This is the Accepted Manuscript version of a Published Work that appeared in final form in Spectrochimica Acta Part B: Atomic Spectroscopy 65(4) : 279-286 (2010). To access the final edited and published work see https://doi.org/10.1016/j.sab.2010.01.009
© 2010. This manuscript version is made available under the CC-BY-NC-ND 4.0 license https://creativecommons.org/licenses/by-nc-nd/4.0/

Classification of archaeological pieces into their respective stratum by a chemometric model based on the soil concentration of 25 selected elements

J.A. Carrero*, N. Goienaga, S. Fdez-Ortiz de Vallejuelo, G. Arana and J.M. Madariaga.

Department of Analytical Chemistry, University of the Basque Country, 48080 Bilbao, Spain.

ABSTRACT

The aim of this work was to demonstrate that an archaeological ceramic piece has remained buried underground in the same stratum for centuries without being removed. For this purpose, a chemometric model based on Principal Component Analysis (PCA), Soft Independent Modelling of Class Analogy (SIMCA) and Linear Discriminant Analysis (LDA) classification techniques was created with the concentration of some selected elements of both soil of the stratum and soil adhered to the ceramic piece. Some ceramic pieces from four different stratigraphic units, coming from a roman archaeological site in Alava (North of Spain), and its respective stratum soils were collected. The soil adhered to the ceramic pieces was removed and treated in the same way as the soil from its respective stratum. The digestion was carried out following the US Environmental Pollution Agency EPA 3051A method. A total of 54 elements were determined in the extracts by a rapid screening ICP-MS method. After rejecting the major elements and those which could have changed from the original composition of the soils (migration or retention from/to the buried objects), the following elements (25) were finally taken into account to construct the model: Li, V, Co, As, Y, Nb, Sn, Ba, La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu, Au, Th and U. A total of 33 subsamples were treated from 10 soils belonging to 4 different stratigraphic units. The final model groups and discriminate them in four groups, according to the stratigraphic unit, having both the stratum and soils adhered to the pieces falling down in the same group.

Keywords: Archaeological pieces, ICP-MS, Linear Discriminant Analysis, Principal Component Analysis, Archaeometry.

* Corresponding author. E-mail: joseantonio.carrero@ehu.es, phone (+34) 946015551, Fax. (+34) 946013500

1. Introduction

Scientific studies in art and archaeology are in vogue among other reasons due to its interest in cultural heritage area [1,2]. A wide range of archaeological and art-historical objects are commonly analyzed including ceramics [1], artworks [2,3] and historical paintings [4-6] among others. The interest of chemical analyses in archaeological samples lies in the knowledge of material composition in order to characterize and understand the context in which the artifacts were created, to avoid falsifications, or to determine their real historical value [7]. In this sense there are always some doubts about the origin or dating of ancient archaeological objects when they are discovered in excavations. There are other economical or political interests behind a new archaeological discovery which makes it prone to falsifications. Quantitative elemental compositional data plays a key role in solving questions concerning dating, provenance, technology used and the relationship of ancient cultures with the environment. Inductively coupled plasma mass spectrometry (ICP-MS) has been used for this purpose [8,9]. Laser ablation (LA-ICP-MS) is a non-destructive alternative for elemental analysis in archaeometry [10]. The combination of elemental analysis by means of energy-dispersive X-ray spectrometry (ED-XRF) and ICP-MS with molecular analysis such as Raman spectroscopy has been of great interest in archaeological ceramics analysis in order to establish differences or similarities among them. The combined use of these techniques proved beneficial, as Raman data gives structural information on the compounds present whereas ICP-MS and XRF data give quantitative values [11,12].

ICP-MS is nowadays a useful technique of great interest due to its multi-element capabilities, low detection limits, high sample throughput, wide linear dynamic range and the possibility to obtain isotopic information [13]. For some applications, extremely low detection limits are not required, but a fast survey of all present elements is preferred instead. ICP-MS offers different quantification procedures depending on the accuracy and precision required. Quantitative analysis allows accurate and precise results to be obtained, but it is time consuming, especially in a context of multi-element determination of analytes at very different concentration levels. Although semi-quantitative analysis in ICP-MS used to be considered only as rapid multi-element survey tool (within 3 min, over 70 elements can be determined in one sample) with accuracies in the range of 30–50% and detection limits (3σ criterion) for most of the elements below 1 μg/L [14], the study carried out by Laborda et al. [15] demonstrate that semi-quantitative analysis is an alternative to quantitative analysis when productivity factors are important or the most significant sources of uncertainty arise from steps of the analytical process different from the measurement and data treatment (i.e. sampling or sample pre-treatment). Semi-quantitative analysis has been successfully applied to different types of samples including: environmental, biological, geological, industrial or food [14,16-20]. It is based on the use of a predefined response table for all the elements and a computer program that can interpret the mass spectrum and correct isotopic interferences. Response factors must be updated to the actual sensitivity of the instrument by measuring a standard of a few selected elements (3–5) of known concentration (external calibration). This updating may also be performed by adding these elements to the samples, resulting in response factors which also include the matrix effects (internal calibration). Since errors of 30–50% are accepted in semi-quantitative procedures, the precision associated with the measurements has little interest and just one replicate per sample is common practice in this type of analysis to save time. The software developed to perform semi-quantitative analysis has evolved in parallel with the instrumentation and, at the present time, accuracy values better than 10% for some elements have been reported [20], which values fall not very far from those typically obtained by quantitative analysis [15].

On the other hand, Raman spectroscopy is a useful technique to obtain structural information. It makes it possible to differentiate several species of the same element. Moreover, Raman spectroscopy has the advantage to measure the sample practically without any pre-treatment and thus species can be identified in the solid state. Raman analysis complements elemental ICP-MS analysis and allows knowing in which form is present the metal of interest (sulphate, carbonate, nitrate…), which in several cases is the key to understand the context in which the artifact was created or if it is a natural or artificial compound. Due to its non destructive character, it has been mainly used in the study of many different kinds of artworks, such as wall paintings [21], glass and ceramics [22] and architectural paint analysis [23] among others. In the last years, Raman spectroscopy has been also used by some authors to search for mineral composition and characterize metallic contaminants in sediments and soils [24-26].

Thanks to the development of modern analytical instrumental techniques a lot of problems concerning all kinds of disciplines, and often we can obtain very numerous data for the analyzed samples, obtained by means of analytical techniques. Chemometric methods (also known as multivariate statistical techniques) are increasingly in use and can identify the natural clustering pattern and group variables on the basis of similarities between the samples. Depending on the knowledge about the category or class membership of the data set, two approaches can be applied: supervised or unsupervised pattern recognition. Supervised methods develop rules for the classification of samples of unknown origin on the basis of a group of samples with known categories (known set). Unsupervised learning methods instead do not assume any known set, and the goal is to find clusters of objects, which may be assigned to classes. The most common methods of chemometric methods for classification are namely, cluster analysis (CA) and principal component analysis (PCA) with factor analysis (FA) among unsupervised ones, and linear discriminant analysis (LDA) as supervised classification tool. The applications of different chemometric methods i.e., cluster analysis (CA), principal component analysis (PCA), factor analysis (FA) and linear discriminant analysis (LDA) aids in reducing the complexity of large data sets and offers better interpretation and understanding of data sets [27-29]. Thus the effort of the analysts on the elaboration of a high number of data is worth it. We can find a typical example in the classification and study of archaeological finds. In the archaeometry, that is the application of scientific methods and analysis techniques to archaeological issues, one of the most important steps is the statistical elaboration of multivariate data obtained by physical and chemical analysis of ancient artifacts. In the bibliography some articles can be found concerning the application of chemometric multivariate techniques in pottery [1,30,31], trying to characterize and classify samples into groups according with their chemical composition and origin.

In this work it was demonstrated that an archaeological ceramic piece has remained buried underground in the same stratum for centuries and it has not been removed or altered, so it belongs to that specific stratigraphic layer where it was found. An archaeological ceramic piece, which has been buried underground in the same stratum for centuries, when it is removed in the excavation it will have soil remains necessarily associated to the soil where it was buried. This soil adhered to the ceramic piece and the stratum soil should have similar composition for some chemical elements. Based on this hypothesis the soil adhered to a ceramic piece should group with the soil of the stratum where it was buried. For this purpose and by means of Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) classification techniques we have built a chemometric model which groups the corresponding soils and discriminate the different ones in base of the concentration of some

selected elements determined by semi-quantitative ICP-MS and supported by μ-Raman and μ-ED-XRF analysis.

2. Experimental section

2.1. Reagents and Instrumentation

All plastic and glassware material in contact with samples or ICP-MS solutions were soaked in a 10% HNO_3 bath at least 24 h, then rinsed twice with Elix (Millipore, USA) quality water and finally rinsed with pure water of resistance 18.2 MΩ.cm obtained from a Milli-Q Element A10 purification system (Millipore, USA). After drying the material in a laminar air flow hood inside a class 100 clean room, it was stored in clean plastic bags until use.

Soil samples were ground in a planetary ball mill Pulverisette 6 (Fritsch, Germany) for their homogenization and prior to the acid digestion in a microwave digestion system Multiwave 3000 (Anton Paar, Graz, Austria) provided with a 8XF-100 microwave digestion rotor and 100 mL fluorocarbon polymer (PTFE) microwave vessels. Nitric acid (69%) and hydrochloric acid (36%) were Tracepur grade and supplied by Merck. All solutions were prepared using Milli-Q water. ICP-MS standard solutions were prepared from Alfa Aesar (Specpure®, Plasma standard solution, Germany) stock solutions.

Elemental analysis was achieved using a Perkin Elmer SCIEX 9000 ICP-MS (Toronto, Canada) in a class 100 clean room equipped with quartz torch, 2.0 mm id alumina injector tube, nickel sampler and skimmer cones, a peristaltic pump (maintaining a 1 mL min^{-1} sample uptake rate), a cross-flow type pneumatic nebulizer and a double pass Scott-type spray chamber. The argon used for the plasma was supplied by Praxair (99.995%, Madrid, Spain). The operating conditions for sample introduction are shown in Table 1.

A portable Renishaw RA100 (Renishaw, Gloucestershire, UK) microprobe system that implements a 785 nm excitation laser, a Peltier-cooled CCD and a mobile diffraction grating of 1200 lines mm^{-1} was used for the soil characterization. The laser has a nominal output power of 150mW at the source and some filters allow working at 1%, 10% or 100% of that power. The microprobe was coupled with 20× objective that allows focusing on a single grain by means of the micro-TV camera implemented on the probe.

X-ray fluorescence spectroscopy analyses were carried out to complete molecular results. With this aim a portable ArtTAX μ-ED-XRF equipped with an X-ray tube with a molybdenum anode working at a maximum voltage of 50 kV and a maximum current of 0.6 mA was used. The X-rays are collimated by a tantalum collimator with a diameter of 0.65 mm and the beam diameter in the sample's surface is 200 μm^2. The measurement head of the equipment implements a CCD camera that enables focusing on the sample by a motorized XYZ positioning unit controlled by a computer.

2.2. Samples Identification

Samples were acquired from Roman settlements of the 3rd century in Araba (North of Spain) that have been recently excavated. Soil samples for four different strata were collected and named SS 1, SS 2, SS 3 and SS 4. In each stratum a ceramic piece was also taken, except in the stratum 1 where three different pieces were collected. The soil adhered to these ceramic

pieces was treated as samples named PS 1.1, PS 1.2, PS 1.3, PS 2.1, PS 3.1 and PS 4.1. Four replicates of each soil sample were analyzed (a, b, c and d), so a total of ten different soil samples were analyzed in quadruplicate. Samples were mainly clay soils, as this is the typical composition of the Araba´s plain.

2.3. Analytical procedure

Soil samples were carefully separated from their archaeological ceramic piece with the help of a brush and a spatula without scratching the piece. Strata soils were sieved to < 2mm and different objects such as bones or stones were removed. After air dried in a fume hood during 24 h soils were milled in a planetary ball mill and stored at 4ºC until analysis.

The analytical determination of heavy metals in soil samples were carried out following the US Environmental Pollution Agency EPA 3051A method which describes the Microwave Assisted Acid Digestion of Sediments, Sludge, Soils, and Oils. 0.5 g of soil were weighed directly into each pre-cleaned vessel in an analytical Mettler-Toledo balance AJ150 model (±0.0001g) and 9 mL HNO_3 – 3 mL HCl acid mixture was added. A first ramp of 1000 watts was applied during 7 minutes to reach the temperature established in the method (170ºC) and then it was maintained by means of a plateau at 800 watts during 4.5 minutes. After cooling, the digests were filtered through a 0.45µm PVDF filter and quantitatively transferred to a polyethylene bottle with a final volume of 50 mL. A vessel containing the same acid mixture as used for the samples was also prepared and utilized as the analytical blank. The digests were stored at 4ºC until analysis.

Sample solutions were diluted to 1% HNO_3 concentration and transported into the argon plasma via a peristaltic pump. Data were acquired using Perkin Elmer TotalQuant III software for multi-element semi-quantitative analysis. This program performs an automated interpretation of the entire mass spectrum, meaning that you do not need to specify the individual elements you want to determine. Based on a response curve, all the intensities are subsequently converted into concentrations, taking into account the isotopic abundance of the nuclides involved. Common isobaric interferences are pre-programmed and corrections are automatically applied. Each element is measured at its most abundant isotope if not other isotope is manually defined. The reference response table of the different elements, which is included in the software, was updated and matched to the sample matrix in order to improve the accuracy, spiking samples with few selected elements with known concentration and covering all the mass range (Be, Sc, In, Re, Bi at 20 ng g^{-1}) [20]. The performance of an ICP-MS instrument strongly depends on the operating conditions. The plasma operating conditions such as the nebulizer flow rate, the position of the torch and the ion lens voltages of the instrument were optimized prior to any experiment continuously nebulising a 10 ng mL^{-1} standard solution of Mg, Rh, In, Ba, Pb and U. The nebulizer gas flow rate was optimized to obtain a good compromise between high sensitivity and low oxide levels (lower than 3% for CeO/Ce). The ion optics in the ELAN 9000 consists of a single lens arrangement. The autolens mode was used. The voltage was optimized aspirating a solution of 10 ng mL^{-1} of Be, Co, In and Bi. Dual detection was employed for all measurements. Obtained metal concentration data were statistically treated with the chemometric software "The Unscrambler 9.2" (Camo Process, Norway) and SPSS 15.0 (SPSS Inc., Chicago, USA). For quality assurance purposes soil certified reference materials (SRM 2711 and BCR 142R) and a freshwater containing trace elements (SRM 1640) were routinely analyzed in each sample batch.

To complement the ICP-MS analysis a preliminary screening by μ-ED-XRF was done to determine the elemental composition, which helps in the knowledge of the nature of the materials present in the samples. Then, Raman microprobe spectroscopy analysis was performed in order to obtain the molecular composition. Combination of these complementary results enables identification of the compounds present in the objects under analysis. The calibration of the Raman spectrometer was carried out daily with the 520 cm^{-1} silicon band. The measurement conditions vary depending on the gathered signal-to-noise ratio and the obtained fluorescence, but exposure times from 5 to 30 s and 1–10 accumulations were used in most of the cases. The spectra were collected with the WIRE (Renishaw, UK) software between 2200 and 200 cm^{-1} (spectral resolution of 1 cm^{-1} and cut-off of the system around 200 cm^{-1}) and processed with Omnic 7.1 (Nicolet, Madison, WI, USA) software. The interpretation of the results was performed by comparing the obtained spectra with standard spectra contained in our home-made Raman database [32].

3. Results and discussion

A total of 54 elements were determined in the extracts by the semi-quantitative ICP-MS method. Multivariate statistical analysis allows differentiating soils on the basis of their chemical composition. In contrast to the use of single-element concentrations, multivariate statistical methods allow verifying the contribution of each variable to the model, and its capacity to discriminate one category from another. In order to find structures in a data set or to reveal similarities of different samples by way of multivariate analysis, the features need to be comparable. The most common possibility is autoscaling. This means that the data are related to measures of their own distribution, namely the mean, and the standard deviation, of a verified normal distribution. This is achieved by subtracting the mean of each feature from each individual value and dividing by the standard deviation of the respective feature. After autoscaling the data, in a first round, metal concentration raw data of all analyzed samples were introduced in "The Unscrambler 9.2" chemometric software and Principal Component Analysis (PCA) was carried out. PCA is a technique widely used for reducing the dimensions of multivariate problems. As an unsupervised pattern recognition technique, it makes no assumptions about the underlying statistical data distribution. It reduces the dimensionality of the data set by explaining the correlation among a large number of variables in terms of a smaller number of underlying factors (principal components or PCs) without losing much information. PCA analyze the structure of multivariate data by projecting them into a reduced hyperspace, which is defined by the first principal component. These are linear combinations of the original variables with the first principal component explaining the largest portion of the variance, the second principal component the second-largest portion, and so on.

As can be seen in Figure 1, soil replicates of different samples are mixed among them, without any group formation. Considering that there are ten different soil samples, ten different groups should be obtained. Besides, PC1 only explains 34% of the total variance of the data.

In order for the grouping to occur, metal concentration data were normalized to a major element of the soil composition, such as Fe, Mg, Al, Mn, Ti, Si or Ca, so-called conservative metal [33]. In this way, variations generated in the microwave extraction efficiency due to different final temperatures reached in each microwave vessel are counteracted. The best grouping was obtained normalizing data to Mn and representing PC1 vs PC 3 (Figure 2) after eliminating some outliers. The replicates of the same soil sample are shown in different circles.

The next step was to try to group soil samples from the same stratum following the hypothesis that the chemical composition of both stratum soil and soil adhered to ceramic pieces buried in that stratum must be the same. However, not all the elements should be included in the chemometric model. Metals whose concentration in the soil could change because of the dissolution or interaction of metal objects with the archaeological ceramic pieces inside the stratum should be discarded. For this purpose, Raman and ED-XRF spectroscopy data together with those obtained by means of ICP-MS give this information about elements that would distort the chemometric model.

Abnormally elevated amounts of some metals compared with the average value found in stratum soil were detected in some soils belonging to different soil strata by ICP-MS. That is the case of metals shown in Table 2, where concentrations of Pb and Zn results forty times higher than the average value were obtained.

The Raman inspection of the soils adhered to the ceramic surfaces showed a big amount of calcite ($CaCO_3$) covering the inner surface of the ceramic pieces. Its presence can only be explained by the in-situ reaction of percolated bicarbonate (HCO_3^-) with available calcium. The bicarbonate ion is formed by the reaction of the carbonate from the top soil surfaces with the acid gases of the atmosphere [34] and then it migrates down the stratum, reacts and forms the white deposits on the ceramic surface (see the microphotography in Figure 3). But over there, different micro-crystals can be observed. As an example, the XRF analysis of the one shown in Figure 4 gives a big amount of Sb, Zn, S with traces of Pb, together with the expected elements in Roman ceramic ware: Si, Fe, Ca, Sr, Ti and Mn. K_α emission lines of the identified elements are shown. For some elements K_β lines also appear. Lines of the Mo source are indicated since they should be considered as the background of the spectrum.

The Raman spectra of such a grain showed the presence of more than one compound. Figure 5 shows one of those spectra collected by the Raman microprobe focused on the whole volume of the microcrystal. The background of calcite appeared in all the spectra. The other signals suggest that the antimony compound could be a mixed sulphate-antimoniate compound because the spectrum of standard Peretaite ($CaSb_4O_4(SO_4)_2(OH)_2 \cdot 2H_2O$) match quite well (the four main Raman signals) with the unknown spectra. Moreover, the basic carbonate of zinc also appears although in minor intensity than the antimoniate compound. Probably this zinc carbonate is spread in the calcite layer covering the surface of the ceramic and since the laser penetrates into the surface of the piece, the signals of both compounds were considered as the background when we worked further in other micro-crystals. Due to the fluorescence effect it is not possible to see the other characteristic peaks of these compounds. This fluorescence effect could be attributed to the clay matrix of the samples.

Therefore the alkaline metals and major components of the soil (Na, K, Mg, Ca, Al, Si, Ti, Mn, Fe and P) were eliminated due to its high mobility and the possibility of precipitate in the surface of the archaeological piece, resulting in a highly variable distribution of these elements along different points of the stratum. Most of the elements listed in Table 2 (Pb, Zn, Hg, Cr, Ni, Cu, Cd) are common heavy metals used in Roman age for the daily life like different metal tools or pigments used for decorative purposes. Others like Sb must have an anthropogenic origin from the pollution of urban areas nearby and have arrived by atmospheric deposition and migrated down forming those synthesis products. This is a clear distorting map of the original composition of the stratum, showing the importance of the referred migrations of elements and the subsequent formation of chemical artifacts.

Once the elements to include in the chemometric model were selected, a new principal component analysis was performed with the following elements: Li, V, Co, As, Y, Nb, Sn, Ba, La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu, Au, Th and U. As can be shown in Figure 6 (PC1 59% vs. PC3 4%) samples from the same stratum form different groups. Samples coming from the stratum 3 differ well from the rest and appear on the right side of the graph. The others are closer among them but in no case can be grouped together. With this model soft independent modelling of class analogies (SIMCA) supervised pattern recognition method was run. This classification procedure uses linear discriminant functions derived from principal component analysis of the data. One set of functions is derived for each category studied by computing the category-mean and a specified number of the principal components. Objects are classified into the category whose principal component model best reproduces the data. Only data points which are members of a given category are used in determining the model functions for that category. The importance of each feature in classification is determined by its contribution to the category covariance matrices. Weights of the elements included in the model are shown in Figure 7 for each one of the 4 groups created. Each group is explained by different elements and therefore they can be separated in PCA. Consequently, knowing the concentration of these metals for different soils, they could be classified in one group or other running a SIMCA analysis.

Linear discriminant analysis (LDA) is a supervised procedure that maximizes the variances between categories and minimizes the variances within categories. LDA is a common linear and parametric classification method for classification purposes. The classification of new objects into groups or the reclassification of the learning data set is carried out by means of the values of non-elementary discriminant functions which are calculated as a linear combination of an optimum separation set of the original features. So, the next step was to perform a LDA in order to demonstrate that the four found groups can be differentiated based on the concentration of the elements. The selection of variables containing the most powerful information for the correct classification was carried out on the basis of the canonical analysis of data, using the forward stepwise inclusion option. In this strategy the variables are included one by one into the discriminant function and, in this way, it is possible, on one hand, to construct a function with the variables which really result useful for the classification, and on the other hand, the individual contribution of each variable to the discriminant model can be evaluated. The statistic parameter chosen for the inclusion of one variable or another was the variance not explained. The variable that decreases is included in the model. The values of F parameter to enter or to remove were fixed at 3.84 and 2.71, respectively. The number of steps was fixed at 50 and no variable was forced to enter into any model. Prior probabilities were established in proportion to the number of samples in each group. A leave-one-out cross-validation procedure was performed for assessing the performance of the classification rule.

The forward stepwise LDA allows 97.2% discrimination amongst soils from four studied strata using only eight discriminant variables: Li, Co, As, Ba, Ho, Yb, Th, and U for distinguishing between soils from different strata. Three canonical discriminant functions were calculated with high canonical correlation (97.2%, 89.0% and 80.2%, respectively). Figure 8 shows the distribution of samples on plane given by first two canonical discriminant functions accounting for 92.1% of total variance (function 1 accounts for 75.5% and function 2 accounts for 16.6%). As occurs in the PCA, samples coming from different stratum appear in different parts of the plane described by the first two canonical discriminant functions. Circles, triangles, diamonds and asterisks represent samples whereas squares are the centroid of the group. Samples from the stratum two and four appear again close and samples from the

third stratum far away from the rest. Each sample fall around the centroid in differentiate groups. Only some samples from the second and fourth strata could mix up. These samples could be in origin close to each other and belong to adjacent stratigraphic layers. In table 3 the classification results of LDA and the percentages of correctly classified samples are presented (97.2%).

4. Conclusions

One of the most important contributions of chemical analyses in archaeological samples lies in the knowledge of material composition in order to avoid falsifications. The combination of μ-ED-XRF and micro-Raman spectroscopy techniques has shown to be a powerful analytical method to characterize archaeological ceramics.

In this work we have treated with the soils adhered to the ceramic pieces when they are removed in the excavation. For our knowledge is the first time that multivariate statistical methods are used directly in metal concentration data of the soil adhered to ceramic pieces, with the aim of classifying them according to their provenance stratum. A chemometric model, based on PCA, SIMCA and LDA classification techniques, was designed. In this way, it is possible to demonstrate that the piece has remained buried for centuries in the same stratum and it has not been manipulated.

Semi-quantitative ICP-MS analysis of the soils allows us knowing its elemental composition (major and trace elements). More than 50 elements were determined in less than 2 minutes per sample. These data were complemented with μ-ED-XRF and Raman analysis in order to find artifact compounds originated in soils as a consequence of the migration of elements from the upper layers or interaction and dissolution of buried objects with the ceramics. In this sense, it was found a crystalline grain of an antimony compound on a calcite ($CaCO_3$) layer which covers the ceramic inner surface. This calcite can only be explained by the in-situ reaction of the available calcium with percolated bicarbonate (HCO_3^-). This is a clear distorting map of the original composition of the stratum and therefore, these compounds should not be taken into account in the model. The final chemometric model groups the corresponding soils and discriminate the different ones according to the concentration of 25 elements: Li, V, Co, As, Y, Nb, Sn, Ba, La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu, Au, Th and U.

Acknowledgements

This work was financially supported by Arabako Foru Aldundia (AFA/DFA). S. Fdez-Ortiz de Vallejuelo research contract is funded by the project IT07-245 (Basque Goverment). Jose Antonio Carrero and Naiara Goienaga are grateful to the University of the Basque Country for their fellowships.

References

[1] F. Bellanti, M. Tomassetti, G. Visco, L. Campanella, A chemometric approach to the historical and geographical characterisation of different terracotta finds, Microchem. J. 88 (2008) 113-120.

[2] K. Castro, M. Perez-Alonso, M.D. Rodriguez-Laso, J.M. Madariaga, Raman fiber optic approach to artwork dating, Spectrochim. Acta, Part A. 60A (2004) 2919-2924.

[3] H.G.M. Edwards, D.L.A. de Faria, Infrared, Raman microscopy and fibre-optic Raman spectroscopy, Compr. Anal. Chem. 42 (2004) 359-395.

[4] H.G.M. Edwards, Raman spectroscopy of inorganic materials in art and archaeology: spectroscopic analysis of historical mysteries, Spectrosc. Prop. Inorg. Organomet. Compd. 40 (2009) 16-48.

[5] P. Vandenabeele, K. Lambert, S. Matthys, W. Schudel, A. Bergmans, L. Moens, In situ analysis of medieval wall paintings: a challenge for mobile Raman spectroscopy, Anal. Bioanal. Chem. 383 (2005) 707-712.

[6] K. Castro, M. Perez-Alonso, M.D. Rodriguez-Laso, N. Etxebarria, J.M. Madariaga, Non-invasive and non-destructive micro-XRF and micro-Raman analysis of a decorative wallpaper from the beginning of the 19th century, Anal. Bioanal. Chem. 387 (2007) 847-860.

[7] P. Ricciardi, P. Colomban, A. Tournie, V. Milande, Nondestructive on-site identification of ancient glasses: genuine artefacts, embellished pieces or forgeries?, J. Raman Spectrosc. 40 (2009) 604-617.

[8] P. Mirti, M. Pace, M.M.N. Ponzi, M. Aceto, ICP-MS analysis of glass fragments of Parthian and Sasanian epoch from Seleucia and Veh Ardasir (central Iraq), Archaeometry. 50 (2008) 429-450.

[9] E. Marengo, M. Aceto, E. Robotti, M.C. Liparota, M. Bobba, G. Pantò, Archaeometric characterisation of ancient pottery belonging to the archaeological site of Novalesa Abbey (Piedmont, Italy) by ICP-MS and spectroscopic techniques coupled to multivariate statistical tools, Anal. Chim. Acta. 537 (2005) 359-375.

[10] B. Giussani, D. Monticelli, L. Rampazzi, Role of laser ablation-inductively coupled plasma-mass spectrometry in cultural heritage research: A review, Anal. Chim. Acta. 635 (2009) 6-21.

[11] B. Wagner, A. Nowak, E. Bulska, J. Kunicki-Goldfinger, O. Schalm, K. Janssens, Complementary analysis of historical glass by scanning electron microscopy with energy dispersive X-ray spectroscopy and laser ablation inductively coupled plasma mass spectrometry, Microchim. Acta. 162 (2008) 415-424.

[12] M. Resano, E. Garcia-Ruiz, R. Alloza, M.P. Marzo, P. Vandenabeele, F. Vanhaecke, Laser ablation-inductively coupled plasma mass spectrometry for characterization of pigments in prehistoric rock art, Anal. Chem. 79 (2007) 8947-8955.

[13] A. Montaser, Inductively Coupled Plasma Mass Spectrometry, Wiley-VCH, New York, 1998.

[14] Y. Hu, F. Vanhaecke, L. Moens, R. Dams, Semi-quantitative panoramic analysis of industrial samples by inductively coupled plasma mass spectrometry, Anal. Chim. Acta. 355 (1997) 105-111.

[15] F. Laborda, J. Medrano, J.R. Castillo, Quality of quantitative and semiquantitative results in inductively coupled plasma mass spectrometry, J. Anal. At. Spectrom. 16 (2001) 732-738.

[16] M.M. Bayon, J.I. Garcia-Alonso, A. Sanz-Medel, Semiquantitative elemental analysis of water samples using double focusing inductively coupled plasma mass spectrometry, J. Anal. At. Spectrom. 13 (1998) 1027-1032.

[17] D. Amarasiriwardena, S.F. Durrant, A. Lasztity, A. Krushevska, M.D. Argentine, R.M. Barnes, Semiquantitative analysis of biological materials by inductively coupled plasma-mass spectrometry, Microchem. J. 56 (1997) 352-372.
[18] M. Barbaro, B. Passariello, S. Quaresima, A. Casciello, A. Marabini, Determination of rare earth elements in rock samples by inductively coupled plasma-mass spectrometry (ICP-MS), Microchem. J. 51 (1995) 312-318.
[19] J.R. Castillo, M.S. Jimenez, L. Ebdon, Semiquantitative simultaneous determination of metals in olive oil using direct emulsion nebulization, J. Anal. At. Spectrom. 14 (1999) 1515-1518.
[20] J. Soldevila, M. El Himri, A. Pastor, M. de la Guardia, Evaluation of operational parameters affecting semiquantitative multi-elemental analysis by inductively coupled plasma mass spectrometry, J. Anal. At. Spectrom. 13 (1998) 803-807.
[21] H.G.M. Edwards, D.W. Farwell, C.J. Brooke, Raman spectroscopic study of a post-medieval wall painting in need of conservation, Anal. Bioanal. Chem. 383 (2005) 312-321.
[22] D. Barilaro, G. Barone, V. Crupi, M.G. Donato, D. Majolino, G. Messina, R. Ponterio, Spectroscopic techniques applied to the characterization of decorated potteries from Caltagirone (Sicily, Italy), J. Mol. Struct. 744-747 (2005) 827-831.
[23] E. Kendix, O.F. Nielsen, M.C. Christensen, The use of micro-Raman spectroscopy in architectural paint analysis, J. Raman Spectrosc. 35 (2004) 796-799.
[24] U. Villanueva, J.C. Raposo, K. Castro, A. de Diego, G. Arana, J.M. Madariaga, Raman spectroscopy speciation of natural and anthropogenic solid phases in river and estuarine sediments with appreciable amount of clay and organic matter, J. Raman Spectrosc. 39 (2008) 1195-1203.
[25] A. Birkefeld, R. Schulin, B. Nowack, A New In Situ Method to Analyze Mineral Particle Reactions in Soils, Environ. Sci. Technol. 39 (2005) 3302-3307.
[26] M. Filippi, B. Dousova, V. Machovic, Mineralogical speciation of arsenic in soils above the Mokrsko-west gold ore deposit, Czech Republic, Geoderma. 139 (2007) 154-170.
[27] P.R. Kannel, S. Lee, S.R. Kanel, S.P. Khan, Chemometric application in classification and assessment of monitoring locations of an urban river system, Anal. Chim. Acta. 582 (2007) 390-399.
[28] B. López, M.J. Latorre, M.I. Fernández, M.A. García, S. García, C. Herreroa, Chemometric classification of honeys according to their type based on quality control data, Food Chem. 55 (1996) 281-287.
[29] G. Lucia, V. Daria Di, B. Giorgio, Classification of monovarietal Italian olive oils by unsupervised (PCA) and supervised (LDA) chemometrics, J. Sci. Food Agric. 83 (2003) 905-911.
[30] R.G. Giménez, R.V.d.l. Villa, P.R.d.l. Rosa, M.D.P. Domínguez, M.I. Rucandio, Analytical and multivariate study of roman age architectural terracotta from northeast of Spain, Talanta. 65 (2005) 861-868.
[31] R. Aruga, P. Mirti, A. Casoli, Application of multivariate chemometric techniques to the study of Roman pottery (terra sigillata), Anal. Chim. Acta. 276 (1993) 197-204.
[32] K. Castro, M. Perez-Alonso, M.D. Rodriguez-Laso, L.A. Fernandez, J.M. Madariaga, On-line FT-Raman and dispersive Raman spectra database of artists' materials (e-VISART database), Anal. Bioanal. Chem. 382 (2005) 248-258.
[33] S. Fernández, U. Villanueva, A. de Diego, G. Arana, J.M. Madariaga, Monitoring trace elements (Al, As, Cr, Cu, Fe, Mn, Ni and Zn) in deep and surface waters of the estuary of the Nerbioi-Ibaizabal River (Bay of Biscay, Basque Country), J. Marine Syst. 72 (2008) 332-341.

[34] D.J. Brigss, P. Smithson, Soil Formation, in: D.J. Brigss, P. Smithson (Eds.), Fundamentals of physical geography Rowman & Littlefield Pub Inc, USA, 1986, p. 456.

FIGURE CAPTIONS

Fig. 1. PCA representation of the raw data of metal concentration in soil adhered to archaeological pieces and stratum soil. Subsamples are mixed among them. No clear grouping can be seen.

Fig. 2. New PCA representation with metal concentration normalized to Mn. In this case subsamples appear forming groups marked inside a circle.

Fig. 3. Microphotography showing the crystal grain on the calcite deposit and the surface of the ceramic on the left side.

Fig. 4. μ-ED-XRF spectrum in helium atmosphere of a crystalline grain on the calcite deposit which covers the surface of a Roman archaeological ceramic piece. K_α and K_β lines are identified.

Fig. 5. Raman spectra of a) a soil sample taken from a Roman archaeologic ceramic piece; b) Standard spectrum of Peretaite ($CaSb_4O_4(SO_4)_2(OH)_2 \cdot 2H_2O$); c) Standard spectrum of the solid Aurichalcite ($Zn_5(CO_3)_2(OH)_6$); and d) Standard of Calcite ($CaCO_3$). The spectrum of the sample is shown without any smoothing treatment.

Fig. 6. PC1 vs PC3 representation of the model for metal composition of soil adhered to archaeological pieces and stratum soil. Groups of soil samples from the same stratum appeared inside a circle.

Fig. 7. Weights of the different metals that describe each group in SIMCA analysis.

Fig. 8. Distribution of samples on plane given by first two canonical discriminant functions, accounting for 92.1% of total variance (function 1: 75.5% and function 2: 16.6%). ■: group centroid; ○, ◊, *, Δ: samples. Soil samples from the same stratum appear close to each other around the centroid without mixing among them.

Table 1. ICP-MS experimental conditions.

RF power	1100 W
Plasma gas flow rate	15 L/min
Auxiliary gas flow rate	1 L/min
Nebulizer gas flow rate	0.9 L/min
Sample uptake rate	1 mL/min
Mass range	6-240 amu
Data acquisition	Peak Hopping
Dwell time	50 ms
Sweeps per reading	6
Readings per replicate	1
Replicates	1

Table 2: ICP-MS obtained metal concentration (μg/g) after microwave acid digestion of the archaeological soils.

Metal	Pb	Zn	Mo	Ag	Sb	Hg	Cr	Ni	Cu	Cd
Maximum	3100	3400	7.6	6.4	1.0	1.3	82	42	164	5.0
Average	70	95	0.4	1.0	0.35	0.2	15	10	40	0.2

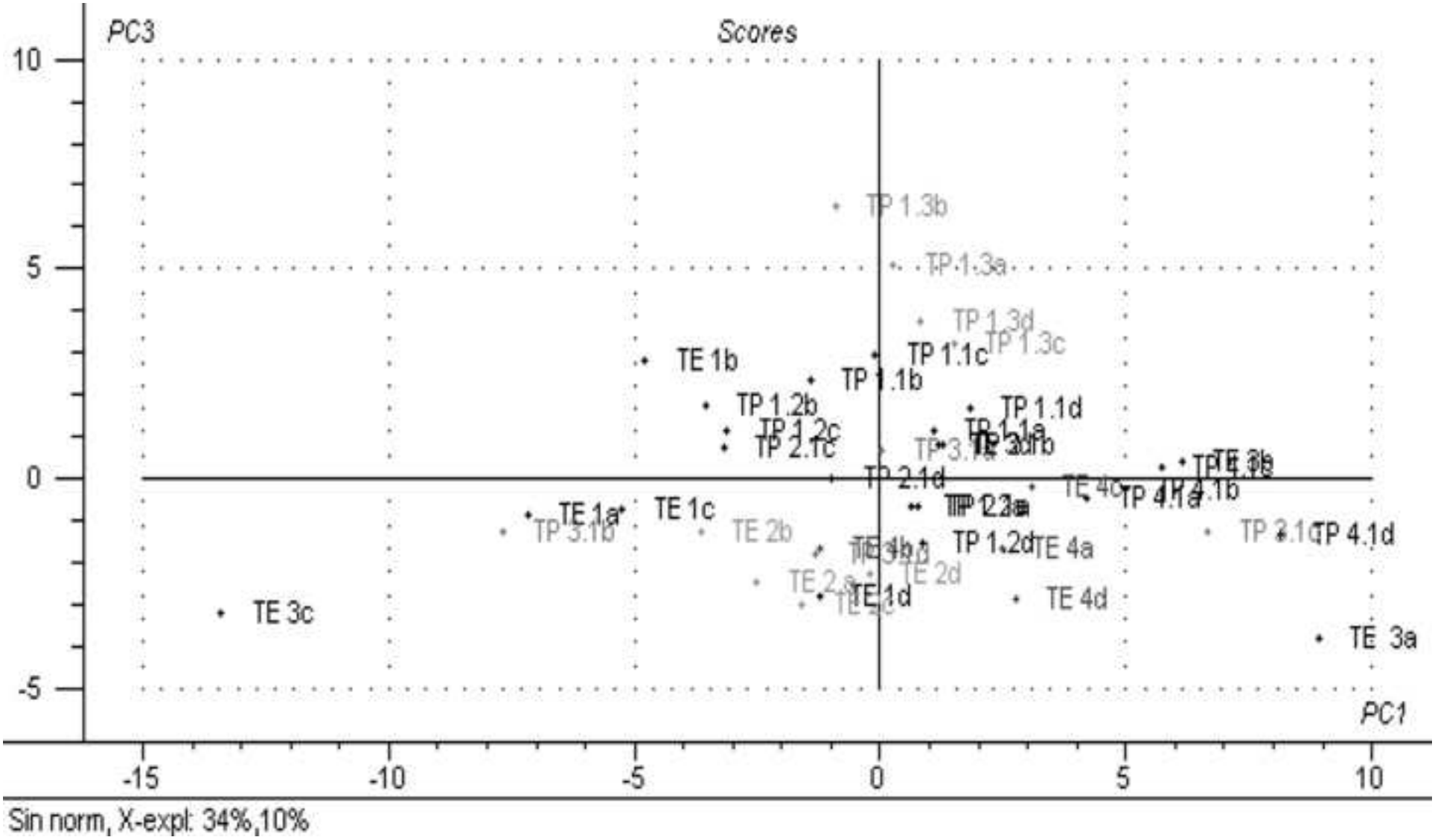
Table 3: Percentages of correctly classified samples corresponding to the LDA classification developed to differentiate between different strata provenance.

Category	Classification (%)	Prediction (leave-one-out) (%)
Stratum 1	100	100 (14/14)
Stratum 2	100	85.7 (6/7)
Stratum 3	100	100 (7/7)
Stratum 4	100	100 (8/8)
Total	100	97.2

The number of correct/total classifications (predictions) appears in parentheses

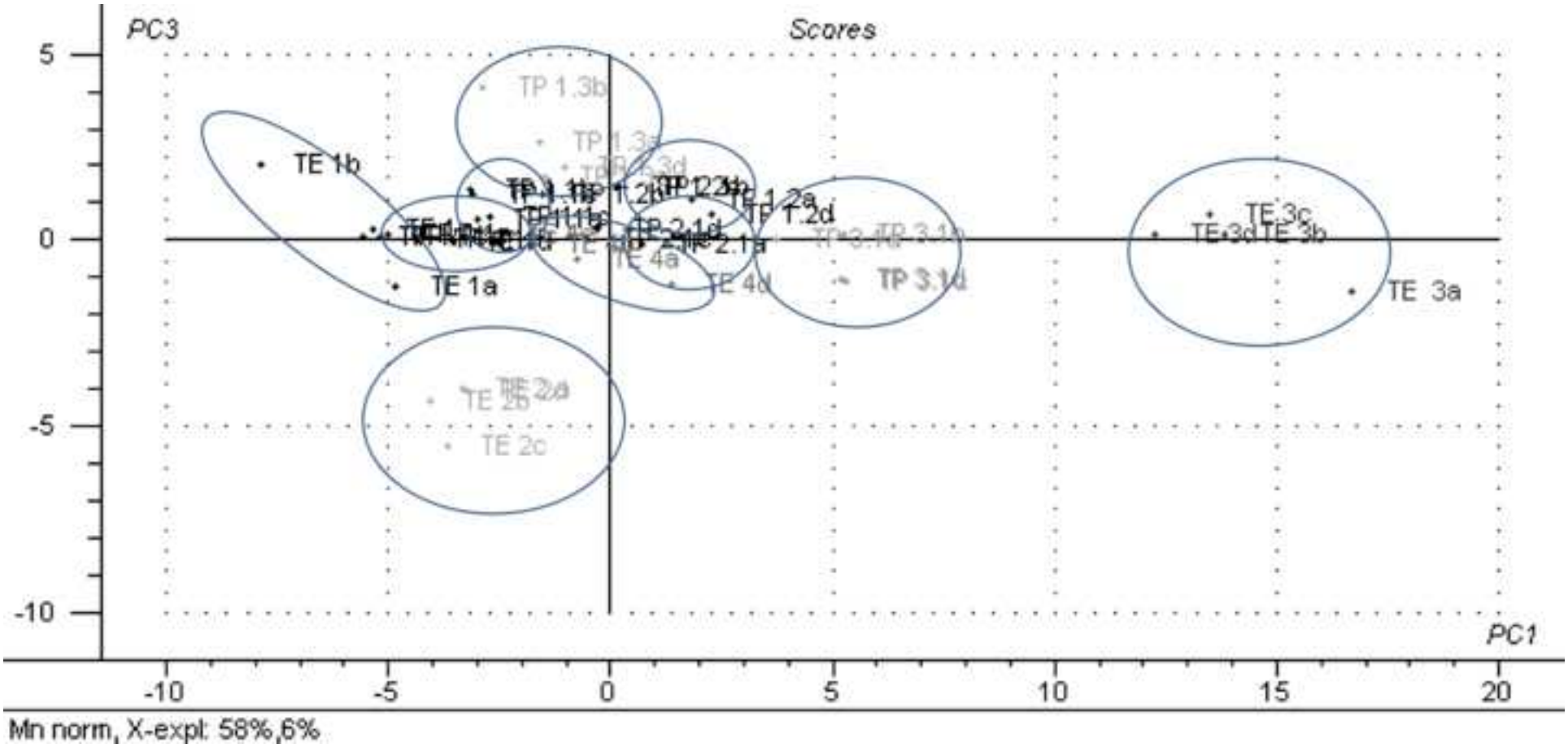
Figure(s)

Click here to download high resolution image



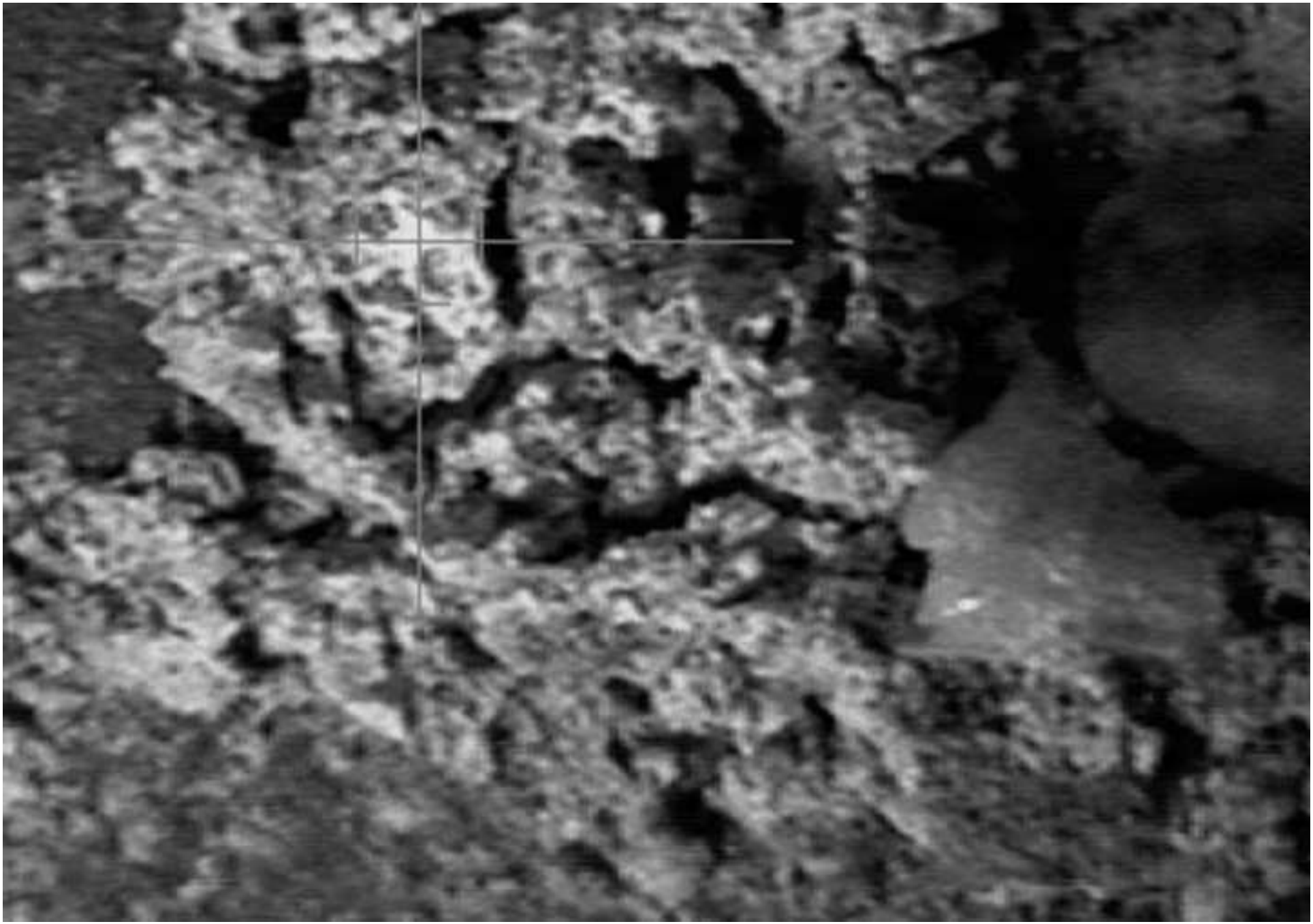
Figure(s)

Click here to download high resolution image



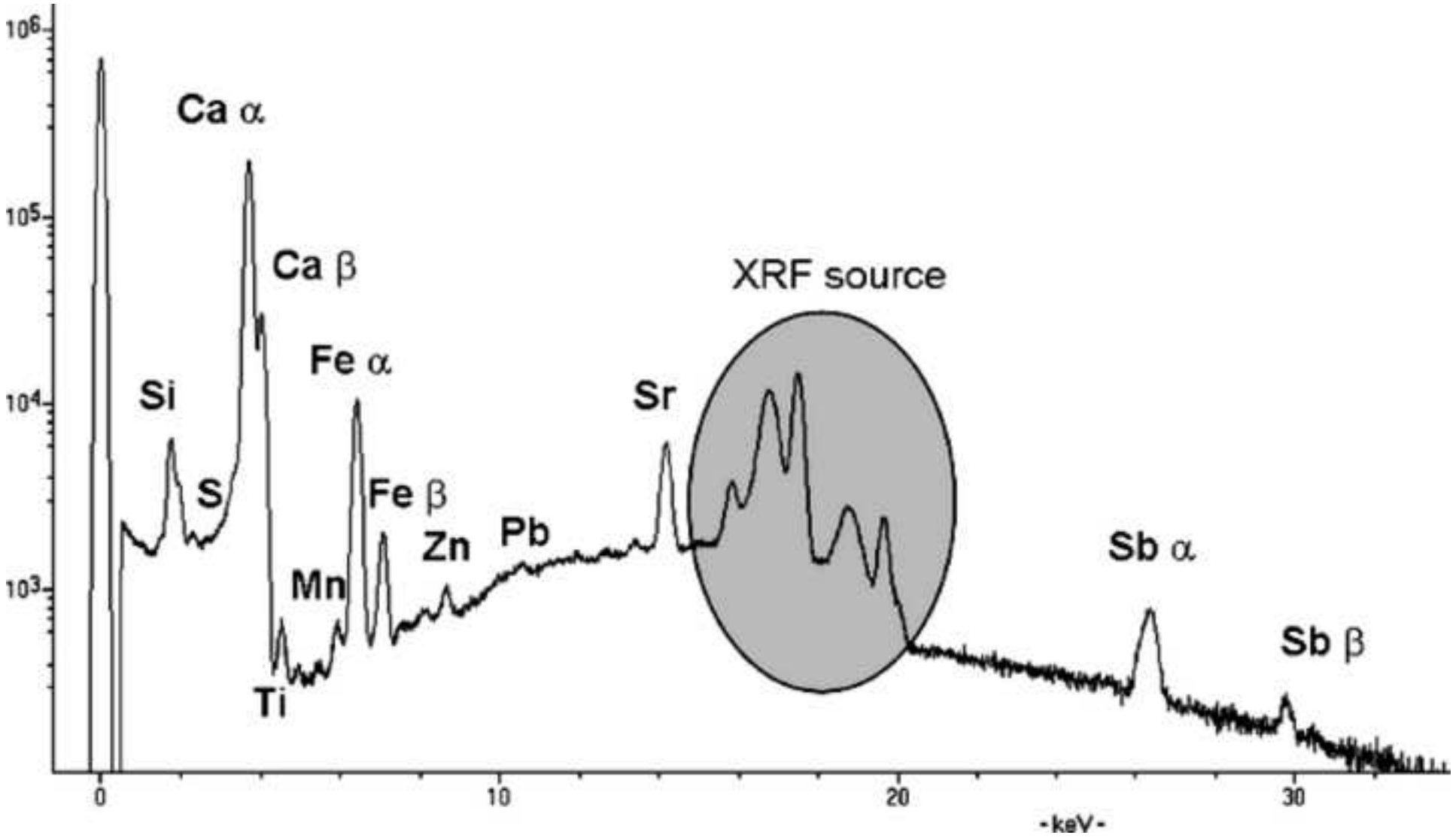
Figure(s)

Click here to download high resolution image



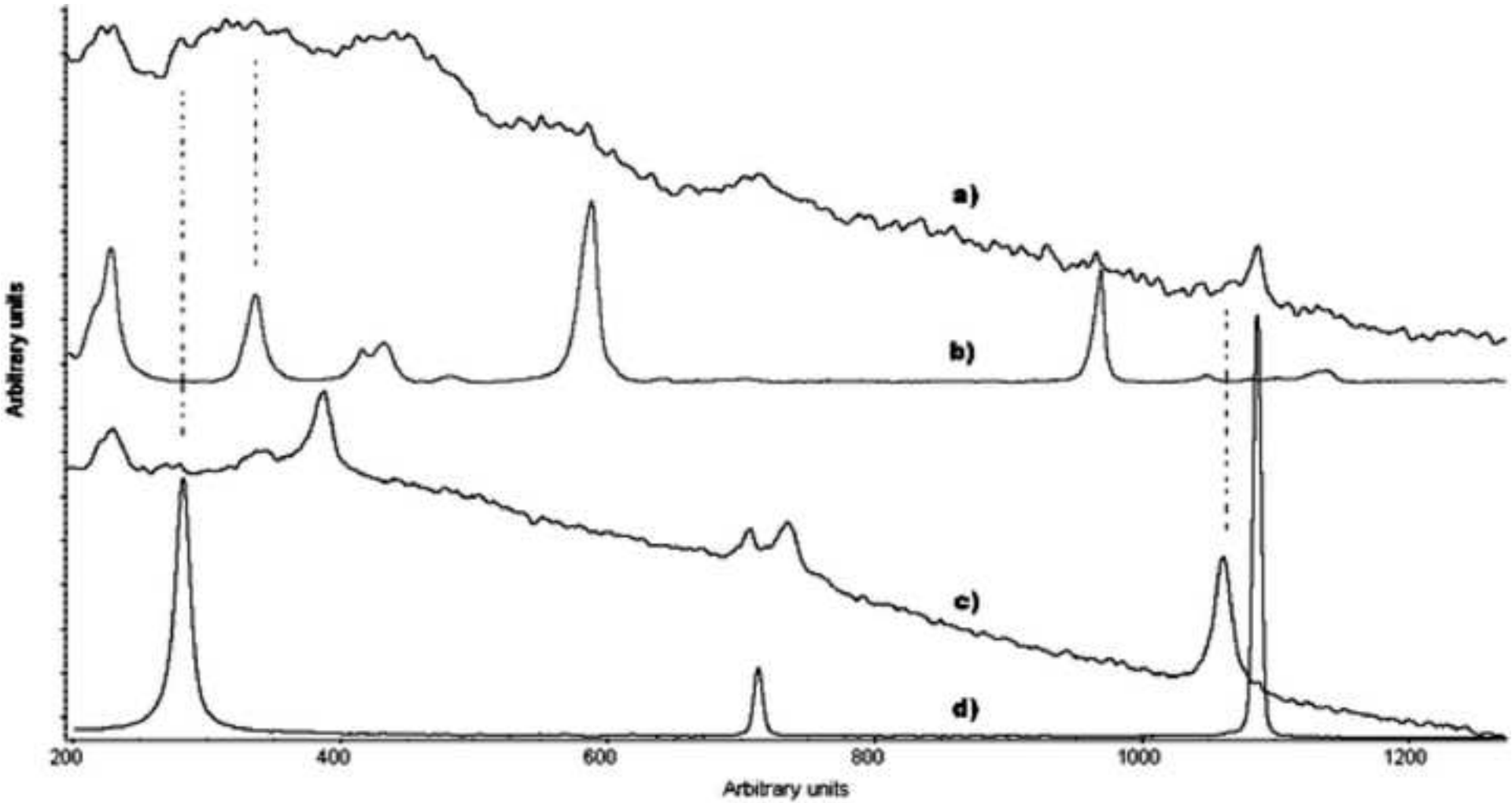
Figure(s)

Click here to download high resolution image



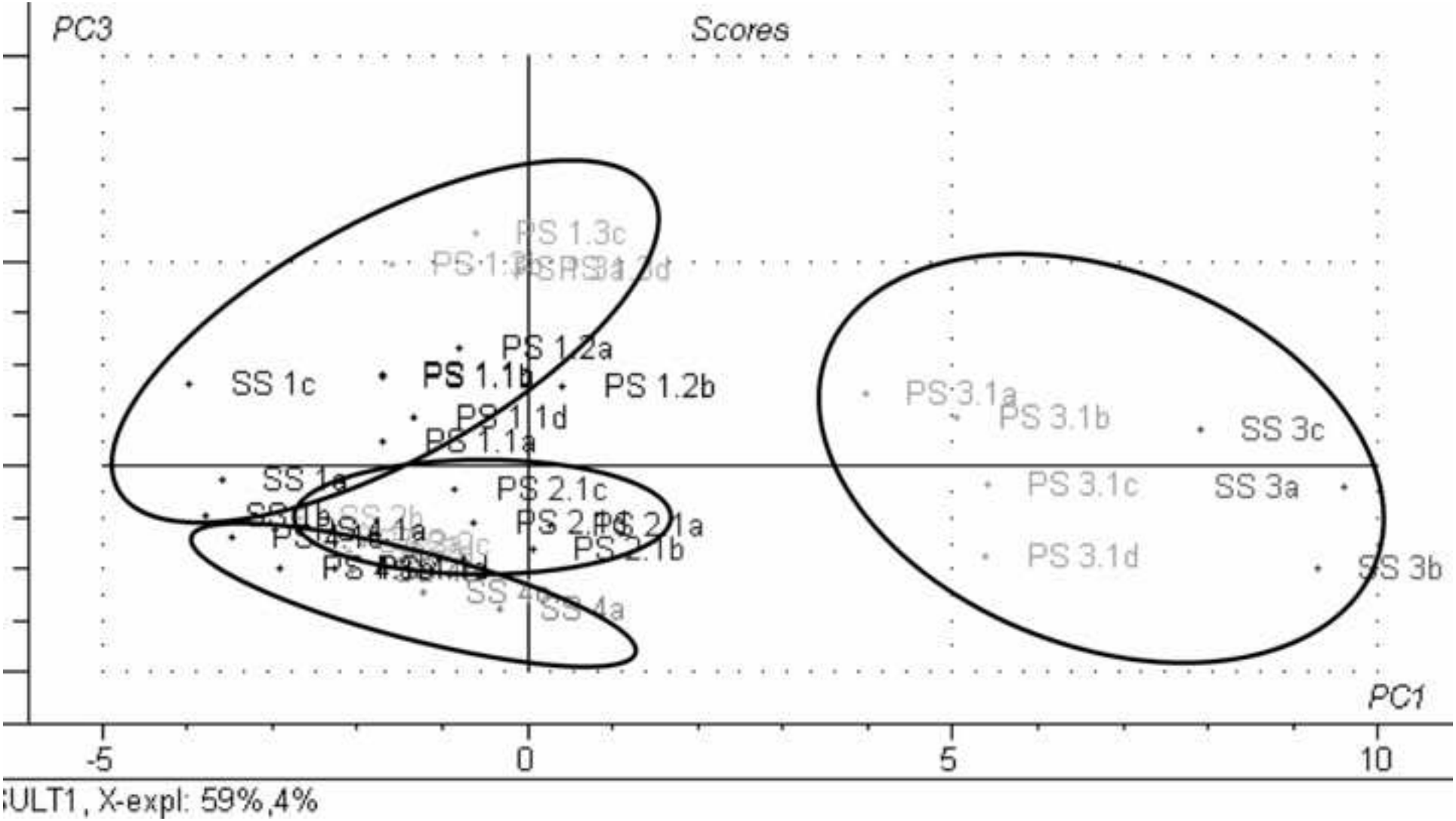
Figure(s)

Click here to download high resolution image



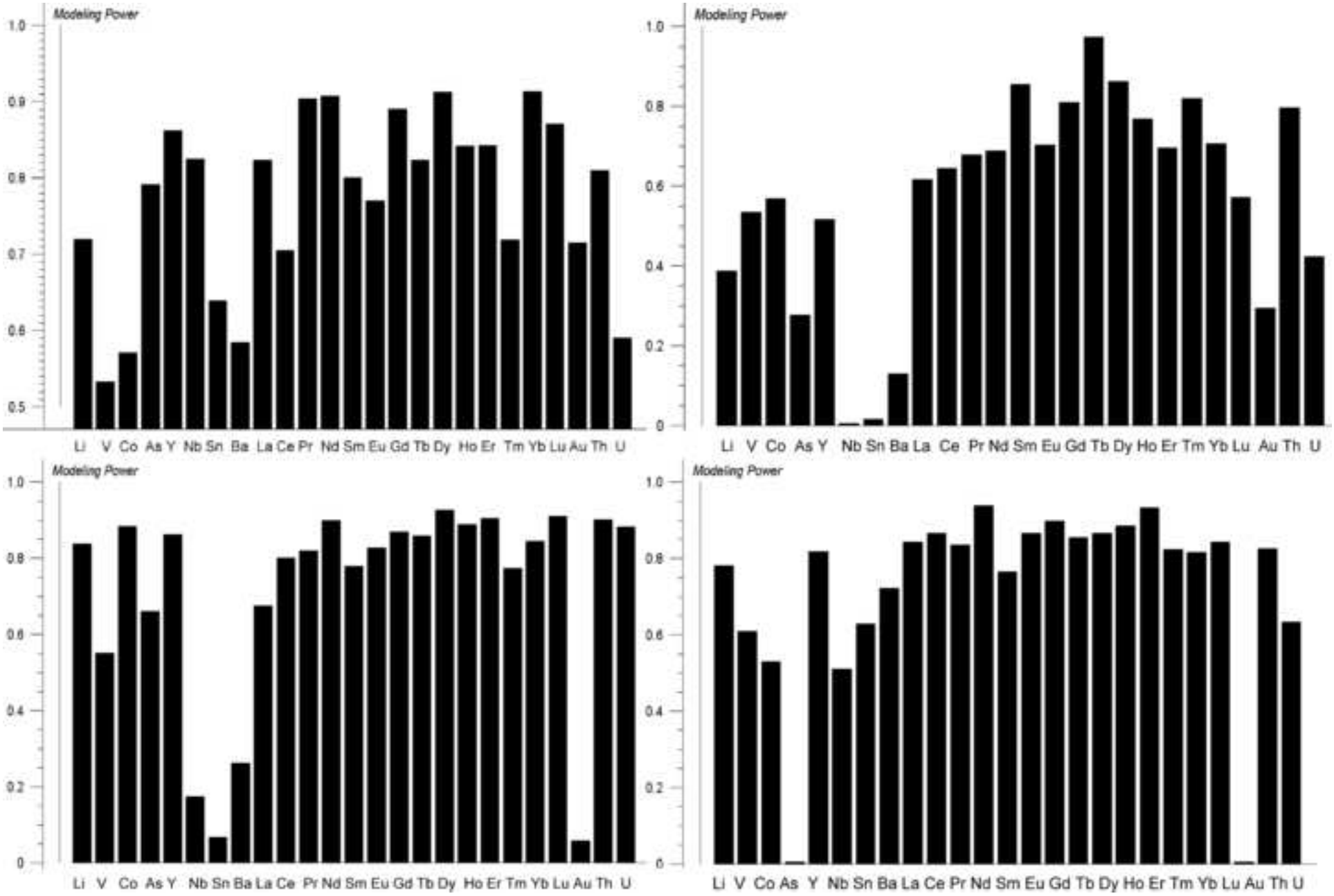
Figure(s)

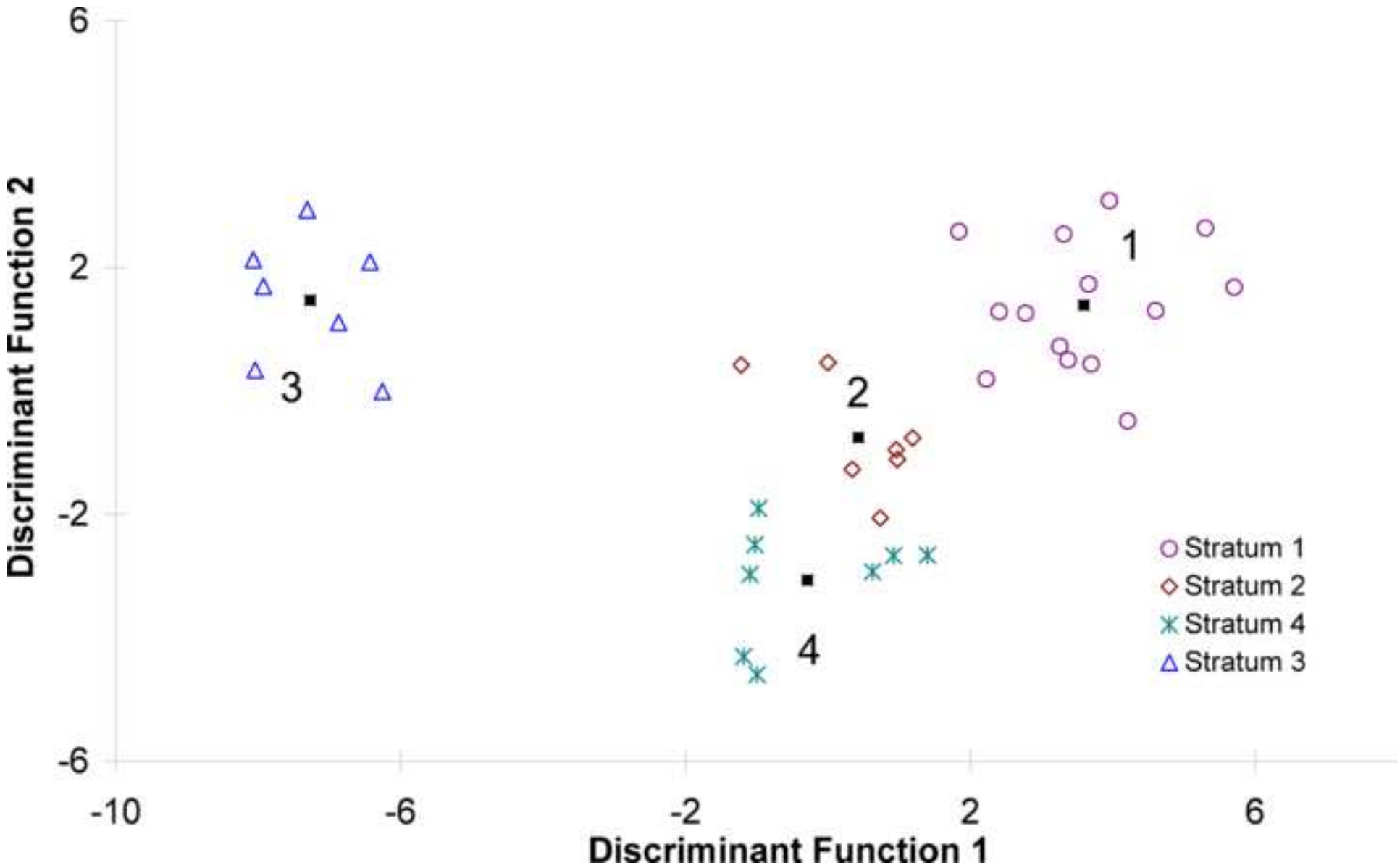
Click here to download high resolution image



Figure(s)

Click here to download high resolution image



Figure(s)

Click here to download high resolution image