Relative variation of ionic composition in a karstic lake under the theory of unweighted vs weighted logratio analysis

Emel Cankaya, E. Rıdvan Sıvacı, Sabri Kılınc and Sükran Dere

ABSTRACT

The use of logratio analysis in limnological studies has proved to be effective for solving the problems of the constrained nature of compositional data. The method offers a graphical tool, the relative variation biplot, to explore relative changes of the ions. However, recent methodological developments have shown that the results can be perturbed by low-valued ions with high variances and proposed downweighting their influences as in correspondence analysis. Additional to all properties of the unweighted version, this weighted logratio analysis extends the previous work and has the advantage of the principle of *distributional equivalence*.

As a motivating application, we chose a karstic lake with dominating ions calcium and sulphate causing other ions to be present relatively in low absolute levels. Besides, one of the collected compositional samples was suspected to be unusual which, in part, contributes to high relative variances. This paper is therefore concerned with the choice of the best method for the analysis of such extremely saline water systems by comparing performances of both unweighted and weighted logratio analyses.

We concluded that introducing weights captured more features of ionic relationships with almost all compositional variability explained. We observed that the ratio of calcium to sulphate, ammonium or phosphorus (to a lesser extent) was particularly valuable in understanding the natural chemical process of the lake. A constant log-contrast model based on calcium, ammonium, nitrate and total soluble phosphorus appeared as an equilibrium equation.

Key words | biplot, ionic composition, karstic lake, ratio map, weighted logratio transformation

Emel Cankaya
Department of Statistics, Art and Science Faculty,
Sinop University, Sinop,
Turkey

E. Rıdvan Sıvacı (corresponding author)
Department of Biology, Art and Science Faculty,
Sinop University, Sinop,
Turkey
E-mail: *rsivacii@gmail.com*

Sabri Kılınc
Department of Biology, Art and Science Faculty,
Cumhuriyet University, Sivas,
Turkey

Sükran Dere
Department of Biology, Art and Science Faculty,
Uludag University, Bursa,
Turkey

INTRODUCTION

Recent limnological studies of water systems have emphasized that any numerical analysis used for discovering hydrochemical parameters and characterizing relationships among them must take into account the relative nature of water chemicals, i.e. the amount of the compound is irrelevant as each component only carries information relatively on the whole, which puts constant sum constraint on each sample. Thus they have agreed that the use of logratio analysis developed by Aitchison (1986) for the compositional data is the suitable statistical approach and effective for solving the problems of this constrained nature (Buccianti *et al.* 2003, 2005; Buccianti & Pawlowsky-Glahn 2005; Otero *et al.* 2005*a*, *b*). In such studies, the interest often

centers on whether there are particular chemical relationships within the composition and whether, for example, these can be related to different origins. Logratio analysis extends the utility of biplots by providing a graphical tool called the relative variation biplot which serves for such purposes. In addition, it reveals the full structure of compositional variability, constant log-contrast models and independence. With the help of this graphical representation, it is also possible to detect unusual compositions as reported in the study of Otero *et al.* (2005*b*).

Theoretical development of logratio methodology in the context of correspondence analysis however showed that a low-valued component of a composition with a very high

ratio across the samples may distort the results (Greenacre 2002; Greenacre & Lewi 2005). Such components may (or may not) be influential enough to cause a compositional sample to appear as an outlier. However, if they produce high logratio variances the suggestion is that some form of weighted logratio analysis is needed. Such a form was provided by Greenacre (2002) and stated to have some advantages over unweighted logratio analysis. In the applications of logratio analysis of which we are aware, there was only one case study of weighted logratio analysis conducted on the chemical composition of ceramics with just one influential rare chemical (Greenacre & Lewi 2005) and this new method has not yet been applied in the limnological area.

In the present study, the available data come from a karstic lake classified as an extremely saline system which has distinctive water chemistry, very rich in calcium and sulphate but comparatively all other ions are at very low concentrations. Besides, a compositional sample seemed to be unusual presumably due to natural reasons. This paper therefore raises the question of whether the results of unweighted logratio analysis become influenced by not only a high number of low-valued ions but also the unusual measurement, and if so whether downweighting this influence is essential or not.

Our main goal is to compare the performances of unweighted and weighted logratio analyses for the evaluation of such extreme saline water systems so as to a) monitor the most recent state of the lake using the correct statistical methodology b) define and model the possible hydrochemical parameters which may be used as indicators of environmental conditions in the study area.

In the following sections, we first give a brief geological description of the karstic lake of our study. Then basic definitions of composition and logratio transformation followed by theoretical background for relative variation biplots and ratio map, graphical tools of the unweighted and weighted logratio analyses respectively, are introduced. The results from the application of both methods on the ionic composition of the lake followed by conclusions are finally presented.

STUDY AREA

The lake Great Lota (39°48′ N, 37°23′W) is located in the middle of Turkey, with a surface area of $980 \, m^2$ and an average depth of approximately $3 - 4 \, m$. It is formed by the karstic subsidence on the east-west orientated gypsum plateau. The local geology consists of gravels, conglomerates, limestones ($CaCO_3$), gypsums ($CaSO_4 + H_2O$), marl and mudstones (Gökçe & Ceyhan 1988), dissolution of which mainly controls the ionic composition of this lake. Depending on the amount of substance solved in the water, such karstic waters are nearly saturated with carbonate and sulphate, derived from dissolution of the limestone, dolomite, anhydrite, and gypsum in the platform rocks (Rejmánková *et al.* 1996). The hydrogeological formation of our karstic lake is also hydrate calcium sulphate ($CaSO_4$ $2H_2O$). It should be noted that in natural freshwaters the expected amount of sulphate is between $3 - 30 \, mg/L$ and of calcium $6 - 78 \, mg/L$ (Moss 1973; Robert *et al.* 1999). The measured values of those in our lake were however extreme as they ranged between $45 - 530 \, mg/L$ and $164 - 640 \, mg/L$ respectively during the study period. This extremity masked the importance of the other ions in the lake as they were present comparatively in very small quantities. Therefore karstic lakes have their own limnological character and should be treated with a special care.

METHODS

Composition and logratio transformation

Any vector $x = [x_1, x_2, \dots, x_D]$ of positive proportions, or *components*, of some whole is termed a *D – part composition* and any collection of such vectors is *compositional data*. In our case study, the relative concentrations of 8 ions to the total (in mg/L) are considered as ionic composition. The special and intrinsic feature of compositional data is that the components of a composition must sum to unity or 100%. This means that a D–part composition is essentially a $d = D - 1$ dimensional vector and so can only be represented in a closed sampling space, the *simplex* S^d, a subset of $\Re^D$. For the analysis of compositional data, applications of standard multivariate analysis methods ignoring the constrained nature of the simplex have been known to be entirely inappropriate since the nineteenth century (Pearson 1897).

A new methodology based on the idea that composition provides information on relative rather than absolute magnitudes of its components was developed by Aitchison

(1981, 1986). He introduced *pairwise*, *additive* and *centered* logratio transformations of compositional data. First one defines logtransformed ratios of every possible pair of components. The additive transformation is based on logtransformed ratios of d components with a common divisor x_D. The centered transformation is defined by the logarithms of the components for each sample relative to their geometric mean. Although, stay-in-the simplex procedures are also now available, Aitchison's logratio analysis is easy to use and opens up all available standard multivariate techniques for analysing the phenomena on the simplex.

One of the governing principles of his approach is what is called *subcompositional coherence*. That is any inference made about some parts of a composition from the analysis of full composition must coincide with the inference made about these parts by the analysis of subcomposition. In other words, if only some parts, say p, of a composition within the full D-part composition ($p < D$) is of interest, then the subcomposition can be formed as

$$(s_1 \ldots s_p) = (x_1 \ldots x_p)/(x_1 + \ldots + x_p)$$

Then, it is clear that ratios are invariant under the formation of subcomposition: $s_j/s_{j'} = x_j/x_{j'}$ (Aitchison 1986).

Unweighted logratio analysis: relative variation biplot

In practice, first step towards the analysis of compositional data is to assess the pattern of variability within the data set, either by the covariance structure characterized by the means and variances of pairwise logratios or the graphical aids. Chemical composition of waters is often shown using *ternary* diagrams, for example as in the composite *Piper* diagram (Piper 1944). These are however constrained systems from a geometrical point of view and suitable for the three-part simplex. Relative variations of compositions with more than three components can be assessed by means of the *Relative Variation Biplot* recently developed by Aitchison (1990). It is based on his idea of the logratio transformation and an adaptation of the biplot methodology first introduced by Gabriel (1971) for representing rows and columns of a matrix in the same graph approximately in lower dimensions.

Suppose that $X = [x_{ij}]$ denotes a compositional data matrix with N rows (samples) and D columns (ion components). Taking the logarithm of its elements produces

$L = [l_{ij} = \log(x_{ij})]$. Then, the relative variation biplot can be constructed by double centering the matrix L with respect to averages of the rows and columns, followed by singular value decomposition (SVD) to obtain least-squares matrix approximations. Aitchison & Greenacre (2002) showed that this is equivalent to analysing all the pairwise logratios or the centered logratios. Then in a low-dimensional subspace, a row is expressed as a dot and a column as a ray defined from the origin to its vertex v_i. The line joining two vertices is defined to be a *link*, specifying the relative variance. The complete set of links therefore determines the compositional covariance structure and gives information about subcompositional variability and independence. Interpretation of the relative variation biplot should be done in the light of the following properties:

(a) Length of a link between v_i and v_j gives an estimate for $Var\,(\log(x_i/x_j))$ in the covariance biplot.

(b) Angle cosines between two links estimate correlations between the corresponding pair of logratios, *i.e.* B showing the intersect, $\cos(v_i B v_k) = corr\left[\log(x_i/x_j), \log(x_k/x_l)\right]$. If two links are orthogonal, then independence is implied.

(c) Two or three vertices laying on the same link give the indication of high correlations between the associated logratios, in other words, of subcomposition.

(d) Four vertices forming a parallelogram suggest a constant log-contrast as

$$\log(x_i) - \log(x_j) + \log(x_k) - \log(x_l) = constant$$

Thus there appears to be a proportionality relationship as $x_i/x_j \propto x_l/x_k$ or equivalently $x_i/x_l \propto x_j/x_k$. That is the ratio of any adjacent components in the parallelogram would reveal a strong linear relationship with the ratio of other two.

Weighted logratio analysis: ratio maps

Relative variation biplots provide substantial insights into the nature of compositional variability. It reveals not only relationships between individual observations and the compositional parts but also the existence of subgroups. As Greenacre & Lewi (2005) showed, however, this biplot can be perturbed by low-valued components with high variances. In such cases, they suggested introducing weights

for each component proportional to their means in the same way as correspondence analysis where rows and columns are weighted by the row and column masses of a contingency table. This simple modification not only rectifies the problem of high variance but also allows the method to obey, in addition to the subcompositional coherence, the principle of distributional equivalence, which is a fundamental property of correspondence analysis. This principle assures invariance in the results when rows or columns with identical relative values are amalgamated into one row or one column (Greenacre 2002). For the analysis of compositional data, being able to combine the components which introduce no new information apart from their constant ratio is a significant improvement over unweighted logratio analysis.

Let r_i and c_j denote respectively row and column masses of the compositional data matrix $X_{N \times D}$ (note that row masses are equal, being $1/N$ for each sample). D_r and D_c are diagonal matrices with the corresponding r_i and c_j elements. Then first step towards the weighted biplot analysis is to double centre the matrix L with respect to its weighted row and column averages. The first centering results a matrix as $L^* = [l_{ij}^* = l_{ij} - l_{i.}]$, where $l_{i.} = \sum_{j=1}^{D} c_j l_{ij}$. This matrix is then again centered according to weighted averages of columns using the row masses to weight each element: $Z = \left[z_{ij} = l_{ij}^* - l_{.j}^* \right]$ where $l_{.j}^* = \sum_{i=1}^{N} r_i l_{ij}^*$. The resultant matrix essentially has elements $z_{ij} = l_{ij} - l_{i.} - l_{.j} + l_{..}$, where $l_{..} = \sum_j \sum_i r_i c_j l_{ij}$.

The weighted least squares approximation of the matrix Z is then obtained by multiplying z_{ij} by $(r_i c_j)^{1/2}$, i.e. $S = D_r^{1/2} Z D_c^{1/2}$, followed by the Singular Value Decomposition (SVD) as $S = U \Gamma V^T$, where Γ has diagonal elements of singular values in descending order $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_D > 0$. To obtain a biplot, S is written as the product of two matrices, $S = FG^T$ where either $F = D_r^{-1/2} U \Gamma$ and $G = D_c^{-1/2} V$ producing a *form biplot*, or $F = D_r^{-1/2} U$ and $G = D_c^{-1/2} V \Gamma$ producing a *covariance biplot* (Greenacre 1984). In practice, using the rank-2 approximation, two columns of F gives the coordinates for the samples (dots) and two columns of G gives the coordinates for the components (rays). The explained part of the whole compositional variability in rank-2 is assessed as a percentage defined by $(\lambda_1^2 + \lambda_2^2)/\sum_{i=1}^{D} \lambda_i^2$, a parameter useful to estimate the quality of the fit.

This weighted version of the biplot called ratio map carries all the properties of the unweighted version, relative variation biplot. In addition, due to the weighting procedure the biplot now obeys the *principle of distributional equivalence*.

RESULTS AND DISCUSSION

This paper deals with the logratio analysis of 8 major ions, or components: TSP, NH_4^+, NO_3^-, SiO_3, SO_4^{2-}, Ca^{2+}, Cl^- and oxygen of the lake. Measurements were taken between October 2000 and October 2001 in sampling periods of approximately 15 days (no measurement available in Feb due to freezing), making up 19 samples. The original unit of measurement was mg/L. For the logratio analysis, the relative values, i.e. percentages, of each ion to the total ion concentration were recalculated, which formed the 8 part compositional data summing to 100. Basic statistical summaries of each component are given in Table 1. The principal chemical constituents are calcium and sulphate, with high median percentages of 60% and 32% respectively, reflecting their abundant release through mainly chemical dissolvement of gypsum bedrock of the lake. Total soluble phosphorus (TSP) and ammonium appear to have high variances in comparison with their low median percentages.

Results of unweighted logratio analysis

Relative variation biplot conducted on the matrix of logtransformed ionic compositions explained 77.5% of the total compositional variability (Figure 1). In the diagram, the first axis shows the contrast between {TSP, SO_4^{2-}} and {NO_3^-, SiO_3, Cl^-, oxygen}. The second axis reflects the calcium-ammonium contrast. Recall that the interpretation of the graph is done through the links connecting two rays of the components. Through the inspection of the length of the links, it can be deduced that logratios of NO_3/TSP and Ca^{2+}/NH_4^+ are more variable than any other logratios. An interesting feature drawn from this biplot is that the link of Ca^{2+} and NH_4^+ (link a) is almost orthogonal to the link of NO_3^- and TSP (link b), suggesting independence between the associated logratios.

Table 1 | Basic statistical properties of major chemical proportions (Q_1 = first quartile, Q_2 = median, Q_3 = third quartile)

Chemicals (%)	Abbr	Min	Q_1	Q_2	Q_3	Max	Stdev
Total soluble phosphorus	TSP	0.0014	0.0038	0.0052	0.0064	0.0183	0.0043
Ammonium	NH_4	0.0029	0.0051	0.0079	0.0259	0.0510	0.0135
Nitrate	NO_3	0.00001	0.00002	0.00005	0.00007	0.00016	0.00004
Silicate	SiO_3	0.0001	0.0003	0.0003	0.0005	0.0006	0.0001
Sulphate	SO_4	9.1526	21.3843	32.523	52.566	63.823	17.480
Calcium	Ca	26.1023	35.7450	60.790	66.803	74.352	16.424
Chlorite	CI	0.0747	0.0940	0.1368	0.2196	0.3679	0.0937
Dissolved oxygen Con.	Oxygen	5.3423	6.5644	8.3396	14.302	19.372	4.2869

The SO_4^{2-}, SiO_3 link and the SO_4^{2-}, Cl^- link indicate high correlations between the corresponding pair of logratios. Such observed collinear links thus imply three subcompositions as: {SO_4^{2-}, SiO_3, Cl^-, oxygen}, {Ca^{2+}, NH_4^+} and {NO_3^-, TSP}. Another feature is that the constituents NH_4^+, NO_3^-, Ca^{2+} and TSP lie approximately on a parallelogram which offers a simple constant log-contrast of the form:
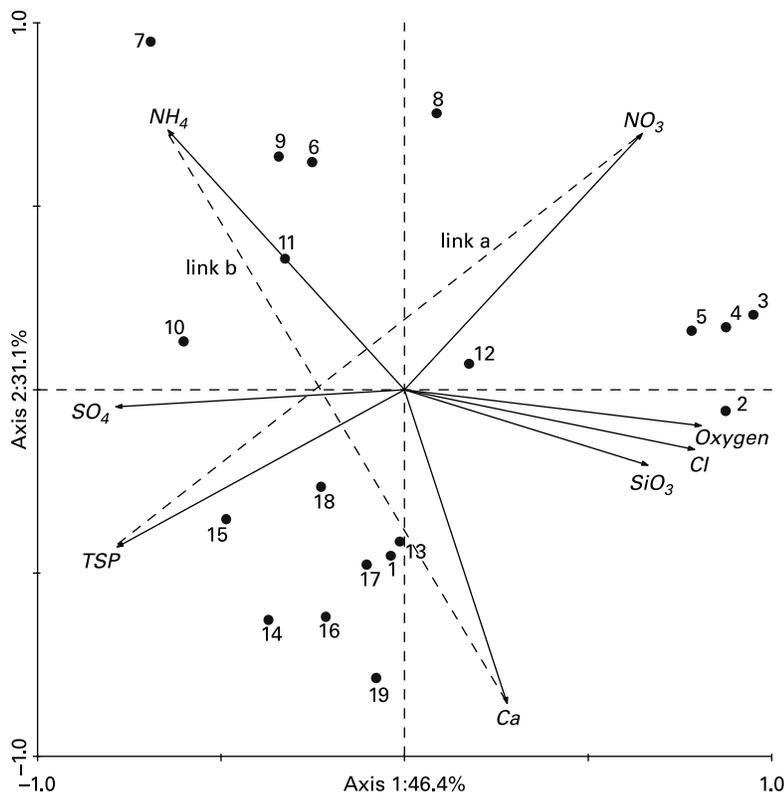
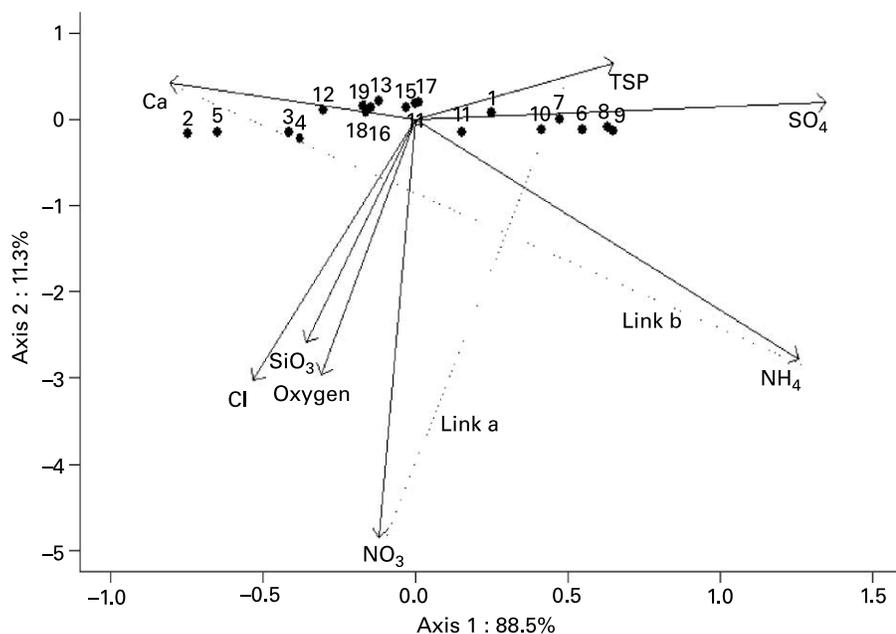**Figure 1** | Relative variation biplot for the ionic compositions.

Figure 2 | Weighted logratio biplot, *i.e.* ratio map, of the ionic composition.

$$\log(NH_4{}^+) - \log(NO_3{}^-) + \log(Ca^{2+}) - \log(TSP) = constant$$

That is to say, $(NH_4{}^+/NO_3{}^-) \propto (TSP/Ca^{2+})$ or equivalently $(NH_4{}^+/TSP) \propto (NO_3{}^-/Ca^{2+})$.

From the diagram it can also be seen that the composition 7, which has by far the largest NH_4^+ proportion, is well out on the calcium-ammonium link, giving the suspicion that it is an outlier. The high relative variations emphasized in the logratio analysis may derive from this unusual composition as well as the low concentrations of most of ions.

Results of weighted logratio analysis

Downweighting the influence of unusual compositions is possible through the use of weighted logratio analysis which attributes importance to the components proportional to their average weights. In the application of our karstic lake, let $X_{N \times D}$ consists of N = 19 vectors of D = 8 part ion compositions, then row weights (r_i) would be equal to a constant of 1/19 and column weights (c_i) would be the average proportions of the ions across the samples. Weighted logratio biplot, namely Ratio map (Figure 2) now explains almost all compositional variability (99.8%) in two dimensions, 88.5% of which belongs to the first axis

reflecting the contrast between Ca^{2+} and SO_4^{2-}, NH_4^+ along with TSP to a lesser extent. This opposition plays a principal role on the sampling points as they all lie along this axis. The second, vertical, axis is under the influence of NO_3^-, SiO_3, Cl^-, oxygen ions.

Compared with the unweighted biplot (Figure 1), calcium, sulphate, ammonium and nitrate are still the main sources of variation, but in contrast, the variation of TSP seems to be diminished. This is an interesting however expected result since TSP had a high variance (Table 1), the effect of which is now downweighted by the new approach. Besides, the opposite behavior of Ca^{2+} and SO_4^{2-} is more highlighted here. Apart from these, all the results deduced from the relative variation biplot can be obtained from the ratio map: approximately orthogonal link of NO_3^- and TSP (link a) to the link of NH_4^+ and Ca^{2+} (link b) indicating the independence of corresponding logratios, diagnosis of the constant log-contrast model between NH_4^+, NO_3^-, Ca^{2+} and TSP which still lie almost on a parallelogram.

CONCLUSIONS

The present study accepts the use of logratio analysis as the correct theoretical framework for the studies of

compositions in chemical monitoring of freshwater systems. It enables us to visualize relative variations of the ionic compositions by means of so called relative variation biplots. In the presence of low-valued constituents with high variances, however, the logratio analysis may not recover the true structure of the compositional relationships as the method does not involve standardization for the variances. Taking advantage of the method being subcompositionally coherent, a solution could be to combine such rare components and treat them as the "other" component. However, for water systems like our karstic lake, out of the investigated eight constituents six appeared with low absolute values (Table 1). Thus merging them is not a proper solution. Maintaining the principles of logratio analysis, an alternative is to downweight their influences by giving small weights proportional to their means which is the idea behind the ratio map approach first developed by Greenacre (2002).

In our case study, due to the geochemical signature of the gypsum bedrock of the lake, abundant existences of calcium and sulphate constituents govern the ionic composition, which causes many other ions to appear as rare constituents. Unweighted logratio analysis of the ionic compositions (Figure 1) revealed high logratio variances of NO_3^-/TSP and Ca^{2+}/NH_4^+ which may derive from the low absolute values of TSP and NH_4^+ the latter of which has an extreme measurement causing sample 7 to be graphically condemned as unusual.

The contrast between calcium and sulphate is more accentuated in the weighted logratio biplot (Figure 2). In this instance, by retaining four ions: Ca^{2+}, SO_4^{2-}, NH_4^+, TSP it is possible to display more than 88% of the compositional variability, so the most important compositional relationships are represented in the water of this karstic lake. Relative to the calcium, the rest exhibits strong relationships which have a huge effect on the positioning of sampling points. Subgroup of NO_3^-, Cl^-, SiO_3 and oxygen dominating the second axis has little role on the compositions as it contributes an extra 11% to the variance. A constant logcontrast model based on calcium, ammonium, nitrate and total soluble phosphorus also still appeared as an equilibrium equation in this natural water system.

In addition to the achievement of a very high percentage of explained compositional variability, this weighted logratio biplot now obeys the principle of *distributional*

equivalence which can be used in merging the components whenever necessary. This is a desirable property and its importance becomes more apparent by the statement made by Buccianti & Pawlowsky-Glahn (2005) "two or more components of a composition are frequently amalgamated or summed, i.e. $Na^+ + K^+$, an operation that may conceal or distort compositional relationships in the data and require therefore particular attention in their analysis".

Unweighted version of the analysis enabled us to study the variations of all ions. However, for the studies of extreme freshwater systems (karstic lakes, brakish waters etc.) like the present one where the interest is not centered on the exploration of the sources of variation, one could consider downweighting the influence of possible variance increasing factors like unusual compositions or components at low proportions. Although results from both analyses were similar, the hydrochemical factor appeared as the opposition between Ca^{2+} and SO_4^{2-}, NH_4^+ and TSP was more emphasized and clearly determined by weighted logratio analysis. We believe such a knowledge on ionic relationships could be used in further studies which may be conducted not only to explore the relationships between ions and the distribution of biological populations but also to investigate the ecological environments (e.g. rivers, agricultural areas etc.) interacting in such extremely saline systems. As a result, this study concludes to be cautious about the high variances in the logratio analysis of hydrochemical variables. Weighted logratio analysis may reveal a better understanding of ionic relationships in extreme water systems like karstic lakes.

REFERENCES

Aitchison, J. 1981 A new approach to null correlations of proportions. *Math. Geology* **13**, 175–189.

Aitchison, J. 1986 *The Statistical Analysis of Compositional Data*. Chapman and Hall, London, UK.

Aitchison, J. 1990 Relative variation diagrams for describing patterns of variability in compositional data. *Math. Geology* **22**, 487–512.

Aitchison, J. & Greenacre, M. J. 2002 Biplots of compositional data. *Appl. Stat.* **51**, 375–392.

Buccianti, A. & Pawlowsky-Glahn, V. 2005 New perspectives on water chemistry and compositional data analysis. *Math. Geology* **37**(7), 703–727.

Buccianti, A., Vaselli, O., Nisi, B., Minissale, A. & Tassi, F. 2003 Monitoring procedures in environmental geochemistry and compositional data analysis theory. Compositional Data

Analysis Workshop – CoDaWork'03 Proceedings. Universtat de Girona, http://ima.udg.es/Activitats/CoDaWork03.

Buccianti, A., Vaselli, O. & Nisi, B. 2005 New insights on river water chemistry by using non-centered simplicial principal component analysis. A case study. Compositional Data Analysis Workshop – CoDaWork'05 Proceedings. Universtat de Girona, http://ima.udg.es/Activitats/CoDaWork05.

Gabriel, K. R. 1971 The biplot-graphical display with applications to principal component analysis. *Biometrika* **58**, 453–467.

Gökçe, A. & Ceyhan, F. 1988 Sivasın güneydoğusundaki Miyosen yaşlı Jipsli çökellerin stratigrafisi, yapısal özellikleri ve oluşumu. *Mühendislik dergisi* **5**, 91–112.

Greenacre, M. J. 1984 *Theory and Applications of Correspondence Analysis*. Academic Press, London, UK.

Greenacre, M.J. 2002 Ratio maps and correspondence analysis. Research report 598, Departament d'Economia i Empresa, Universitat Pompeu Fabra (submitted for publication).

Greenacre, M.J. & Lewi, P.J. 2005 Weighted Logratio Biplots, Correspondence Analysis and Spectral Maps. Compositional Data Analysis Workshop – CoDaWork'05, Proceedings. Universtat de Girona, http://ima.udg.es/Activitats/CoDaWork05.

Moss, B. 1973 The role of pH and the carbon dioxide-bicarbonate system. *J. Ecol.* **61**, 157–177.

Otero, N., Tolosana-Delgado, R., Soler, A., Pawlowsky-Glahn, V. & Canals, A. 2005*a* Relative vs. absolute statistical analysis of compositions: a comparative study of surface waters of a Mediterranean river. *Water Res.* **39**, 1404–1414.

Otero, N., Tolosana-Delgado, R. & Soler, A. 2005*b* A factor analysis of hidrochemical composition of llobregat river basin. Compositional Data Analysis Workshop – CoDaWork'05 Proceedings. Universtat de Girona, http://ima.udg.es/Activitats/ CoDaWork05.

Pearson, K. 1897 Mathematical contributions to the theory of evolution. On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proc. R. Soc. LX*, 489–502.

Piper, A. M. 1944 A graphic procedure in the geochemical interpretation of water analysis. *Trans. Am. Geophys. Union* **25**, 914–923.

Rejmánková, E., Pope, K. O., Post, R. & Maltby, E. 1996 Herbaceous wetlands of the Yucatan peninsula: communities at extreme ends of environmental gradients. *Int. Revue der Gesamten Hydrobiol.* **81**, 223–252.

Robert, R., Charles, N. A. & Robert, O. R. 1999 Stable Isotope Systematics of Sulfate Minerals. *Rev. Minerol. Geochem (Series: Sulfate Minerals – Crystallography, Geochemistry and Environmental Significance)* **40**, 540–582.