

# Thermodynamics and log–contrast analysis in fluid geochemistry

Antonella Buccianti, Barbara Nisi and Orlando Vaselli

Dipartimento di Scienze della Terra, Università degli Studi di Firenze  
Firenze, Italy.

**Mailing address:**

A. Buccianti

Dip. Scienze della Terra, Università degli Studi di Firenze

Via G. La Pira 4, 50125, Firenze, Italy

e-mail:antonella.buccianti@unifi.it

fax: +39-055-284571

**Keywords:** water chemistry, compositional data analysis, Nitrogen, log–contrast analysis, geochemical modeling

## Abstract

There are two principal chemical concepts that are important for studying the natural environment. The first one is thermodynamics, which describes whether a system is at equilibrium or can spontaneously change by chemical reactions. The second main concept is how fast chemical reactions (kinetics or rate of chemical change) take place whenever they start. In this work we examine a natural system in which both thermodynamics and kinetic factors are important in determining the abundance of  $NH_4^+$ ,  $NO_2^-$  and  $NO_3^-$  in superficial waters. Samples were collected in the Arno Basin (Tuscany, Italy), a system in which natural and anthropic effects both contribute to highly modify the chemical composition of water. Thermodynamical modelling based on the reduction-oxidation reactions involving the passage  $NH_4^+ \rightarrow NO_2^- \rightarrow NO_3^-$  in equilibrium conditions has allowed to determine the  $E_h$  redox potential values able to characterise the state of each sample and, consequently, of the fluid environment from which it was drawn. Just as  $pH$  expresses the concentration of  $H^+$  in solution, redox potential is used to express the tendency of an environment to receive or supply electrons. In this context, oxic environments, as those of river systems, are said to have a high redox potential because  $O_2$  is available as an electron acceptor.

Principles of thermodynamics and chemical kinetics allow to obtain a model that often does not completely describe the reality of natural systems. Chemical reactions may indeed fail to achieve equilibrium because the products escape from the site of the reaction or because reactions involving the transformation are very slow, so that non-equilibrium conditions exist for long periods. Moreover, reaction rates can be sensitive to poorly understood catalytic effects or to surface effects, while variables as concentration (a large number of chemical species can coexist and interact concurrently), temperature and pressure can have large gradients in natural systems. By taking into account this, data of 91 water samples have been modelled by using statistical methodologies for compositional

data. The application of log–contrast analysis has allowed to obtain statistical parameters to be correlated with the calculated  $E_h$  values. In this way, natural conditions in which chemical equilibrium is hypothesised, as well as underlying fast reactions, are compared with those described by a stochastic approach.

## Introduction

Continental freshwaters are very important to humans since they are the only reliable source of drinking water. The chemical composition of rivers, lakes and groundwaters varies widely and is governed predominantly by three factors: element chemistry, weathering regimes and biological processes (Berner and Berner, 1996). In addition, human perturbations may have a major effect (Skinner et al., 1997; Bouwman, 1998; van Breemen, 2002; Okafor and Ogbonna, 2003). In river systems Nitrogen chemistry is complex because the element is present in several oxidation states, of which  $N(0)$  nitrogen gas ( $N_2$ ), ( $N^{3-}$ ) ammonium ( $NH_4^+$ ) and ( $N^{5+}$ ) nitrate ( $NO_3^-$ ) are the most important. The nitrogen cycle is dominated by reactions involving biological material. All the reactions in the series  $N_2(\text{dinitrogen}) \rightarrow NH_3(\text{ammonia}) \rightarrow NO_2^-(\text{nitrite}) \rightarrow NO_3^-(\text{nitrate}) \rightarrow \text{aminoacids} \rightarrow \text{proteins}$ , and the reverse reactions back to  $N_2$ , can be carried out by microorganisms (Krug and Winstanley, 2002).

Nitrogen gas dissolved in river water cannot be utilised as a nitrogen source by most plants and algae because they cannot break its strong triple bond. Specialised nitrogen-fixing bacteria and fungi do exist to exploit  $N_2$ , but it is not an energetically efficient way of obtaining nitrogen. Hence, these microorganisms are only abundant when  $N_2$  is the only available nitrogen source. Nevertheless, along with the fixation of  $N_2$  by lightning, nitrogen-fixing microorganisms provide the major natural source of nitrogen for rivers. Biological processes use nitrogen in the 3–oxidation state, particularly as amine groups in

proteins. This is the preferred oxidation state for algal uptake and also the form in which nitrogen is released during organic matter decomposition, largely as  $NH_4^+$ . Once released into soil or water,  $NH_4^+$  being cationic, may be adsorbed onto negatively charged organic and inorganic coatings such as soil particles and clay mineral surfaces. Ammonium is also taken up by plants or algae, or oxidised to  $NO_3^-$ , a process that is usually catalysed by bacteria. In contrast to  $NH_4^+$ ,  $NO_3^-$  is anionic and not retained in soils. Therefore,  $NO_3^-$  from rainwater or fertilisers or derived from the oxidation of soil organic matter and animal wastes will wash out of soils and into rivers. Apart from biological uptake, denitrification in low-oxygen environments is the most important way for which nitrate is removed from soils, rivers and groundwaters. Seasonal variations of  $NO_3^-$  concentrations in rivers from temperate climate are caused by fluctuation in supply of  $NO_3^-$  from soil, depending on the intensity and quantity of rainfall. Increases in both the area and the intensity of agricultural activity are probably responsible for the increased  $NO_3^-$  concentrations in several rivers of Europe and North America (Codispoti and Christensen, 1985; Berner and Berner, 1996). There are entire ecosystems ranging from forests to coastal waters that are now being overwhelmed by nitrogen compounds and the attention on this problem by the European Countries is increasing (Ekholm et al., 2000; Wade et al., 2002). Excess nitrate ion in wastewater flowing into seawater (for example the Baltic Sea) has resulted in algal blooms that pollute the water once they die. Excess nitrate ion in drinking water is a potential health hazard since it can result in methemoglobinemia in newborn infants as well as in adults, with a particular enzyme deficiency. Recently, an increase in the risk of acquiring non-Hodgkin's lymphoma has been found for persons consuming drinking water having high levels (long-term average of 4 ppm nitrogen as nitrate) of nitrate in drinking water. However, since epidemiological investigations to establish statistically significant positive relationship between nitrate levels in drinking water and health problems are difficult, the expenditure of public money on nitrate level reductions has become a controversial subject (Baird, 2001). However, mapping of the

areas in which this type of problems is present, *i.e.* river water and ground-waters, may be important to manage the use of the land and its resources. This paper aims to give a contribution in the investigation of the abundances of  $NH_4^+$ ,  $NO_2^-$  and  $NO_3^-$  in river waters. A methodological approach based on thermodynamical and statistical bases is proposed to obtain parameters able to monitor natural systems in time and space. Data were collected in the Arno Basin (Tuscany, central–northern Italy), an area in which natural weathering phenomena and anthropical input characterise the chemical composition of the rivers. Due to the compositional nature of the collected data, their investigation has been performed in the light of the recent developments of the compositional data analysis theory (Aitchison, 1986a, b).

## Geological background

The Arno River Basin, with a drainage surface of about  $8,228 \text{ km}^2$  is located within the mountain belt of the Northern Apennines (Tuscany, central-northern Italy). It originates at about 1,650 m from the SW flank of Mt. Falterona and flows for 242 km into the Tyrrhenian Sea after crossing highly urbanised, industrialised and cultivated areas. Florence, the biggest city in Tuscany, together with Arezzo, Pisa, Pistoia and Prato are the main cities contributing to major pollution. The catchment can be divided into six sub-basins from East to West: Casentino (CA), Chiana Valley (CH), Sieve (SI), Upper Valdarno (UV), Middle Valdarno (MV) and Lower Valdarno (LV) (Figure 1). The Arno Basin mainly consists of Oligocene-Miocene arenaceous-calcareous-marly flysch, structurally complex clay and Neogene marine and fluvio-lacustrine deposits (Abbate and others, 1982; Moretti, 1994). Mesozoic and Paleozoic clastic and carbonaceous rocks of the Tuscan Series (Triassic and evaporitic rocks) are exposed in the upper part of the Elsa and Era sub-basins, while Paleozoic quartzites widely occur in the Zambra sub-basin. Anthropogenic contam-

inants, such as nitrates and nitrites, in the shallow aquifers of the alluvial plain of the Middle Valdarno were associated with the infiltration of sewage effluents near the surface by Bencini and others (1993). Moreover, Bencini and others (1995) suggested that nitrate contamination largely affect the plain with nitrate concentrations as high as 159 mg/L, probably due to domestic sewage disposal of cesspools, septic tanks and leaking sewers. Conversely, a recent investigation on the Chiana Valley surface waters (Ramaldi and others, 2004) found concentrations of  $NH_4^+$  and  $NO_2^-$  up to 18 (mean value: 2.2 mg/l) and 5.3 (mean value: 1.2 mg/l) mg/l, respectively, whereas  $NO_3^-$  concentrations are generally <10 mg/l

Starting from 2002 a wide monitoring campaign was developed with the aim to characterize the nature of the main factors influencing the chemistry of the superficial waters collected in the Arno Basin. Beside the concentrations of the major anions and cations, the abundance of  $NH_4^+$ ,  $NO_2^-$  and  $NO_3^-$ , by using spectrophotometric methodologies, was also determined for more than 450 samples. Compositional changes in the sub-composition of the nitrogen species have been modeled by both thermodynamical and statistical approach, choosing a representative group of 91 samples (Figure 1).

## Methodological approach and results

The branch of physical chemistry known as thermodynamics is focused on the study of the transformations of energy and, in particular the transformation of energy from heat into work and vice versa. Thermodynamics allows to describe natural systems in terms of certain measurable properties of matter, such as volume, pressure, temperature and chemical composition of the constituent components. Geochemical modelling can be based on the fundamental laws of thermodynamics and kinetics to answer questions at the basis of everyday chemistry, such as why reactions reach equilibrium, what their composition is

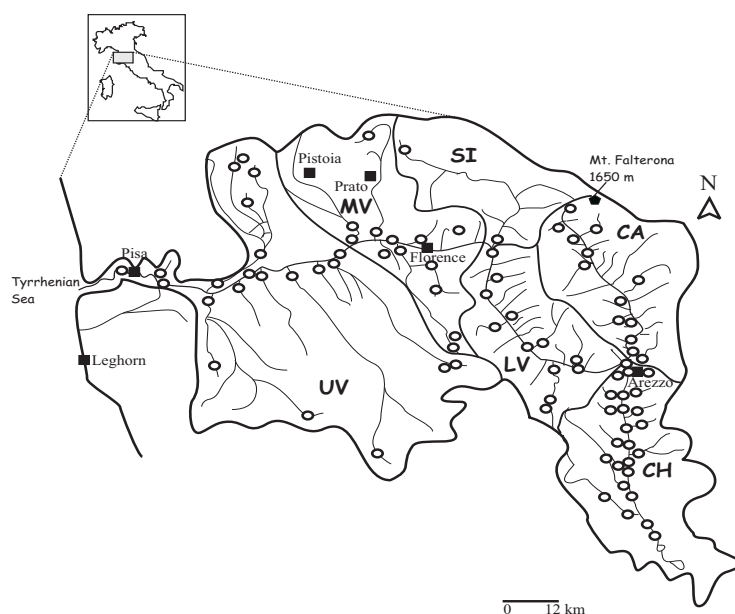


Figure 1: Map of the Arno basin with the location of the sampling sites (Tuscany, central-northern Italy): CA = Casentino, CH = Chiana Valley, SI = Sieve, UV = Upper Valdarno, MV = Middle Valdarno, LV = Lower Valdarno.

at equilibrium and so on. As mineral assemblages and fluids in nature are complicated, equilibrium thermodynamics is essential to Earth scientists and is usually the starting point for any geochemical investigation (Anderson and Crerar, 1993; Anderson, 1996). This point was established in the nineteenth century by Jacobus Van't Hoff (1852–1911) who used the Gibbs phase rule to reconstruct the composition of ancient seawater from the mineral assemblages found in evaporite deposits. Stated in terms of the primary thermodynamic functions of temperature, energy and entropy, the laws of thermodynamics can succinctly be stated for any actual process as follows: first law, the total energy is conserved (i.e. remains unchanged); second law, the total entropy increases; third law, the absolute (Kelvin) temperature remains above zero degrees. When used appropriately, these laws provide the necessary and sufficient conditions for geochemists to 1) establish the equilibrium in a system that includes chemical reactions, heat transfer and mechanical

strain; 2) discharge as impossible some processes; 3) convert thermodynamic data such as the energy and entropy changes into an equilibrium constant; 4) predict the effects of changes in temperature, pressure or composition on the reaction equilibrium (Kraynov, 1997).

The most widely used geochemical modeling programs consist of a computer code plus a related file of data called *database*. The latter contains thermodynamic and kinetic parameters to be used by the code together with concentrations or other constrains as input with the aim to produce results that describe a geochemical model for a particular chemical system (Lichtner, 1996). Databases associated with geochemical modeling programs consist of a list of the basic species, and a list of secondary or auxiliary species, minerals and gases, each described in terms of the basis species, and with the equilibrium constant of the reaction linking the secondary species or mineral to the basis species. Basically the algorithms solve a set of mathematical equations describing chemical equilibria and chemical mass balance (Westall et al., 1976; Nordstrom et al., 1979; Bethke, 1996). The quality of the information reported in the data base is crucial to obtain good results and a discussion of data accuracy is reported in Nordstrom and Munoz (1986) and Criscenti et al. (1996). A summary of the most reliable codes is presented by Mangold and Tsang (1991) and Paschke and van der Heijde (1996). We have used EQ3/6 code, Version 7.2c, supported by the EQLIB library (Wolery, 1992), in order to obtain the values of  $E_h$ , the *emf* (electromotive force) of the oxidation-reduction reactions characterising the chemical composition of each water sample. In natural systems  $E_h$  is a parameter whose values reflect the ability of the environment to be an electron donor or acceptor relative to the standard  $H_2$  electrode as reference. Ecosystems in contact with the atmosphere (rain-water, streams, oceans and mine waters) have positive  $E_h$  values and work as electron acceptors (oxidizing agents). On the other hand, surroundings that are isolated from the atmosphere (waterlogged soils, euxenic marine basins, organic-rich brines) have negative  $E_h$  values, and work as electron donors (reducing agents). The probability plot of the



$E_h$  values for the 91 water samples from the Arno Basin is reported in Figure 2. By comparing the continuous line related to the normal model, the probability distribution can be considered as Gaussian (Kolmogorov-Smirnov test,  $\alpha = 0.05$ ). This means that data have been drawn from the same population and that the mean, equal to 0.6567, and the standard deviation, equal to 0.0366, are good estimators of its parameters. The  $E_h$  values reflect our samples from a thermodynamical point of view and consequently the oxidizing conditions of environments in contact with the atmosphere. The probability distribution say us that conditions more oxidant that the mean, as well as more reductent, are present with the same probability and that  $E_h$  values are given by the sum of numerous independent events.

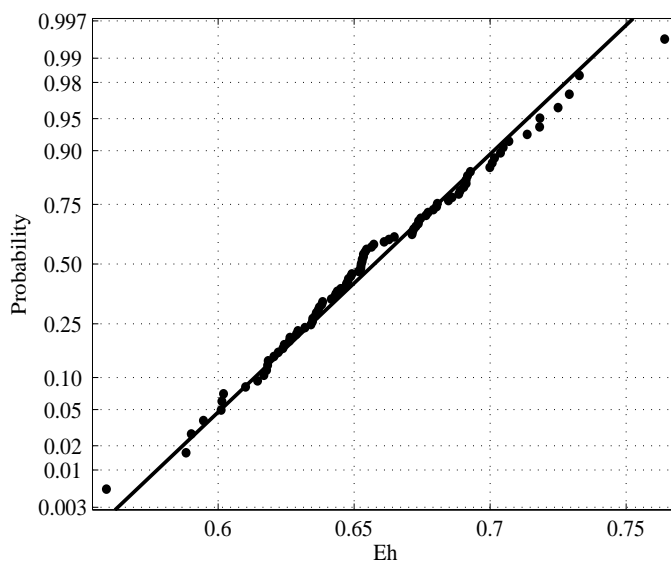


Figure 2: Normal probability plot of  $E_h$  values determined by EQ3/6 code for the Arno Basin running waters

In order to correlate this parameter with another obtained by statistical investigation, a log-contrast analysis for compositional data has been performed (Aitchison, 1983). Compositional data, consisting of vectors of proportions, are difficult to be handled statisti-

cally because of the awkward constraint that the components of each vector must sum to unity. Such data sets frequently display marked curvature so that linear techniques such as standard principal component analysis are likely to prove inadequate. Aitchison (1981; 1982) has proposed a resolution of the constrain difficulty in the analysis of compositional data. In practical terms this consists of first transforming each compositional vector  $x^{(d)} = (x_1, \dots, x_d)$  in  $S^{d-1}$  (the simplex space) to a vector  $y^{(d-1)} = (y_1, \dots, y_{d-1})$  in  $R^{d-1}$ . The log ratio transformation  $y^{(d-1)} = \ln(x_{-j}/x_j)$  is then applied, where  $x_{-j}$  denotes the vector  $x^{(d)}$  with  $x_j$  omitted. Transformed vectors are finally analysed by standard multivariate methods available in  $R^{d-1}$ . The nonlinearity of the logarithmic function opens up the possibility to capture the curvature in data sets and its approximate linearity over parts of its range also makes feasible the modelling of linear data sets. However, in this case a major difficult remains since different choices of the divisor  $x_j$  lead to a different principal components. We can solve this problem by working with a more symmetric approach if the denominator is given by the geometric mean of the  $d$  components of the composition  $g(x) = (x_1, \dots, x_d)^{1/d}$ . The covariance matrix that is investigated in this case is given by:

$$\Omega = cov\{\ln(x^{(d)}/g(x))\}. \quad (1)$$

The matrix  $\Omega$  is positive-semidefinite, its one zero eigenvalue having an associated eigenvector  $u_d$  consisting of  $d$  units. The other  $d - 1$  eigenvalues  $\lambda_1, \dots, \lambda_{d-1}$ , labelled in descending order of magnitude, are positive and the corresponding eigenvectors  $a_1, \dots, a_{d-1}$ , being orthogonal to  $u_d$ , yield log linear contrasts of the proportions. They are expressed as:

$$\sum_{i=1}^d a_i \ln x_i = \sum_{i=1}^d a_i \ln(x_i/g(x)), \quad (2)$$

where  $a_1, \dots, a_d = 0$ . The first log constraint, explaining about the 74.8% of the data variability, is given by

$$-0.39 \ln(NH_4^+) - 0.43 \ln(NO_2^-) + 0.81 \ln(NO_3^-) = \kappa_1. \quad (3)$$

If the equation is simplified and written in exponential form we obtain

$$\frac{(NO_3^-)}{(NH_4^+)^{0.5} \cdot (NO_2^-)^{0.5}} = \exp(\kappa_1), \quad (4)$$

leading to the relationship  $NH_4^+ : NO_2^- : NO_3^- = 1 : 1 : 2$ . The  $\kappa_1$  variable follows the Gaussian curve (Kolmogorov-Smirnov test,  $\alpha = 0.05$ ). Its normal probability plot is reported in Figure 3. The second log constraint, explaining about the 25.2% of the data variability is given by:

$$0.72 \ln(NH_4^+) - 0.69 \ln(NO_2^-) - 0.02 \ln(NO_3^-) = \kappa_2, \quad (5)$$

or simplifying:

$$\frac{(NH_4^+)}{(NO_2^-)} \sim \exp(\kappa_2), \quad (6)$$

due to the low value of  $NO_3^-$  coefficient. In this case the relationship  $NH_4^+ : NO_2^- = 1 : 1$  is obtained. The  $\kappa_2$  variable is again describable with the Gaussian model (Kolmogorov-Smirnov test,  $\alpha = 0.05$ ) and its probability plot is reported in Figure 4.

The two log constraints describing relationships among the variables of the sub composition  $NH_4^+ - NO_2^- - NO_3^-$  can be reported on a ternary diagram as in Figure 5. Another way to see the reciprocal weight of the components of the composition is the variation diagram as those reported in Figure 6. The compositional lines represent the reciprocal relationships among the members of the composition as described by the first log contrast. The linear

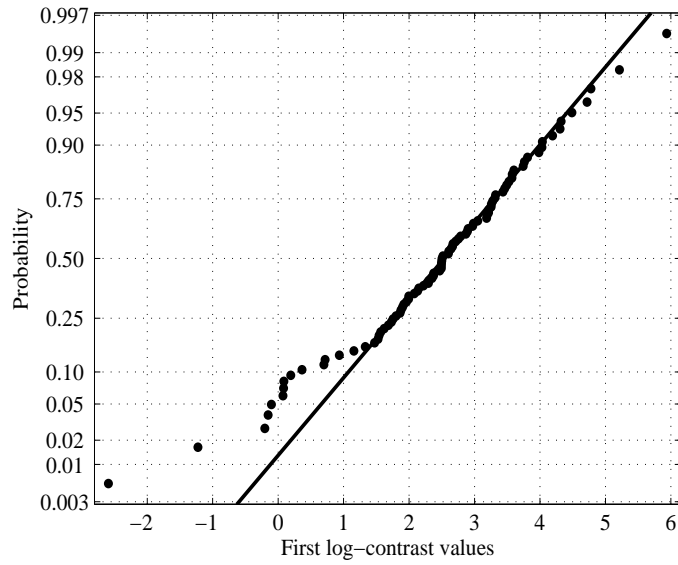


Figure 3: Normal probability plot of the first log-contrast values

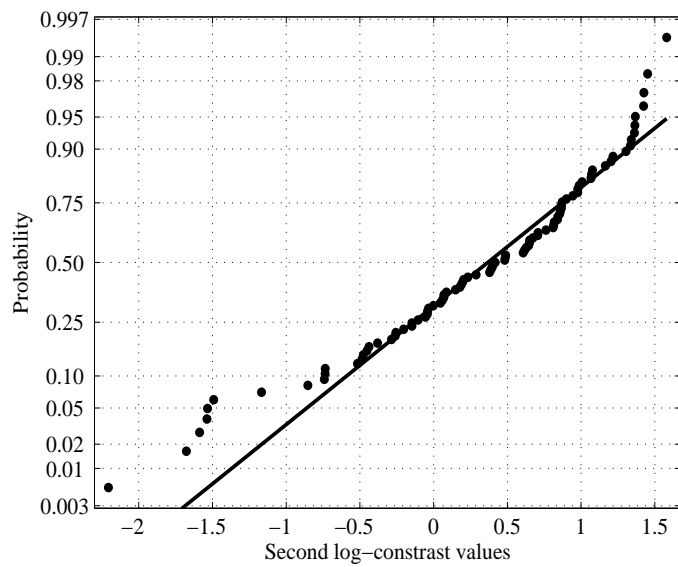


Figure 4: Normal probability plot of the second log-contrast values

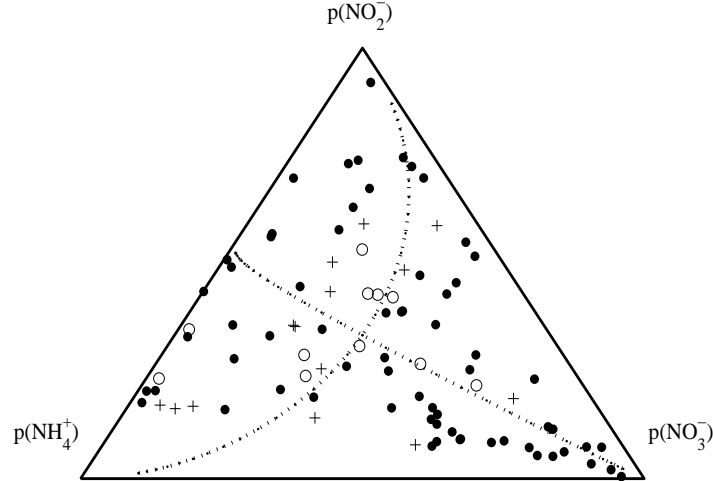


Figure 5: Ternary diagrams of the system  $NH_4^+ - NO_2^- - NO_3^-$  (+ = Chiana, o = Canale Maestro, • = Arno valley); data have been perturbed by the vector  $\mathbf{p} = [35.9161.672.41]$  to obtain a better visualisation (the original barycentre of the data moved toward the barycentre of the triangle).

patterns of Figures 5 and 6 are obtained by using perturbation and power transformation operations, both able to give to the simplex a  $(d - 1)$  dimensional vector space structure (Barceló-Vidal et al., 2001; Pawlowsky-Glahn and Egozcue, 2002). They are given by  $\alpha \odot \mathbf{p}$  where  $\mathbf{p}$  is the unitary perturbation vector obtained from log contrast analysis,  $\alpha$  is a scalar value that ranges from  $-3$  to  $+3$  and  $\odot$  indicates the power transformation operation. The  $\alpha$  values needed to post observations on the plot can be obtained from the log contrast scores or by projecting each observation onto the linear trend given by the log contrast (Pawlowsky-Glahn and Egozcue, 2001). The behaviour of  $NO_3^-$  is well represented by the model for all the  $\alpha$  values while those of  $NH_4^+$  and  $NO_2^-$  show a wide scattering of the real data around the models, particularly far from the barycentre (0 value), towards negative values. The  $\kappa_1$  values have been correlated with the  $E_h$  values as obtained by the thermodynamical modelling. The results of the linear fitting are reported

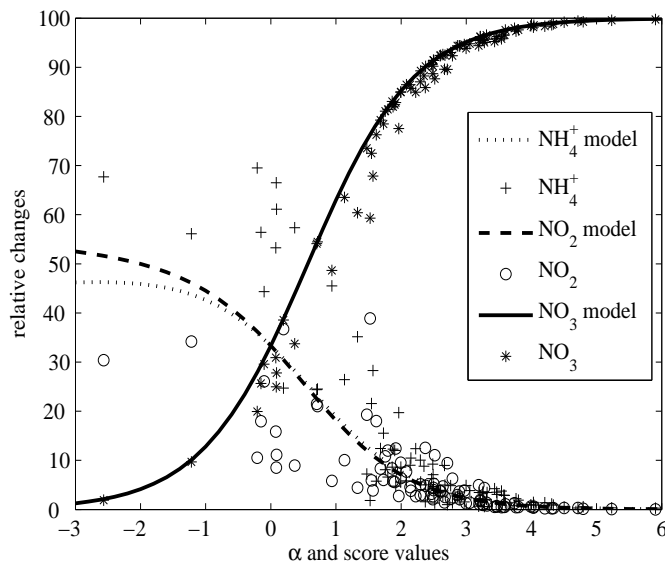


Figure 6: Relative variation diagram of the system  $NH_4^+-NO_2^- - NO_3^-$  as modelled by  $\alpha \odot \mathbf{p}$ ;  $\alpha$  is a scalar value that ranges from  $-3$  to  $+3$  while  $\mathbf{p}$  is the unitary perturbation vector obtained from log contrast analysis.

in Figure 7. In Figure 8 the bar plot of the residulas is also reported. Notwithstanding the scarce goodness of the linear fitting (correlation between the variables is 0.44), the line can be considered as a reference to characterise data more and more far from it. The bar plot of Figure 8 help us to visualise the difference between the  $E_h$  calculated values and those obtained by the model. If only the  $n = 59$  samples inside the area between  $-0.03 \div +0.03$  (limits corresponding to the first and third quartile) are considered, the correlation among the two variables move from 0.44 to 0.70. These samples are characterised by electrical conductivity ranging in the interval  $0.12 \div 6.6$  mS/cm (median 0.84 mS/cm) and nitrogen species in the intervals  $0.026 \div 6.32$  mg/l for  $NH_4^+$  (median 0.22),  $0.01 \div 7.28$  mg/l for  $NO_2^-$  (median 0.20),  $0.10 \div 32.5$  mg/l for  $NO_3^-$  (median 32.5). All the samples out from the area are characterised for the lower values of conductivity ( $0.22 \div 1.86$  mS/cm, median 0.73 ms/cm) as well as of  $NH_4^+$  ( $0.040 \div 4.39$  mg/l, median  $9 \times 10^{-2}$ ),  $NO_2^-$  ( $0.01 \div 1.97$

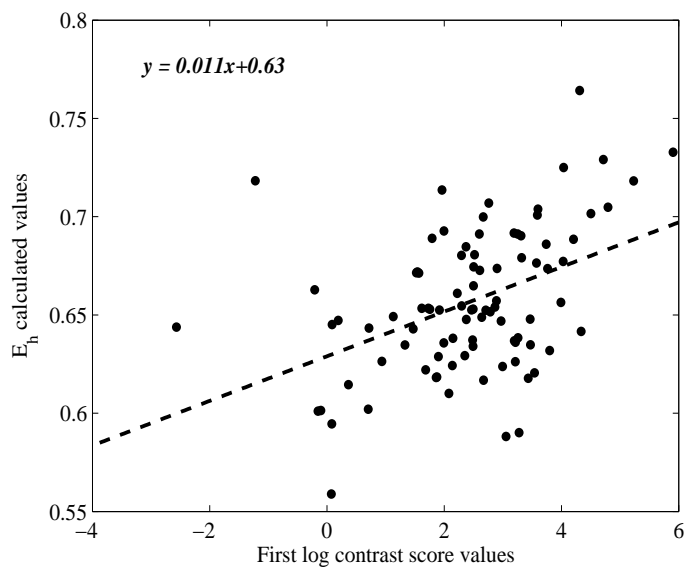


Figure 7: Linear relationship between the first log contrast values and the  $E_h$  calculated data.

mg/l, median  $6.39 \times 10^{-2}$ ) and  $NO_3^-$  ( $0.12 \div 30.5$  mg/l, median 2.33). In Figure 9 data have been mapped by considering the discrimination in two groups, according to the results of Figure 8. Full black circles are related to samples pertaining to difference values between the calculated and the modelled  $E_h$  in the range  $-0.03 \div +0.03$ . For them an equilibrium condition is hypothesised and deduced by the first log contrast  $k_1$  values. On the other hand, open circles are associated to samples characterised by high negative or positive values of the difference. On the whole these samples are related to situations characterised by the lowest values of nitrogen species and conductivity of the investigated area.

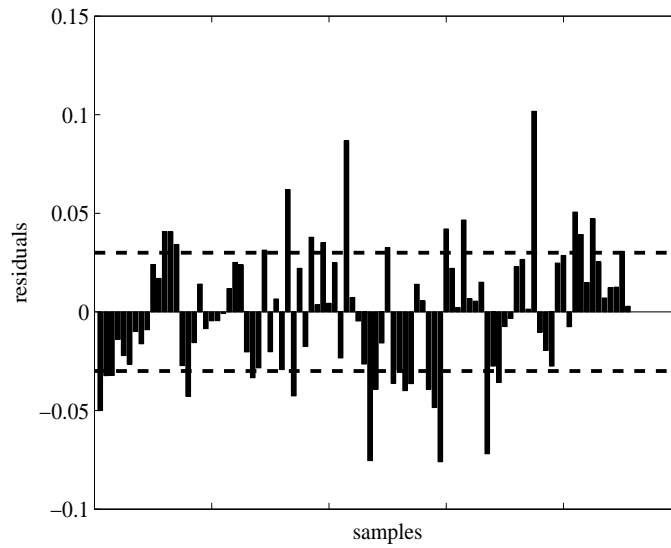


Figure 8: Bar plot of the residuals of the linear fitting.

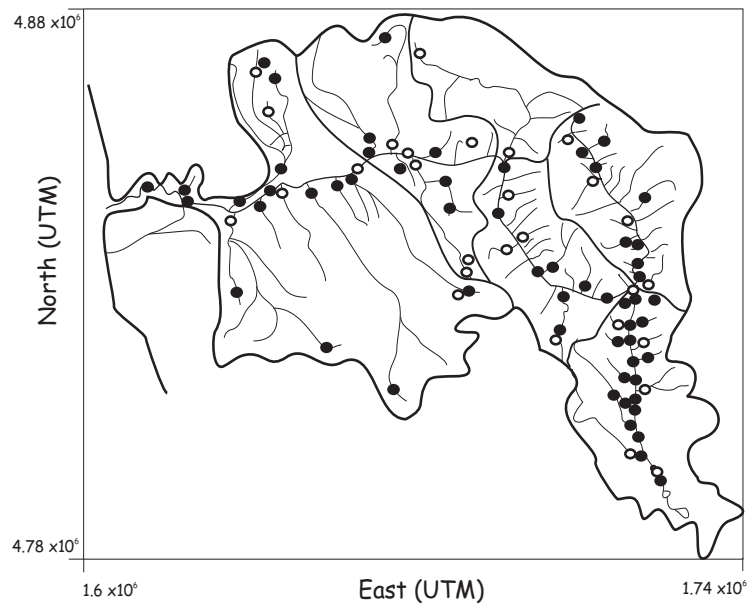


Figure 9: Spatial variation of the difference values between the calculated and the modelled  $E_h$ . Full black circles are related to samples pertaining to difference values in the range  $-0.03 \div +0.03$ ; open circles are associated to samples characterised by high negative or positive values of the difference.



## Discussion

The redox potential of a natural system is determined by the physico-chemical features of the system itself. Geochemists use  $E_h - pH$  diagrams to predict the oxidation state of various constituents in natural environments.  $E_h$  determines what modes of microbial activity are possible in a given environment. Few oxic environments have redox potential of less than +0.6 V and a progressive decrease in redox potential occurs when soils are flooded. Redox potential drops as heterotrophic respiration of organic carbon depletes the soil of  $O_2$  and in most cases, organic matter contributes a large amount of reducing power that lowers the redox potential (Bartlett, 1986). The results of many studies suggest that a particular sequence of reactions is expected, as progressively lower redox potentials are achieved (Achnich et al., 1995; Lovley, 1995; Peters and Conrad, 1996). In our case the  $E_h$  calculated values (EQ3/6 code) ranges in the interval  $0.559 \div 0.764$  V with the mean representative value of 0.6567 V, and from a thermodynamical point of view describe chemical equilibrium conditions. The fact that  $E_h$  values are modelled by a probability density function as the Gaussian reveals a sort of *statistical equilibrium*. Values higher or lower with respect to the mean are expected with the same probability and represent the fluctuations due to the sum of several independent factors affecting the oxidation–reduction conditions of the investigate ecosystem. In order to link thermodynamical and statistical modelling,  $E_h$  values have been correlated with the  $\kappa_1$  scores of the first log-contrast (74.8% of the data variability). On the whole they represent the relationship  $NH_4^+ : NO_2^- : NO_3^- = 1 : 1 : 2$ , quantitatively shifted towards  $NO_3^-$  abundance, the final species of the oxidation processes. The linear relationship between  $E_h$  and  $\kappa_1$  is not so good due to the wide scattering of the data. Apparently this result may be attributable to the behavior of  $\kappa_1$  (Figure 3) showing some anomalous values in the tails of the frequency distribution. Thus, it is evident that the variability described by log contrast analysis it is not equivalently revealed by thermodynamical modelling. The

distance value of each sample from the best fitted linear model in the plot  $\kappa_1 - E_h$  has been used to point out cases where correlation is sustainable. Cases for which the distance values are in the range  $-0.03 \div +0.03$  show a correlation between  $\kappa_1$  and  $E_h$  equal to 0.70. In this case, the relationships among the variables are not quantitatively disturbed for absence (scarse abundance) of one of the components. All the observations pertain to the right part of the variation diagram of Figure 6, thus indicating that the system is fastly moving towards  $NO_3^-$ .

Cases for which the linear relationship between  $\kappa_1$  and  $E_h$  is not a validate model represent situations of discordance between thermodynamical and statistical results. In this case observations are characterised by low values of  $NH_4^+$ ,  $NO_2^-$ ,  $NO_3^-$ , as well as of conductivity, all indications of low recent and ancient pollution levels. Data pertaining to the two groups ( $\kappa_1$  and  $E_h$  correlated or uncorrelated) have been mapped and the results indicate us that different conditions (chemical equilibrium described or not described by log contrast analysis) can closely be associated in space. The system  $NH_4^+ - NO_2^- - NO_3^-$  is thus dominated by the presence of punctual phenomena with low spatial correlation overlapped to a general situation of chemical equilibrium. This circumstance can lead, in general, to serious problems in the mapping of interpolated values. A way to solve this question may be the plotting of data, for which thermodynamical equilibrium is also correspondent to the statistical description of the system, separately. Equilibrium data can reasonably represent a continuous phenomenon in space and interpolation may have sense.

## Conclusions

Most urban and agricultural-zootechnical areas are affected by the problem of nitrate pollution in both surface and ground waters. Despite the increasing efforts to reduce  $NO_3^-$

inputs from intensive agriculture at national and European (EC Directive 91/676/CEE) levels, nitrate is still one of the main contaminants. Many aquatic environments have elevated  $NO_3^-$  levels as the result of anthropogenic activities that involve nitrogen compounds deriving from mineral fertilizers, septic systems, animal manures and so forth. Development of both effective management practices to preserve water quality and remediation plans for contaminated sites require the identification of pollutant sources and consequently of the processes affecting local  $NO_3^-$  concentrations. In this framework the combination of thermodynamical and statistical modelling may represent a powerful tool to describe natural ecosystems and to map phenomena showing a spatial continuity with respect to those dominated by punctual factors.  $E_h$  values determined for waters collected in the Arno basin by EQ3/6 code, under the reasonable hypothesis of chemical equilibrium, have been matched with the scores of the first log contrast summarising the reciprocal relationships among  $NH_4^+$ ,  $NO_2^-$  and  $NO_3^-$ . Results indicate that the approach can be used to discriminate data for which the concept of chemical equilibrium is associated with the statistical modelling of the variability, thus identifying potentially continuous spatial phenomena. These samples are closely associated in space with those characterised by local high variability and low spatial correlation. If sets of data with these features are considered together, the interpretation of maps obtained by interpolation procedures can be, in general, highly compromised.

## Acknowledgements

This research has been financially supported by Italian MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca Scientifica e Tecnologica), PRIN 2004, through the GEOBASI project (prot. 2004048813–002).

## References

- [1] Abbate, E., Boccaletti, M., Braga, G., Coli, M., Dallan Nardi, L., Marchetti, G., Nardi, R., Pochini, A., and Puccinelli, A., 1982. *Le Unit Toscane*. In: Note illustrative della carta strutturale dell'Appennino Settentrionale, C.N.R., Prog. Fin. Geodinamica, L.A.C., Florence (Italy), 46-65.
- [2] Achtnich, C., Bak, F. and Conrad, R., 1995. *Competition for electron donors among nitrate reducers, ferric iron reducers, sulfate reducers, and methanogens in anoxic paddy soil*. *Biology and Fertility of Soils*, 19, 65-72.
- [3] Aitchison, J., 1981. *A new approach to null correlation of proportions*. *Mathematical Geology*, 13(5), 175-189.
- [4] Aitchison, J., 1982. *The statistical analysis of compositional data (with discussion)*. *Journal of the Royal Statistical Society*, B44(2), 139-177
- [5] Aitchison, J., 1986a. *Principal component analysis of compositional data*. *Biometrika*, 70(1), 57-65.
- [6] Aitchison, J., 1986b. *The Statistical Analysis of Compositional Data*. Chapman and Hall, London, 416 pp.
- [7] Anderson, G. M., 1996. *Thermodynamics of Natural Systems*. New York, John Wiley & Sons, 382 pp.
- [8] Anderson, G. M. and Crerar, D. A., 1993. *Thermodynamics in Geochemistry - The Equilibrium Model*. New York, Oxford University Press, 588 pp.
- [9] Baird, C., 2001. *Environmental Chemistry*. W. H. Freeman and Company, New York, 557 pp.
- [10] Barceló-Vidal, C., Martín-Fernández, J. and Pawlowsky-Glahn, V., 2001. *Mathematical foundations of compositional data analysis*. In: Ross, G. (Ed.), *Proceedings*

- of IAMG'01, The sixth annual conference of the International Association for Mathematical Geology, CD-ROM, 20 pp.
- [11] Bartlett, R. J., 1986 *Soil redox behaviour*. In Sparks, D. L. (Ed.), *Soil Physical Chemistry*, CRC Press, Boca Raton, Florida, 179-207.
- [12] Bencini, A. Gargini, A., and Pranzini, G., 1993. *Relazione fra qualit delle acque di sottosuolo, vulnerabilit delle falde ed ubicazione dei centri di pericolo nel bacino di Firenze-Pistoia*. Proceedings 2<sup>nd</sup> International Congress on Geohydrology 29 November 3 December 1993, Florence (Italy), 261-283.
- [13] Bencini, A., Cazzaroli, G. Gargini, A. and Pranzini, G., 1995. *La qualit des eaux souterraines et sa relation avec la vulnerrabilit la pollution des aquifres un exemple en Toscane (Italie): la plaine de Florence*. *Hydrogeology*, 3, 59-72
- [14] Berner, E. K. and Berner, R. A. *Global Environment. Water, Air, and Geochemical Cycles*. Prentice Hall, Inc., Upper Saddle River, New Jersey, 376 ppp.
- [15] Bethke, C. M., 1996. *Geochemical Reaction Modelling*. New York, Oxford University Press, 397 pp.
- [16] Bouwman, A. F., 1998. *Nitrogen oxides and tropical agriculture*. *Nature*, 392,866-867.
- [17] Codisposti, L. A. and Christensen, J. P., 1985. *Nitrification, denitrification and nitrous oxide cycling in the eastern tropical south Pacific Ocean*. *Marine Chemistry*, 16, 277-300.
- [18] Criscenti, L. J., Laniak, G. F. and Erikson, R. L., 1996. *Propagation of uncertainty through geochemical code calculations*. *Geochimica and Cosmochimica Acta*, 60, 3551-3568.
- [19] Ekholm, P., Kallio, K., Salo, S., Pietilainen, O. P., Rekolainen, S., Laine, Y. and Joukola, M., 2000. *Relationship between catchment characteristics and nutrient concentrations in an agricultural river system*. *Water Resources*, 15, 3709-3716.

- [20] Kraynov, S. P., 1997. *Thermodynamic models for groundwater chemical evolution versus real geochemical properties of groundwater: a review of capabilities, errors, and problems*. *Geochemical International*, 35, 639-655.
- [21] Krug, E. C. and Winstanley, D., 2002. *The need for comprehensive and consistent treatment of the nitrogen cycling and mass balance studies: I. Terrestrial nitrogen cycle*. *The Science of the Total Environment*, 293, 1-29.
- [22] Lichtner, P. C., 1996. *Continuum formulation of multicomponent-multiphase reactive transport*. In: Lichtner, P. C., Steefel, C. I. and Oelkers, E. H. (Eds.), *Reactive transport in Porous Media*, Washington. D. C., *Review in Mineralogy*, Mineralogical Society of America, 1-82.
- [23] Mangold, D. C. and Tsang C. F., 1991 *A summary of subsurface hydrological and hydrochemical models*. *Review in Geophysics*, 29, 51-79.
- [24] Moretti, S., 1994. *The Northern Apennines* Proceeding 76th Summer Meeting of the Italian Geological Society, 3, 739-956.
- [25] Nordstrom, D. K. and Munoz, J. L., 1986. *Geochemical Thermodynamics*. Palo Alto, Blackwell Scientific Publications, 477 pp.
- [26] Nordstrom, D. K., Plummer, L. N., Wigley, T. M. L., Wolery, T. J. and Ball, J. W., 1979. *A comparison of computerized chemical models for equilibrium calculations in aqueous systems*. In: Jenne, E. A. (Ed.), *Chemical modelling in Aqueous Systems*, Symposium Series 93, American Chemical Society, 857-892.
- [27] Okafor, P. N. and Ogbona, U. I., 2003. *Nitrate and nitrite contamination of water sources and fruit juices marketed in South-Eastern Nigeria*. *Journal of Food Composition and Analysis*, 16, 213-218.
- [28] Paschke, S. S. and van der Heijde, P., 1996. *Overview of chemical modeling in groundwater and listing of available geochemical models*. *International Ground-*

water Modeling Center, report GWMI-96-01, Golde, Co., Colorado School of Mines.

- [29] Pawlowsky-Glahn, V and Egozcue, J. J., 2001. *Geometric approach to statistical analysis on the simplex*. Stochastic Environmental Research and Risk Assessment (SERRA), 15(5), 384-398.
- [30] Pawlowsky-Glahn, V and Egozcue, J. J., 2002. *BLU estimators and compositional data analysis*. Mathematical Geology, 34(3), 259-274.
- [31] Peters, V. and Conrad R., 1996. *Sequential reduction processes and initiation of CH<sub>4</sub> production upon flooding of oxic upland soils*. Soil Biology and Biogeochemistry, 28, 371-382.
- [32] Ramaldi, E. Nisi, B. Buccianti, A. Vaselli, O. Tassi, F., and Bencini, A., 2004. *Geochemical variations in running waters from Chiana Basin (Tuscany, Central Italy)*. 32<sup>nd</sup> International Geological Congress, 20-28 August, 2004, Florence (Italy), 637.
- [33] Skinner, J. A., Lewis, K. A., Bardon, K. S., Tucker, P., Catt, J. A. and Chambers, B. J., 1997. *An overview of the Environmental Impact of Agriculture in the U.K.*. Journal of Environmental Management, 50, 111-128.
- [34] van Breemen, N., 2002. *Natural inorganic tendency*. Nature, 415, 381-382.
- [35] Wade, A. J., Whitehead, P. G. and O'Shea, L. C. M., 2002 *The prediction and management of aquatic nitrogen pollution across Europe: an introduction to the Integrated Nitrogen in European Catchments project (INCA)*. Hydrology and Earth System Sciences, 6(3), 299-313.
- [36] Westall, J. C., Zachary, J. L. and Morel, F. M. M., 1976. *MINEQL: A computer program for the calculation of chemical equilibrium composition of aqueous system*. Technical Note 18, department of Civil Engineering, Cambridge, MA, Mas-

sachusetts Institute of Technology, 91 pp.

- [37] Wolery, T. J., 1992. *EQ3/EQ6, a software package for geochemical modeling aqueous of systems*, package overview and installation guide (version 7.0), Lawrence Livermore Laboratory, Report UCRL-MA-110662(1).