenance studies of New Zealand obsidian artifacts. *Nuclear Instruments and Methods in Physics Research* B3: 419–423.

Duranni, S., Khan, M.T., and Renfrew, C. 1971 Obsidian source identification by fission track analysis. *Nature* 233: 242–245.

Elam, J.M. 1993 *Post-Formative Obsidian Exchange in Oaxaca, Mexico.* Ph.D. dissertation, University of Missouri: Columbia.

Ericson, J.E. 1981 *Exchange and Production Systems in Californian Prehistory: The Results of Hydration Dating and Chemical Characterization of Obsidian Sources.* British Archaeological Reports, International Series 110.

Freter, A.C. 1992 Chronological research at Copan. *Ancient Mesoamerica* 3: 117–133.

Fuller, R.E. 1927 The mode of origin of the color of certain varicolored obsidians. *Journal of Geology* 35: 570–573.

Gabriel, KR. 1971 The biplot-graphic display of matrices with application to principal components analysis. *Biometrika* 58: 453–467.

Glascock, M.D. 1992 Characterization of archaeological ceramics at MURR by neutron activation analysis and multivariate statistics. *In* Neff. H., ed., *Chemical Characterization of Ceramic Pastes in Archaeology.* Madison, Prehistory Press.

Glascock, M.D. and Anderson M.P. 1993 Geological reference materials for standardization and quality assurance of instrumental neutron activation analysis. *Journal of Radioanalytical and Nuclear Chemistry, Articles* 174: 229–242.

Glascock, M.D., Elam, J.M., and Aoyama, K. 1990 Provenience analysis of obsidian artifacts from the La Entrada Region, Honduras. *In* Pernicka, E. and Wagener, G.A., eds., *Archaeometry '90.* Basel, Birkhauser Verlag: 365–374.

Glascock, M.D., Elam, J.M., and Cobean, R.H. 1988 Differentiation of obsidian sources in Mesoamerica. *In* Farquhar, R.M., Hancock, R.G.V., and Pavlish, L.A., eds., *Archaeometry '88.* Toronto, Archaeometry Laboratory, University of Toronto: 245–251.

Glascock, M.D., Neff, H., Stryker, K.S., and Johnson, T.N. 1994 Sourcing archaeological obsidian by an abbreviated NAA procedure. *Journal of Radioanalytical and Nuclear Chemistry, Articles* 180:29–35.

Graham, C.C., Glascock, M.D., Carni, J.J., Vogt, J.R., and Spalding, T.G. 1982 Determination of elements in National Bureau of Standards' geological standard reference materials by neutron activation analysis. *Analytical Chemistry* 54: 1623–1627.

Harbottle, G. 1976 Activation analysis in archaeology. *In* Newton, G.W.A., ed., *Radiochemistry.* London, The Chemical Society: 33–72.

_____ 1982 Chemical characterization in archaeology. *In* Ericson, J.E. and Earle, T.K., eds., *Contexts for Prehistoric Exchange.* New York, Academic Press: 13–51. .

Hughes, R.E. 1982 Age and exploitation of obsidian from the Medicine Lake Highland, California. *Journal of Archaeological Science* 9: 173–185.

_____ 1984 Obsidian sourcing studies in the Great Basin: problems and prospects. *In* Hughes, R.E. ed., *Obsidian Studies in the Great Basin.* Contributions of the University of California Archaeological Research Facility No. 45: Berkeley, California: 1–19.

_____ 1986 *Diachronic Variability in Obsidian Procurement Patterns in Northeastern California and Southcentral Oregon.* Berkeley, University of California Press.

_____ 1988 The Coso Volcanic Field reexamined: implications for obsidian sourcing and hydration dating research. *Geoarchaeology* 3: 253–265.

Huntley, D.J. and Bailey, D.C. 1978 Obsidian source identification by thermoluminescence. *Archaeometry* 20: 159–170.

Hurtado de Mendoza, L. and Jester, W.A. 1978 Obsidian sources in Guatemala: a regional approach. *American Antiquity* 43: 424–435.

Jackson, T.L, and Love, M.W. 1991 Blade running: middle preclassic obsidian exchange and the introduction of prismatic blades at La Blanca, Guatemala. *Ancient Mesoamerica* 2: 47–59.

Klecka, W.R. 1980 Discriminant analysis. *Sage University Paper, Quantitative Analysis in the Social Sciences*, Series No. 07–019. Beverly Hills, California, Sage Publications.

Lachenbruch, P.A. 1975 *Discriminant Analysis*. New York, Hafner.

Leach, F. 1973 Obsidian from the Chatham Islands. *New Zealand Archaeological Association Newsletter* 16: 104–106.

Leach, F. and Manly, B. 1982 Minimum Mahalanobis distance functions and lithic source characterization by multi-element analysis. *New Zealand Journal of Archaeology* 4: 77–109.

Leach, F. and Warren, S.E. 1981 Neutron activation analysis of New Zealand and oceanic obsidians: towards a simple screening technique. *In* Leach, F. and Davidson, J., eds., *Archaeological Studies of Pacific Stone Resources*. Oxford, BAR International Series 104, Oxford, England: 151–166.

Leach, B.F., Warren, S.E., and Fankhauser, B. 1978 Obsidian from the far north of New Zealand: A method of sourcing based on natural radioactive emission. *New Zealand Journal of Science* 21: 123–128.

Longworth, G. and Warren, S.E. 1979 The application of mössbauer spectroscopy to the characterisation of western Mediterranean obsidian. *Journal of Archaeological Science* 6: 179–193.

McDougall, J.M., Tarling, D.H., and Warren, S.E. 1983 The magnetic sourcing of obsidian samples from Mediterranean and Near Eastern sources. *Journal of Archaeological Science* 10:441–452.

Moholy-Nagy, H. and Nelson, F.W. 1990 New data on sources of obsidian artifacts from Tikal, Guatemala. *Ancient Mesoamerica* 1: 71–80.

Neff, H. 1990 A series of routines written in GAUSS language (unpublished).

_____1994 RQ-mode principal components analysis of ceramic compositional data. *Archaeometry* 36: 115–130.

Neff, H., Bishop, R.L., and Arnold, D.E. 1988 Reconstructing ceramic production from ceramic compositional data: an example from Guatemala. *Journal of Field Archaeology* 15: 339–348.

Neff, H., Bove, F.J., Johnson, T.N., and Arroyo, B. 1993 Fachimentos a traves de hidratacion de obsidiana en la costa sur de Guatemala. *Apuntes Arqueologicos* 2: 55–79.

Nelson, F.W. 1985 Summary of the results of analysis of obsidian artifacts from the Maya lowlands. *Scanning Electron Microscopy* II: 631–649.

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

Picon, M., Carre, C., Cordoliani, M.L., Vichy, M., Hernandez, J.A., and Mignard, J.L. 1975 Compositions of the La Graufesenque, Banassac and Montans terra sigillata. *Archaeometry* 17: 191–199.

Pires-Ferreira, J.W. 1973 *Formative Mesoamerican Exchange Networks with Special Reference to the Valley of Oaxaca*. Memoirs of the Museum of Anthropology No. 7, Ann Arbor, University of Michigan.

Reeves, R.D. and Armitage, G.C. 1973 Density measurements and chemical analysis in the identification of New Zealand archaeological obsidians. *New Zealand Journal of Science* 16: 561–572.

Sappington, R.L. 1981 The archaeology of the Lydle Gulch Site (10-AA-72): prehistoric occupation in the Boise River Canyon, southwestern Idaho. University of Idaho Anthropological Research Manuscript Series 66. Moscow, Idaho, University of Idaho.

_____ 1984 Procurement without quarry production: examples from southwestern Idaho. *In* Ericson, J.E. and Purdy, B.A., eds., *Prehistoric Quarries and Lithic Production*. Cambridge, Cambridge University Press: 23–34.

Sayre, E.V. 1975 Brookhaven procedures for statistical analyses of multivariate archaeometric data. Brookhaven National Laboratory Report BNL-21693, New York.

Shackley, M.S. 1988 Sources of archaeological obsidian in the Southwest: an archaeological, petrological, and geochemical study. *American Antiquity* 53: 752–772.

Skinner, C. 1983 *Obsidian Studies in Oregon: An Introduction to Obsidian and Investigation of Selected Methods of Obsidian Characterization Utilizing Obsidian Collected at Prehistoric Quarry Sites in Oregon*. M.A. thesis, University of Oregon, Eugene.

Sneath, P.H.A. and Sokal, R.R. 1973 *Numerical Taxonomy*. San Francisco, W.H. Freeman and Company.

Stark, B.L., Heller, L., Glascock, M.D., Elam, J.M., and Neff, H. 1992 Obsidian-artifact source analysis for the Mextequilla region, south-central Veracruz, Mexico. *Latin American Antiquity* 3: 221–239.

Stevenson, C.M. and McCurry, M.O. 1990 Chemical characterization and hydration rate development for New Mexican obsidian sources. *Geoarchaeology* 5: 149–170.

Stocker, T.R. and Cobean, R.H. 1984 Preliminary report on the obsidian mines at Pico de Orizaba, Veracruz. *In* Ericson, J.E. and Purdy, B.A., eds., *Prehistoric Quarries and Lithic Production*, Cambridge, Cambridge University Press: 83–95.

Stross, F.H., Sheets, P., Asaro, F., and Michel, H.V. 1983 Precise characterization of Guatemalan obsidian sources, and source determination of artifacts from Quirigu. *American Antiquity* 48: 323–346.

Vogt, J.R., C.C. Graham, M.D. Glascock, and R.H. Cobean 1982 A study of Mesoamerican obsidian sources using activation analysis. *Journal of Radioanalytical Chemistry* 69: 271–289.

Ward, G.K. 1974 A systematic approach to the definition of sources of raw material. *Archaeometry* 16: 41–53.

Wheeler, M.E. and Clark, D.W. 1977 Elemental characterization of obsidian from the Koyukuk River, Alaska, by atomic absorption spectrophotometry. *Archaeometry* 19: 15–31.

Wilson, A.L. 1978 Elemental analysis of pottery in the study of its provenance: a review. *Journal of Archaeological Science* 5: 219–236.