Chemometric Valuation of Some Rivers and Lakes in Eastern Slovakia

[a]D. TRUCKENBRODT, [a]J. EINAX, [b]M. MATHERNY, and [c]J. RIGAS

[a]*Institute of Inorganic and Analytical Chemistry, Faculty of Chemical and Geological Sciences, Fr. Schiller University, D-077 40 Jena*

[b]*Department of Chemistry, Faculty of Metallurgy, Technical University, SK-042 00 Košice*

[c]*Slovak Inspection of Environment, Inspection of the Protection of Waters, SK-040 53 Košice*

Received 29 March 1993

The purpose of this study was the chemometric characterization of the water quality of the rivers Torysa and Hornád. For comparison the waters of the lakes Ružín on the river Hornád and the lake Bukovec were used. In this paper it is demonstrated that the use of the multivariance and discriminant analysis and also the factor analysis is sufficient for the complex chemometric valuation of the chemical character and pollution of waters compared.

The investigation of the pollution of rivers and lakes should require the periodic sampling and analysis of various parameters, like the content of main, minor and trace elements and the chemometric valuation of the analytical results as well. The analysis of the pollution situation and the following statistical treatment is necessary because of the large number of samples, involving a multielement characterization of each sample. For instance, a sampling monthly and the determination of 20 parameters yield a matrix with 240 components. Such sample is rather insufficient to interpret as a univariate sample.

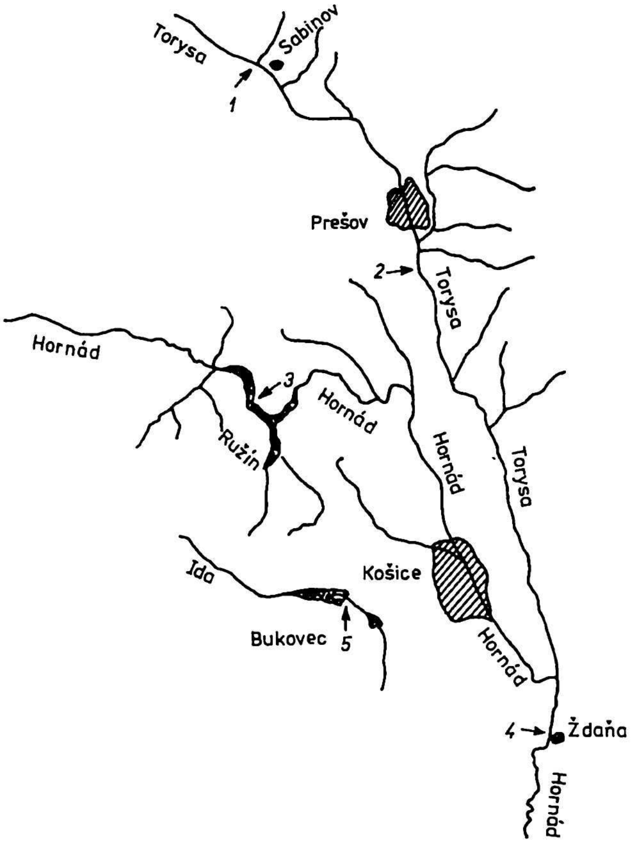

Fig. 1. Sampling places of the investigated lakes and rivers. *1*. Torysa — Sabinov; *2*. Torysa — Prešov; *3*. Hornád — Ružín; *4*. Hornád — Ždaňa; *5*. Bukovec — Ida. ~~~~~~~~ Stream, river, ⃠ village, town, ⬬ water reservoir.

EXPERIMENTAL

Sampling

Sampling started in May 1988, and samples were taken monthly until April 1989. The sample locations were selected by assumed points of environmental exposure [1]. These locations are shown in Fig. 1.

It is known that Torysa is due to the lower load with inorganic pollutants a moderately contaminated river. The degree of contamination was determined by an analysis according to the Czechoslovak water norm [2, 3]. According to this standard the water from the drinking water reservoir Bukovec is the purest from the whole Eastern Slovakia.

Former investigation of the river Hornád showed that the concentrations of copper, iron, manganese, and mercury differ more than the concentrations of chromium, lead, and vanadium within a year [1, 4]. Elements silver, boron, and nickel at least form the so-called "trace element background". Further studies [5] showed that copper, iron, lead, zinc, and nickel are of greatest interest; a main reason for this choice was the nonferrous metallurgical plant in Krompachy, which is a possible contamination source for heavy metals. Furthermore calcium and magnesium were determined because water hardness is an important feature for the description of natural waters.

Sample Preparation and Analysis

All water samples were stabilized by addition of 5 cm^3 of HNO_3 in 1 dm^3 of each sample and in laboratory they were filtered immediately over a G4 glass frit. In the following procedure impregnated polyurethane foams were used as sorbents for preconcentration and separation processes to improve the ratio of trace elements and matrix [6]. The used method reached a quasi-constant precision from 6 up to 15 % by mass concentration gradients from 0.01 up to 10 $\mu g\ dm^{-3}$. For final determination the energy dispersive X-ray fluorescence analysis was applied [7, 8].

Reliability and accuracy were examined with atomic absorption spectroscopy [9, 10]. The relative precision of the concentration determination of analytical elements lies in the interval $s(\rho_{X,r}) \in \langle 0.50\ \%, 0.85\ \% \rangle$. Accuracy of the analytical results was tested by comparison of concentration data obtained using tungsten atomizer [10, 11], graphite cuvette, and Ta-foil atomization. It has been proved [11] that the atomization from Ta-foil and W-tube are both equally valuable.

RESULTS AND DISCUSSION

Multivariate statistics are frequently used in social sciences and biology as well, but more and more also in environmental and analytical chemistry [12, 13]. For data exploration analysis the cluster analysis and the principal component analysis in commercial software package SPSS and the program for multidimensional variance and discriminant analysis according to *Henrion* [14, 15] were used. Pattern cognition and pattern recognition methods as statistical procedures for

a) detecting structures in multivariate feature spaces,

b) converting multivariate data sets into classes and,

c) reducing the number of variables

were used here. In contrast to an analysis of the univariate samples they permit to consider all characteristics with their interactions simultaneously.

As an unsupervised learning technique the hierarchical cluster analysis according to Ward [16] seems to be a useful tool finding pattern between the sampling locations (Fig. 2) and the content of the analyzed metals. The principle of unsupervised learning techniques consists in the partition of objects in different groups, unknown before [17]. In fact a separation of the sampling locations at the investigated rivers is not possible in this way.

More suitable recognizing pattern in this case is the multivariate variance and discriminant analysis (MVDA), a supervised learning procedure with pos-

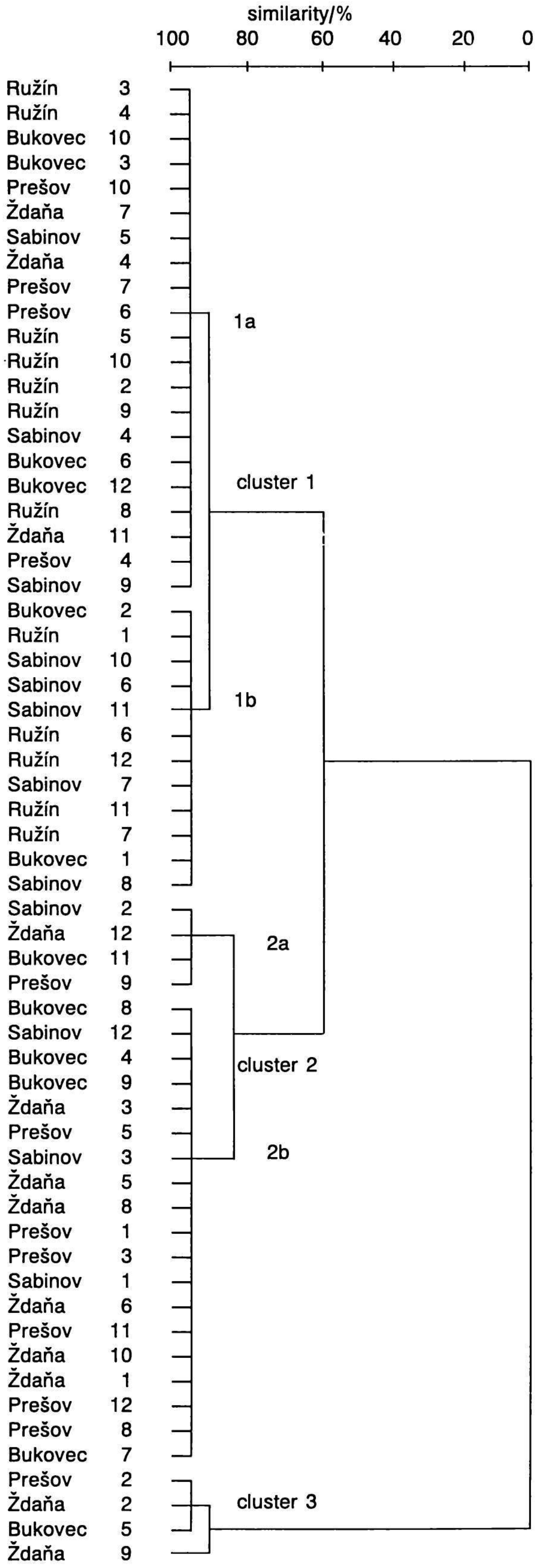

Fig. 2. Dendrogram of the hierarchical cluster analysis according to Ward.

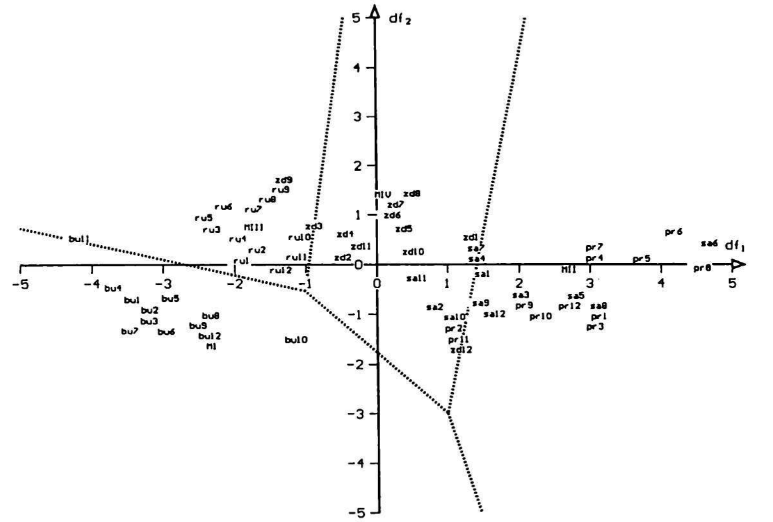

Fig. 3. Two-dimensional display of the first two non-elementary discriminant functions df_1 and df_2.

tulated a priori classification to achieve best discrimination of multivariate data matrices [15, 18]. So a weighting of the parameters according to their separation information is possible.

The model postulates a priori that the loads in Prešov are not so high, so the sample locations Sabinov and Prešov are classified as one class (Torysa), whereas lakes Ružín, Bukovec and the river Hornád are regarded as different classes because of their different origin, geology, and anthropogene load.

As a first result the variables calcium and magnesium have the greatest discriminating power so that the best distinction is achieved by means of geogene influences. The misclassification rate for the rearranging procedure according to Lachenbruch is 10 %, six objects are misclassified [15]. In case of randomly distributed objects the theoretical classification error amounts 68 %. Fig. 3 shows the two-dimensional display of the first two most discriminating non-elementary discriminant functions for the four-class model.

It shows a good separation, especially of the samples from Bukovec. This agrees with the fact that the water from Bukovec originates in an old mining region with specific geochemical background.

The results of the MVDA are impressively confirmed by the principal component analysis (PCA) [14, 19]. The purpose of this procedure is description of changes of the observed variables in complex environmental compartments by finding summarized factors. By the help of factors the main information of the data set can be extracted. In this case 2 common factors result, explaining 59.1 % of the variance; Table 1 contains the results. Factor 1 is formed by water hardness, factor 2 by iron and lead.

Table 1. Factor Analysis of the Investigated Rivers and Streams. Factor Loadings

Element	Factor 1	Factor 2
Ca	0.93	– 0.09
Mg	0.94	– 0.12
Cu	0.31	– 0.01
Fe	0.05	0.88
Pb	– 0.03	0.81
Zn	0.28	– 0.05

Figs. 4 and 5 show the factor scores of all sampling locations.

Correlation analysis of the factor scores in Table 2 demonstrates that after junction of Torysa and Hornád factor 1 Ždaňa is highly correlated with factor 1 Ružín and factor 2 Ždaňa is correlated especially with factor 2 Prešov, so after confluence the samples are characterized by patterns of the water hardness of the river Hornád and the iron/lead pattern from Torysa river. An estimation concerning the anthropogene of geogene origin of the heavy metal pattern is not possible, but these investigations have detected that the origin is situated above Sabinov.

CONCLUSION

It was demonstrated that multivariate variance, discriminant analysis and factor analysis are useful tools for data elucidation in environmental science. The concentrations of calcium and magnesium are legitimate with regard to basic information whereas the contents of copper, zinc, and lead represent here only limited information. So the state of these rivers and streams is mainly coined by the geological background. In the investigated territory the pollution is

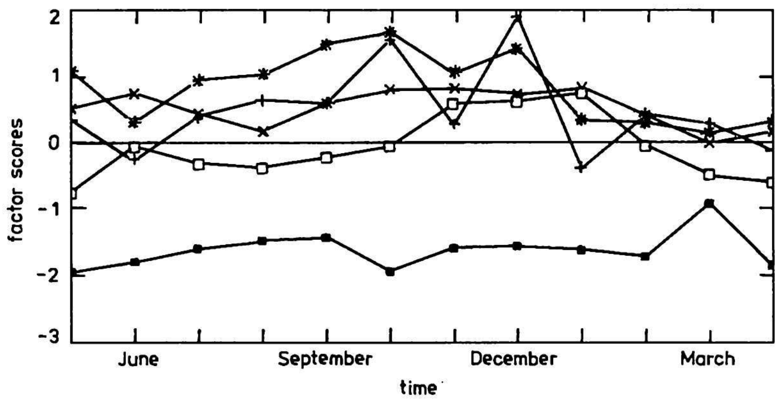

Fig. 4. Factor analysis. Factor scores of the first factor. ■ Bukovec, + Sabinov, * Prešov, □ Ružín, × Ždaňa.

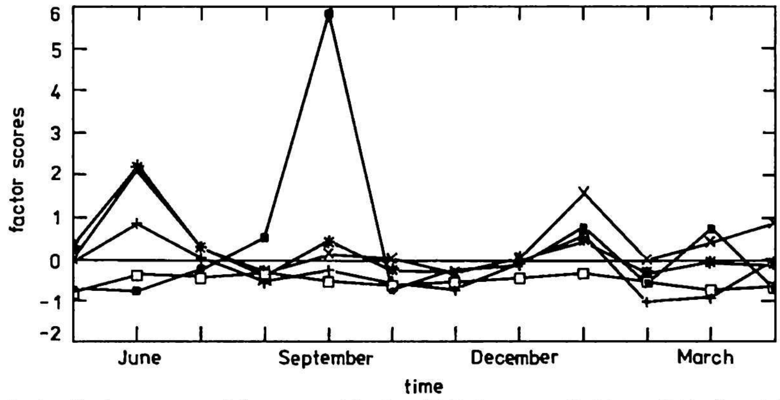

Fig. 5. Factor analysis. Factor scores of the second factor. ■ Bukovec, + Sabinov, * Prešov, □ Ružín, × Ždaňa.

Table 2. Correlation Analysis of Factor Scores

Correlated sampling places	Correlated factor scores of factor	Correlation coefficient r
Sabinov/Prešov	1	0.886
Sabinov/Prešov	2	0.790
Prešov/Ždaňa	1	0.426
Prešov/Ždaňa	2	0.802
Ružín/Ždaňa	1	0.713
Ružín/Ždaňa	2	0.403

not so heavy with respect to the analyzed heavy metals. It can be concluded that the pollution state of Torysa and Hornád with the investigated metals is not so ready.

REFERENCES

1. Flórián, K., Matherny, M., and Pliešovská, N., *Vodní hospodářství 42*, 52 (1992).
2. *CSN 75 7221*, Klasifikace jakosti povrchových vod. (Classification of the Quality of Surface Waters.) Prague, 1990.
3. *CSN 75 7111*, Pitná voda. (Drinking Water.) Prague, 1990.
4. Pliešovská, N. and Seszták, J., *Transaction of the Technical University of Košice 2*, 79 (1992).
5. Pliešovská, N. and Flórián, K., *Chem. Listy 86*, 589 (1992).
6. Rigas, J., Palágyi, Š., Holéczyová, G., and Matherny, M., *Chem. Papers 45*, 509 (1991).
7. Balgavá, V. and Matherny, M., *J. Radioanal. Nucl. Chem. 170*, 171 (1993).
8. Balgavá, V. and Matherny, M., *Fresenius J. Anal. Chem. 346*, 162 (1993).
9. Holéczyová, G. and Matherny, M., *Chem. Papers 46*, 385 (1992).
10. Krakovská, E. and Sučik, G., *Chem. Listy 85*, 312 (1991).
11. Krakovská, E., *CLB Chem. Lab. Biotech. 42*, 247 (1991).
12. Brereton, R. G., *Chemometrics*. Horwood, New York, 1990.
13. Einax, J., *GIT Fachz, Lab. 8*, 815 (1992).
14. *Statistik-Programm-System für die Sozialwissenschaften:* Beutel, P. and Schubö, W., *Eine Beschreibung der Programmversionen 8 und 9*. 4th Edition. Fischer Verlag, Stuttgart, 1983.
15. Henrion, G., Henrion, R., and Henrion, A., *Beispiele zur Datenanalyse mit Basic-Programmen*. Deutscher Verlag der Wissenschaften, Berlin, 1988.
16. Henrion, A., Henrion, R., Urban, P., and Henrion, G., *Z. Chem. 27*, 56 (1987).
17. Vermuza, K., *Pattern Recognition in Chemistry*. Springer-Verlag, Berlin, 1980.
18. Ahrens, H. and Läuter, J., *Mehrdimensionale Varianzanalyse*. Akademie-Verlag, Berlin, 1981.
19. Weber, E., *Grundriss der Biologischen Statistik*. Fischer Verlag, Jena, 1986.