Journal of Archaeological Science: Reports 13 (2017) 428–434

Contents lists available at ScienceDirect

Journal of Archaeological Science: Reports

journal homepage: www.elsevier.com/locate/jasrep

Using Fordisc software to assign obsidian artifacts to geological sources: Proof of concept

Marin A. Pilloud*, Derek J. Reaux, Geoffrey M. Smith, Kristina M. Wiggins

Department of Anthropology, University of Nevada, Reno, 1664 No. Virginia Street/MS0096, Reno, NV 89557, USA

ARTICLE INFO

Keywords:
Multivariate statistics
Great Basin
Provenance studies
X-ray fluorescence

ABSTRACT

In recent years, source provenance studies employing portable X-ray fluorescence (pXRF) technology have become commonplace in archaeology; however, they are not without critiques. Concerns center on the capability of instruments to produce valid results and researchers' abilities to accurately interpret those results and make correct source assignments. In this paper, we focus on the latter issue with a look towards statistical means of assigning artifacts to obsidian types using data provided by pXRF spectrometers. Using a sample of 677 obsidian artifacts from the northwestern Great Basin, we evaluate the ability of various approaches (principal components, cluster, and discriminant function analyses) to correctly assign artifacts to particular obsidian types. These multivariate methods generally work well to separate artifacts into different groups (i.e., obsidian types); however, they are less well-suited to assign individual artifacts to an obsidian source or type. We therefore tested the ability of the statistical program Fordisc, commonly used in forensic anthropology, to assign individual artifacts to specific geochemical obsidian sources or types. Our results indicate that Fordisc made accurate source assignments. Furthermore, because Fordisc provides probability values for different possible matches, it offers an advantage over other methods.

1. Introduction

Applications of portable X-ray fluorescence (pXRF) technology to address archaeological questions have increased dramatically in recent years and many academic institutions and cultural resource management (CRM) firms now possess units. Although relatively expensive to purchase, pXRF instruments offer numerous benefits: (1) a non-destructive method to determine trace elements; (2) the ability to conduct in-field analyses (important when artifacts may not be collected); (3) the ability to characterize large numbers of artifacts in a relatively short amount of time; and (4) the elimination of commercial lab fees (Shackley, 2011, 2012). Their rapid and widespread adoption by researchers lacking previous experience in geochemical characterization techniques has led some experienced analysts to express concern that some applications of pXRF technology have "no real foundation in science" (Shackley, 2012:2). Such concerns primarily center on issues related to *repeatability* (agreement between measurements collected under identical conditions at different times), *reproducibility* (agreement between measurements collected at different times under different conditions), *accuracy* (agreement between measurements collected using different instruments; for example, between pXRF and conventional wavelength-dispersive [WDXRF] and energy-dispersive [EDXRF] systems), and *validity* (the ability to collect and analyze data to differentiate raw material types and assign artifacts to those types) (Newlander et al., 2015).

In this paper, we focus on the latter topic – validity – and how trace element data may be used to assign artifacts to geologic sources of raw material. We briefly review the range of approaches that analysts may use when making source assignments. We then present a novel method of data analysis that draws from the subfield of forensic anthropology. In this approach, we use the computer program Fordisc to assign artifacts to obsidian types. Analysts typically use Fordisc to help in establishing the biological profile for a set of unknown skeletal remains. To the best of our knowledge, our study represents the first time that Fordisc has been used in a source provenance study. Fordisc is easy to learn and use, provides custom-order discriminant functions, allows flexibility in analyses, and generates probabilities for individual group assignment as well as model performance. Although our sample of artifacts is small and we recognize potential limitations to the approach, using Fordisc to make source assignments represents an improvement over, or alternative to, other means of comparing univariate and bivariate trace element data.

---

* Corresponding author.
E-mail addresses: mpilloud@unr.edu (M.A. Pilloud), dreaux@nevada.unr.edu (D.J. Reaux), geoffreys@unr.edu (G.M. Smith), krismw81@gmail.com (K.M. Wiggins).

http://dx.doi.org/10.1016/j.jasrep.2017.04.023
Received 29 November 2016; Received in revised form 17 April 2017; Accepted 23 April 2017

2. Source provenance studies and source assignment practices

Source provenance studies are a routine component of many archaeological projects. Such studies use various techniques (e.g., XRF, neutron activation analysis, inductively coupled plasma mass spectrometry) to determine trace elements in artifacts (e.g., ceramics, obsidian and chert tools) and match their source profiles or signatures to those of geologic sources of raw material (e.g., clay, temper, stone and glass outcrops). Source provenance data help researchers calculate the distances and directions that artifacts (or the raw materials used to produce them) were conveyed. In turn, this information is used to interpret how and where prehistoric populations traveled (e.g., Jones et al., 2003; Shackley, 1990, 1996, 2002; Smith, 2010), how groups organized their lithic technology (e.g., Smith et al., 2013), and whether raw materials were obtained via exchange or procured directly (e.g., Beck and Jones, 2011; Kelly, 2011; King et al., 2011).

There are a number of ways that researchers can compare and correlate trace element data collected from artifacts and geologic sources. The simplest is to analyze elemental data one variable at a time to identify similarities and differences between source samples and an artifact. However, these univariate methods may result in the misclassification of an artifact's source (Glascock et al., 1998). An improvement is to compare two variables at a time, typically through the use of graphical methods (e.g., bivariate scatterplots or ternary diagrams). Combinations of various trace element data are displayed and artifacts are visually compared to geological reference sources. Scatterplots may be generated using widely-available programs such as Microsoft's Excel or IBM's SPSS, although this can be time-consuming and could potentially lead to incorrect source assignments. More sophisticated graphical methods, such as those available through the free GAUSS software developed by the Archaeometry Laboratory at the University of Missouri Research Reactor (MURR), may also be used. The GAUSS program allows analysts to generate multiple bivariate scatterplots of various trace element combinations and quickly calculate confidence ellipses. In many cases, especially involving regions where only a few obsidian types were available, this approach may be sufficient to correctly assign artifacts to geological sources (Glascock et al., 1998). In other cases, including regions where chemically distinct obsidian types are numerous, bivariate scatterplots or other basic graphical methods may be inadequate to differentiate raw material types (Glascock et al., 1998).

A variety of multivariate methods including principal components analysis (PCA), cluster analysis, and discriminant function analysis (DFA) may also be used to aid in source assignments; however, these methods have some shortcomings. Using diagnostic trace element data, cluster analysis assigns individual samples to distinct groups (i.e., obsidian types) based on any number of variables. There are many algorithms to calculate a cluster analysis but hierarchical techniques are most commonly used. In this method, dendrograms are created that graphically illustrate the arrangement of clusters and the distances between groups (Manly, 2005). While dendrograms may accurately characterize differences between members *within* clusters, Glascock et al. (1998) note that because cluster analysis generally assumes that trace elements are uncorrelated, it can misrepresent differences *between* clusters. Clearly, this is a problem when the ultimate goal is to assign artifacts to particular obsidian sources.

Principal components analysis is a multivariate statistical technique that linearly transforms a set of variables into a set of uncorrelated indices or components. An advantage of this approach is that if the variation can be adequately represented in a few components, large datasets can be effectively described with fewer variables (Manly, 2005). Principal components analysis can be used to identify patterns in the data and the components generated can be used in further analyses that require uncorrelated variables (e.g., cluster analysis); however, PCA is not a classification or distance statistic, which is ultimately what is needed to classify an unknown sample in provenance studies.

Discriminant function analysis is a means to address which variables separate two or more defined groups. In the case of source provenance data, the elemental variables would be used to predict group assignment (i.e., source). A common approach in DFA uses Mahalanobis distance as a means to calculate group centroids from which individual cases can be classified. Once a model is created, it is possible to allocate unknown individuals to one of the groups in the model, with the assumption that the reference sample is representative of the origin of the unknown artifact. The error rate of the overall model (i.e., the accuracy of the assignment of known individuals to the correct groups) can be used to evaluate individual assignment (Manly, 2005).

A major advantage of using multivariate statistical methods, rather than univariate or bivariate methods, is that they capture more of the sample variation in addition to providing some level of certainty (i.e., p values, probabilities, error rates) that the models are working correctly. However, as outlined above, they are not without problems, particularly in the case of assignment of an unknown artifact. For example, cluster and principal components analyses are not well-suited for the assignment of an individual artifact; rather, they explore similarities between already identified groups. Conversely, DFA can provide equations to be used to assign an unknown artifact to a reference sample. However, depending on the statistical package being used, the result does not necessarily provide probabilities of correct assignment of that individual artifact to an obsidian type; it only provides probabilities of the overall model performance. Moreover, calculating these equations can be cumbersome if done by hand and there are many possible source assignments.

Here, we offer an alternative approach: we explore the use of the computer program Fordisc to assign individual artifacts to obsidian types based on a reference sample. Fordisc is an interactive software package widely used by forensic anthropologists to estimate ancestry, sex, and stature of a set of unknown skeletal remains in a medicolegal context; however, because it offers the ability to import databases to be used within its statistical framework, it can be used on any type of continuous data including the elemental composition of obsidian artifacts.

3. Materials and methods

To evaluate Fordisc's utility in assigning artifacts to particular obsidian types, we used a sample of 677 artifacts from various sites in the northwestern Great Basin currently housed at the University of Nevada, Reno (Table 1). These artifacts were previously characterized by the Northwest Research Obsidian Studies Laboratory (NWROSL) in Corvallis, Oregon. Over the past decade or so, NWROSL staff characterized the artifacts using either a Spectrace 5000 EDXRF (pre-2012) or ThermoElectron QuanX EC EDXRF spectrometer (2012–2016). They determined the concentrations of various trace elements (e.g., Ti, Mn, Fe_2O_3, Zn, Ga, Rb, Sr, Y, Zr, Nb, Ba) in each artifact and compared them

Table 1
Sample distribution of obsidian sources used in this study.

Sample	N
Badger Creek	15
Beatys Butte	111
BSPPFM	20
Buck Spring	59
Cowhead Lake	14
Coyote Spring	19
Craine Creek	32
Horse Mountain	35
MLGV	357
Tank Creek	15
Total	677