Journal of Hydroinformatics | 11.2 | 2009

Model of behaviour of conductivity versus pH in acid mine drainage water, based on fuzzy logic and data mining techniques

A. Jiménez, J. Aroba, M. L. de la Torre, J. M. Andujar and J. A. Grande

ABSTRACT

Acid Mine Drainage is a water pollution type characterized by several topics such as high acidity, sulfate and heavy metal concentrations. One of the chemical characteristics is the absence of correlation between pH and conductivity, as could be expected. This last parameter is well correlated with other variables such as sulfate concentration and can be used as a field assessment. The absence of pH/conductivity correlation is largely discussed by several authors. In this work, the use of fuzzy logic algorithms in a large temporal database (over 20,000 records) has allowed us to study the "hidden" relation between them. This work finds this correlation, with some conditions such as the range of pH where it happens. Maybe the study of the usual range of pH values in previous studies has disturbed the correlation because of other chemical processes.

Key words | AMD, conductivity, data mining, fuzzy logic, Iberian Pyrite Belt, pH

A. Jiménez (corresponding author)
M. L. de la Torre
J. A. Grande
Grupo de Geologia Costera y Recursos Hídricos,
Universidad de Huelva,
E-21819 Palos de la Frontera,
Huelva,
Spain
Tel.: +34 959 217343
Fax: +34 959 217304
E-mail: *antonio.jimenez@dimme.uhu.es*

J. Aroba
J. M. Andujar
Grupo de Control y Robótica,
Universidad de Huelva,
E-21819 Palos de la Frontera,
Huelva,
Spain

INTRODUCTION

Acid Rock Drainage (ARD) is a natural process whereby sulfuric acid is produced when sulfides in rocks are exposed to air and water. Acid Mine Drainage (AMD) is essentially the same process, greatly magnified. When large quantities of rock containing sulfide minerals are excavated from an open pit or opened up in an underground mine, they react with water and oxygen to create sulfuric acid. When the water reaches a certain level of acidity, a naturally occurring type of bacteria called *Thiobacillus ferroxidans* may kick in, accelerating the oxidation and acidification processes, leaching even more trace metals from the wastes (EMCBC 1997).

The problems associated with Acid Mine Drainage are contaminated drinking water, disrupted growth and reproduction of aquatic plants and animals, and corroding effects of acid on parts of infrastructure such as bridges. It is not only an ecological concern for the states but an economic concern as well (USEPA 2006). Acid generation can last for decades, centuries or longer, and its impacts can travel many miles downstream (EMCBC 1998).

doi: 10.2166/hydro.2009.015

Both sulfate and conductivity are useful indicators of acid mine drainage contamination. Unlike pH, they are both extremely sensitive to AMD even where large dilutions have occurred. The advantage of using sulfate to trace AMD is that conductivity is especially sensitive to sulfate ions. Therefore, as sulfate analysis is difficult in the field, conductivity can be used to predict sulfate concentration in both AMD and contaminated surface waters using regression analysis. The most accurate predictions are achieved by using equations given for specific conductivity ranges or AMD sources. There is also potential in conductivity to be used to predict approximate concentrations of key metals when the water pH is within their respective solubility ranges (Gray 1996).

Supported by this work (Gray 1996), Heikkinen *et al.* (2002) argued that the pH of AMD contaminated water has only weak negative correlation with heavy metals.

Grande *et al.* (2005*b*) establish a comparison, using classical statistical techniques, between the geochemical behaviour of two streams, one of them contaminated by AMD (Chorrito stream) and another uncontaminated one (Higuereta stream), both located in the Iberian Pyrite Belt (Huelva, Spain). In this paper, the authors prove that the pH in the AMD uncontaminated stream has a high correlation with conductivity and sulfates, unlike the case of an AMD contaminated stream, where there is an absence of correlation between conductivity and pH, the former being closely linked to the sulfate content. Later, Grande *et al.* (submitted), using a continuous measurement multi-parameter probe placed in the Chorrito stream channel, with which they obtain a data mass of over 21,000 records, prove that in this case, despite being the same AMD contaminated stream, there is a high negative correlation between pH and conductivity. This change is explained by the water sampling, in the first case, taking place at time intervals longer than the rainfall/pH/conductivity response.

In this paper we intend to go further and, by applying fuzzy logic and data mining techniques to the data obtained from the multiparameter probe, to obtain a simulation that allows the creation of a qualitative model of behaviour describing the relationships between pH and conductivity in an AMD contaminated channel.

LOCATION AND DESCRIPTION OF THE STUDY AREA

The fluvial network affected by AMD processes in the Iberian Pyrite Belt is a unique study framework for a better knowledge of the phenomenon. Since this network is affected by more than 100 abandoned sulfide mines, this gives rise to extreme environments where water keeps a pH around 2 and sulfate concentrations are higher than 20 g/l. This problem has been largely described by different authors (Borrego 1992; Elbaz-Poulichet *et al.* 1999, 2000, 2001; Sáinz 1999; Davis *et al.* 2000; Grande *et al.* 2000, 2003*a*,*b*, 2005*a*,*b*, 2009; Leblanc *et al.* 2000; Borrego *et al.* 2002; Sáinz *et al.* 2002, 2003*a*,*b*, 2004; Fernández-Remolar *et al.* 2003, 2005; González-Toril *et al.* 2003; Aroba *et al.* 2007).

The most important feature of the Iberian Pyrite Belt, of a metalogenic nature, is the existence of Paleozoic massive sulfide deposits, described in several works such as the one by Sáez *et al.* (1999). These massive deposits contain pyrite, spheralite, galena and chalcopyrite in their paragenesis, and have been exploited for over 2000 years.

One of the Iberian Pyrite belt mining operations that determines the chemical makeup of the Chorrito stream water is the Herrerías Mine (Figure 1). According to Pinedo (1963), 1,820,000 tonnes of sulfides had been obtained from it up until 1960. The mine was under exploitation until February 1989, when environmental laws in Spain were almost non-existent. As a result, large amounts of all kinds of waste were stored in the surroundings of the operations without preventive measures against water pollution. This caused the development of AMD processes in the Chorrito stream and, subsequently, in the Cobica River.

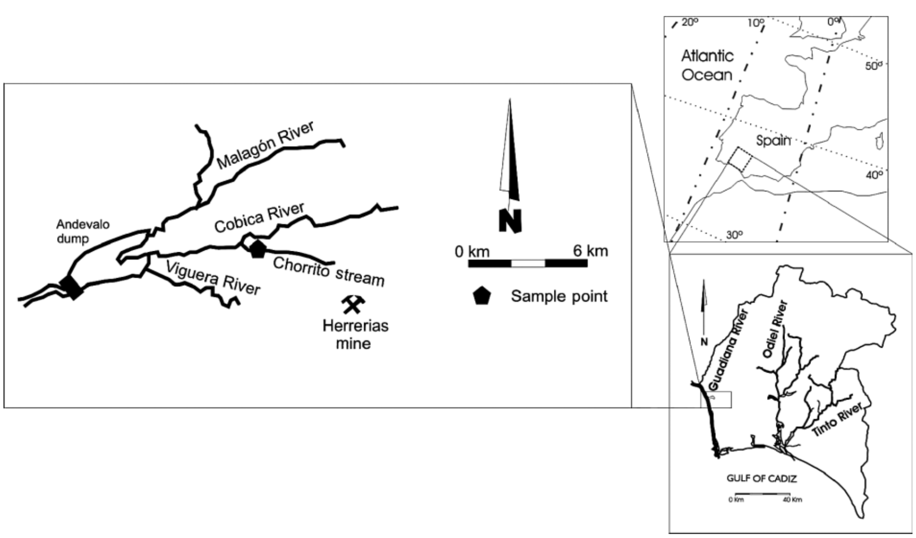

Figure 1 | Location map.

Water pollution from mining, which arrived in the Cobica watercourse coming from the Chorrito stream, is produced by mine underground water, waste dump leachate and general washout of all pyrite waste dispersed through the numerous mining operations (Grande et al. 2005a).

The channel regime in this environment is directly associated with rainfall in the presence of impermeable outcropping rocks in the basin, which gives it practically torrential characteristics, with large floods in winter and almost no inflow during low water (Grande et al. 2003b).

MATERIALS AND METHODS

A continuous measurement OTT Hydrometrie minisonde 4b multiparameter probe, calibrated to measure pH and conductivity every 15 min with a precision of 0.2 units for pH and a resolution of 0.01 was used. Precision for conductivity was 1% of the measured value and resolution was 0.001 mS/cm.

For its location, a platform built in the Chorrito channel itself was used. This platform collects the waters coming from the Herrerías mine with the aim of measuring the instantaneous flow of the Chorrito channel. At the centre of the platform a 0.25 m long and square section channel was opened, where the probe was installed (Figure 2).

The probe was maintained on a weekly basis (cleaning of sensors, calibration and battery replacement) and data were dumped into a laptop computer.

The reading period was from 3 October 2003 to 12 May 2004 (during the time water was running along the stream). This makes a total of 223 d in which readings were conducted every 15 min, therefore resulting in a total of 21 408 data on pH and conductivity. This large data mass was later treated with fuzzy logic and data mining techniques by means of the tool PreFuRGe. The computer tool developed, PreFuRGe (Aroba 2003; Aroba et al. 2007), is based on the previously described methodology (Sugeno & Yasukawa 1993). This initial methodology has been adapted and improved in the following aspects:

1. It allows working with quantitative databases, with n input and m output parameters.
2. The different variables that are the object of study can be weighted by assigning them weights for the calculation of distances between points of the space being partitioned.
3. The achieved fuzzy clusters are processed by another algorithm to obtain graphic rules trapeziums (Figures 3 and 4).
4. An algorithm processes and solves cases of multiple projections in the input space (mounds).
5. The output provided in the original method has been improved with a graphic interface showing graphics of the achieved rules.

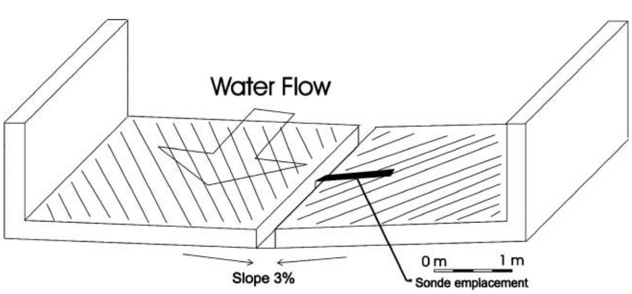

Figure 2 | Multiparametric probe emplacement in the Chorrito Stream.

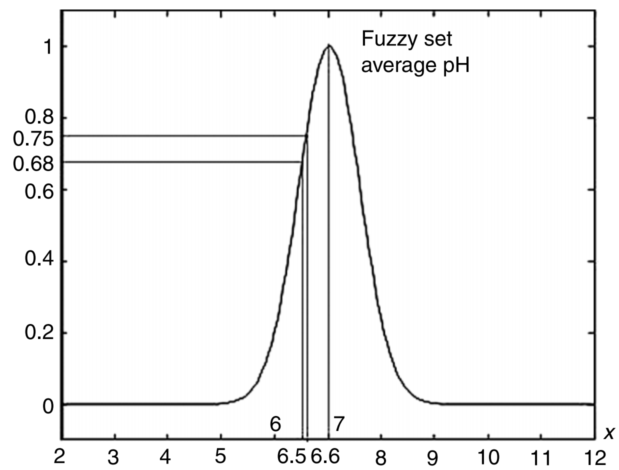

Figure 3 | Example of the membership function average pH.

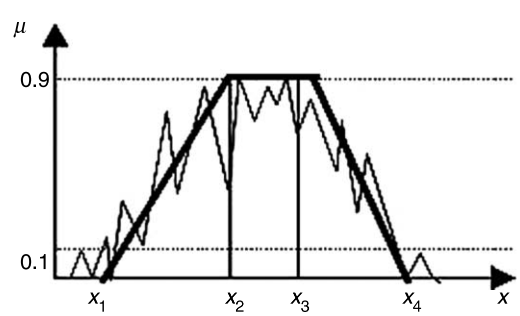

Figure 4 | Trapezium approximation of a fuzzy set.

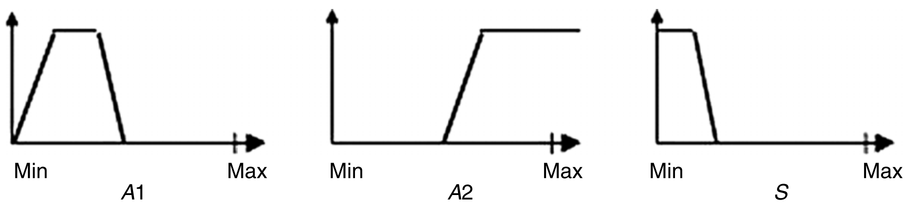

Figure 5 | Example of fuzzy rules generated with PreFuRGe.

6. An algorithm automatically provides the interpretation of the fuzzy graphic rules in natural language.

Figures 5 and 6 show two examples of rules generated by means of the tool PreFuRGe.

In the rule of Figure 5, the fuzzy set assigned to each parameter is represented by a polyhedron. The parameter values are represented on the x axis of each fuzzy set and the value of membership to a cluster on the y axis. This fuzzy rule would be interpreted as follows:

IF $A1$ is *small* and $A2$ is *bigger* or *equal* to *average* THEN S is *very small*.

When applying the fuzzy clustering algorithm (Aroba *et al.* 2001) to the generated databases, it is possible to obtain multiple projections in the input parameters (mountain). In the fuzzy rule of Figure 6, a multiple projection (mountain) is represented in the input parameter $A1$. In this case we observe how the parameter $A1$ can take different types of values for a certain kind of output. This fuzzy rule can be interpreted as follows:

IF $A1$ is *small* or *big* and $A2$ is *average* THEN S is *very small*

Recently, one of the authors of this paper has investigated the stability of fuzzy logic control systems, as well as their advanced industrial applications (Andújar *et al.* 2004, 2006; Andújar & Barragán 2005; Andújar & Bravo 2005).

RESULTS AND DISCUSSION

In order to model the behaviour of conductivity versus pH, 21,408 records were analysed by applying the PreFurGe tool (Aroba 2003). This tool allowed the creation of Figures 7–9, which show a set of fuzzy rules in a graphic form, where the consequents are the different pH values obtained in the samples and the antecedents are the corresponding fuzzy clusters for conductivity.

Figure 7 represents pH variation from a minimum value of 0.98 up to a maximum value of 4.44, as well as the evolution of conductivity as compared to these variations from a value of 0 up to a value of 11 mS/cm, in six different fuzzy rules. Rules R1–R6 globally show the evident negative correspondence between pH and conductivity evolution, i.e. as pH increases, conductivity gradually decreases. However, with this tool it can also be observed that, occasionally, the lowest conductivity values are produced for the highest pH values (rule R6). In contrast, R1, which represents the lowest pH values, matches the highest conductivity values.

A more detailed analysis of Figure 7 will enable us to verify that, although rules R1 and R6 have a clear correlation, this is not the case for rules R2 and R3, where a large dispersion of conductivity values can be seen. As a result, to conduct a more thorough analysis of this stretch of

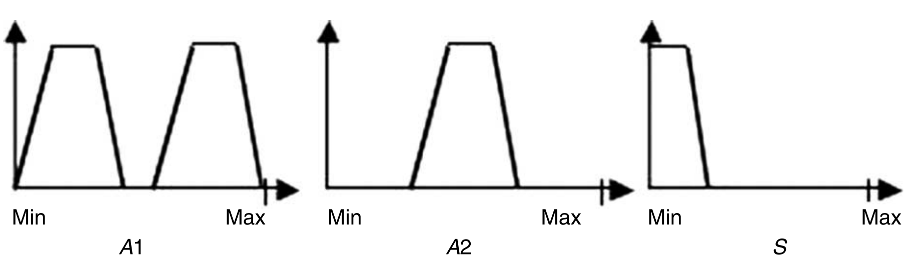

Figure 6 | Example of fuzzy rules generated with PreFuRGe.

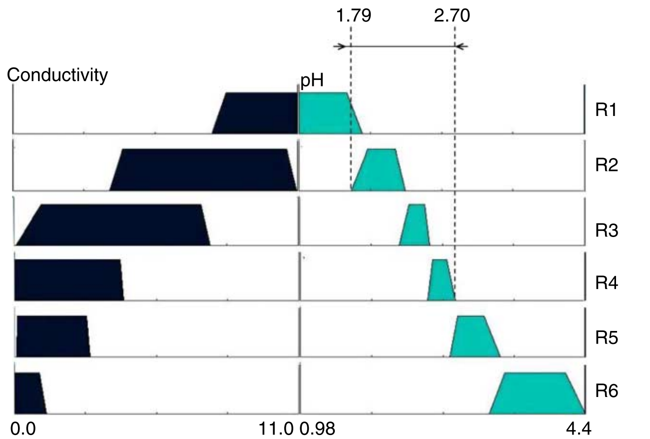

Figure 7 | Graphic fuzzy rules for conductivity response as regards pH for 21,408 records (pH between 0.98 and 4.44).

pH values, a filter was applied to the database, focusing on the pH values that involve rules R2, R3 and R4 (when pH values are between 1.79 and 2.70), which is the interval at which the passing of minimum to maximum conductivity takes place. This filter has allowed us to know that, out of 21,408 records, 17,028 (practically 80%) are between these values. The results are shown in Figure 8, where the minimum value (pH = 1.79) is given by the minimum value of rule R2 in Figure 7, and the maximum value in Figure 8 (pH = 2.70) is given by the maximum value of rule R4 in Figure 7.

In Figure 8, it is reflected how, by the time pH is approximately between 1.79 and 2.15, conductivity may have from an average to a maximum value (rule R1 in Figure 8). However, when pH has a set of values clustered around 2.60 (rule R5 in Figure 8), conductivity clearly changes from average to very low, and stays there in a range of practically similar values in spite of pH going up to 2.70 in rule R7. Also, an evident qualitative change can be observed between rules R1 and R2. In the former (R1), the centre of gravity of the 'cluster' is found on average to maximum conductivities, and in the latter (R2), the centre of gravity is on average to low conductivities. As a result, to analyse in more detail this change in behaviour, we re-filtered the database to analyse the data involving rules R1 and R2 in Figure 8. More specifically, we analysed the measurements corresponding to pH values between 1.79 and 2.47, which gave us the analysis of 2,787 records (13% of the total data), which resulted in Figure 9.

This last Figure (Figure 9) shows how from rules R1–R3 conductivity decreases as compared with pH, basically as compared with the minimum values, to show a drastic change of behaviour in the gap from rule R3 to R4. Thus, as pH increases from values around 2.20, conductivity clearly changes its behaviour, being able to reach the lowest values. That is, it is clearly observed that, when pH is around the values of 2.15 to 2.40 (Figure 9), conductivity undergoes an evident change in trend.

This can be explained because practically all metals within this pH range of values are in dissolution and conductivity is closely linked with sulfate content, as stated by Grande *et al.* (2005*a*,*b*). These authors refer to the work by Singh & Rawat (1985) and argue that metal precipitation as hydroxides takes place at higher pH, especially in the case of Fe, which has the highest concentrations in these

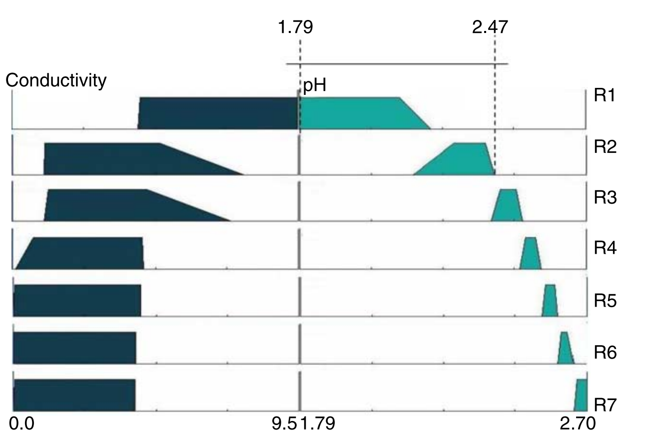

Figure 8 | Graphic fuzzy rules for conductivity response as regards pH for 17,028 records (pH between 1.79 and 2.70).

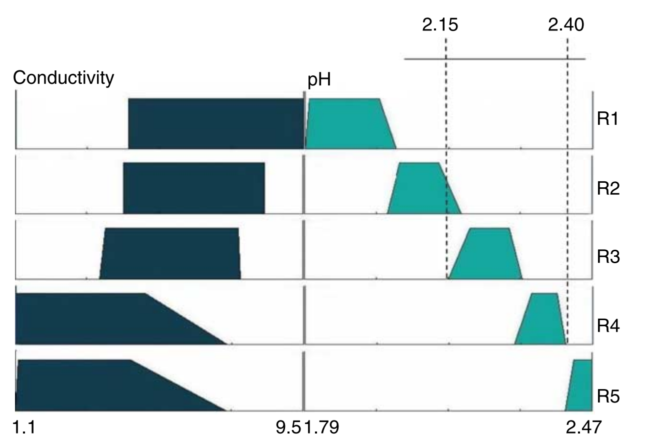

Figure 9 | Graphic fuzzy rules for conductivity response as regards pH for 2,787 records (pH between 1.79 and 2.47).

waters and, as a result, its dissolution or precipitation affects sulfate content and, therefore, conductivity more directly.

CONCLUSIONS

By using the PreFurGe tool, it could be checked that conductivity and pH are inversely related in an AMD contaminated medium (as pH increases, conductivity gradually decreases), as already reported by Grande *et al.* (submitted) for a large data mass, contrary to what happens when correlations are established between both parameters in water samples taken at time intervals longer than the rainfall/pH/conductivity response.

On the other hand, by using this technical tool based on fuzzy logic and data mining in the pure investigation, we could also determine that within the pH interval between 1.79 and 2.70 this correlation is missing, with conductivity covering a much wider range of values. It was determined in more detail (see Figure 9) that the change in conductivity behaviour in relation to pH occurs within pH values from 1.79–2.47.

ACKNOWLEDGEMENTS

The present study is a contribution of the followings projects funded by the Spanish Ministry of Science and Technology: CICYT REN2002-01897/HID, CICYT TIN2007-67843-C06-03, CICYT DPI2005-01065 and CICYT DPI2007-62336

REFERENCES

Andújar, J. M. & Barragán, A. J. 2005 A methodology to design nonlinear fuzzy control systems. *Fuzzy Sets Syst.* **154** (2), 157–181.

Andújar, J. M., Barragán, A. J., Córdoba, J. M. & Fernández, I. 2006 Diseño de sistemas de control borroso: modelado de la planta. *Revista Iberoamericana de Automática e Informática Industrial* **3** (1), 75–81.

Andújar, J. M. & Bravo, J. M. 2005 Multivariable fuzzy control applied to the physical-chemical treatment facility of a cellulose factory. *Fuzzy Sets Syst.* **150** (3), 475–492.

Andújar, J. M., Bravo, J. M. & Peregrín, A. 2004 Stability analysis and synthesis of multivariable fuzzy systems using interval arithmetic. *Fuzzy Sets Syst.* **148** (3), 337–353.

Aroba, J. 2003 *Avances en la toma de decisiones en proyectos de desarrollo de software* PhD Thesis, University of Sevilla.

Aroba, J., Grande, J. A., Andújar, J. M., de la Torre, M. L. & Riquelme, J. C. 2007 Application of fuzzy logic and data mining techniques as tools for qualitative interpretation of acid mine drainage processes. *Environ. Geol.* **53**, 135–145.

Aroba, J., Ramos, I. & Riquelme, J. C. 2001 Application of machine learning techniques to software project management. In: *ICEIS 2001 (III International Conference on Enterprise Information Systems), Setubal (Portugal)*. Congress Acts, Portugal, p 433.

Borrego, J. 1992 *Sedimentología del estuario del río Odiel, Huelva, S.O. España*. PhD Thesis, University of Sevilla.

Borrego, J., Morales, J. A., de la Torre, M. L. & Grande, J. A. 2002 Geochemical characteristics of heavy metal pollution in surface sediments of the Tinto and Odiel river estuary (southwestern Spain). *Environ. Geol.* **41**, 785–796.

Davis, R. A., Jr, Welty, A. T., Borrego, J., Morales, J. A., Pendón, J. G. & Ryan, J. G. 2000 Rio Tinto estuary (Spain): 5,000 years of pollution. *Environ. Geol.* **39**, 1107–1116.

Elbaz-Poulichet, F., Braungardt, C., Achterberg, E., Morley, N., Cossa, D., Beckers, J., Nomérange, P., Cruzado, A. & Leblanc, M. 2001 Metal biogeochemistry in the Tinto-Odiel rivers (Southern Spain) and in the Gulf of Cadiz: a synthesis of the results of TOROS project. *Cont. Shelf Res.* **21** (18–19), 1961–1973.

Elbaz-Poulichet, F., Dupuy, C., Cruzado, A., Velasquez, Z., Achterberg, E. & Braungardt, C. 2000 Influence of sorption processes by iron oxides and algae fixation on arsenic and phosphate cycle in an acidic estuary (Tinto river, Spain). *Water Res.* **34** (12–15), 3222–3230.

Elbaz-Poulichet, F., Morley, N. H., Cruzado, A., Velasquez, Z., Achterberg, E. P. & Braungardt, C. B. 1999 Trace metal and nutrient distribution in an extremely low pH (2.5) river-estuarine system, the Ria of Huelva (south-west Spain). *Sci. Total Environ.* **227**, 73–83.

EMCBC 1997 *Acid Mine Drainage. Mining & Water Pollution Signs in BC*. Environmental Mining Council of BC, Vancouver, Canada.

EMCBC 1998 *EMCBC Mining and the Environment Primer: Acid Mine Drainage*. Available at: http://www.miningwatch.ca/index.php?/251/AMD_primer

Fernández-Remolar, D. C., Morris, R. V., Gruener, J. E., Amils, R. & Knoll, A. H. 2005 The Río Tinto Basin, Spain: mineralogy, sedimentary geobiology, and implications for interpretation of outcrop rocks at Meridiani Planum, Mars. *Earth Planet Sci. Lett.* **240**, 149–167.

Fernández-Remolar, D. C., Rodriguez, N., Gómez, F. & Amils, R. 2003 Geological record of an acidic environment driven by iron hydrochemistry: the Tinto River system. *J. Geophys. Res.* **108** (7), 16–19.

González-Toril, E., Gómez, F., Rodríguez, N., Fernández-Remolar, D., Zuluaga, J., Marín, I. & Amils, R. 2003 Geomicrobiology of the Tinto River, a model of interest for biohydrometallurgy. *Hydrometallurgy* **71**, 301–309.

Grande, J. A., Andújar, J. M., Aroba, J., de la Torre, M. L. & Beltrán, R. 2005 Precipitation, pH and metal load in AMD river basins: an application of fuzzy clustering algorithms to the process characterization. *J. Environ. Monit.* **7**, 325–334.

Grande, J. A., Beltrán, R., Sáinz, A., Santos, J. C., de la Torre, M. L. & Borrego, J. 2005 Acid mine drainage and acid rock drainage processes in the environment of Herrerías Mine (Iberian Pyrite Belt, Huelva-Spain) and impact on the Andevalo Dam. *Environ. Geol.* **47**, 185–196.

Grande, J. A., Borrego, J., de la Torre, M. L. & Sáinz, A. 2003 Geochemistry of the estuary waters in the Tinto and Odiel rivers (Huelva, Spain). *Environ. Geochem. Health* **25**, 233–246.

Grande, J. A., Borrego, J. & Morales, J. A. 2000 Study of heavy metal pollution in the Tinto-Odiel estuary in Southwestern Spain using spatial factor analysis. *Environ. Geol.* **39** (10), 1095–1101.

Grande, J. A., Borrego, J., Morales, J. A. & de la Torre, M. L. 2003 A description of how metal pollution occurs in the Tinto-Odiel rias (Huelva-Spain) through the application of cluster analysis. *Mar. Pollut. Bull.* **46** (4), 475–480.

Grande, J. A., Jiménez, A., Borrego, J. & de la Torre, M. L. 2009 Relationships between conductivity and pH in channels exposed to a acid mine drainage processes. Functioning models based on large data masses. *Environ. Geol.* (submitted).

Gray, N. F. 1996 Field assessment of acid mine drainage contamination in surface and ground water. *Environ. Geol.* **27** (4), 358–361.

Heikkinen, P. M., Korkka, K., Lahti, M. & Salonen, V. P. 2002 Groundwater and surface water contamination in the area of the Hitura nickel mine, Western Finland. *Environ. Geol.* **42**, 313–329.

Leblanc, M., Morales, J. M., Borrego, J. & Elbaz-Poulichet, F. 2000 4500 year-old mining pollution in Southwestern Spain: long-term implications for modern mining pollution. *Econ. Geol.* **95**, 655–662.

Pinedo, I. 1963 *Piritas de Huelva*. Summa, Madrid.

Sáez, R., Pascual, E., Toscano, M. & Almodóvar, G. R. 1999 The Iberian type of volcano-sedimentary massive sulphide deposits. *Miner. Deposita* **34**, 549–570.

Sáinz, A. 1999 *Estudio de la contaminación química de origen minero en el río Odiel*. PhD Thesis, University of Huelva.

Sáinz, A., Grande, J. A. & de la Torre, M. L. 2004 Characterisation of heavy metal discharge into the ria of Huelva. *Environ. Int.* **30**, 557–566.

Sáinz, A., Grande, J. A. & de la Torre, M. L. 2003 Analysis of the impact of local corrective measures on the input of contaminants from the Odiel river to the ria of Huelva (Spain). *Water Air Soil Pollut.* **144**, 375–389.

Sáinz, A., Grande, J. A. & de la Torre, M. L. 2003 Odiel river, acid mine drainage and current characterisation by means of univariate analysis. *Environ. Int.* **29**, 51–59.

Sáinz, A., Grande, J. A., de la Torre, M. L. & Sánchez-Rodas, D. 2002 Characterisation of sequential leachate discharges of mining waste rock dumps in the Tinto and Odiel rivers. *J. Environ. Manage.* **64** (4), 345–353.

Singh, G. & Rawat, N. S. 1985 Removal of trace elements from acid mine drainage. *Int. J. Mine Water IMWA* **2**, 817–827.

Sugeno, M. & Yasukawa, A. 1993 A fuzzy-logic based approach to qualitative modeling. *IEEE Trans. Fuzzy Syst.* **1**, 7–31.

USEPA 2006 *What is Acid Mine Drainage*. US EPA Region 3 Water Protection Division Nonpoint Source Pollution Program. Available at: http://www.epa.gov/Region3/acidification/what_is_amd.htm

First received 13 February 2008; accepted in revised form 16 September 2008.