Journal of Archaeological Science: Reports xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Journal of Archaeological Science: Reports

journal homepage: http://ees.elsevier.com/jasrep

An experimental design for the classification of archaeological ceramic data from Cyprus, and the tracing of inter-class relationships

Elisavet Charalambous [a],*, Maria Dikomitou-Eliadou [b], Georgios M. Milis [c], Georgios Mitsis [a,d], Demetrios G. Eliades [e]

[a] *Department of Electrical and Computing Engineering, University of Cyprus, Nicosia, Cyprus*
[b] *Archaeological Research Unit of the Department of History and Archaeology, University of Cyprus, Nicosia, Cyprus*
[c] *G.M. EuroCy Innovations Ltd, Nicosia, Cyprus*
[d] *Department of Bioengineering, McGill University, Montreal, QC, Canada*
[e] *KIOS Research Center for Intelligent Systems & Networks, University of Cyprus, Nicosia, Cyprus*

ARTICLE INFO

Article history:
Received 4 November 2014
Received in revised form 22 June 2015
Accepted 13 August 2015
Available online xxxx

Keywords:
Experimental design
Ancient pottery analysis
Classification
Compositional analysis
Statistical testing

ABSTRACT

This paper proposes an experimental design for the compositional classification of 177 ceramic samples deriving from domestic and tomb contexts in Cyprus dated to the Early and Middle Bronze Age. In this design, ceramic sample classification is achieved with three well-known methods, a standard statistical learning method termed k-Nearest Neighbours (k-NN), a method using Decision Trees (C4.5) and a more complex neural network based method known as Learning Vector Quantisation (LVQ). It is shown that the examination of classification patterns through confusion matrices allows the exploitation of inter-class relationships and the ability to provide extra information to the researcher about the compositional categorisation of samples; which could not be grouped (with certainty) into classes with the employment of ceramic petrography. Due to the compositional heterogeneity of ceramics, the effectiveness of classification using only chemical elements with mean concentrations lower than 0.1% is also evaluated to illustrate their potential significance. The developed design follows a systematic approach and well-established methods, such as bootstrapping with replacement and the 5 × 2 cross validation (paired t-test and F-test) tests, to ensure that the results are statistically significant.

1. Introduction

Archaeology ultimately aims at investigating social causation through the examination of gathered residue evidence (Barceló, 2008). Pottery analysis, in particular, has been proven cross-culturally an indispensable tool for indirectly approaching past people and societies through their cultural remnants, allowing inferences about their technology, and their interaction with their surrounding physical and social environments. For this reason, compositional (mineralogical and chemical) and microstructural analyses have become an integral part of interdisciplinary archaeological research, underlining the importance of compositional and technological comparative studies. Nonetheless, any pottery analysis is not a straightforward process, and there are various parameters (i.e. contextual, spatial, chronological, compositional, technological) that the researchers need to consider while defining their research design, their sampling strategy, and later while evaluating their research results. Among these parameters, the inherent heterogeneity characterising ceramic composition sets significant challenges when trying to utilise the greatest amount of information possible, especially considering that, generally, the most highly variable elements have the greatest of the impacts on the multivariate data ensemble and that they do not necessarily depict elements with high concentrations (Reimann et al., 2012).

Multivariate statistical methods have a long track in archaeometric data analysis of which the most common are cluster analysis and dimensionality reduction techniques (both supervised and unsupervised). The analysis of archaeometric data imposes problems that are not easily handled, if at all, by classical methods. Due to this, over the past couple decades there has been a great interest in alternatives to the standard statistical methods of analysis (Baxter, 2006). The size of the produced datasets coupled with the multiple influencing analysis and contextual factors impose both analytical and computational problems, while it is important to note that the selection of the most effective analysis method depends on the characteristics of the data.

Statistical analysis on data may allow the study of their internal structure and reveal interesting technological and compositional patterns. As such, archaeological data classification is concerned with the application of classification methods on archaeological data, while also taking into account the characteristics of the artefacts under study.

---

* Corresponding author.

http://dx.doi.org/10.1016/j.jasrep.2015.08.010
2352-409X/

Please cite this article as: Charalambous, E., et al., An experimental design for the classification of archaeological ceramic data from Cyprus, and the tracing of inter-class relations..., Journal of Archaeological Science: Reports (2015), http://dx.doi.org/10.1016/j.jasrep.2015.08.010

Classification aims at identifying to which element of a set of categories, a new uncategorised artefact belongs, on the basis of a training set of artefacts the class/type of which is known.

Classification results are not unique; they depend on the deployed classification method, their parameterisation, as well as the manner in which the raw data are treated to form the input dataset. Evaluation of the validity and the plausibility of classification results is not only necessary, but critical. Statistical hypothesis testing methods allow the inference of a hypothesis ensuring that the predicted result is unlikely to have occurred by chance alone, according to a pre-determined threshold probability (Coolican, 1999).

The roots of classification analysis of archaeometric data are traced back multiple decades ago with the contribution of Kowalski in 1972 (Kowalski et al., 1972) being an early landmark. In subsequent years, classification methods have been used in a number of studies (Mussumarra et al., 1995; Fermo et al., 2008; Kowalski and Bender, 1973; Baxter, 1994; Lopez-Molinero et al., 2000). A clear milestone in the analysis of archaeometric data is Baxter's work in 2006, where he reviews the application of classification methods (among others) on the chemical composition of glass artefacts (Baxter, 2006). The effectiveness of a variety of classification methods was evaluated; among them also the three methods deployed in this paper. However, despite the similarities between Baxter's review and this paper, the results of the two works cannot be straightforwardly comparable due to the different experimental data and deployed methodology.

In this paper, an experimental design is proposed for the classification of chemical compositional data obtained from a sample of utilitarian pottery. The aim of the experiment is neither to achieve perfect classification, nor to discriminate the origin of each artefact. The target is rather to develop a plausible, unbiased and statistically valid methodology for classification, which takes into consideration the idiosyncrasies of archaeometric data, in general, and chemical compositional in particular, and also to examine the validity of the produced categorisation. The proposed methodology is subsequently used to differentiate a series of ceramic specimens based on their fabric, and investigate the degree of similarity between discriminated types. For demonstration purposes and for the needs of this paper, classification is achieved with three well-known methods, a standard statistical learning method termed k-Nearest Neighbours (k-NN) (Duda et al., 2012), a method based on Decision Trees (C4.5) (Quinlan, 1993) and a more complex structure, based on neural networks, known as Learning Vector Quantisation (LVQ) (Kohonen, 2001). The selection of these three algorithms was driven by the need to test the effectiveness of different types of algorithms on the analysis of archaeological data in an effort to exploit different artefact attributes. Despite our selection of classification methods, the proposed design may be realised in combination with any classification method.

The deployment of established methods allows the evaluation of the validity of the results through the use of a special form of cross validation testing. The developed design follows a systematic approach and well-established methods, such as bootstrapping with replacement (Efron and Tibshirani, 1986) and the 5×2 cross validation (paired t-test and F-test) tests in order to ensure that the results are statistically significant. The proposed scheme is tested with the use of a sample of Early and Middle Bronze Age utilitarian pottery from Cyprus. The statistical experiment involved two analytical datasets deriving from the mineralogical and chemical characterisation of 177 ceramic samples, with the respective employment of petrography on ceramic thin sections and ED-XRF on pressed-powder pellets (Dikomitou, 2012).

1.1. Archaeometric analysis of ceramic data

Archaeological data is often characterised as complex data due to the large number of involved influencing factors during the analysis procedures. Much attention is given upon the gathering of the archaeological artefacts during excavations and their subsequent micro-structural, mineralogical and/or chemical analysis. However, many parameters influence the reliability of the produced data. Different archaeologists implement the same procedures in different ways, thereby increasing the within-class variance. This problem intensifies by taking into account that apart from variations generated due to the human factor, the acquired variability is also caused due to the deterioration of the source material because of its natural ageing as well as the environment of deposition.

The analysis of archaeological data is not a straightforward task. The aim of the archaeologist is to make inferences by taking into consideration as many parameters as possible. Ceramic classification remains the principal approach to the study of pottery in identifying patterns in the data. The most common way to categorise pottery is primarily based on the macroscopic observation of technological attributes and morphological types; extra attention is given to the shape, size and surface treatment. An alternative to this method is the characterisation based on their chemical composition by isolating ceramic ground of similar chemical profiles and statistically testing the validity of those groups (Garcia-Heras et al., 2001). The use of techniques achieving a clearer separation in different groups also results in increasing their interpretability. The validity of the emerging groupings can be further evaluated through typological and potentially mineralogical comparisons with data bearing known fingerprints, so as to address different aspects of ancient ceramic production and distribution.

2. Classification of chemical compositional archaeological data

2.1. Chemical compositional analysis of ceramics in archaeology

Chemical compositional data are defined as vectors of strictly positive components, usually expressed as percentages or parts-per-million (ppm), with constant sum — a restriction not always maintained. Quantitative chemical analysis is not involved in measuring, but in enumerating, or counting, the number of each type of atoms in a sample (Buxeda, 2008). Chemical compositional data do not vary independently and concentration based approaches to data analysis can lead to misleading conclusions (Reimann et al., 2012).

Chemical compositional data lay in the constrained Simplex Space (Aitchison et al., 1982; Buxeda, 2008), where correlation analysis and the Euclidean distance are not mathematically meaningful concepts (Reimann et al., 2012). Furthermore, graphical depiction of raw or log-transformed data should only be used in an exploratory data analysis sense, to detect unusual data behaviour or candidate subgroups of samples (Acton, 2013).

The chemical constituents of an archaeological artefact, or any other object, can be categorised into major and trace elements. Major elements comprise large proportions of the artefact under analysis, while trace elements are present in concentrations less than 0.01%. As ceramics are heterogeneous in composition with the majority of their major elements present in most artefacts, the discrimination of objects into groups makes necessary the utilisation of trace elements in determining the fingerprint of a deposit (Mirti et al., 1994).

2.2. The classification problem

Classification is a procedure that aims to assign items to a number of (possibly pre-known) target categories or classes based on statistical/machine learning principles (Bishop, 2006). It is an instance of supervised/unsupervised learning, whereby the former assumes that a training set of correctly identified observations is available (Bishop, 2006). Classification of archaeological ceramics deals with the categorisation of ceramic specimens of similar chemical profiles, given a number of artefacts of known fabric identify.

The supervised classification problem may be broken down into two separate stages: 1) the inference stage where training data is used to learn a model and 2) the decision stage in which the trained model is

Please cite this article as: Charalambous, E., et al., An experimental design for the classification of archaeological ceramic data from Cyprus, and the tracing of inter-class relations..., Journal of Archaeological Science: Reports (2015), http://dx.doi.org/10.1016/j.jasrep.2015.08.010