The use of a combined portable X ray fluorescence and multivariate statistical methods to assess a validated macroscopic rock samples classification in an ore exploration survey

J. Figueroa-Cisterna [a], M.G. Bagur-González [b,c,*], S. Morales-Ruano [a,c], J. Carrillo-Rosúa [c,d], F. Martín-Peinado [e]

[a] *Department of Mineralogy and Petrology, Faculty of Sciences, Avda. Fuentenueva s/n, University of Granada, 18071 Granada, Spain*
[b] *Department of Analytical Chemistry, Faculty of Sciences, Avda. Fuentenueva s/n, University of Granada, 18071 Granada, Spain*
[c] *Instituto Andaluz de Ciencias de la Tierra (University of Granada-CSIC), Faculty of Sciences, Avda. Fuentenueva s/n, University of Granada, 18071 Granada, Spain*
[d] *Department of Didactics of Experimental Sciences, Faculty of Education Sciences, Campus de Cartuja s/n, University of Granada, 18071 Granada, Spain*
[e] *Department of Soil Science, Faculty of Sciences, Avda. Fuentenueva s/n, University of Granada, 18071 Granada, Spain*

ARTICLE INFO

ABSTRACT

The combination of "ex situ" portable X ray fluorescence with unsupervised and supervised pattern recognition techniques such as hierarchical cluster analysis, principal components analysis, factor analysis and linear discriminant analysis have been applied to rock samples, in order to validate a "in situ" macroscopic rock samples classification of samples collected in the Boris Angelo mining area (Central Chile), during a drill-hole survey carried out to evaluate the economic potential of this Cu deposit. The analysed elements were Ca, Cu, Fe, K, Mn, Pb, Rb, Sr, Ti and Zn. The statistical treatment of the geological data has been arisen from the application of the Box–Cox transformation used to transform the data set in normal form to minimize the non-normal distribution of the data. From the statistical results obtained it can be concluded that the macroscopic classification applied to the transformed data permits at least, to distinguish quite well in relation to two of the rock classes defined (70.5% correctly classified ($p < 0.05$)) as well as for four of the five alteration types defined "in situ" (75% of the total samples).

1. Introduction

The extraction of metals from the earth crust initially requires the identification of the areas in which they have anomalous concentration in relation to the host rock of the ore mineralization and, in general sense, to the background in the mining zone. In this sense, the geological characterization of the potential host rocks of ore mineralization is crucial and must be the preliminary objective in any exploration survey.

In relation with this fact, two fundamental stages must be covered, the establishment of the geological cartography and the drill-hole survey. The former because permits the knowledge of the main geological features (lithologies, structures, mineralization evidences, etc.) and the latter because gives an invaluable set of data over the geology under the surface. From the study of the information obtained in these stages it is possible to get a three dimensional idea about the existing rocks and their characteristics. Thus recognition and classification of the different rock types, as well as its alteration pattern, play an important role which could be critic in the selection of the areas that could be adequate to explore host ore bodies with economic interest.

During the initial field campaign necessary in order to obtain the data, a lot of samples are generated. In this sense, they must be classified attending criteria closely related to the type of deposits to be exploited, e.g. type of lithologies, hydrothermal alteration patterns among others. In the most of the cases these criteria are applied "in situ" in remote areas without confirmatory analytical information from a laboratory, and, in the best of the cases, using basic equipment like a magnifying glass or some safe and easily portable chemical reagent. In this way, it could be helpful to dispose of qualimetric tools that could validate this macroscopic rock samples classification in order to facilitate and accelerated the remained work necessary to determine the goodness of the ulterior mining exploration of the zone investigated.

Bearing in mind these reasons, the use of analytical techniques as portable X-ray fluorescence (P-XRF) combined with statistical pattern recognition techniques can be offered as an adequate tool in order to obtain a feasible model that could permits the assessment

* Corresponding author at: Department of Analytical Chemistry, Faculty of Sciences, Avda. Fuentenueva s/n, University of Granada, 18071 Granada, Spain. Tel.: +34 958243327; fax: +34 958243328.
E-mail address: mgbagur@ugr.es (M.G. Bagur-González).

of a validated macroscopic rock samples classification in an ore exploration survey.

Up to the present day, the use of field portable X-ray fluorescence (P-XRF) analysers [1–5] has been demonstrated be adequate in order to solve questions related with a great variety of deals, e.g. for the assessment of the composition of painting materials in order to offer information about their conservation and/or restoration procedures [6,7], for archaeological studies [8,9], for the screening and assessment studies about metalloids and/or heavy metals in contaminated or potentially contaminated areas [3,4,10–14], FDA regulated products [15], or metal contents in waters [16], among others. On the other hand, the relatively low cost of these devices permits the possibility of their use in lab for routine analysis in quality control assessments.

In parallel, it has been demonstrated that the use of unsupervised and supervised pattern recognition techniques permits to extract reliable information from analytical parameters for exploratory assessment of geological sets [17–19], mainly due to they allow (a) to verify associations among variables, (b) to group or to cluster samples with respect to comparable chemical or geological descriptors, and (c) to search multivariate data classification on the basis of known class membership of those objects.

Nowadays, copper is one of the most demanded materials on the metal market showing a growing demand perspective at the present such as the future. Together with this growing demand, the exploration of this metal has been widespread for the entire world to satisfy the copper supply. In this context, Chile is the first copper-producing country holding a 36% of the world production of this metal.

Boris Angelo Cu–(Ag) deposit is located in the "stratabound Cu–(Ag) belt" [20] in the Costal Cordillera, Central Chile. It corresponds to practically unknown deposit in this belt, thus the study carried out in the area can be considered as a typical case study of an exploration survey of a copper deposit. In this paper, a normalized data matrix obtained from P-XRF measurements of rock samples from a "preliminary ore exploratory survey" has been subjected to different pattern recognition techniques in order to confirm the rock classification parameters of samples taken during the drill-hole survey made in the Boris Angelo area.

2. Materials and methods

2.1. Studied area and macroscopic classification defined

The Boris Angelo Cu–(Ag) deposit is located in the easternmost Coastal Cordillera, in Central Chile, between 32°30′S and 70°40′W (Fig. 1). It is part of the Cretaceous stratabound Cu–(Ag) deposits belt, which are also known as "Chilean Manto-type" Cu–(Ag) deposits. The geology of the deposit area is characterized by volcanoclastics sequences intruded by different small subvolcanic bodies. Table 1 shows the four different lithologies recognized in the zone and its most representative characteristics. As well as, Table 1 included the coded values assigned to them in the macroscopic classification made "in situ".

From the point of view of the alteration patterns, the area of the metallic deposit is affected by hydrothermal alteration, caused by the interaction between hot and slightly acidic fluids and the host rocks [21]. These fluids can leach metals (with economic interest) and re-concentrate them. As mentioned above, the recognition and cartography of alteration patters in the rocks is a useful tool used by exploration geologist as evidence to localize enriched-metals areas with economic potential. The most common classification method, and the simplest visual method too, is that which defined the type of alteration as a function of the most abundant or most obvious mineral in the altered rock. Table 1 shows the five different hydrothermal alterations recognized in the zone, on the basis of the occurrence of certain "key minerals" or "key mineral assemblages" product of the hydrothermal alteration, and its most representative characteristics. In order to facilitate the analysis of the data, a second numerical code has been assigned to the alteration types used in the study. These codes have also been included in Table 1.

2.2. Sampling preparation and measurement

During the field campaign 44 rock samples, corresponding to ore grade zones and barren zones, were taken from five different drill-hole cores selected (see Fig. 1). The samples were coded and placed into sealed plastic bags in order to their preservation and transportation to the "Minera Las Cenizas S.A." mining facilities where they were powdered (until <100 microns particle size) and homogenised using standard procedures before their transportation to the laboratory.

The monitored parameters were the concentration of Ca, Cu, Fe, K, Mn, Pb, Rb, Sr, Ti and Zn. The measurements were made in the laboratory to select the better measure conditions; the equipment used in this study was a field portable X-ray fluorescence analyser NITON XLt 792 (Niton, Billerica, USA), with a 40 kV X-ray tube with Ag anode target excitation source and a Silicon PIN-diode with a Peltier cooled detector. As part of the standard set-up routine, variables as type of holder (zip sealed plastic bag or polyethylene sample cups with Mylar X-Ray Fil (TF-160-255; Gauge 0.00024″–6 μm, 2.5′ diameter) obtained from the supplier, source count time (60, 90 and 120 s) and matrix effects among others were tested.

In relation to the holder to be used in the procedure, the influence of the type of material used was studied analysing a set of 15 holders for each type of containers without sample. No statistical differences ($p = 0.95$) were found between the holders supplied by Niton and the plastic bags used in the exploratory survey to storage the samples. In all the cases, the content of the elements were lower than those expected in the samples, not being necessary to used the average element content to correct the measurements. For ulterior analysis the zip sealed plastic bags were chosen.

In relation with the influence of the source count time, the best results were obtained using 90 s. These variables were then kept fixed for the rest of measurements. On the other hand, no matrix effect was detected using the program algorithm included in the analyser software. The analyser was calibrated using the silver and tungsten shielding on the inside shutter. After data acquisition, the results were downloaded to a portable PC for further processing. The results obtained for the rock samples analysed (expressed as the arithmetic means of five replicates of each sample) are shown in Table 2.

The Relative percentage difference (RPD) found for each measured element in the five replicated analysis of the samples has been: $5.9 < RPD_{Ca} < 7.8$; $6.9 < RPD_{Cu} < 9.9$; $5.8 < RPD_{Fe} < 9.2$; $7.1 < RDP_K < 9.5$; $9.2 < RPD_{Mn} < 11.9$; $12.6 < RPD_{Pb} < 15.2$; $13.9 < RPD_{Rb} < 15.1$; $7.8 < RPD_{Sr} < 9.9$; $6.9 < RPD_{Ti} < 9.3$; $13.1 < RPD_{Ca} < 15.0$.

The accuracy of the method for all the elements except Rb, was corroborated analysing nine replicates of two Certified Materials: CRM052-050 (RT Corporation, Salisbury, United Kingdom) and RTS-1 (Canadian Certified Reference Methods Project, CANMET, Ottawa, Canada). According to the US EPA Method 6200 recommendations for soil samples [22], the accuracy was estimated by the RPD between the concentration in the reference material and the concentration measured (expressed as arithmetic mean of the nine replicates) by P-XRF, in all the cases results were in good agreement with the quality US EPA Method (RPD < 10 for Cu, Fe, Mn, Ti and Zn, 10< RPD < 25 for the rest of the elements).
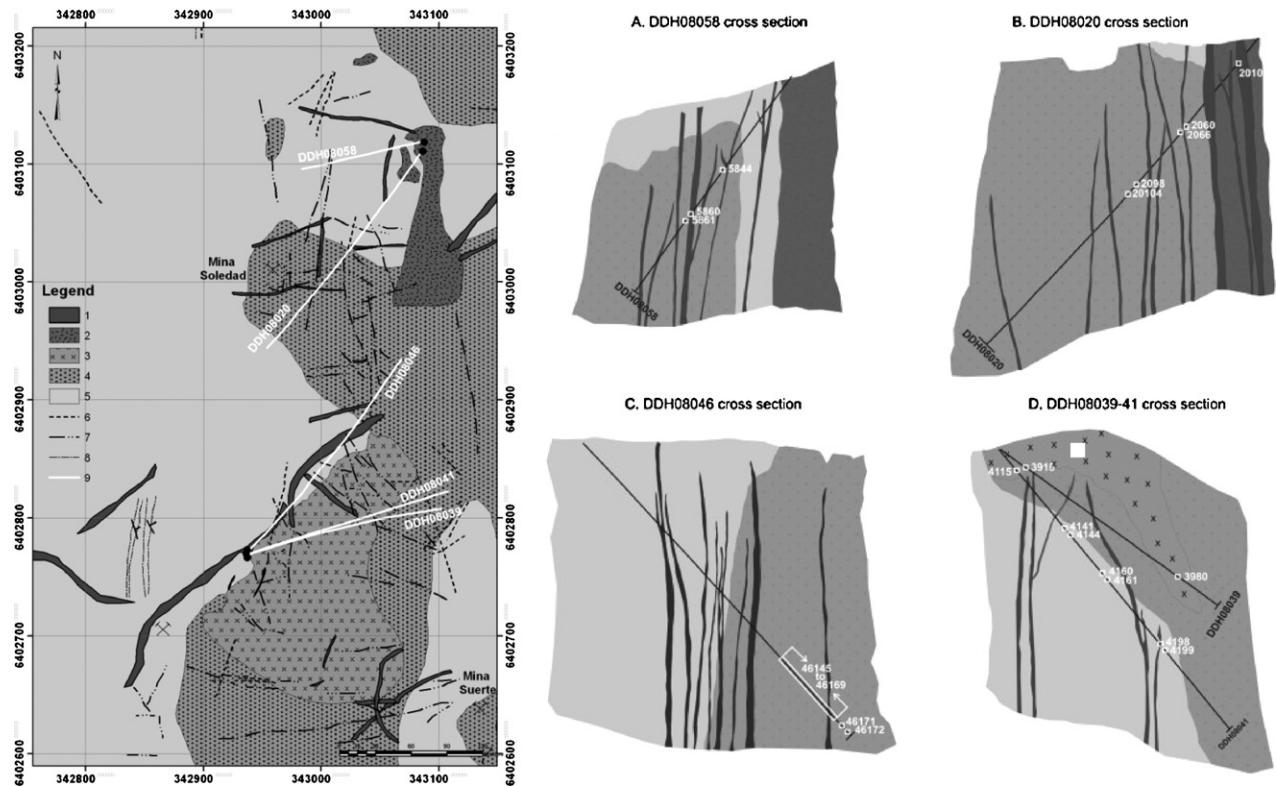
Fig. 1. Geological map of the Boris Angelo area. (1) Porphyritic dykes. (2) & (3) Brecciated porphyritic sub-volcanic rocks. (4) Porphyritic sub-volcanic rocks. (5) Volcanoclastic rocks. (6) Fault. (7) Mineralized vein-fault. (8) Contact. (9) Drill-hole A, B, C and D. Cross sections showing the drill holes position and samples location.

The establishment of the accuracy in the determination of Rb was made by means of an "in house validation protocol". Thus, three sets of spiked matrix matched samples (nine replicates) containing known Rb concentrations (one level for each set of spiked samples) were measured. The RPD estimated was 12, which is in good agreement with those obtained for the rest of the elements.

The precision was estimated as intermediate precision by the RPD of the nine measurements of each reference materials or spiked-matrix samples for Rb. In all the cases the obtained RPD values are lesser than 15. In order to estimate the detection limits sets of nine replicate samples that contained the target elements at concentration levels close to the detection limit estimated by US EPA Method 6200 [22].

2.3. Data treatment and statistical methods

Fig. 2 shows the statistical procedure used to the data treatment. Initially, to check the fit of the data to a normal distribution, Kolmogorov–Smirnov, Shapiro–Wilks and skewness score normality tests were applied. In all the cases, statistical evidences at 95% confidence interval showed that not all variables are normally distributed.

In order to transform the data set into normal form, Box–Cox transformation was used [17,23–25]. To study the correlation structure between normalized variables the Spearman rank-order correlations (Spearman R coefficient) were used, due to this one is less sensitive to outliers than the Pearson coefficients.

Table 1
Lithological types (A) and hydrothermal alteration types (B) identified in the studied zone.

Classes	Characteristics	Coded values
(A)		
Porphyritic dykes	Small and tabular bodies with andesitic composition. Porphyritic to aphanitic texture with plagioclase and occasionally amphibole phenocrysts.	11
Porphyritic sub-volcanic rocks	Sub-volcanic intrusive body (stock). Porphyritic texture with plagioclase and amphibole phenocrysts.	12
Brecciated porphyritic sub-volcanic rocks	Brecciated texture sub-volcanic intrusive body (stock). Brecciated and porphyritic texture with plagioclase and amphibole phenocrysts.	13
Volcanoclastic rocks	Tuff, breccias and agglomerated sequences. Homoclinal structure.	14

Classes	Characteristics	Coded values
(B)		
Propylitic	Chlorite (Chl), epidote (Ep), calcite (Cac) and hematite (Hem) assemblages.	21
Carbonatization	Calcite veins.	22
Albitic	Strong replacement of plagioclase by albite or albite veins presence.	23
Argillic assemblages	Assemblages of undetermined clay minerals, Fe oxy-hydroxide. This alteration is mainly related with fault zones.	24
Chloritic	Veins and replacement of rock components by chlorite.	25

Table 2
Analyte concentrations in the rock samples analysed.

Sample	Code		Content of elements analysed (expressed as mg/kg)									
	Rock	Alteration	Ca	Cu	Fe	K	Mn	Pb	Rb	Sr	Ti	Zn
2010	11	25	65,394	90	59,518	8918	1255	<l.o.d.	20	554	4698	77
20,104	12	22	62,591	356	35,454	11,934	1508	31	38	396	1908	118
2060	12	21	33,347	20,899	44,028	17,251	1509	146	48	372	1978	<l.o.d.
2066	11	21	34,559	376	63,901	15,345	1747	14	39	595	5742	79
2098	12	23	30,490	782	39,927	19,037	1789	26	52	327	2233	179
3918	13	24	39,603	130	45,353	15,216	1183	177	46	487	3015	94
3980	13	21	36,430	455	46,006	8933	1149	57	23	362	2939	34
41,100	14	21	23,879	10,143	64,067	17,028	1800	60	85	410	3955	103
4115	13	21	33,144	798	46,765	14,652	1225	86	38	449	3251	109
4141	14	21	33,634	95	57,738	17,780	1253	40	52	608	4681	96
4160	14	21	54,627	271	42,096	7456	1126	49	24	908	4727	78
4161	14	21	43,300	244	54,505	12,096	1389	34	34	704	5196	98
4198	11	21	35,326	649	60,672	11,675	1523	79	31	519	5320	84
4199	14	21	24,932	21,193	74,146	12,254	2118	81	62	391	3011	182
46,145	12	23	30,026	1109	39,501	18,823	1405	36	57	388	2249	125
46,146	12	23	28,999	1372	36,744	19,900	1641	152	54	341	2239	132
46,147	12	23	31,156	1342	42,245	19,365	2003	44	54	402	1970	147
46,148	12	23	29,162	1457	43,010	21,714	1868	27	60	345	2124	135
46,149	12	23	27,233	985	38,604	20,166	1475	14	59	321	2403	114
46,150	12	23	28,988	741	33,599	19,807	1429	15	58	350	2071	135
46,151	12	21	29,941	165	33,961	19,946	1395	20	56	384	2130	110
46,152	12	23	29,218	478	37,989	19,713	1378	19	53	421	2083	92
46,153	12	23	28,530	446	34,185	18,404	1381	21	49	405	2154	95
46,154	12	23	27,669	1250	38,656	18,521	1600	24	51	388	2190	93
46,155	12	23	27,794	1075	43,310	20,612	1598	18	58	406	2082	104
46,156	12	23	30,885	1658	40,289	19,044	1568	29	45	420	2254	112
46,157	12	23	29,943	876	33,427	18,759	1359	22	53	346	2121	106
46,158	12	23	27,928	379	35,874	17,373	1203	21	46	331	2159	82
46,159	12	23	27,100	348	33,360	17,850	1331	29	49	441	2100	75
46,160	12	23	26,597	813	32,892	21,555	1292	19	59	577	1979	76
46,161	11	21	34,623	203	42,404	14,304	1613	19	42	376	4617	131
46,162	12	23	30,826	1960	37,317	17,308	1336	32	52	286	2299	110
46,163	12	23	31,330	1244	37,513	17,092	1452	27	57	295	2151	86
46,164	12	23	28,131	350	39,190	20,512	1606	50	60	342	2254	94
46,165	12	23	32,414	2239	34,170	17,103	1216	22	47	320	2071	74
46,166	12	23	30,772	2109	36,361	17,134	1464	36	53	394	2389	128
46,167	12	23	24,912	978	37,848	18,888	1329	27	46	448	2017	68
46,168	12	23	29,855	743	36,164	16,095	1251	24	41	402	1979	97
46,169	12	21	26,345	525	40,945	18,021	1435	33	51	435	2064	102
46,171	12	21	25,233	164	37,103	19,325	1103	30	54	422	2010	79
46,172	12	24	25,949	274	36,263	17,169	1122	139	56	369	2155	105
5844	12	21	36,269	21,691	43,426	17,828	1385	152	58	411	1963	<l.o.d.
5860	11	25	49,599	108	51,867	13,772	1266	19	31	678	4819	67
5861	12	21	47,808	514	58,262	14,273	2177	92	40	444	5686	161

Geochemical Data non normal distributed

Box-Cox Transformation

Box-Cox Normalized Data Matrix

Pattern Recognition Methods

Unsupervised

HCA   PCA   FA

Supervised

LDA

Validate the macroscopic classification made "in situ"
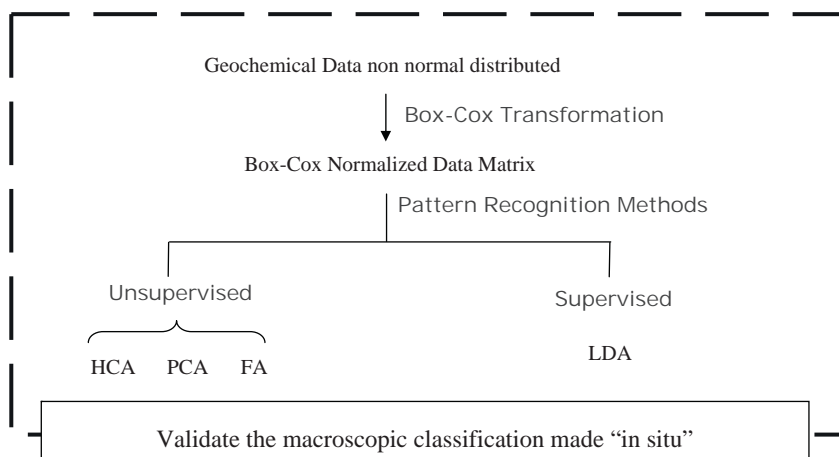
Fig. 2. Scheme of the statistical procedure used for data treatment.

Fig. 3. Dendrogram resulting from HCA of the Box–Cox normalized data set (R: rock code; A: alteration code; S: sample).
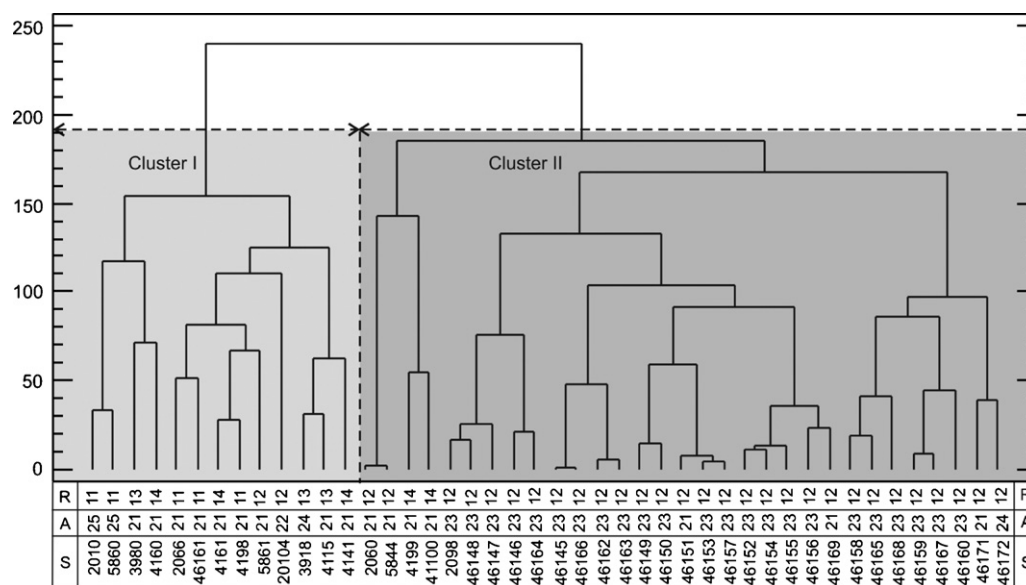
In all the cases, the univariate and multivariate statistical treatment of the data was performed using: (a) Statgraphic Centurion XV (15.2.05 version) for Windows (Statpoint Technologies Inc, Warrenton, USA), Matlab Version 7.0.4 R14 (The Mathworks, Inc.) and the PLS Toolbox Version 3.0.4 (eigenvector Research, Inc.).

2.4. Unsupervised pattern recognition methods: cluster analysis (HCA), principal component analysis (PCA) and factor analysis (FA)

Hierarchical agglomerative (HCA) was performed on the normalized data set by means of Manhattan (city-block) distance – a particular case of Minkowski distance (taxicab geometry) – as similarity measurement and Ward's method as amalgamation rule. These criterions have been selected with two objectives: (i) to find at each stage those two clusters whose merger gives the minimum increase in the total within group error sum of squares (Ward objective) and, (ii) to dampen the effect of outliers bearing in mind using city-block distance the average differences across dimensions are not squares. It was applied to the Box–Cox transformed monitoring matrix data set in order to observe the relationship between natural grouping observed and the two criteria of macroscopic classification made (lithologies and alteration patterns). On the other hand, in order to verify the natural grouping obtained in HCA, a PCA was applied to the standardized normalized data set. Finally, to reduce the interdependence of the data set of standardized normalized variables and to obtain knowledge of the underlying structure of the data, FA was applied. In this case the factorization type used was a principal component which supposes that all of the variability of the data corresponds exclusively to common factors. The orthogonal rotation of the axis defined by PCA, and obtained maximizing (Varimax rotation) produces new groups of variables called varifactors (VFs), which usually group the studied variables in accordance with common features which can include unobservable, hypothetical and/or latent variables [26].

2.5. Supervised pattern recognition methods: linear discriminant analysis (LDA)

This method has been applied in order to obtain a discriminant model that permits us to validate the "in situ" macroscopic rock samples classification of the sample assuming the number of groups

or classes, as well as, the group membership of each sample taken. Thus, by means of linear discriminant analysis, a discriminant function has been built up for each group on raw data. The classification functions associated to each group defined could be used to determine to which group each sample most likely belongs. In this study, LDA were performed on the Box–Cox transformed measured data.

3. Results and discussion

3.1. Macroscopic classification made on the basis of lithology criteria

The Dendrogram resulting from HCA of normalized data set is represented in Fig. 3. It can be seen that assuming a $D_{linkage} = 0.8 D_{maximum}$, the samples could be grouped in two different clusters: (i) Cluster I, constituted by 29.7% of the samples. According the rock codes, it contents samples of the four defined lithologies. From which, there are included 100% of the rock samples defined like porphyritic dykes (rock code 11) and brecciated porphyritic sub-volcanic rocks (rock code 13). (ii) Cluster II, constituted by the 70.3% of the samples, it can be observed that it is characterized by the biggest city-block distance (high significant clustering). Belonging to this cluster it can be found 93.5% rock samples defined as porphyritic sub-volcanic rocks (rock code 12) and 40% of the samples defined as volcanoclastic rocks (rock code 14).

The PCA (made on the basis of eigenvalues > 1 criterion) facilitated the explanation of the original 10 geochemical variables in a reduced space by three sets of the calculated principal components (PCs) which explaining about 73.9% of the total variance. The analysis of the data (see Table 3) shows that PC1 (41.6% of the total variance) is mainly influenced positively by the normalized concentration of Ca, Sr and Ti, and negatively by the normalized concentration of K and Rb. PC2 (19.9% of the total variance) is manly influenced positively by the normalized concentration of Fe, Mn and Zn, and PC3 (which explains the 12.4% of the total variance) influenced positively by Cu and Pb and negatively by Zn.

On the other hand, from the scattered plot obtained, considering PC1 and PC2 (Fig. 4(a)) two group of samples (A and B) can be observed. Group A is mainly characterized by negative values of the PC1 and PC2 whereas Group B is characterized by positive values of the PC1. In relation to the samples lithological classification made, Group A is formed by one lithology class, i.e. porphyritic sub-
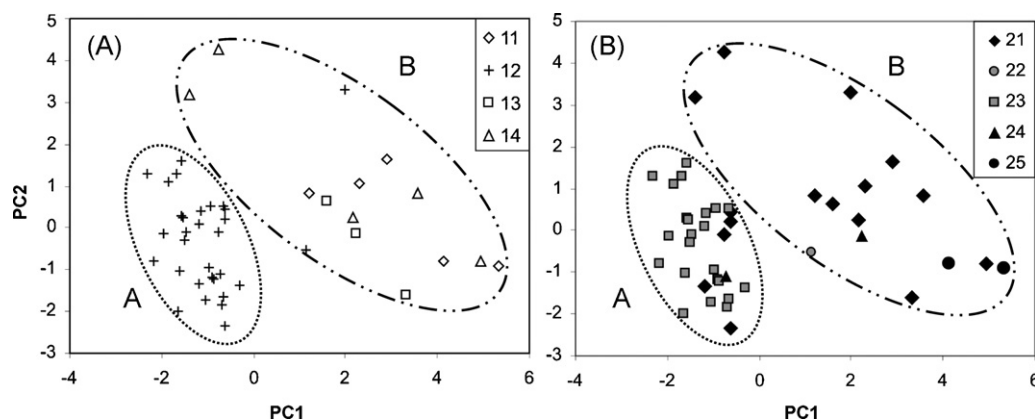
Fig. 4. Scatterplots obtained from PCA of the Box–Cox normalized data set using: (a) rock codes and (b) alteration codes.

Table 3
Standardized coefficients of weights of Box–Cox transformed variables in principal component and factorial analysis functions.

Variables	Principal components			Varimax rotated varifactors		
	PC1	PC2	PC3	VF1	VF2	VF3
WCa	0.4	0.02	−0.1	0.4	0.02	−0.1
WCu	−0.3	0.3	0.4	−0.3	0.3	0.4
WFe	0.3	0.5	0.05	0.3	0.5	0.05
WK	−0.4	−0.1	−0.2	−0.4	−0.1	−0.2
WMn	−0.2	0.6	−0.2	−0.2	0.6	−0.2
WPb	0.04	0.3	0.7	0.04	0.3	0.7
WRb	−0.4	0.2	−0.01	−0.4	0.2	−0.01
WSr	0.3	0.02	−0.1	0.3	0.02	−0.1
WTi	0.4	0.3	−0.1	0.4	0.3	−0.1
WZn	−0.2	0.4	−0.4	−0.2	0.4	−0.4

volcanic rocks and contents 93.5% of all of the samples coded as 12. Group B, more heterogeneous than the one, collects 100% of the samples classified as porphyritic dykes (rock code 11), brecciated porphyritic sub-volcanic (rock code 13) and volcanoclastic rocks (rock code 14), as well as two samples coded as rock code 12.

In relation with the previous HCA it can be observed that, although the natural behaviour of the samples is quite similar, the lineal combination of the normalized values of concentrations of the elements made by PCA permit us the grouping of the samples into two groups well defined: Group A, constituted by porphyritic sub-volcanic rocks, and Group B, constituted mainly for the rest of the lithologies classes.

When the FA was applied, the Varimax rotation of the axes defined by the PCs explains again about 73.9% of the total variance of the normalized data, but modify the weight of the normalized variables in the Three VFs obtained by means of the eigenfactor > 1 criterion (see Table 3). The scattered plot of the first two varifactors is shown in Fig. 5(a). It can be observed again two different groups of rock samples (A and B) which can be considered as identical as those obtained by PCA.

In general, the unsupervised pattern recognition methods applied to the Box–Cox transformed concentrations of the analysed elements by P-XRF permit assess that in all cases 61.4% of the samples are "naturally" grouped as belonging to the macroscopic group named porphyritic sub-volcanic rocks (rock code 12). The samples macroscopically classified as porphyritic dykes and brecciated porphyritic sub-volcanic rocks (rock codes 11 and 13 respectively) present a slightly defined patter of distribution between them. The dispersion shown by rocks classified like volcanoclastics rocks (rock code 14) could be attributed to the compositional heterogeneity characteristics of them.

In order to verify the natural grouping obtained and to validate the macroscopic classification, by means a statistical significant dis-

criminant model, a LDA was performed. The discriminant functions were calculated considering the Box–Cox normalized metals concentrations that mainly affect to VF1. On these bases, i.e. WCa, WFe, WK, WRb, WSr and WTi, three discriminant orthogonal functions were calculated. DF1 and DF2 contribute 99.20% to the total discriminant power being both functions statistically significant at 95% percentage. Eqs. (1) and (2) represent the expression of these functions.

$$DF1_{lithol.} = -0.9\,WK - 0.4\,WCa + 0.2\,WFe + 0.5\,WRb$$
$$+ 0.6\,WSr + 0.8\,WTi \tag{1}$$

$$DF2_{lithol.} = -0.9\,WK - 0.1\,WFe + 1.6\,WRb + 0.6\,WSr - 0.2\,WTi$$
$$\tag{2}$$

In order to evaluate the adequation of the classification model obtained, the samples were split into six groups and two levels for a 6-fold cross validation inner group by means of Venetian blinds. The percentage of variance captured by the model was 72.80%. Fig. 6(a) shows the scatter plots of observations in the space of discriminant functions.

Table 4(a) shows the classification percentages obtained from each class macroscopically defined. It can be observed that the predicted model with DF1 and DF2, considering the normalized geochemical information obtained with P-XRF measurements, classify satisfactory the 70.45% of the data. These facts could be closely related to (i) the heterogeneous composition of the "volcanoclastics rocks" (rock code 14), (ii) the similarity of the original composition between the "brecciated porphyyritic sub-volcanic rocks" (rock code 13) and "porphyritic sub-volcanic rocks" (rock code 12), as well as the matrix composition of the first rock type, (iii) the state of aggregation (powdered particles with <2 mm) in which the samples are measured that does not permit to appreciate the brecciated structure.

3.2. Macroscopic classification made on the basis of hydrothermal alteration criteria

In relation with the macroscopic classification made on basis of hydrothermal alteration criteria, from the analysis of the Dendrogram (see Fig. 3) can be observed that, in this occasion Cluster I is also the most heterogeneous. It is constituted by 56.25% of samples coded as propylitic alteration (alteration code 21), 100% of samples classified as "chloritic alteration" and "carbonatization" (alteration codes 25 and 22) and 50% of samples coded as "argillic assemblages" (alteration code 24). Whereas Cluster II is constituted by 100% of samples macroscopically defined as "albitic alteration" (alteration
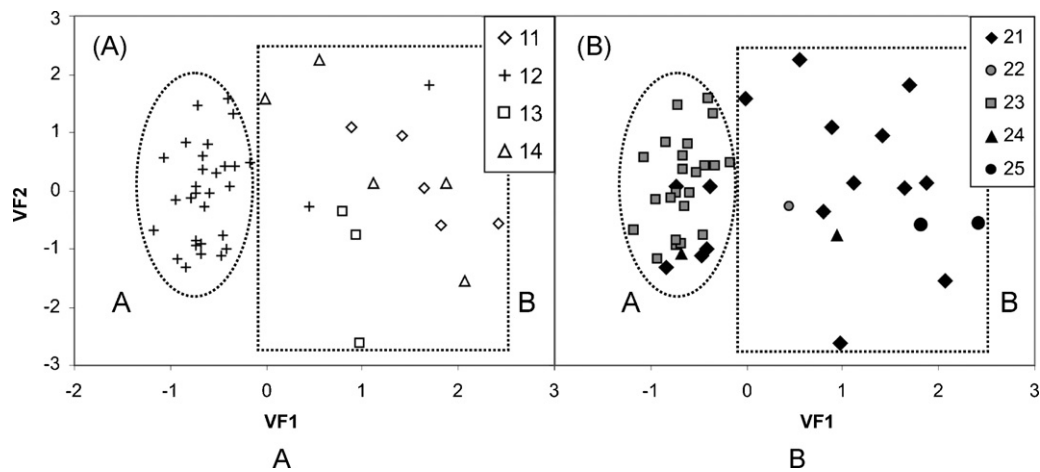
Fig. 5. Scatterplots obtained from FA of the Box–Cox normalized data set using (a) rock codes and (b) alteration codes.
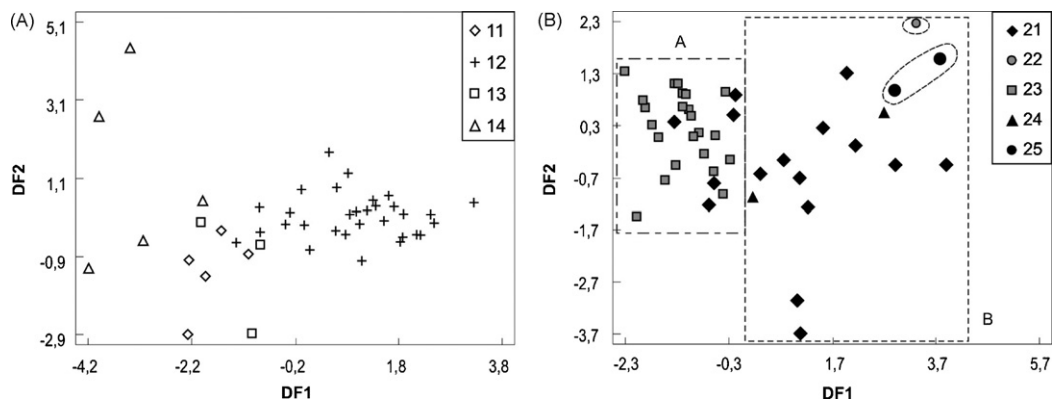

Fig. 6. Scatterplots obtained from discriminant functions of the Box–Cox normalized data set using (a) rock codes and (b) alteration codes.

Table 4
DLA classifications obtained for the rock samples analysed.

Classes	Code	Samples	Predicted			
			11	12	13	14
(a) Criteria: lithology of the samples						
Porphyritic dykes	11	5	0 (0.00%)	5 (100%)	0 (0.00%)	0 (0.00%)
Porphyritic sub-volcanic rocks	12	31	0 (0.00%)	*31* (100%)	0 (0.00%)	0 (0.00%)
Brecciated porphyritic sub-volcanic rocks	13	3	0 (0.00%)	3 (100%)	0 (0.00%)	0 (0.00%)
Volcanoclastic rocks	14	5	0 (0.00%)	5 (100%)	0 (0.00%)	0 (0.00%)

Alteration	Code	Samples	Predicted			
			21	22	23	24	25
(b) Criteria: alteration types						
Propylitic	21	16	6 (37.50%)	1 (6.25%)	3 (18.75%)	4 (25%)	2 (12.50%)
Carbonatization	22	1	0 (0.00%)	*1* (100%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Albitic	23	23	1 (4.35%)	0 (0.00%)	*22* (95.65%)	0 (0.00%)	0 (0.00%)
Argillic assemblages	24	2	0 (0.00%)	0 (0.00%)	0 (0.00%)	*2* (100%)	0 (0.00%)
Chloritic	25	2	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	*2* (100%)

The values denoted as italics correspond to the classes predicted correctly.

code 23), 50% of samples defined as "argillic assemblages" (alteration code 24) and by 43.75% of samples defined like propylitic assemblages (alteration code 21).

The results arisen from PCA (see Fig. 4(b)) permit us consider again two groups: Group A constituted by 100% of samples denoted as "albitic alteration" (alteration code 23), 45.45% of rock samples assigned as "propylitic alteration" (alteration code 21) and 50% of rock samples assigned as "argillic assemblages" (alteration code 24). Whereas, Group B includes 68.75% of samples assigned to "propylitic alteration" (alteration code 21), as well as samples defined as carbonatization (100%), argillic assemblages (50%) and chloritic (100%) alterations respectively.

In relation with the information obtain from FA, the scattered plot of the first two varifactors is shown in Fig. 5(b). It can be observed again two different groups of rock samples (A and B) which can be defined as those obtained by PCA. Thus, Group A is mainly constituted by 100% of "albitic alteration" (alteration code 23), meanwhile Group B is mainly constituted by samples belonging to "propylitic alteration" (alteration code 21), most of them characterized by high values of VF1 and VF2 and widespread distributed.

On the other hand, VF1 could be considered in close relation with alteration process, because most of the elements with highest weight, i.e. Ca, K, Rb and Sr could be involved in hydrothermal process [21]. These facts are in good agreement with positive influence of WCa, WSr and WTi, and negative influence of WK and WRb in VF1 found in Group B. It is constituted by samples with propylitic assemblages and chloritic and carbonatization alterations. Additionally, this is concordant with the mineralogy (possibly unaltered plagioclase like main mineral phase) of porphyritic (andesitic) dykes. The effect of Ti could be attributed to the presence of ilmenite in the ground mass of dykes. The samples classified as "albitic alteration" has a higher negative influence of WK and WRb, this fact could be due to Na widespread replacement involved in the plagioclase albitization process. This last mineral could be constituted (as initial composition) by significant amounts of K an Rb in its structure.

VF2 is influenced by the normalized concentrations of base metals as Mn and Zn, while VF3 are influenced by Cu and Pb. These elements could be related with ore mineralization process, nevertheless all the samples show a similar scattered distribution pattern in the diagrams without significance statistical differences inter samples.

Finally, in order to obtain a feasible classification model a PLS-LDA was made. In this case, considering geochemical features, i.e. (a) elements closely related to hydrothermal alterations, i.e. WCa, WK, WRb, WSr and (b) the element with economical interest, WCu. On these bases, four discriminant orthogonal functions were calculated. DF1 and DF2 contribute 89.13% to the total discriminant power being both functions the most statistically significant at 95% percentage. Eqs. (3) and (4) show the expressions of these functions,

$$DF1_{alt} = 0.6\,WCa - 1.1\,WK + WRb + 0.3\,WSr - 0.5\,WCu \qquad (3)$$

$$DF2_{alt} = WCa + 0.8\,WK - 0.2\,WRb - 0.6\,WSr - 0.1\,WCu \qquad (4)$$

On the other hand, in order to evaluate the adequation of the classification model obtained, the samples were split into six groups and two levels for a 6-fold cross validation inner group by means of Venetian blinds. The percentage of variance captured by the model was 72.80% for 73.32%. Fig. 6(b) shows the scatter plots of observations in the space of discriminant functions.

Table 4(b) shows the classification percentages obtained from each type of alteration macroscopically considered. It can be observed that the predicted model considering (i) geological features and (ii) chemical information obtained with P-XRF measurement classify satisfactory the 75% of the data. The heterogeneity in the predicted types obtained for the samples assigned

to "propylitic alteration" could be attributed to the fact that this kind of alteration could content, at micro-scale level, small bodies corresponding to a different alteration process (e.g. carbonatization, chloritic, etc.) not detectable with the equipment used in field campaign sampling.

4. Conclusions

The statistical treatment of the Box–Cox transformed geological data obtained from "ex situ" portable X ray fluorescence measurements of ore exploratory samples with unsupervised and supervised pattern recognition techniques such as HCA, PCA, FA and LDA has been shown as a helpfully tool for validate the "in situ" macroscopic rock samples classification applied to an exploratory survey in a potential mining area. From the LDA it can be concluded that in relation with the macroscopic rock samples classification based on lithology classes assuming a probability level of 80% the discriminant model obtained confirms correctly 81.8% of the analysed samples. When the alteration types are considered, the discriminant model obtained permits to confirm four of the five alteration types defined "in situ" (75% of the total samples). Thus, it can be pointed out that the classification assessed could be applied to facilitate and accelerated the remained work necessary to determine the goodness of the ulterior mining exploration.

On the other hand, the proposed approach could be apply directly "in situ" without pre-treatment of the rock samples during exploratory in those cases in which the characteristic of the samples could be well assessed.

References

[1] S. Piorek, Trends in Analytical Chemistry 13 (1994) 281–285.
[2] A. Argyraki, M.H. Ramsey, P.J. Potts, Analyst 122 (1997) 743–749.
[3] P.J. Potts, M.H. Ramsey, J. Carlisle, Journal of Environmental Monitoring 4 (2002) 1017–1024.
[4] J. Chou, G. Clement, B. Bursavich, D. Elbers, B. Cao, W. Zhou, Environmental Pollution 158 (2010) 2230–2234.
[5] P.J. Potts, in: J. Potts, M. West (Eds.), Portable X-ray Fluorescence Spectrometry: Capabilities for In-situ Analysis, RSC Publishing, London, 2008, pp. 1–12.
[6] F. Rosi, A. Burnstock, K.J. Van den Berg, C. Miliani, B.G. Brunetti, A. Sgamellotti, Spectrochimica Acta. Part A, Molecular and Biomolecular Spectroscopy 71 (2009) 1655–1662.
[7] G. Van der Snickt, K. Janssens, O. Schalm, C. Aibéo, H. Kloust, M. Alfelda, X-Ray Spectrometry 39 (2010) 103–111.
[8] S.C. Phillips, R.J. Speakman, Journal of Archaeological Science 36 (2009) 1256–1263.
[9] A.J. Nazaroff, K.M. Prufer, B.L. Drake, Journal of Archaeological Science 37 (2010) 885–895.
[10] P.L. Drake, N.J. Lawryk, K. Ashley, A.L. Sussell, K.J. Hazelwood, R. Song, Journal of Hazardous Materials 102 (2003) 29–38.
[11] C. Kilbride, J. Poole, T.R. Hutchings, Environmental Pollution 143 (2006) 16–23.
[12] T. Radu, D. Diamond, Journal of Hazardous Materials 171 (2009) 1168–1171.
[13] K. Hürkamp, T. Raab, J. Völkel, Geomorphology 110 (2009) 28–36.
[14] M. Ramsey, in: J. Potts, M. West (Eds.), Portable X-ray Fluorescence Spectrometry: Capabilities for In-situ analysis, RSC Publishing, London, 2008, pp. 39–55.
[15] P.T. Palmer, R. Jacobs, P.E. Baker, K. Ferguson, S. Webber, Journal of Agricultural and Food Chemistry 57 (2009) 2605–2613.
[16] F.L. Melquiades, P.S. Parreira, M.J. Yabe, M.Z. Corazza, R. Funfas, C.R. Appoloni, Talanta 73 (2007) 121–126.
[17] M.G. Bagur, S. Morales, M. López-Chicano, Talanta 80 (2009) 377–384.
[18] F. Martín-Peinado, S. Morales-Ruano, M.G. Bagur-González, C. Estepa-Molina, Geoderma 159 (2010) 76–82.
[19] M.G. Bagur-González, C. Estepa-Molina, F. Martín-Peinado, S. Morales-Ruano, Journal of Soils and Sediments 11–1 (2011) 281–289.

[20] V. Maksaev, M. Zentilli, in: T.M. Porter (Ed.), Hydrothermal Iron Oxide Copper–Gold & Related Deposits: A Global Perspective, vol. 2, PGC Publishing, Adelaide, 2002, pp. 185–205.

[21] H.L. Barnes, Geochemistry of Hydrothermal Ore Deposits, third ed., John Wiley & Sons, 1997.

[22] U.S. EPA, 2006. XRF technologies for measuring trace elements in soil and sediment. Niton XLt 700 Series XRF Analyzer. Innovative technology verification report EPA/540/R-06/004.

[23] G.E.P. Box, D.R. Cox, Journal of Royal Statistical Association Series B 26 (1964) 211–250.

[24] M. Meloun, J. Militký, M. Forina, PC-Aided Statistical Data Analysis, vol. 1, Ellis Horwood, Chichester, United Kingdom, 1992.

[25] M. Meloun, M. Sâňka, P. Němec, S. Krītikova, K. Kupta, Environmental Pollution 137 (2005) 273–280.

[26] C. Pérez-López, Estadística Práctica con STATGRAPHICS, Pearson Education S.A., Madrid, Spain, 2002.