Spectrochimica Acta Part B 172 (2020) 105966

Contents lists available at ScienceDirect

Spectrochimica Acta Part B

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

Analytical note

Graph clustering and portable X-Ray Fluorescence: An application for in situ, fast and preliminary classification of transport amphoras

E. Odelli[a,b], V. Palleschi[c], S. Legnaioli[c], F. Cantini[a], S. Raneri[c,*]

[a] University of Pisa, Department of Civilization and Forms of Knowledge, Via dei Mille 19, 56126 Pisa, Italy
[b] Department of Archaeology, Ghent University, Sint-Pietersnieuwstraat 35, 9000 Ghent, Belgium
[c] Applied and Laser Spectroscopy Laboratory, Institute of Chemistry of Organometallic Compounds, Research Area of National Research Council, Via G. Moruzzi, 1 – 56124 Pisa, Italy

ARTICLE INFO

Keywords:
Potteries
Portable ED-XRF
Graph Clustering
Provenance
Non-destructive analysis

ABSTRACT

In the last decade, numerous papers have been delivered on the potential of portable X-Ray Fluorescence (XRF) in archaeological ceramics. Additionally, new chemometric methods have been proposed to manage chemical dataset and facilitate the use of geochemical discrimination for provenance classification of ancient ceramics. In this contribute, the potential of portable Energy Dispersive X-Ray Fluorescence (ED-XRF) analysis and chemical data processing by Graph Clustering is evaluated for provenance classification of archaeological potteries, discussing possible merits and limits of the employed routine. A ceramic assemblage represented by seventy-three transport amphorae classified by typological analysis have been used as testing materials; spectra have been collected on samples simulating in situ analysis conditions (*e.g.* on fresh cut surfaces without any preparation) and Graph Clustering method has been applied in chemical data processing; comparison with classical Cluster Analysis (CA) and Principal Component Analysis (PCA) is also evaluated. The obtained results favor the use of Graph Clustering for a preliminary classification of ceramics, which can be chemically analyzed in easy, fast and non-destructive way. With a 75.35% of correct attribution, the study shows the suitability of portable ED-XRF in rapid screening of a large number of ceramic samples usually recovered in the framework of archaeological excavation. Misclassifications have been mostly verified for samples exhibiting a coarse-grained clay paste, suggesting that the method is particularly suitable for fine-grained ceramic materials.

1. Introduction

Potteries represent among the most numerous records in archaeological excavation; archaeologists use potteries to create statistic compilation, periodization and typologies based on the similarities and differences between types, styles or features that are relatively continuous in time and/or geographic areas. Typological analysis usually support interpretation regarding trade routes and exchanges of goods among sites; however, certain provenance attribution often requires the support of compositional and geochemical analysis. The technological advancements and the development of affordable portable analytical methods has made possible the characterization of large number of artifacts quickly, easily and in non-destructive and non-invasive way [1–6], meeting the requirements of the archeologists to obtain the maximum amount of information from minute samples or directly in situ from intact objects.

In this perspective, the use of portable X-Ray Fluorescence analysis for pottery analysis has drawn a great attention in the last decades; numerous papers have been in fact published trying to understand the affordability of portable chemical methods in ceramic classification and provenance investigation [7,8]. The literature in the field seems to be split in two main factions, including researchers considering the application of portable ED-XRF in ceramic and sediment studies as a challenge, although recognizing its limits and drawbacks [9–16], and others which warns against the limits of a technique that cannot substitute the classical chemical investigations [17]. Actually, limitations due to detection limits in light elements, density of analyzed materials, surfaces vs. bulk composition, heterogeneity of ceramic materials, variability in measured intensity and matrix effects have to be taken in great consideration in ED-XRF analysis of potteries [18]. The replacement of chemical analytical methods with portable XRF is not possible; however, the application of carefully tailored measurement and processing protocols might provide good statistical results, returning classifications similar to the ones obtained by laboratory methods such

* Corresponding author.
E-mail address: simona.raneri@pi.iccom.cnr.it (S. Raneri).

https://doi.org/10.1016/j.sab.2020.105966
Received 5 May 2020; Received in revised form 20 August 2020; Accepted 22 August 2020
Available online 25 August 2020
0584-8547/

Table 1
Chemical data obtained by portable ED-XRF analysis. Elemental concentrations are reported in arbitrary units.

Sample ID	Si	K	Ca	Ti	Mn	Fe	Ba
DIS16	29.1	9.4	53.3	4.1	0.6	17.6	0.9
DIS17	58.4	6.1	4.1	7.0	0.3	37.5	0.8
DIS36	35.8	8.3	40.3	4.2	0.8	23.9	1.3
DIS54	21.4	8.2	53.8	5.0	0.3	24.8	0.9
DIS56	24.5	4.4	59.8	3.5	0.2	15.8	0.4
DIS58	24.3	5.9	59.9	4.0	0.2	15.9	0.8
DIS97	63.7	5.4	3.5	5.9	0.4	32.8	0.4
DIS100	24.9	5.1	63.4	2.9	0.3	11.7	0.6
DIS123	29.5	6.4	45.0	3.8	0.6	25.4	0.8
DIS141	31.0	4.6	55.5	3.0	0.3	13.5	0.6
DIS142	20.2	6.2	64.7	3.0	0.4	15.1	0.6
DIS143	25.3	5.6	51.4	4.2	0.4	23.3	0.7
DIS144	37.0	3.4	49.6	3.2	0.4	13.3	0.6
DIS168	29.3	6.3	42.1	4.1	0.7	28.6	0.5
DIS170	43.5	4.6	33.4	3.1	0.5	23.1	0.7
DIS173	21.1	5.5	66.2	3.1	0.4	12.7	0.8
DIS174	26.6	7.5	54.2	4.2	0.7	19.2	0.6
DIS232	32.3	6.0	43.5	4.4	0.3	24.2	0.9
DIS249	32.8	5.2	49.5	3.6	0.2	17.7	0.5
DIS260	24.7	4.7	57.5	3.1	0.1	17.7	0.6
DIS261	36.5	6.5	36.4	4.0	0.7	27.1	0.9
DIS262	37.2	8.7	11.3	6.9	0.8	51.5	0.5
DIS263	29.8	8.7	49.6	4.3	0.5	20.6	0.6
DIS264	27.2	8.1	52.8	4.3	0.5	20.0	0.8
DIS265	34.8	6.8	39.1	4.2	1.0	26.1	0.7
DIS268	33.4	6.9	45.4	3.6	0.5	21.1	0.8
DIS277	36.1	9.1	35.3	4.4	0.7	28.6	1.0
DIS278	37.0	8.0	34.7	4.0	0.7	28.3	0.7
DIS282	30.4	7.1	49.0	5.0	0.3	20.6	0.4
DIS283	49.0	5.4	24.3	4.2	0.3	26.6	1.0
DIS284	32.8	7.7	41.0	4.3	0.8	26.2	1.0
DIS287	36.2	5.9	39.9	3.8	0.5	23.9	0.8
DIS288	33.6	6.6	33.8	4.3	0.8	32.6	1.0
DIS293	30.5	7.9	52.1	4.3	0.4	17.4	0.7
DIS320	33.8	7.7	33.6	4.4	1.0	32.6	0.9
DIS321	59.6	9.5	3.0	6.0	0.3	37.5	1.3
DIS322	30.2	8.9	34.9	4.5	0.9	34.9	0.9
DIS323	55.0	9.2	2.9	5.4	0.2	42.0	0.6
DIS324	48.6	7.0	6.5	5.0	0.5	44.9	1.4
DIS325	29.4	7.1	50.4	3.7	0.8	20.2	1.1
DIS326	29.9	5.1	52.8	3.6	0.3	17.3	0.6
DIS399	23.1	4.2	52.8	3.1	0.3	17.8	0.6
DIS401	23.7	7.6	61.9	3.7	0.4	14.4	0.7
DIS424	33.7	3.9	50.5	2.9	0.3	15.8	0.5
DIS427	36.4	6.2	33.5	5.2	0.5	30.1	0.7
DIS450	49.5	7.7	13.6	5.4	0.8	36.9	0.6
DIS451	21.5	3.9	69.6	2.5	0.3	8.9	0.5
DIS504	31.9	4.0	28.7	2.7	0.6	39.4	0.8
DIS506	31.2	6.8	47.7	3.5	0.6	21.0	0.9
DIS507	18.4	4.5	61.2	3.1	0.7	20.3	0.5
DIS508	46.7	8.6	8.0	5.6	0.7	45.2	1.2
DIS521	32.6	8.4	38.6	4.6	0.6	28.9	1.1
DIS522	35.4	7.2	35.2	3.9	0.6	29.4	0.9
DIS527	23.6	5.8	60.3	2.8	0.4	16.2	0.9
DIS529	25.6	4.6	63.0	2.7	0.4	11.4	0.8
DIS530	25.2	4.5	55.1	3.7	0.8	19.7	0.9
DIS548	27.8	5.9	52.8	3.1	0.6	19.4	0.8
DIS549	26.4	6.0	44.1	4.1	0.6	29.5	0.7
DIS550	31.6	5.7	49.0	2.9	0.5	19.4	0.9
DIS559	22.2	6.3	67.0	2.9	0.4	10.8	0.7
DIS560	33.1	8.1	41.0	4.2	0.6	26.0	0.5
DIS561	20.8	8.5	9.8	8.6	0.7	69.3	1.3
DIS562	40.2	6.2	25.5	4.4	0.5	34.3	1.0
DIS583	51.1	7.8	11.3	5.5	0.9	37.6	1.1
DIS584	16.9	6.6	72.9	3.3	0.4	10.2	1.0
DIS585	31.1	7.0	49.7	3.6	0.5	19.2	0.8
DIS586	33.8	7.0	47.4	3.1	0.5	18.9	0.7
DIS587	40.8	8.0	31.3	4.0	0.7	27.8	0.7
DIS592	30.9	6.7	48.8	3.4	0.5	20.3	0.7
DIS594	18.6	6.5	71.6	2.8	0.4	9.8	0.4
DIS623	33.9	6.3	45.3	3.6	0.5	20.9	0.9
DIS649	41.9	5.3	6.4	5.8	0.3	51.7	1.5
DIS657	36.6	3.8	47.9	3.0	0.3	15.5	0.8

as instrumental neutron activation analysis (INAA), Inductively coupled plasma mass spectrometry (ICP-MS), Wavelength Dispersive X-ray Fluorescence (WD-XRF), scanning electron microscopy coupled with Energy Dispersive X-ray Analysis (SEM-EDS), etc. [19–21].

Criticism is particularly decisive when the method is applied to obtain a detailed quantitative composition of ceramics for provenance studies; in that case, the grain-size and the possible heterogeneity of the paste might play a fundamental role in returning significantly different results moving from one analysis spot to another, as well as from the surface (sometimes interested by slips, glazes or engobes) to the bulk. These limits can be minimized by working on fine-grained and depurated shards, which are usually characterized by an homogeneous paste in which very tiny clasts are dispersed [18,22]. Portable XRF method might be employed for a preliminary discrimination among ceramic classes in support of archeological interpretation, especially when shapes and morphological analysis already provide clues on provenance and manufacturing (such as for transport amphorae), by supplying some *a priori* information on the studied potteries [21].

Another crucial aspect in evaluating the not trivial results supplied by portable XRF investigation is related to the data processing; in provenance studies, chemical data are usually processed by chemometric methods [8,23,24], including PCA [25], dendrograms [26], cluster analysis [27], artificial neural networks [28], etc. However, the results given back by these statistical approaches might introduce flaws - independently on the specific experimental methods used - related to the amplification of small differences among samples due to possible merely intensity fluctuations in the analysis of heterogeneous materials [29]; the interferences which may be caused by the surface morphology, by coatings applied on the surface and by mineralogy itself, have been previously discussed [30].

An alternative processing approach can be represented by the application of the Graph Clustering (GC) method [31,32]. The GC method is particularly suited for classification of spectra (and then samples, since the spectra are hypothesized representative of the samples) based on their similarities. The parameter used for measuring the degree of similarity is the correlation between the spectra, mathematically expressed as the normalized scalar product between the two vectors representing the intensities of the XRF signal at the different energies. The correlation may vary from 1 (identical spectra) to 0 (completely uncorrelated spectra). The correlations between the n XRF spectra are organized in a $n \times n$ symmetric adjacency matrix, which is interpreted in the framework of the Graph Theory as an undirected Graph with the spectra as nodes and the mutual correlation as similarity parameter. Usually, a threshold is set, below which the correlation is considered null. The advantage of the GC method, with respect to other statistical approaches, is the robustness of the classification (spectra strongly correlated between them are usually clustered in well separated groups) and the possibility of spotting easily outliers, which appear as single nodes not correlated with any group.

In this scenario, this study aims to explore the information potentially offered by the application of GC analysis on portable ED-XRF data in supporting the preliminary screening of archaeological ceramics, with a focus on transport amphorae. The forecast application intends to provide a tentative grouping of archaeological fragments, without replacing the laboratory analysis, which is in any case essential for the geochemical characterization of the samples.

This approach appears quite useful in the framework of large archaeological excavations that often return a huge number of ceramic fragments. The usual protocol includes drawing and archiving of the shards, and the subsequent microscopic observation of shapes and pastes. Coupling this practice with a portable XRF investigation might help for a better classification of doubtful fragments.

In order to test this routine, 73 amphorae fragments representative of different production centers through the Mediterranean area between 1st BC–5th–7th AD and found in the Acropolis of Volterra (Italy) excavation have been selected and analysed by portable ED-XRF.

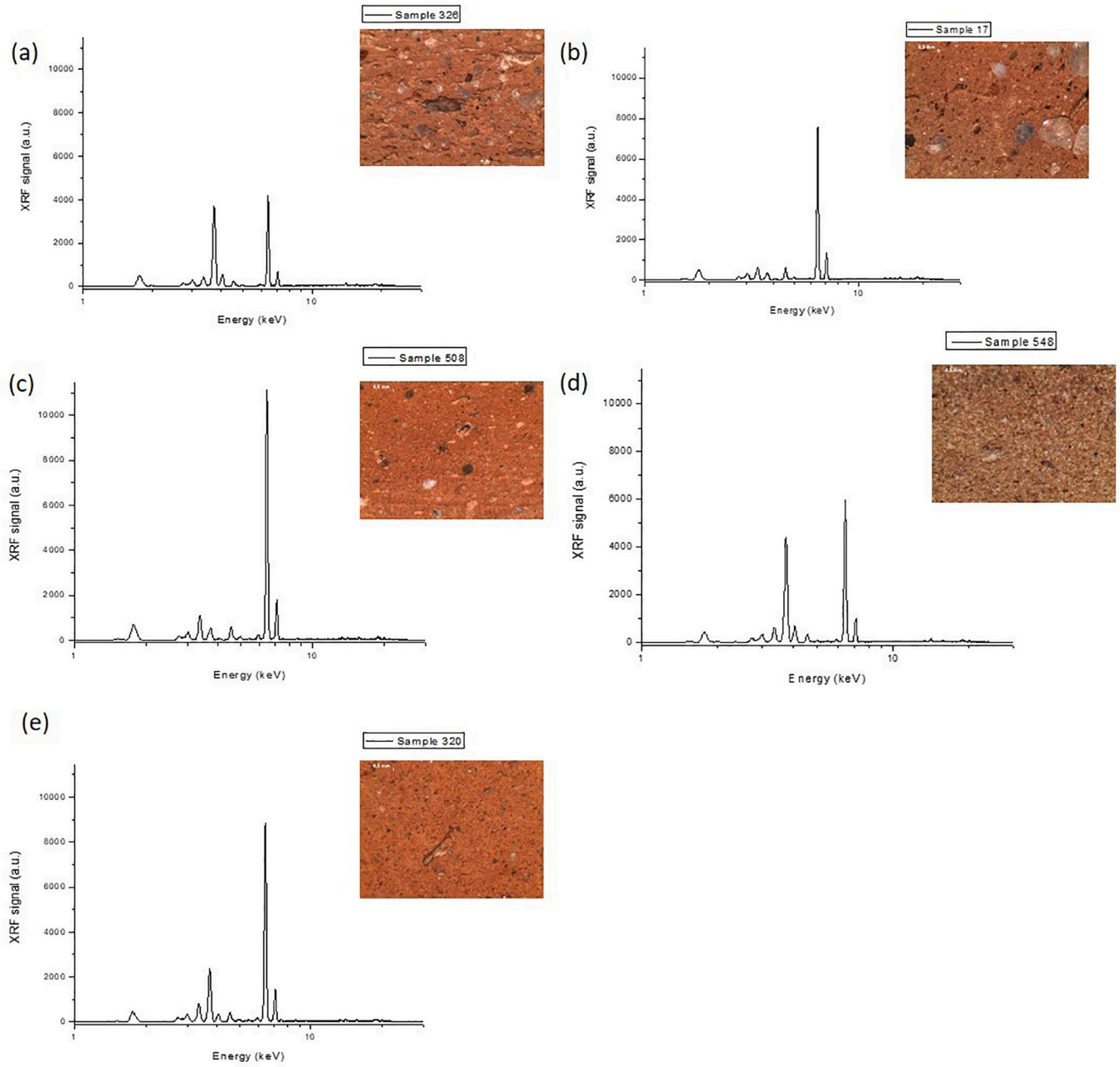

Fig. 1. ED-XRF spectra collected on amphoras from (a-b) African regions (sample 326, sample 17); (c) Iberian Peninsula (sample 508); (d) Gallia provinces (sample 548) and (e) Italy (sample 320), as example of the different clay ceramic pastes.

The collected XRF spectra were thus processed using the GC method to evaluate limits and merits of the proposed approach, comparing the results with the archaeological classification of the shards. The results indicate that the application of a protocol based on portable XRF is a tangible treat for pottery screening in field archaeology, making it possible the reliable classification of the samples in different groups related to different productions centers and characterized by peculiar clay paste features.

2. Experimental

2.1. Materials

Seventy-three fragments of transport amphorae, dated from the 2nd BC to the 8th AD, from the archaeological excavations of Volterra (Italy) Acropolis (1967–1979) have been selected for this study (Table S1). The typological classification of the transport amphoras has been provided according to the literature [33–43]. Overall, the main provenance areas are represented by Western Mediterranean area - including Iberian Peninsula and Gallia-, North-Africa and Italy.

2.2. Methods

XRF analysis have been carried out by using the Bruker ELIO portable ED-XRF spectrometer. The instrument has a weight of about 2 kg and is equipped with an SDD detector and an X-Ray tube with Rhodium anode. The measuring spot on the surface is about 1 mm in diameter. The spectra were acquired using a 30kV tension of the X-ray tube, 90 μA current, 90 s acquisition time. The spectra have been acquired on fresh cut surfaces of ceramic body, gently cleaned before the analysis to

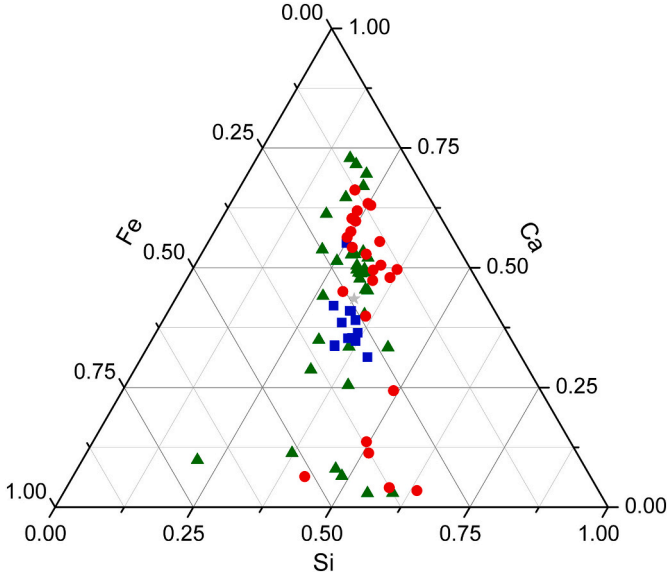

Fig. 2. Ternary diagram for the major elements (Si, Ca and Fe) in the measured samples. Red circles: Africa, Green triangles: North West Europe, Blue squares: Italy; Gray stars: Not Classified(For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

simulate in situ working conditions; the point of analysis was on the matrix, trying to avoid interferences due to mineral grains, even fine-grained, dispersed in the paste. Care was taken to acquire several spectra in different points of the same sample and in different moments, to assess the reproducibility of the analysis. In Fig. Supplementary Fig. S1 is shown the comparison of several XRF spectra, acquired in different points on one sample selected as example. Two spectra were acquired on the fresh cut surface, and two other points were acquired on the raw surface, two months after the first acquisitions. The reproducibility (and, consequently, the classification) of the spectra was extremely good; for this reason, in the following analysis we acquired a single spectrum per sample, and this was enough for obtaining the classification discussed in the paper.

The acquired spectra were used to build the adjacency matrix; the corresponding Graph was graphically displayed using the Gephi software (ver. 0.8.2, Force Atlas Layout). The threshold for the correlation was optimized for having well distinct clustering, but also for minimizing the unclassified samples (samples with correlation under the threshold for all the other samples). The optimum classification was obtained with a threshold in the correlation fixed at 0.99. The parameter used for measuring the similarity among spectra is the correlation between the spectra, mathematically expressed as the normalized scalar product between the two vectors representing the intensities of the XRF signal at the different energies.

$$C(S1, S2) = \frac{\langle S1 \mid S2 \rangle}{\sqrt{\langle S1 \mid S1 \rangle \langle S2 \mid S2 \rangle}}$$

The correlation may vary from 1 (identical spectra) to 0 (completely uncorrelated spectra).

For having a comparison with the results of conventional statistical analysis, the samples were also analysed in terms of their chemical composition using the Fundamental Parameter method [44]. According to this method, the composition of the sample is determined by linking the intensity of the fluorescence lines to the concentration of the relevant element, using the known cross-sections for absorption of the X-rays from the element, the probability of ionization in the corresponding shell, the probability of having a radiative transition of the electrons from the upper shells, etc. These are the fundamental parameters which give the name to the technique. For an absolute determination of the element concentrations, the calibration of the system would be necessary, using a standard (at least) having the same matrix of the ones to be analysed. In our case, the concentration of the elements of interest was not absolutely determined, since the low-Z elements are not measurable with portable ED-XRF, but the relative variations in concentration from sample to sample were correctly recovered, and a classification was attempted using this information.

3. Results

Chemical data showed that, among the detected elements, Si, Ca and Fe seem to express the major variance (Table 1, Fig. 1), even if the comparison of their abundance prevents any clear classification and/or grouping (Fig. 2).

Provenance attribution based on chemical composition has been tentatively attempted by the application of standard chemometric methods as CA and PCA; however, the obtained results do not show any systematic pattern able to create reference groups according to provenance and/or clay paste common features (Fig. 3). On the contrary, the GC method allows to provide interesting insights. Up to the graph clustering classification (Table 2, Fig. 4), five clusters can be identified. The first one (labelled as 1) splits in two branches; the first one including only North-Western European productions, and the second one mainly encompassing African vessels; in the latter one, two North-Western European samples are also included. The second cluster, labelled as 2, mainly includes Iberian Peninsula productions, even if few vessels classified as African productions are also herewith grouped; these misclassifications could be related to the geological similarities between north-African areas and south Iberian Peninsula regions. A third cluster, labeled as 3, groups nine African shapes mainly classified as productions from Zeugitana and Byzacena regions. Another cluster (labeled as 4) groups 23 samples, the 78% of which is due to North-Western European productions (Iberian Peninsula and Gallia); it is interesting to note that four samples classified as Dressel 1 and Dressel 2-4 Tyrrhenian and previously indicated as Italian productions are included in this cluster. Up to the literature, workshops producing these classes have been also identified in the Iberian Peninsula, thus allowing a new classification of the vessels as North-Western European productions. Five samples due to Keay amphorae type are misclassified in cluster 4. The last cluster (labeled as 5) includes the 65% of Italian productions, mainly located in central Italy, along with the 35% of North-Western European productions and one African amphora. Only one sample was not classified by the GC. Although the Graph Clustering method does not provide a 'prototype' spectrum per each cluster, as other methods [45,46], a comparison can be done between the average spectrum associated to the five clusters (Fig. 5, Fig. Supplementary Fig. S2, Fig. Supplementary Fig. S3, Fig. Supplementary Fig. S4).

4. Discussion and conclusions

In this work we have presented a method based on portable ED-XRF technique and Graph Clustering method for a tentative preliminary classification routine of fine-grained potteries, comparing the results with the one obtained by typological analysis and using standard classification methods (CA and PCA). In spite of the presence of a few clear mismatches, probably related to interference of coarser mineral grains in the spectra acquisition, the obtained results demonstrate the advantages of this approach over classical statistical methods and the feasibility of the use of portable ED-XRF instrumentation for a rapid non-invasive and non-destructive classification of pottery.

The proposed procedure achieved a 75.35% of correct attribution and classification of transport amphoras based on geochemical affinities, also verified by observing the similarity of clay paste and textural features. Transport vessels accounting three regional contexts in the Mediterranean area (North-Africa, North-western Mediterranean, and Italy) have been discriminated in five main groups. The analysis

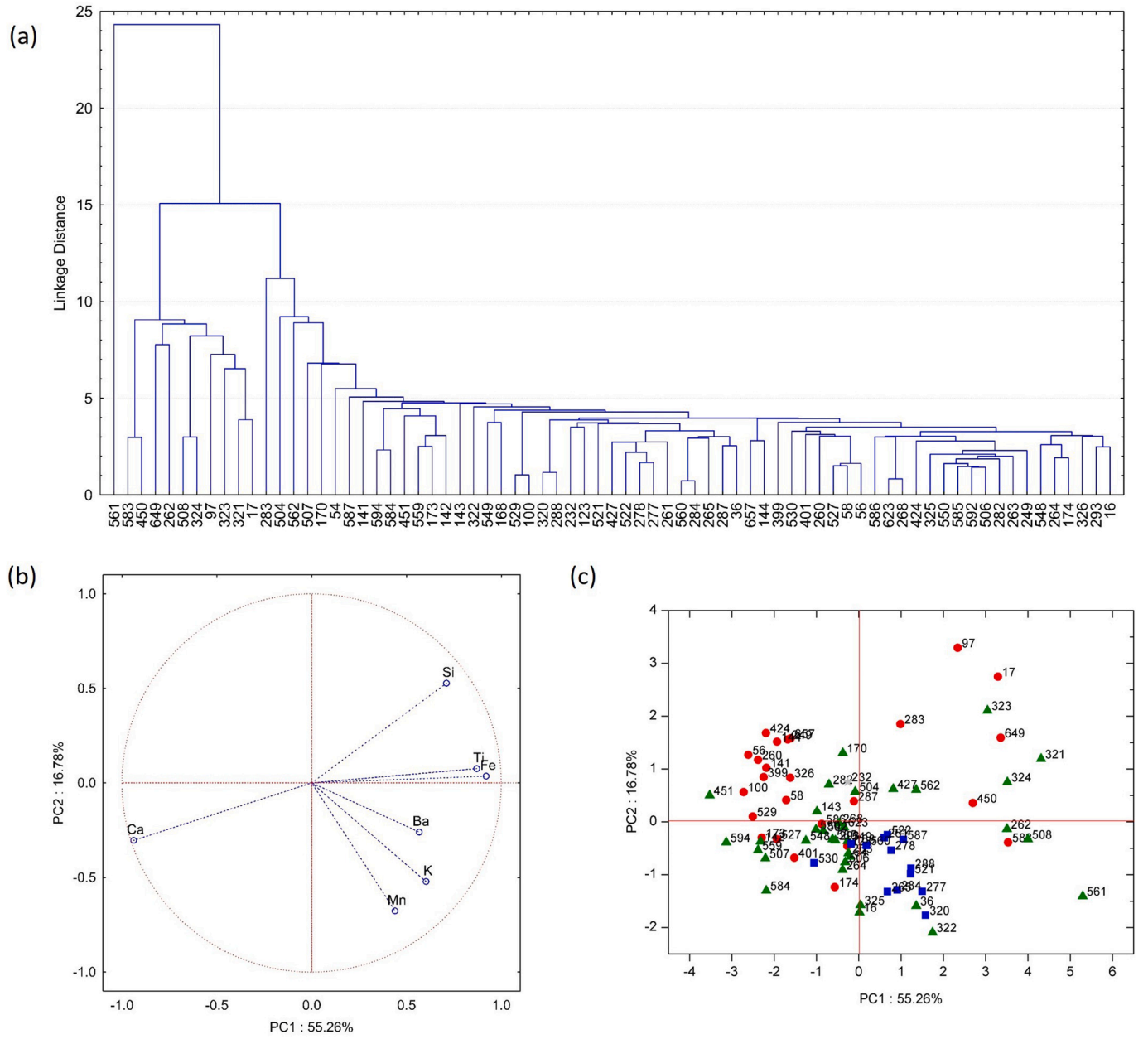

Fig. 3. Application of classical chemometric methods for the ceramic samples studied in this work. Top: Cluster analysis; Bottom left: Factor loadings and Bottom right: PC1 vs. PC2 plot. Red circles: Africa, Green triangles: North West Europe, Blue squares: Italy; Gray: Not Classified(For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

reported the 23.3% of discrepancies between archaeological hypothesis and chemical grouping, while 1.4% of the samples, for which no clear correlations have been evidenced by the Graph, could not be classified. In few cases, even if the typological attribution was doubtful, the Graph Clustering method enabled the discrimination between the possible alternatives. It is noteworthy that discrepancies can be mostly attributed to the interference of mineral grains in the clay pastes, suggesting that the routine is preferable in the case of homogeneous matrix and fine-grain potteries.

In conclusion, this study suggests that portable ED-XRF analysis should not be considered as an alternative to classical laboratory methods (both chemical and petrographic techniques), but can help in a preliminary screening of large amount of fragments, allowing in easy, fast and non-destructive way (in principle directly on site) a first classification of arcaheological potteries on the basis of which further chemical and petrographic analysis can be planned.

The trial and testing of the procedure on different classes of materials, as well as directly in archaeological excavation contexts, might support the development of a more precise protocol, possibly further reducing the discrepancies and mismatching evidenced in this work.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.sab.2020.105966.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Table 2
Provenance attribution based on graph clustering analysis with indication on correct attribution and misclassification.

Sample ID	Archaeological attribution	Graph Clustering	Provenance based on Graph Clustering	Correct attribution (%)	Misclassification (%)
DIS649	Africa (Hammamet gulf)	1 - branch 1	Africa	71%	29%
DIS97	Africa (Byzacena or Zeugitana)	1 - branch 1	Africa		
DIS17	Africa, Tunisia (Nabeul)	1 - branch 1	Africa		
DIS450	Africa (north Tunisia)	1 - branch 1	Africa		
DIS583	Africa (Tripolitania)	1 - branch 1	Africa		
DIS323	Hispania Baetica	1 - branch 1	Western Mediterranean area		
DIS561	Hispania Tarraconensis	1 - branch 1	Western Mediterranean area		
DIS508	Hispania Baetica	1 - branch 2	Western Mediterranean area	100%	0%
DIS262	Gallia Narbonensis	1 - branch 2	Western Mediterranean area		
DIS324	Hispania Terraconensis	1 - branch 2	Western Mediterranean area		
DIS321	Hispania Baetica	1 - branch 2	Western Mediterranean area		
DIS58	Africa (?)	2	Africa	60%	40%
DIS56	Africa (Sahel)	2	Africa		
DIS529	Africa (Buzacena or Zeugitna)	2	Africa		
DIS173	Africa (Nabeul or Sahel)	2	Africa		
DIS142	Hispania Baetica	2	Western Mediterranean area		
DIS559	Hispania Lusitania	2	Western Mediterranean area		
DIS584	Hispania Baetica	2	Western Mediterranean area		
DIS594	Hispania Baetica	2	Western Mediterranean area		
DIS451	Hispania Baetica	2	Western Mediterranean area		
DIS16	Hispania Lusitania	2	Western Mediterranean area		
DIS326	Africa (Byzacena)	3	Africa	100%	0%
DIS657	Africa (Byzacena)	3	Africa		
DIS399	Africa (Byzacena or Zeugitana)	3	Africa		
DIS141	Africa (Byzacena or Zeugitana)	3	Africa		
DIS144	Africa (Byzacena or Zeugitana)	3	Africa		
DIS401	Africa (Sahel)	3	Africa		
DIS424	Africa (Byzacena or Zeugitana)	3	Africa		
DIS100	Africa (Nabeul or Sahel)	3	Africa		
DIS527	Africa (Zeugitana)	3	Africa		
DIS507	Hispania Baetica	4	Western Mediterranean area	78%	22%
DIS506	Hispania Lusitania	4	Western Mediterranean area		
DIS293	Hispania Baetica	4	Western Mediterranean area		
DIS585	Italy (central Italy)	4	Western Mediterranean area		
DIS36	Italy, south-italian regions and Sicily OR Hispania Baetica	4	Western Mediterranean area		
DIS263	Hispania Baetica	4	Western Mediterranean area		
DIS54	Hispania Baetica	4	Western Mediterranean area		
DIS143	Hispania Baetica	4	Western Mediterranean area		
DIS282	Hispania Baetica	4	Western Mediterranean area		
DIS268	Italy (Ager pisanus, volaterranus, Albinia) OR Hispania Baetica	4	Western Mediterranean area		
DIS264	Italy (Ager pisanus, volaterranus, Albinia) OR Hispania Baetica	4	Western Mediterranean area		
DIS548	Gallia Narbonensis	4	Western Mediterranean area		
DIS550	Gallia Narbonensis	4	Western Mediterranean area		
DIS592	Gallia Narbonensis	4	Western Mediterranean area		
DIS623	Gallia Narbonensis	4	Western Mediterranean area		
DIS325	Gallia Narbonensis	4	Western Mediterranean area		
DIS586	Hispania Baetica	4	Western Mediterranean area		
DIS287	Hispania Terraconensis	4	Western Mediterranean area		
DIS174	Algeria?	4	Africa		
DIS123	Africa	4	Africa		
DIS260	Africa (north Tunisia)	4	Africa		
DIS249	Africa, Sahel	4	Africa		
DIS530	Italy, south-italian regions and Sicily (Calabria)	4	Italy		
DIS168	Italy (Ager pisanus, volaterranus, Albinia)	5	Italy	65%	35%
DIS284	Italy (Valdarno, Ager Volterranus)	5	Italy		
DIS265	Italy, south-italian regions and Sicily (Calabria)	5	Italy		
DIS587	Africa OR Italy (Valdarno)	5	Italy		
DIS320	Italy (ager Pisanus, Volterranus, Albinia)	5	Italy		
DIS261	Italy (Tuscany)	5	Italy		
DIS521	Italy (Tuscany, Fine and Cecina Valley)	5	Italy		
DIS560	Italy, Tuscany (Empoli)	5	Italy		
DIS277	Italy (Ager pisanus, volaterranus)	5	Italy		
DIS278	Italy (Ager pisanus, volaterranus)	5	Italy		
DIS288	Italy, south-italian regions and Sicily	5	Italy		
DIS522	Italy, Tuscany (Empoli)	5	Italy		
DIS549	Africa (Nabeul, Sahel) OR Gallia Narbonensis	5	Western Mediterranean area		
DIS322	Gallia Narbonensis	5	Western Mediterranean area		
DIS504	Hispania Lusitania	5	Western Mediterranean area		
DIS562	Hispania Tarraconensis	5	Western Mediterranean area		
DIS427	Hispania Terraconensis	5	Western Mediterranean area		
DIS170	Hispania Baetica	5	Western Mediterranean area		
DIS283	Africa (Tripolitania)	5	Africa		

(*continued on next page*)

Table 2 (*continued*)

Sample ID	Archaeological attribution	Graph Clustering	Provenance based on Graph Clustering	Correct attribution (%)	Misclassification (%)
DIS232	Keay 35A	Not Classified	Not Classified		
			Correct classification	75.35%	
			Misclassification	23.3%	
			Not classified	1.4%	

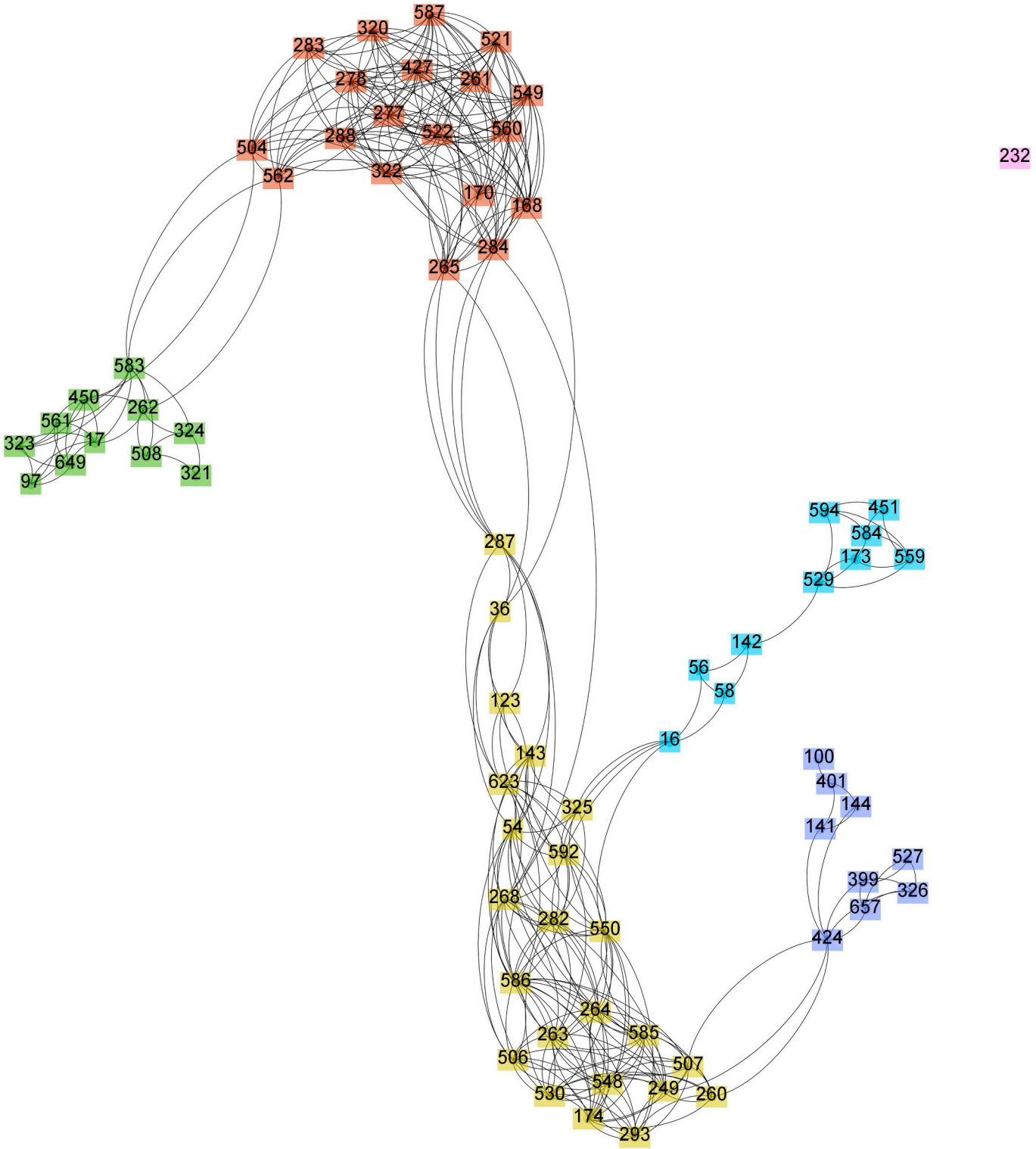

Fig. 4. Graph Clustering analysis with threshold correlation fixed at 0.99. Green: Cluster 1; Cyan: Cluster 2; Blue: Cluster 3; Yellow: Cluster 4; Red: Cluster 5; Pink: Not Classified(For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

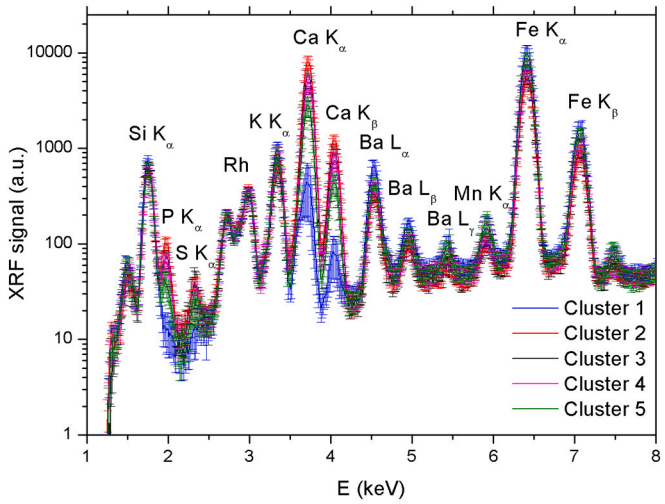

Fig. 5. Comparison between the main fluorescence lines in the average spectra corresponding to the different cluster. The error bars correspond to the standard deviation of the spectra in the clusters.

Acknowledgements

Eleonora Odelli acknowledges the support of Tuscany Region for the "Pegaso" PhD fellowship.

References

[1] P. Vandenabeele, M.K. Donais, Mobile spectroscopic instrumentation in archaeometry research, Appl. Spectrosc. 70 (2016) 27–41.
[2] P.J. Potts, M. West (Eds.), Portable X-Ray Fluorescence Spectrometry: Capabilities for In Situ Analysis, RSC Publisher, 2008.
[3] A.N. Shugar, J.L. Mass, Handheld XRF for Art and Archaeology, 1st ed, Leuven University Press, 2014.
[4] M. Uda, A. Ishizaki, R. Satoh, K. Okada, Y. Nakajima, D. Yamashita, K. Ohashi, Y. Sakuraba, A. Shimono, D. Kojima, Portable X-ray diffractometer equipped with XRF for archaeometry, Nucl. Inst. Methods Phys. Res. Sect. B Beam Interact. with Mater. Atoms. 239 (2005) 77–84, https://doi.org/10.1016/J.NIMB.2005.06.214.
[5] V. Crupi, S. D'Amico, L. Denaro, P. Donato, D. Majolino, G. Paladini, R. Persico, M. Saccone, C. Sansotta, G.V. Spagnolo, V. Venuti, Mobile spectroscopy in archaeometry: some case study, J. Spectrosc. 2018 (2018) 1–11, https://doi.org/10.1155/2018/8295291.
[6] P. Moioli, C. Seccaroni, Analysis of art objects using a portable X-Ray Fluorescence spectrometer, X-Ray Spectrom. 29 (2000) 48–52, https://doi.org/10.1002/(SICI)1097-4539(200001/02)29:1<48:AID-XRS404>3.0.CO;2-H.
[7] E. Frahm, R.C.P. Doonan, The technological versus methodological revolution of portable XRF in archaeology, J. Archaeol. Sci. 40 (2013) 1425–1434, https://doi.org/10.1016/J.JAS.2012.10.013.
[8] V. Panchuk, I. Yaroshenko, A. Legin, V. Semenov, D. Kirsanov, Application of chemometric methods to XRF-data – a tutorial review, Anal. Chim. Acta 1040 (2018) 19–32, https://doi.org/10.1016/j.aca.2018.05.023.
[9] G. Barone, V. Crupi, F. Longo, D. Majolino, P. Mazzoleni, G. Spagnolo, V. Venuti, E. Aquilia, Potentiality of non-destructive XRF analysis for the determination of Corinthian B amphorae provenance, X-Ray Spectrom. 40 (2011) 333–337, https://doi.org/10.1002/xrs.1347.
[10] L. Bonizzoni, A. Galli, M. Gondola, M. Martini, Comparison between XRF, TXRF, and PXRF analyses for provenance classification of archaeological bricks, X-Ray Spectrom. (2013), https://doi.org/10.1002/xrs.2465.
[11] L. Ceccarelli, I. Rossetti, L. Primavesi, S. Stoddart, Non-destructive method for the identification of ceramic production by portable X-rays Fluorescence (pXRF). A case study of amphorae manufacture in central Italy, J. Archaeol. Sci. Rep. 10 (2016) 253–262, https://doi.org/10.1016/J.JASREP.2016.10.002.
[12] F.S. Pirone, R.H. Tykot, T. Pluckhahn, J.G. Ryan, D. Tanasi, N. Vella, N. White, Trade, interaction and change: trace elemental characterization of maltese neolithic to middle bronze age ceramics using a portable X-Ray Fluorescence spectrometer, Open Archaeol 3 (2017), https://doi.org/10.1515/opar-2017-0012.
[13] A.M.W. Hunt, R.J. Speakman, Portable XRF analysis of archaeological sediments and ceramics, J. Archaeol. Sci. 53 (2015) 626–638, https://doi.org/10.1016/J.JAS.2014.11.031.
[14] R.H. Tykot, Using nondestructive portable X-ray fluorescence spectrometers on stone, ceramics, metals, and other materials in museums: advantages and limitations, Appl. Spectrosc. 70 (2016) 42–56, https://doi.org/10.1177/0003702815616745.
[15] T. Trojek, M. Hložek, T. Čechák, X-ray fluorescence analyzers for investigating postmediaeval pottery from Southern Moravia, Appl. Radiat. Isot. 68 (2010) 879–883, https://doi.org/10.1016/J.APRADISO.2009.10.038.
[16] F.P. Romano, G. Pappalardo, L. Pappalardo, S. Garraffo, R. Gigli, A. Pautasso, Quantitative non-destructive determination of trace elements in archaeological pottery using a portable beam stability-controlled XRF spectrometer, X-Ray Spectrom. 35 (2006) 1–7, https://doi.org/10.1002/xrs.880.
[17] M.S. Shackley, Is there reliability and validity in Portable X-Ray Fluorescence Spectrometry (PXRF)? SAA Archaeol. Rec. 10 (2010) 17–20.
[18] L. Bonizzoni, A. Galli, M. Milazzo, XRF analysis without sampling of Etruscan depurata pottery for provenance classification, X-Ray Spectrom. 39 (2010) 346–352, https://doi.org/10.1002/xrs.1263.
[19] C.C. Stremtan, H. Ashkanani, R.H. Tykot, C. Stremtan, H. Ashkanani, R.H. Tykot, M. Puscas, Constructing a Database for pXRF, XRD, ICP-MS and Petrographic Analyses of Bronze Age Ceramics and Raw Materials from Failaka Island (Kuwait), (2012), pp. 274–279.
[20] A.M. Gueli, A. Delfino, E. Nicastro, S. Pasquale, G. Politi, A. Privitera, S. Spampinato, G. Stella, Investigation by pXRF of caltagirone pottery samples produced in laboratory, Open Archaeol. 3 (2017) 235–246, https://doi.org/10.1515/opar-2017-0014.
[21] R.J. Speakman, N.C. Little, D. Creel, M.R. Miller, J.G. Iñañez, Sourcing ceramics with portable XRF spectrometers? A comparison with INAA using Mimbres pottery from the American Southwest, J. Archaeol. Sci. 38 (2011) 3483–3496, https://doi.org/10.1016/J.JAS.2011.08.011.
[22] E. Frahm, Ceramic studies using portable XRF: from experimental tempered ceramics to imports and imitations at Tell Mozan, Syria, J. Archaeol. Sci. 90 (2018) 12–38, https://doi.org/10.1016/J.JAS.2017.12.002.
[23] G. Musumarra, M. Fichera, Chemometrics and cultural heritage, Chemom. Intell. Lab. Syst. 44 (1998) 363–372.
[24] D. Seetha, G. Velraj, Characterization and chemometric analysis of ancient pot shards trenched from Arpakkam, Tamil Nadu, India, J. Appl. Res. Technol. 14 (2016) 345–353, https://doi.org/10.1016/j.jart.2016.08.002.
[25] M. Gajić-Kvaščev, M. Marić-Stojanović, R. Jančić-Heinemann, G. Kvaščev, Vd. Andrić, Non-destructive characterisation and classification of ceramic artefacts using pEDXRF and statistical pattern recognition, Chem. Cent. J. 6 (2012) 102, , https://doi.org/10.1186/1752-153X-6-102.
[26] I. Papageorgiou, I. Liritzis, Multivariate mixture of normals with unknown number of components: an application to cluster neolithic ceramics from aegean and asia minor using portable XRF, Archaeometry. 49 (2007) 795–813, https://doi.org/10.1111/j.1475-4754.2007.00336.x.
[27] R. Ikeoka, C. Appoloni, P. Parreira, F. Lopes, A. Bandeira, PXRF and multivariate statistics analysis of pre-colonial pottery from northeast of Brazil, X-Ray Spectrom. 41 (2012) 12–15, https://doi.org/10.1002/xrs.1378.
[28] G. Barone, P. Mazzoleni, G.V. Spagnolo, S. Raneri, Artificial neural network for the provenance study of archaeological ceramics using clay sediment database, J. Cult. Herit. 38 (2019) 147–157, https://doi.org/10.1016/J.CULHER.2019.02.004.
[29] J. Qi, T. Zhang, H. Tang, H. Li, Rapid classification of archaeological ceramics via laser-induced breakdown spectroscopy coupled with random forest, Spectrochim. Acta Part B At. Spectrosc. 149 (2018) 288–293, https://doi.org/10.1016/J.SAB.2018.09.006.
[30] N. Forster, P. Grave, N. Vickery, L. Kealhoferc, Non-destructive analysis using PXRF: methodology andapplication to archaeological ceramics, X-Ray Spectrom. 40 (2011) 389–398, https://doi.org/10.1002/xrs.1360.
[31] E. Grifoni, S. Legnaioli, G. Lorenzetti, S. Pagnotta, V. Palleschi, Application of Graph Theory to unsupervised classification of materials by Laser-Induced Breakdown Spectroscopy, Spectrochim. Acta Part B At. Spectrosc. 118 (2016) 40–44, https://doi.org/10.1016/j.sab.2016.02.003.
[32] V. Palleschi, L. Pagani, S. Pagnotta, G. Amato, S. Tofanelli, Application of graph theory to the elaboration of personal genomic data for genealogical research, PeerJ. (2015) e27, , https://doi.org/10.7717/peerj-cs.27.
[33] R. Auriemma, Un carico di anfore Keay LII nelle acque dello Ionio, in: L. Saguì (Ed.), Atti Del Convegno Onore Di J.W.Hayes "Ceramica Ital. VI - VII Secolo", All'Insegna del Giglio, Firenze, 1998, pp. 753–760.
[34] T. Bertoldi, Guida alle anfore romane di età imperiale. Forme, impasti e distribuzione, Brossura, 2012.
[35] M. Bonifay, Etudes sur la céramique romaine tardive d'Afrique, Archeopress, Oxford, 2004.
[36] F. Cantini, G. Boschian, M. Gabriele, Empoli, a late antique pottery production centre in the Arno valley (Florence, Tuscany, Italy), in: N. Poulou-Papadimitriou, E. Nodarou, V. Kilikoglou (Eds.), LRCW 4. Late Rom. Coarse Wares, Cook. Wares Amphor. Mediterr. Archaeol. Archaeom. Mediterr. a Mark. without Front, Archeopress, Oxford, 2014, pp. 203–212.
[37] L. Cherubini, A. Del Rio, Officine ceramiche di età romana nell'Etruria settentrionale costiera: impianti, produzioni, attrezzature, Rei Cr. Faut. Acta XXXV (1997) 133–141.
[38] G. Ciampoltrini, A. Andreotti, P. Notini, P. Rendini, C. Spataro, Traffici e consumi ceramici nella valle del Serchio in età teodosiana, in: S. Menchelli, S. Santoro, M. Pasquinucci, G. Guiducci (Eds.), LRCW3. Late Rom. Coarse Wares, Cook. Wares Amphor. Mediterr. Archaeol. Archaeom. Comp. between West. East. Mediterr, Archeopress, Oxford, 2010, pp. 319–328.
[39] A. Costantini, Le Anfore, A. Alberti, E. Paribeni (Eds.), Archeol. Piazza Dei Miracoli. Gli Scavi 2003-2009, Felci Editore, Pisa, 2011, pp. 393–430.
[40] F. Laubenheimer, La production des amphores en Gaule narbonnaise, Les Belles Lettres, Paris, 1985.
[41] G. Murialdo, Le anfore da trasporto, in: T. Mannoni, G. Murialdo (Eds.), S. Antonino Un Insediamento Fortif. Nella Liguria Biz. All'Insegna del Giglio, Bordighera, 2001, pp. 255–307.
[42] F. Pacetti, La questione delle Keay LII nell'ambito della produzione anforica in

Italia, in: L. Saguì (Ed.), Atti Del Convegno Onore Di J.W.Hayes “Ceramica Ital. VI - VII Secolo,”, All'Insegna del Giglio, Firenze, 1995, pp. 185–208.
[43] D. Pieri, Le commerce du vin oriental à l’époque byzantine (Ve-VIIe siècle). Le témoignage des amphores en Gaule, IFPO, Beyrouth, 2005.
[44] V.A. Solé, E. Papillon, M. Cotte, P. Walter, J. Susini, A multiplatform code for the analysis of energy-dispersive X-Ray Fluorescence spectra, Spectrochim. Acta Part B At. Spectrosc. 62 (2007) 63–68, https://doi.org/10.1016/J.SAB.2006.12.002.
[45] C. Arias, S. Bani, F. Catalli, G. Lorenzetti, E. Grifoni, S. Legnaioli, S. Pagnotta, V. Palleschi, X-Ray Fluorescence analysis and self-organizing maps classification of the etruscan gold coin collection at the monetiere of florence, Appl. Spectrosc. 71 (2017) 817–822, https://doi.org/10.1177/0003702816641421.
[46] S. Pagnotta, E. Grifoni, S. Legnaioli, M. Lezzerini, G. Lorenzetti, V. Palleschi, Comparison of brass alloys composition by laser-induced breakdown spectroscopy and self-organizing maps, Spectrochim. Acta Part B At. Spectrosc. 103–104 (2015) 70–75, https://doi.org/10.1016/j.sab.2014.11.008.