The use of multivariate statistical analysis to evaluate spatial and temporal water contamination in Germunde coal mine (Portugal)

RENATA SANTOS[1,2], LUÍS RIBEIRO[1] & AMÉLIA CARVALHO DILL[1,3]

[1]CVRM – Instituto Superior Técnico, Av. Rovisco Pais, 1049 - 001 Lisbon, Portugal.
[2]Instituto Nacional Engenharia, Tecnologia e Inovação, Rua da Amieira, 4466-956 S. Mamede de Infesta, Portugal.
renata.santos@ineti.pt
[3] Universidade do Algarve / Instituto Nacional Engenharia, Tecnologia e Inovação, Estrada da Portela, Zambugal, Apartado 7586, 2720 ALFRAGIDE, Portugal.

Abstract Mine water pollution is a widespread problem in the whole world. In Portugal, since the early nineties, several actions aiming to study and to characterize the seriousness and the extend of pollution due to abandoned mines were performed. Germunde Colliery is an underground coal mine closed in 1994. Since then, some water-quality monitoring (1998, 2003 and 2004) has been made with the aim of study the evolution of surface and groundwater. The abandoned network comprises both mines that have flooded naturally as well as free-draining mines that have not flooded. Mine waste deposit drainage is also analyzed. Different waters reflecting water's evolution are expected to exist. In order to achieve spatial / temporal patterns of hydrochemistry dataset multivariate statistical techniques were applied. Principal Component Analysis (PCA) and Cluster Analysis (CA) allow us the identification of structural relationships (similarities and/or oppositions) between various hydrochemical descriptors as well geological features and mining infrastructures. The results will emphasize the major pollution risk areas providing valuable information to implement groundwater monitoring network. Finally cluster analysis techniques were applied in order to classify groups of variables with similar characteristics.

Key words: Acid Rock Drainage (ARD); Mine water; Multivariate Analysis;

INTRODUCTION

Drainage from coal mines is one of the most important environmental legacies of industrial economics. Several problems are associated with coal mine drainage such as: sedimentation of chemical precipitates, soil erosion, loss of aquatic habitat, corrosion of metal structures due to contact with acid water (Williams *et al*, 2002) and acid water generated by the oxidation of pyrite reacts with parent rock, resulting in the leaching of many elements (Sullivan & Yelton, 1988).

Multivariate data analysis is used here to classify and characterize structural relationships between variables and samples and to study the contribution of each one to the structure of hydrochemical data (Morell *et al*, 1996). Furthermore these techniques are powerful tools to distinguish the different processes of pollution.

MATERIAL AND METHODS

Study Area

The Germunde coal mine is part of the Douro Carboniferous Basin (BCD - Bacia Carbonífera do Douro). It is located in the NW of Portugal and its composed by a narrow NW-SE strip of continental Carboniferous terrains. It corresponds to an intra-mountain basin of the Centro-Iberian Zone.

The study area is located on the left margin of the Douro River and bounded by the Germunde Mine infrastructures in the NW and the Arda River in the SE (Fig. 1). The SW and NE borders of the coal mine are formed by the Upper Precambrian and/or Cambrian Schist Complex ("Complexo Xisto - Grauváquico") and the Ordovician Formations, respectively.

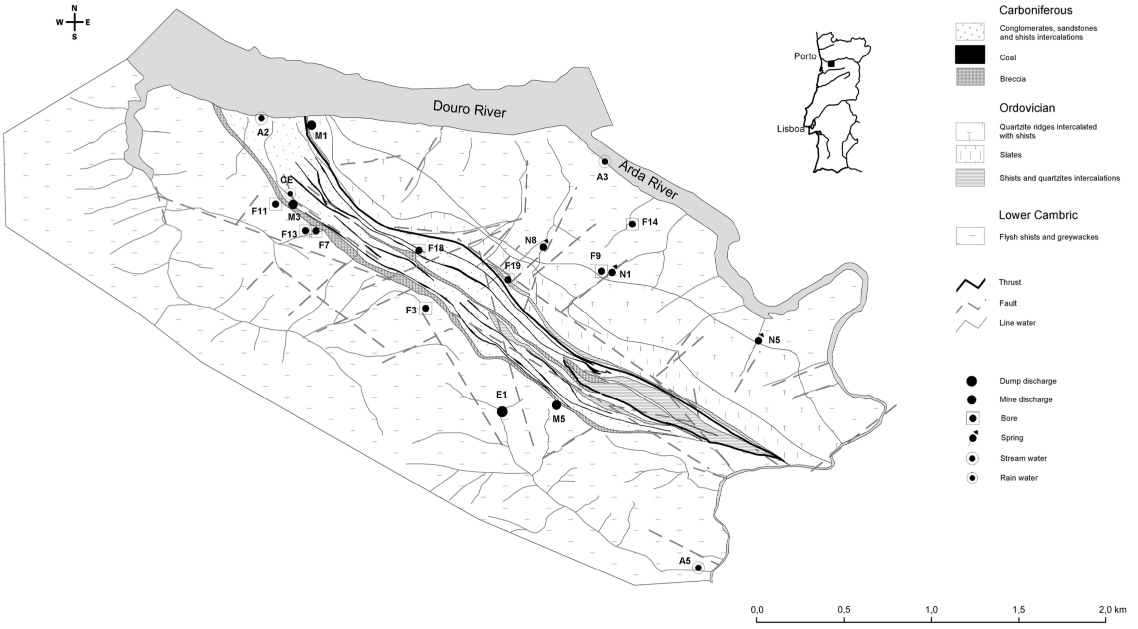

Fig. 1 - Location sampling points overlying a section of “Planta Geológica do Sector Germunde - Arda, E.C.D.” (elaborated by Gaspar, A. M. F. (ECD); Pinto de Jesus, A. (EDM) & Durão, M. (ECD), 1993).

The Carboniferous rocks of this area consist of a basal breccia overlayed by a complex system of coal flakes, intercalated by conglomerates, sandstones and schists. The whole system trends to 140ºE and dips between 60º and 90º towards NE (Pinto de Jesus, 2001). The underlying geological units presented in the area (fig. 1) are (Gaspar *et al*, 1993):

- Quartzite ridges intercalated by Carenigian schists (Ordovician);
- Landeilian Slates belonging to the Valongo Formation (Ordovician);
- Caradocian schists and quartzite intercalations (Ordovician);
- Lower Cambrian Complex – the "Complexo Xisto-Grauváquico" (C.X.G.) – mainly constituted by flysch type schists and greywackes.

Important conjugated NW-SE and NE-SW faults can be observed across the whole basin. The former direction coincides with the axial surface of the Valongo anticline. Minor faults can also be found, their directions varying from NNE-SSW to NNW-SSE, and from ENE-WSW to WNW-ESE.

Vertical discontinuities can also be observed at the surface caused by subsidence (Chaminé & Silva, 1997).

Monitored parameters and analytical methods

The monitoring network was implemented to observe mine waters (shaft, adits and dump discharge) and their impact by acid mine drainage or unpolluted like bores, springs and streams (fig. 1). This sampling strategy was designed in order to cover a wide range of determinants sites, which reasonably represent the water quality in the study area. Under the water-quality monitoring program, six field campaigns (winter / summer) were performed during 1998, 2003 and 2004 in order to proper study the spatial and temporal patterns of both surface and groundwater (Table 1). During summer, some samples were not collected because these had dried up.

Water samples were collected from discharges and the following main field parameters were measured *in situ* - pH, temperature (T), electrical conductivity (EC), reduction-oxidation potential (Eh) and total dissolved salts (TDS). Major, minor ions and trace element had been analysed in laboratory following standard protocols.

Data treatment and multivariate statistical methods

Exploratory data analysis was carried on 6 field water-campaigns. The boxplots of the elements pH, SO_4, Al, Fe, Mn and Ni are shown in fig. 2.

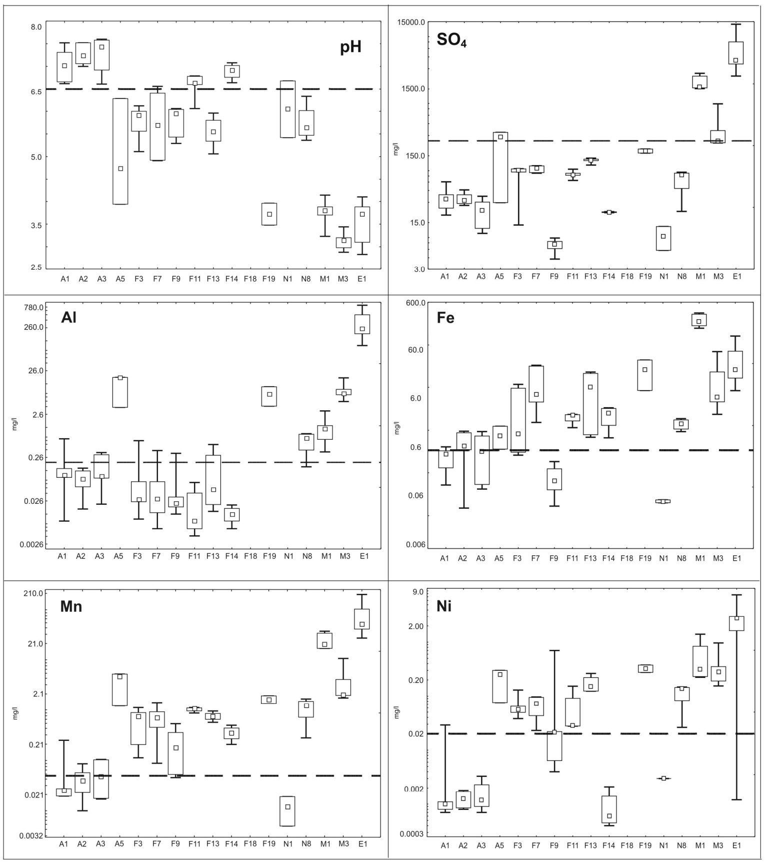

Fig. 2 - Boxplots illustrating distribution of pH, So_4, Al, Fe, Mn, and Ni (y axes in logarithmic scale). Broken line represent the parametric value from Directive 98/83/CE.

In all boxplots it is possible to distinguee 3 kind of groups: superficial waters (streams), groundwaters (springs and bores) and mine /dump discharges. Exceptions inside these major groups (like A5, F19, N8 in Al boxplot) reveal acid drainage. In general, the group of values on each type of station is very close to each median values.

Table 1 – Sampling stations and monitoring plan.

Station	Type	May 98	October 98	April 03	October 03	June 04	October 04
A1	Stream	A	B	D	E	F	G
A2	Stream	A	B	D	E	F	G
A3	Stream	A	B	D	E	F	G
A5	Stream	A	B	D			
F3	Bore	A	B	D	F	F	G
F7	Bore	A	B	D	E	F	G
F9	Bore	A	B	D	E	F	G
F11	Bore			D	E	F	G
F13	Bore			D	E	F	G
F14	Bore			D	E	F	G
F18	Bore			D			
F19	Bore			D		F	
N1	Spring	A	B				
N8	Spring			D	E	F	G
M1	Adit	A	B	D	E	F	G
M3	Adit	A	B	D		F	G
E1	Dump	A	B	D	E	F	G

SO_4 boxplot show clearly the high concentrations of this anion in mine discharges.

Cluster analysis (CA) was used to highlight groups of samples with similar characteristics. Hierarchical methods are the most widely applied clustering techniques in Earth Sciences (Davis, 1986). Various aggregate and distance criteria were used. The best dendogram was obtained with the Euclidean distance.

Principal Component Analysis (PCA) allows us to reduce a set of observed variables into a smaller set of artificial variables called principal components (PC). This technique attempts to reveal the correlation structure of the variables allowing interpretation of geological processes affecting the hydrochemical data.

PC's were extracted using a correlation matrix of 27 variables X 78 water samples. The physical-chemical parameters selected were: pH, EC, alkalinity, hardness, Cl, SO_4, NO_3, F, Na, K, Ca, Mg, SiO_2, HCO_3, Li, Fe, Mn, Zn, Cu, Co, Ni, Ba, Be, Y, Sr, Al and Si. Chemical elements with values very lower or under the detection limit were omitted.

RESULTS AND DISCUSSION

Fig. 3 represents the dendogram obtained by CA. A total of 78 water analyses were grouped into six statistically significant clusters at $(D_{link}/D_{max}) \times 100 < 45$. Each cluster contains a set of samples with similar hydrochemistry facies. Clusters 1, 2 and 3 represent hydrochemical variations of unpolluted waters or relatively low pollution waters by ARD.

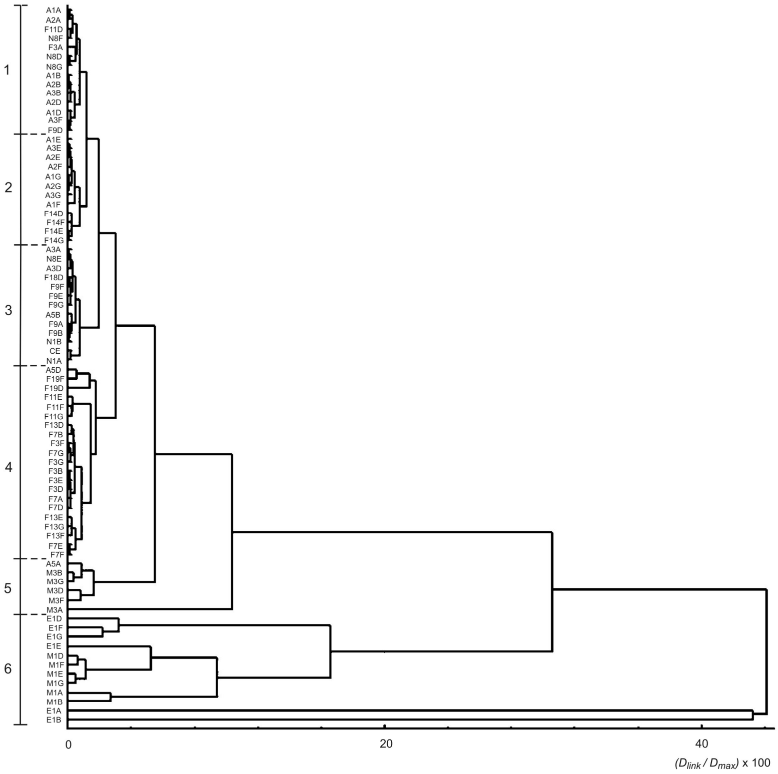

Fig. 3 - Dendogram obtained by the hierarchical cluster analysis using raw data.

Cluster 4, 5 and 6 represent a set of contaminated waters, with concentrations increasing from 4 to 6. Cluster 4 represents bores and stream (A5D) affected by mine flood or drainage. Cluster 5 represent adit M3 which is a free-draining level that is not flooded with seasonal variations and stream A5 is characterized by a higher degree of pollution. Cluster 6 is build with water analysis of mine dump (E1) and by adit M1 which represent the discharge of flooded level. Note that samples collected in the same site (for example F3 or F11) for different campaigns can appear in a different cluster. This reveals that the mixture of waters can occur (by fractures) during the variation of the seasonal level.

In PCA six latent eigenvalues were identified explaining a total of 89, 18% of the total variance of the original matrix dataset (Table 2).

Fig. 4 shows the projection of variables on the 1st factorial plan; build with the 1st (F1) and 2nd (F2) axes.

Table 2 – Eigenvalues and percentage of explained variance for each factorial axis.

	F1	F2	F3	F4	F5	F6
Eigenvalue	13.98	3.43	2.41	1.76	1.37	1.12
% variance	51.78	12.71	8.94	6.53	5.08	4.14
Cumulative % Variance		64.49	73.43	79.96	85.05	89.18

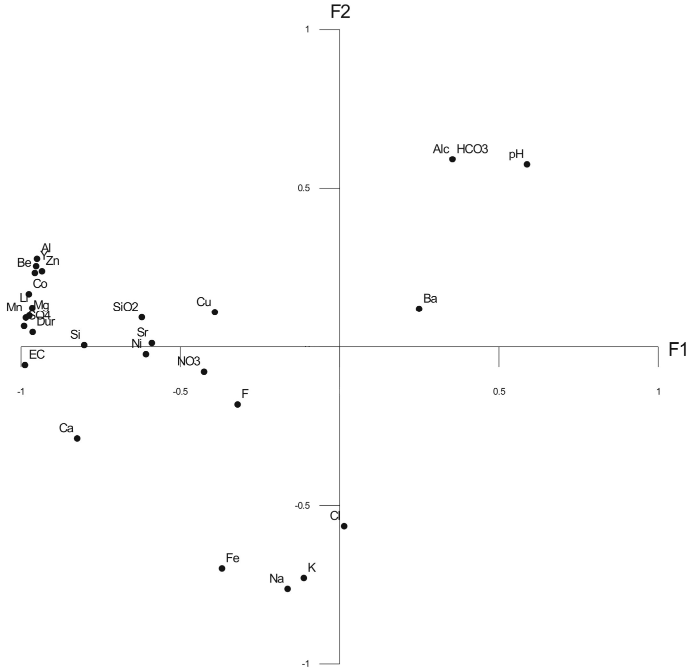

Fig. 4 – Projection of variables on the 1st factorial plan

F1 accounts for 51.78% of the variation of the initial matrix data. This axis shows an opposition between EC, hardness, SO_4, Ca, Mg, SiO_2, Li, Mn, Zn, Co, Ni, Be, Y, Sr, Al, Si (negative side) and pH (positive side). In ARD this negative correlation is typically observed between pH and concentrations of many metals and metalloids, base cations and sulphate (Banks, 2004). This correlation is due to genetic co-variation (generation of protons, sulphate and metals in sulphide weathering reactions) and pH-dependent solubility of many ARD - related metals. The relative position of Fe in this 1st factorial plan is explained by the fact that the relationship between pH and Fe is not straight -forward due to the multiplicity of "sink" reactions which operate when water is very aggressive (Wood *et al*, 1999). F2 explains 12.71% of total variance. Variables alkalinity (Alc) and HCO_3 are located in the positive side in opposition to Cl, Na, K and Fe, projected in negative part of the axis, which can be attributed to water mineralization. The remaining axes (not presented in figure) can be explained by the relative projection of Sr and Cu, (3rd axis), SiO_2 (4th axis), F (5th axis) which in spite of this halogenic origin, its chemical behaviour is different from the other group elements, and NO_3 (6thaxis) a specie that can be attribute to agricultural origin.

Finally fig. 5 shows the projection of samples on the 1st factorial plan. We can distinguish here 3 groups: dump discharges (1), mine discharges (2) and a third group formed by bores, springs, streams and rain water (3).

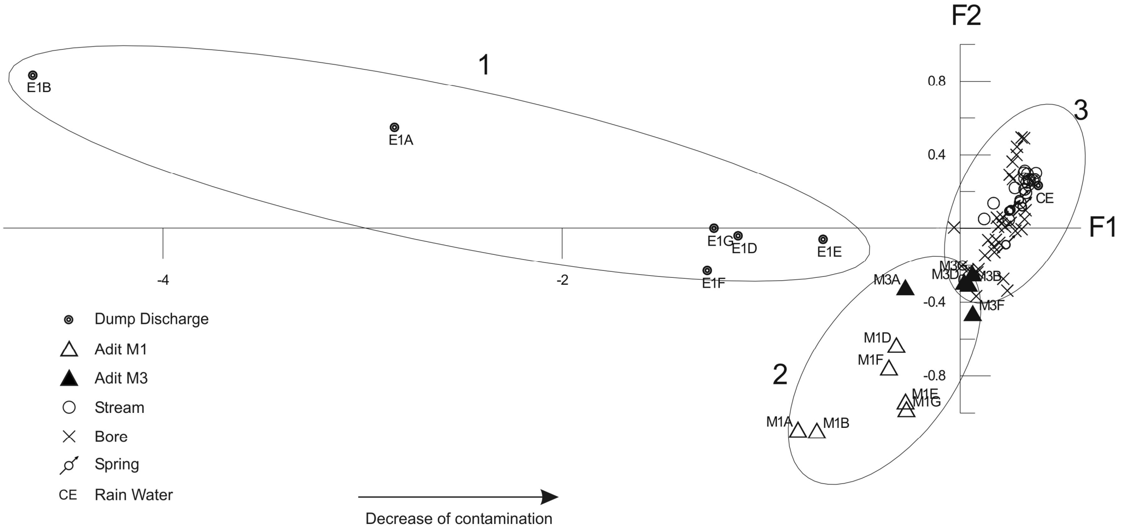

Fig. 5 – Distribution of different water samples on the 1st factorial plan.

As we can see, the contaminated sites are in the negative side of F1, according to interpretation of the 1st factorial plan (see Fig. 4). fig. 5 reveals a decrease of contamination in dump discharge and mine adits groups since 1998. For the other hand bores group don't have the same hydrochemical variations which are associated with similar variations of flow, recharge and piezometer level in the terrain.

FINAL CONSIDERATIONS

Multivariate statistical techniques are a good approach to the identification of hydrochemical variations in a dataset. They allow us to distinguish unpolluted waters from contaminated waters by ARD. They also allow us to distinguish the chemical elements associated with ARD.

For a more complete interpretation of these structural relationships the use of traditional hydrochemical methods should be encouraged.

REFERENCES

Banks, D. (2004) Geochemical processes controlling minewater pollution. In: *Proceedings of the* 2^{nd} *IMAGE-TRAIN Advanced Study Course "Groundwater management in mining areas"*, 17 – 44. Umweltbundesamt / Austrian Federal Environment Agency, *Conference Papers CP – 035*, Wien.

Chaminé, H. I. & Silva, P. B. (1997) A Subsidência Mineira na Mina de Carvão de Germunde – Contribuição da Geologia de Superfície. *GEOlogos, Porto University,* Porto, 1 - 7.

Davis, J.C. (1986) Statistics and Data Analysis in Geology. John Wiley, New York.

EDC (1995) Investigação sobre trabalhos superficiais e subterrâneos para minimizar os efeitos da subsidência de minas de carvão com camadas sub-verticais. Relatório final. Comissão das Comunidades Europeias. *Direcção Geral Energia – D. G. XVII. Investigação Tecnológica Carvão*. Lisboa, 229 p.

Gaspar, A. M. F., Pinto de Jesus, A. & Durão, M. (1993) Planta Geológica do Sector Germunde - Arda (Escala 1: 2500). Empresa Carbonífera do Douro, SA. Unpublish Map.

Gaspar, A. M. F., Silva, P. B. & Chaminé, H. I. (1994) Mining subsidence and its relation with geological features in the Germunde coal mine (NW of Portugal) In: *7th International IAEG Congress* (ed. by Oliveira, R. *et al*), 4441 – 4444. Balkema, Rotterdam.

Pinto de Jesus, A. (2001) Génese e Evolução da Bacia Carbonífera do Douro (Estefaniano C inferior, NW de Portugal): Um Modelo. PHD Thesis. Science Faculty, Porto University, Portugal.