Chemometrical Evaluation of Multispecies-Multichemical Data by Means of Graphical Techniques Combined to Multivariate Analyses

J. DEVILLERS [1*], J. THIOULOUSE [2], AND W. KARCHER [3]

1. CTIS, 21 rue de la Bannière, 69003 Lyon, France.

2. Laboratoire de Biométrie, Génétique et Biologie des Populations, U.R.A. CNRS 243, Université Lyon-I, 43 Bd du 11 Novembre 1918, 69622 Villeurbanne CEDEX, France.

3. Commission of the European Communities, Joint Research Centre, ISPRA Establishment, I-21020 Ispra Varese, Italy.

Running title: Graphical Techniques for Multivariate Analyses

* To whom correspondence may be addressed.

SUMMARY

A new method combining graphical displays with principal component analysis (PCA) has been used to evaluate published data on the toxicity of seven chemicals to 14 species (17 testing procedures) of aquatic biota. The results show the usefulness of simple graphical approaches to analyze the structure of environmental data sets. Thus, the study indicates the importance of endpoint selection and underlines some relationships among the species and chemicals.

INTRODUCTION

The literature and the databanks contain increasing amounts of toxicity data, both in terms of numbers of chemicals and species of biota tested (McCutcheon *et al.*, 1990; Penning *et al.*, 1991; Voigt *et al.*, 1991). There is a need for effective techniques to analyze and present such data. The toxicity of several chemicals to a number of species can be given in chemicals versus species toxicity tables with, for example, species in rows, and chemicals in columns. Such tables can be analyzed by means of various multivariate approaches such as principal component analysis (Laurence *et al.*, 1984; Nendza and Seydel, 1988; Thioulouse *et al.*, 1991), SIMCA method (Wängberg and Blanck, 1988), correspondence factor analysis (Devillers *et al.*, 1988; Devillers and Karcher, 1990), hierarchical cluster analysis (Benigni and Giuliani, 1985; Devillers *et al.*, 1988; Roux, 1991), and minimum spanning tree method (Devillers and Doré, 1989). The aim of this paper is to demonstrate that simple graphical techniques coupled with PCA allow to extract ecotoxicological information from a data matrix of seven chemicals tested on 14 aquatic species (17 testing procedures) occupying different trophic levels in the environment.

MATERIALS AND METHODS

Toxicity data

The animals chosen for the study included four Crustacea (*Daphnia magna*, *Orconectes immunis*, *Daphnia pulex*, and *Ceriodaphnia reticulata*), one Insect (*Tanytarsus dissimilis*), six fish (*Oncorhynchus mykiss*, *Lepomis macrochirus*, *Gambusia affinis*, *Ictalurus punctatus*, *Carassius auratus*, and *Pimephales promelas*), one Amphibia (*Rana catesbeiana*), one echinoderm (*Arbacia punctulata*), and one bacterium (*Photobacterium phosphoreum*) (Table 1). They were selected on the following criteria:

- they present different degrees of evolution and complexicity (i.e.; bacteria to vertebrates),
- they occupy different trophic levels in the environment,
- they are commonly used for aquatic toxicity tests, and are easily cultured or maintained in the laboratory,
- they allow the study of the relationships between different endpoints.

Selection of chemicals (Table 2) was based in part on available toxicity data obtained under standardized conditions, and in part because of different physiological mechanisms of toxic action (Thurston *et al*., 1985).

The toxicity data (Table 3) were obtained from Thurston *et al*. (1985), Elnabarawy *et al.* (1986), Jackim and Nacci (1986), and Nacci *et al.* (1986). The data were expressed as (-log(mol/L)).

Graphical analysis

The basic principles of graphical techniques for multidimensional data analysis have been recently reviewed by Thioulouse *et al*. (1991) and can be summarized as follows:

- a graph should not use conventions (e.g.; linguistic, cultural), and should be based upon the three following fundamental relationships:
    - similarity/dissimilarity relationship
    - order relationship
    - proportionality relationship,
- the relations between the elements of a graph should reflect the relations between the elements which are symbolized,
- the introduction of unnecessary elements should be avoided (e.g.; the representation of a quantity should be done with a circle or a square rather than with a non-geometrical picture),
- a graph should be readable for different levels of perception (e.g.; global, mean, local).

Taking into account these basic principles, Table 3 has been analyzed by means of a collection of graphs generated from the MacMul and GraphMu software (Thioulouse, 1989; 1990).

RESULTS AND DISCUSSION

Figure 1 shows a simple graphical display of Table 3 with the seven chemicals arranged in columns (1 to 7), and the 17 organisms in rows (A to Q). The squares are proportional to the toxicity values. An obvious feature stressed by this display is the strong differences within the seven chemicals. Indeed, pentachlorophenol (compound 7) and hexachloroethane (compound 6) are the most toxic for almost all species, while 2-methyl-2,4-pentanediol (compound 1) and 2-methyl-1-propanol (compound 2) are the least. 2,2,2-tricholoroethanol (compound 3), 2,4-pentanedione (compound 4), and 2-chloroethanol (compound 5) generally present an intermediate toxicity. This effect, which

can be termed a "compound effect", is widely observed in ecotoxicology (Johnson and Finley, 1980; Hudson *et al.*, 1984; Mayer and Ellersieck, 1986; Devillers and Exbrayat, 1992). No "organism effect" can be observed in figure 1, but this may be due to the fact that it is hidden by the compound effect. To remove this compound effect, one can perform a centering by columns of the raw data matrix (Table 3). This is simply achieved by substracting the mean of each column (i.e.; chemical) to the values of the corresponding column (Table 4). Figure 2 shows the result of this operation, with circles for negative values and squares for positive ones. No strong organism effect can be stressed out in this figure, but it is interesting to note that *Arbacia punctulata* (early embryo, [^{3}H] thymidine incorporation), organism K, is little sensitive to almost all chemicals, since the first six compounds are represented by circles.

To remove this slight organism effect, we can perform a centering by rows of Table 4 (i.e.; of the raw data matrix after centering by columns), which leads to Table 5 (i.e.; the doubly centered table). Figure 3 shows the result of this double centering, and we can notice that only "interactions" between chemicals and organisms are visible. Here, "interactions" means the high or low reactivity of one particular species to a particular chemical. These interactions can be analyzed visually by checking the biggest circles and squares. Thus, figure 3 reveals first that *Arbacia punctulata* (sperm cell), organism L, presents a particular ecotoxicological behavior. Comparatively to the other organisms under study (Table 1), *Arbacia punctulata* (sperm cell) is highly sensitive to 2-methyl-2,4-pentanediol (compound 1) and 2,4-pentanedione (compound 4). At the opposite, this organism is not sensitive to the other chemicals, especially compound 5 (2-chloroethanol) and 6 (hexachloroethane). This result is not surprising since it has been advanced that the sea urchin sperm cell tests overestimated the toxicity of compounds that are spermicidal or that use specific mechanisms of toxicity not associated with lethality in standard acute tests (Nacci *et al.*, 1986; Cherr *et al.*, 1987). However, we can note that numerous publications show that the results of sea urchin sperm cell bioassays are affected by a number of factors including temperature, pH, salinity, sperm:egg ratios, sperm exposure times, test materials, echinoid species, and so on (Pagano *et al.*, 1982; Greenwood, 1983; Dinnel *et al.*, 1987; 1989; Ringwood, 1992). Figure 3 also reveals that *Tanytarsus dissimilis* (organism B) and *Orconectes immunis* (organism C) are less sensitive to pentachlorophenol (compound 7) than the other organisms. These results agree with those found towards organisms occupying the same taxonomic level (Jones, 1981; Ewell *et al.*, 1986; Hedtke *et al.*, 1986). Last, figure 3 underlines the opposition between organisms D to J (i.e.; *Rana catesbeiana*, *Oncorhynchus mykiss*, *Lepomis macrochirus*, *Gambusia affinis*, *Ictalurus punctatus*, *Carassius auratus*, and *Pimephales promelas*) and organisms M, and O to Q (i.e.; *Photobacterium phosphoreum*, *Daphnia magna*

(EC50), *Daphnia pulex*, and *Ceriodaphnia reticulata*) which present an inverted pattern. These two groups of organisms allow to oppose the sensitivity of the Vertebrates (especially the fishes) to that of the nonvertebrates. Among this last group, it is interesting to note that the ecotoxicological behaviors of *Daphnia magna* (EC50), *Daphnia pulex*, and *Ceriodaphnia reticulata* are highly correlated. However, the sequence of squares and circles obtained for *Daphnia magna* (LC50) (organism A) is very different of that recorded for *Daphnia magna* (EC50) (organism O). We can advance that this discrepancy is due to interlaboratory variability instead of endpoint selection. At the opposite, the different sequences recorded for the tests performed on *Arbacia punctulata* (K, L, and N) are due to the endpoint selection and stress the usefulness to take into account this parameter for comparative evaluation of toxicity tests.

Figure 3 underlines the heuristic potency of simple graphical approaches to compare the performances of different test systems in ecotoxicology. However, two main drawbacks can be advanced against the methodological approach presented above. First, one can say that it was easy to underline the group D to J because these organisms were not randomly arranged in Table 3. Second, this kind of visual approach is particularly suitable for simple and strong patterns. In other situations, our approach can be improved by using a more objective and mathematically optimized graphical display. Principal component analysis (PCA) of the doubly centered table (Table 5) allows to obtain factor coordinates (FCs) which can be used to overcome the above graphical problems. These FCs have the mathematical property of maximizing the inertia of the projection of compounds and organisms onto each principal axis. Our goal is to rearrange the compounds and the organisms so that interactions are more clearly visible. By plotting the row FCs versus the column FCs, we obtain a graph on which the position of compounds and organisms is optimal from this point of view. Moreover, we can represent on this graph the toxicity values (after double centering) by means of circles and squares (as described above). Figure 4 shows the result of this procedure. On this figure, each item of the data matrix after double centering (Table 5) is placed according to the first FCs of the corresponding chemical and organism. For example, the square marked with an arrow corresponds to the toxicity of compound 1 (2-methyl-2,4-pentanediol) to organism L (*Arbacia punctulata*, sperm cell). The horizontal (x-axis) position of this square on the graph is given by the first FC of compound 1, and its vertical position (y-axis) is given by the first FC of organism L. This representation is the PCA equivalent of the classical canonical map of canonical correlations analysis (Lebart *et al.*, 1984). The size of the square is proportional to the value of the corresponding toxicity after double centering. One can easily see that this representation makes obvious the particular ecotoxicological behavior of *Arbacia punctulata* (sperm cell), and particularly its high sensitivity to 2-methyl-2,4-

pentanediol (compound 1) and 2,4-pentanedione (compound 4), as opposed to its low sensitivity to compound 5 (2-chloroethanol) and 6 (hexachloroethane). On the opposite side, the group of organisms D to J is also in a conspicuous position. Circles in the left part denote a low sensitivity to 2-methyl-2,4-pentanediol (compound 1) and 2,4-pentanedione (compound 4), while squares in the right part denote a high sensitivity to compound 5 (2-chloroethanol) and 6 (hexachloroethane).

At this step, it is still possible to proceed with the analysis, and try to reveal further details of the data set. We must first try to remove the effects that have already been characterized. To achieve this, we use a well-known property of multivariate analyses, which is the possibility of recomputing the initial data table from FCs (e.g.; recompute the data set starting from the first principal components, Lebart *et al.*, 1984, p. 8). The obtained reconstitution of the data set is a model taking into account only the information that was extracted by the principal components. In the present study, we can consider that the first principal component is a good model of the interaction between *Arbacia punctulata* (sperm cell) and compounds 1 and 4, as opposed to compounds 5 and 6. So, we can reconstitute the toxicity table with this first principal component, and compute the residuals between this reconstitution and the doubly centered table. This will remove the corresponding interaction from the data set, allowing us to perform further investigations. The results are shown in figure 5, in which the position of circles and squares is given by the second principal component (for compounds horizontally and organisms vertically), and their size is proportional to the residuals between the doubly centered data table and the reconstitution by the first principal component (Table 6). The most conspicuous feature of this representation is the low sensitivity of organisms B (*Tanytarsus dissimilis*) and C (*Orconectes immunis*) to compound 7 (pentachlorophenol), opposed to the high toxicity of the same compound to organism N (*Arbacia punctulata* (early embryo, fluorometric determination of DNA)).

We could carry on this procedure by computing the residuals between the doubly centered table and the reconstitution with the first two principal components, and plotting these residuals as functions of the third FCs for compounds and organisms. However it is necessary to decide when the remaining structures of the data set are only random noise, and stop the procedure. Two solutions can be used to answer this question. The first consists in looking at the decrease of the successive eigenvalues obtained from PCA (figure 6). The magnitude of eigenvalues represents the percentage of inertia (i.e.; the quantity of information) extracted from the data set by the corresponding principal component. Figure 6 shows that the first two eigenvalues are clearly higher than the following ones. Eigenvalues 3 to 7 have low values and are regularly decreasing. This means that the information remaining in the data set is low and has no strong structure.

The second solution consists in plotting the values of the residuals at the **same scale** as the initial data set. Figure 7a shows the residuals after removing the interactions taken into account by the first two principal components, and figure 7b, the doubly centered data set. One can see that the main differences between organisms and chemicals have disappeared, and that the circles and squares are small and do not present strong structures. The remaining differences (e.g.; the difference in sensitivity of organism M (*Photobacterium phosphoreum*) to compounds 3 (2,2,2-tricholoroethanol) and 4 (2,4-pentanedione)) are taken into account by the next principal components (particularly the third component for organism M). However, these differences are low and comparable to other small interactions.

CONCLUSION

Chemometrical evaluation of multispecies-multichemical data by means of graphical information extraction techniques combined to multivariate analyses is important since it provides a good insight into the structure of the data and draws attention to details that may have gone unnoticed. For example, in the present data, our approach readily indicates the clustering of vertebrate and nonvertebrate species and the importance of endpoint selection (e.g.; *Arbacia punctulata* tests). Interlaboratory variability may have contributed to the differences between *Daphnia magna* EC50 and LC50 endpoints. Even if these ecotoxicological results cannot be generalized due to the reduced size of the data matrix they illustrate the heuristic potency of our method in environmetrics.

ACKNOWLEDGMENTS

The authors wish to thank Dr. V. Zitko (Marine Chemistry Division, Department of Fisheries and Oceans, St Andrews, Canada) for securing the data in Table 3 and reading an early draft of the present paper.

REFERENCES

BENIGNI, R., AND GIULIANI, A. (1985). Cluster analysis of short-term tests: a new methodological approach. *Mutation Res.* **147**, 139-151.

CHERR, G.N., SHENKER, J.M., LUNDMARK, C., AND TURNER, K.O. (1987). Toxic effects of selected bleached kraft mill effluent constituents on the sea urchin sperm cell. *Environ. Toxicol. Chem.* **6**, 561-569.

DEVILLERS, J., AND DORE, J.C. (1989). Heuristic potency of the minimum spanning tree (MST) method in toxicology. *Ecotoxicol. Environ. Safety* **17**, 227-235.

DEVILLERS, J., ELMOUAFFEK, A., ZAKARYA, D., AND CHASTRETTE, M. (1988). Comparison of ecotoxicological data by means of an approach combining cluster and correspondence factor analyses. *Chemosphere* **17**, 633-646.

DEVILLERS, J., AND EXBRAYAT, J.M. (1992) (Eds.). *Ecotoxicity of Chemicals to Amphibians.* Gordon and Breach Science Publishers, Reading, England.

DEVILLERS, J., AND KARCHER, W. (1990). Correspondence factor analysis as a tool in environmental SAR and QSAR studies. In, *Practical Applications of Quantitative Structure-Activity Relationships (QSAR) in Environmental Chemistry and Toxicology* (W. Karcher and J. Devillers, Eds.), pp. 181-195. Kluwer Academic Publishers, Dordrecht, The Netherlands.

DINNEL, P.A., LINK, J.M., AND STOBER, Q.J. (1987). Improved methodology for a sea urchin sperm cell bioassay for marine waters. *Arch. Environ. Contam. Toxicol.* **16**, 23-32.

DINNEL, P.A., LINK, J.M., STOBER, Q.J., LETOURNEAU, M.W., AND ROBERTS, W.E. (1989). Comparative sensitivity of sea urchin sperm bioassays to metals and pesticides. *Arch. Environ. Contam. Toxicol.* **18**, 748-755.

ELNABARAWY, M.T., WELTER, A.N., AND ROBIDEAU, R.R. (1986). Relative sensitivity of three daphnid species to selected organic and inorganic chemicals. *Environ. Toxicol. Chem.* **5**, 393-398.

EWELL, W.S., GORSUCH, J.W., KRINGLE, R.O., ROBILLARD, K.A., AND SPIEGEL, R.C. (1986). Simultaneous evaluation of the acute effects of chemicals on seven aquatic species. *Environ. Toxicol. Chem.* **5**, 831-840.

GREENWOOD, P.J. (1983). The influence of an oil dispersant chemserve OSE-DH on the viability of sea urchin gametes. Combined effects of temperature, concentration and exposure time on fertilization. *Aquatic Toxicol.* **4**, 15-29.

HEDTKE, S.F., WEST, C.W., ALLEN, K.N., NORBERG-KING, T.J., AND MOUNT, D.I. (1986). Toxicity of pentachlorophenol to aquatic organisms under naturally varying and controlled environmental conditions. *Environ. Toxicol. Chem.* **5**, 531-542.

HUDSON, R.H., TUCKER, R.K., AND HAEGELE, M.A. (1984) (Eds.). *Handbook of Toxicity of Pesticides to Wildlife*. U.S. Fish Wildl. Serv., Resour. Publ. 153, Washington, D.C.

JACKIM, E., AND NACCI, D. (1986). Improved sea urchin DNA-based embryo growth toxicity test. *Environ. Toxicol. Chem.* **5**, 561-565.

JOHNSON, W.W., AND FINLEY, M.T. (1980) (Eds.). *Handbook of Acute Toxicity of Chemicals to Fish and Aquatic Invertebrates*. U.S. Fish Wildl. Serv., Resour. Publ. 137, Washington, D.C.

JONES, P.A. (1981) (Ed.). *Chlorophenols and Their Impurities in the Canadian Environment*. Environment Canada, Report SPE 3-EC-81-2F.

LAURENCE, R.A., LAMBERT, N.G., SAINT-GERMAIN, G., DRAPEAU, A.J., AND BOULIANE, M. (1984). Essais de caractérisation des microorganismes par leurs profils métalliques. *Ann. Microbiol. (Inst. Pasteur)* **135 B**, 347-358.

LEBART, L., MORINEAU, A., AND WARWICK K.M. (1984) (Eds.). *Multivariate Descriptive Statistical Analysis. Correspondence Analysis and Related Techniques for Large Matrices.* John Wiley & Sons, New York.

MAYER, F.L., AND ELLERSIECK, M.R. (1986) (Eds.). *Manual of Acute Toxicity: Interpretation and Data Base for 410 Chemicals and 66 Species of Freshwater Animals*. U.S. Fish Wildl. Serv., Resour. Publ. 160, Washington, D.C.

McCUTCHEON, P., NORAGER, O., KARCHER, W., AND DEVILLERS, J. (1990). Sources of data for risk evaluation and QSAR studies. In, *Practical Applications of Quantitative Structure-Activity Relationships (QSAR) in Environmental Chemistry and Toxicology* (W. Karcher and J. Devillers, Eds.), pp. 13-24. Kluwer Academic Publishers, Dordrecht, The Netherlands.

NACCI, D., JACKIM, E., AND WALSH, R. (1986). Comparative evaluation of three rapid marine toxicity tests: sea urchin early embryo growth test, sea urchin sperm cell toxicity test and microtox. *Environ. Toxicol. Chem.* **5**, 521-525.

NENDZA, M., AND SEYDEL, J.K. (1988). Multivariate data analysis of various biological test systems used for quantification of ecotoxic compounds. *Quant. Struct. Act. Relat.* **7**, 165-174.

PAGANO, G., ESPOSITO, A., AND GIORDANO, G.G. (1982). Fertilization and larval development in sea urchins following exposure of gametes and embryos to cadmium. *Arch. Environ. Contam. Toxicol.* **11**, 47-55.

PENNING, W., ROI, R., AND BONI, M. (1991). ECDIN-The European databank on environmental chemicals. In, *Environmental Information and Communication Systems: ECOINFORMA 1* (O. Hutzinger and H. Fiedler, Eds.), pp. 25-38. Gordon and Breach Science Publishers, Reading, England. ECDIN is now available on CD-ROM from J.D. or W.K.

RINGWOOD, A.H. (1992). Comparative sensitivity of gametes and early developmental stages of a sea urchin species (*Echinometra mathaei)* and a bivalve species (*Isognomon californicum*) during metal exposures. *Arch. Environ. Contam. Toxicol.* **22**, 288-295.

ROUX, M. (1991). Basic procedures in hierarchical cluster analysis. In, *Applied Multivariate Analysis in SAR and Environmental Studies* (J. Devillers and W. Karcher, Eds.), pp. 115-135. Kluwer Academic Publishers, Dordrecht, The Netherlands.

THIOULOUSE, J. (1989). Statistical analysis and graphical display of multivariate data on the Macintosh. *CABIOS* **5**, 287-292.

THIOULOUSE, J. (1990). MacMul and GraphMu: two macintosh programs for the display and analysis of multivariate data. *Comput. Geosci.* **16**, 1235-1240. These programs, for the Apple Macintosh® microcomputer, are available upon request from the author in exchange for 5 Macintosh diskettes, or can be freely downloaded on the Internet by anonymous FTP to biomserv.univ-lyon1.fr (134.214.100.236).

THIOULOUSE, J., DEVILLERS, J., CHESSEL, D., AND AUDA, Y. (1991). Graphical techniques for multidimensional data analysis. In, *Applied Multivariate Analysis in SAR and Environmental Studies* (J. Devillers and W. Karcher, Eds.), pp. 153-205. Kluwer Academic Publishers, Dordrecht, The Netherlands.

THURSTON, R.V., GILFOIL, T.A., MEYN, E.L., ZAJDEL, R.K., AOKI, T.I., AND VEITH, G.D. (1985). Comparative toxicity of ten organic chemicals to ten common aquatic species. *Water Res.* **19**, 1145-1155.

VOIGT, K., BENZ, J., MATTHIES, M., AND MÜCKE, W. (1991). Data-sources for environmental chemicals: databank of data-sources and strategy for computer-aided searches. In, *Environmental Information and Communication Systems: ECOINFORMA 1* (O. Hutzinger and H. Fiedler, Eds.), pp. 1-16. Gordon and Breach Science Publishers, Reading, England.

WÄNGBERG, S.A., AND BLANCK, H. (1988). Multivariate patterns of algal sensitivity to chemicals in relation to phylogeny. *Ecotoxicol. Environ. Safety* **16**, 72-82.

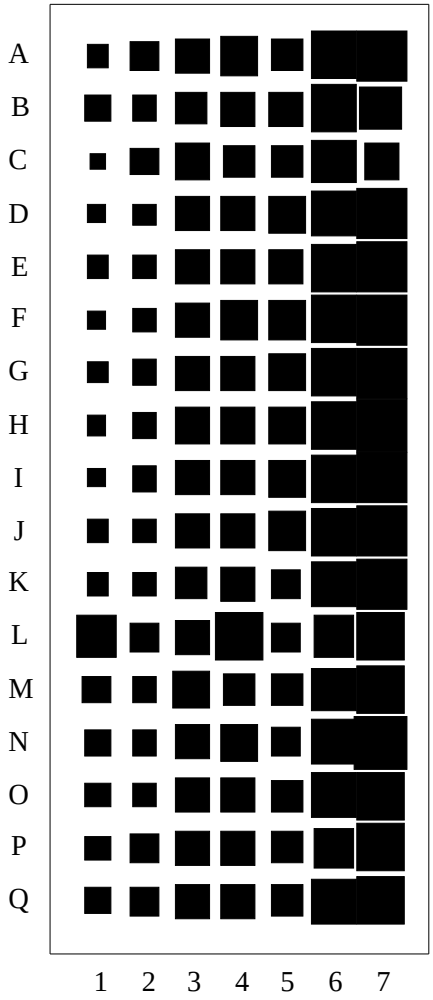

Fig. 1. Graphical display of the raw data matrix. Squares are proportional to the values listed in Table 3. Compounds 1 to 7 are represented in columns and organisms (A to Q) are in rows (for caption see Tables 1 and 2).

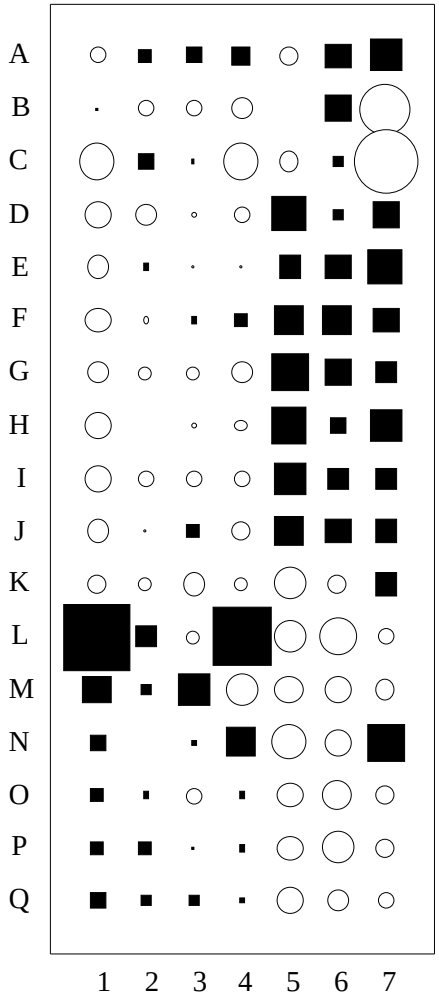

Fig. 2. Graphical display of the data matrix after centering by column (Table 4). Squares are proportional to positive values and circles to negative ones.

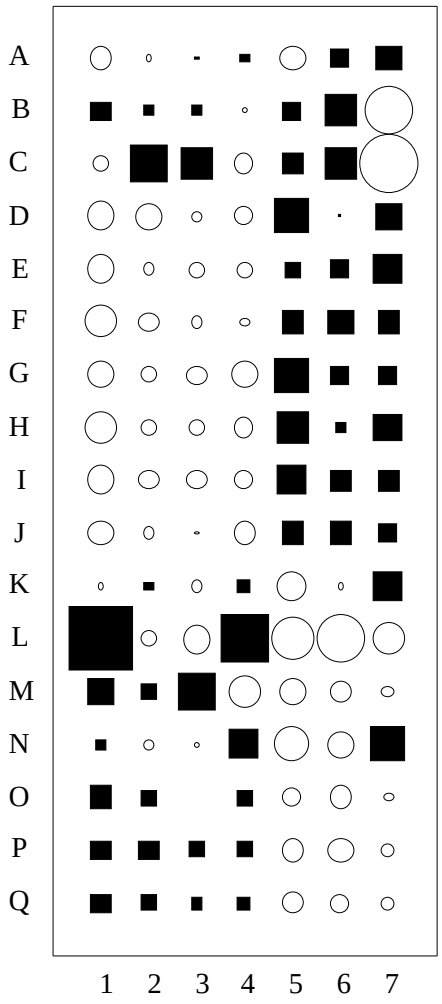

Fig. 3. Graphical display of the data matrix after double centering (Table 5). Squares are proportional to positive values and circles to negative ones.

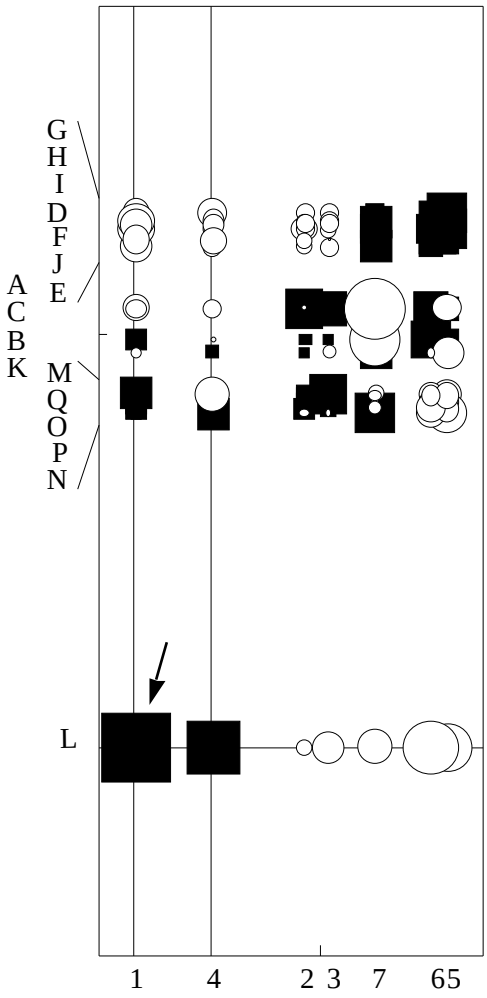

Fig. 4. Graphical display of the data matrix after double centering (Table 5). Squares are proportional to positive values and circles to negative ones. They are arranged according to the first PCA factor coordinates for the corresponding compounds and organisms. For example, the square marked with an arrow corresponds to the toxicity of compound 1 (2-methyl-2,4-pentanediol) to organism L (*Arbacia punctulata*, sperm cell).

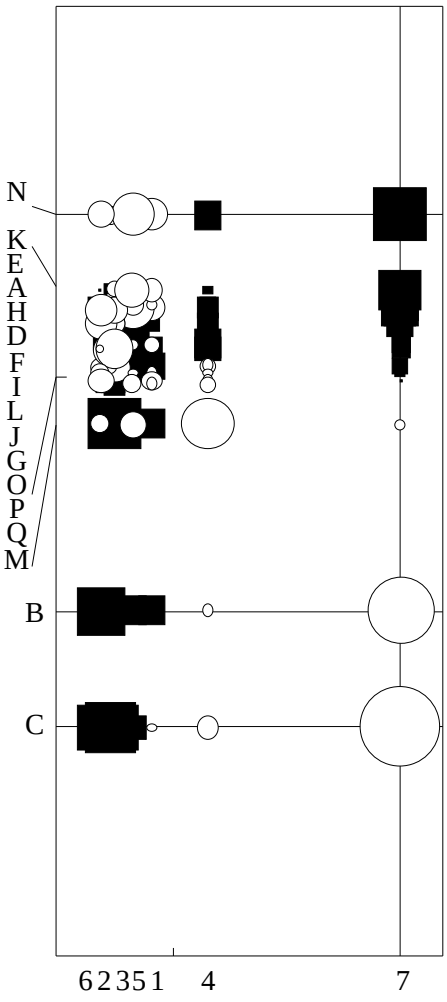

Fig. 5. Graphical display of the residuals between the doubly centered values and the reconstitution of the toxicity table by the first principal component. Squares are proportional to positive values and circles to negative ones. They are arranged according to the second PCA factor coordinates for the corresponding compounds and organisms.

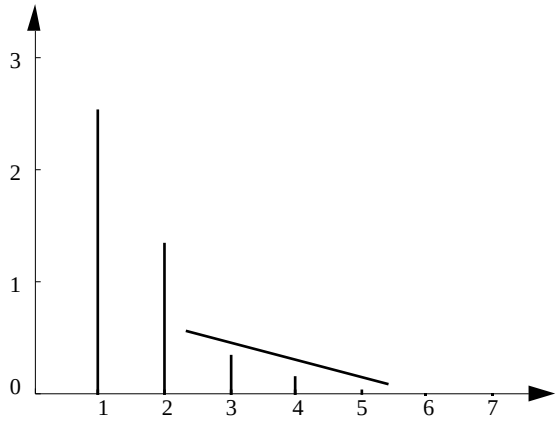

Fig. 6. Bar chart of PCA eigenvalues.

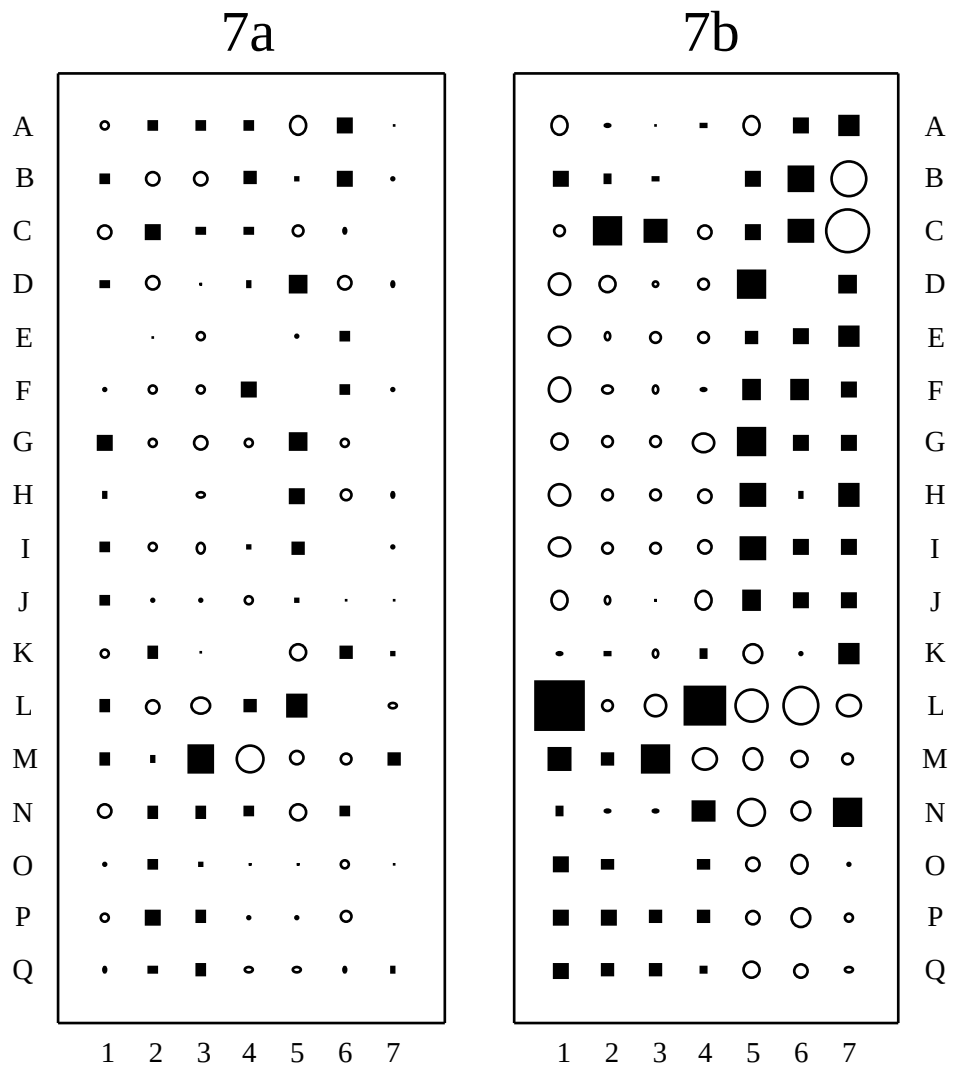

Fig. 7. 7a: Graphical display of the residuals between the doubly centered values and the reconstitution of the toxicity table by the first two principal components. 7b: Graphical display of the doubly centered values (same representation as figure 3, but with a different scale). Squares and circles in graphs 7a and 7b are at the same scale.

Table 1. Identity of the organisms.

Species	code
Daphnia magna (LC50)	A
Tanytarsus dissimilis	B
Orconectes immunis	C
Rana catesbeiana	D
Oncorhynchus mykiss	E
Lepomis macrochirus	F
Gambusia affinis	G
Ictalurus punctatus	H
Carassius auratus	I
Pimephales promelas	J
Arbacia punctulata (early embryo, [^{3}H] thymidine incorporation)	K
Arbacia punctulata (sperm cell)	L
Photobacterium phosphoreum	M
Arbacia punctulata (early embryo, fluorometric determination of DNA)	N
Daphnia magna (EC50)	O
Daphnia pulex	P
Ceriodaphnia reticulata	Q

Table 2. Test chemicals.

Chemical	code
2-methyl-2,4-pentanediol	1
2-methyl-1-propanol	2
2,2,2-trichloroethanol	3
2,4-pentanedione	4
2-chloroethanol	5
Hexachloroethane	6
Pentachlorophenol	7

Table 3. Data matrix.

	1	2	3	4	5	6	7
A	1.224	1.824	3.004	3.323	2.579	5.241	6.264
B	1.438	1.54	2.63	2.801	2.833	5.284	4.024
C	0.679	1.893	2.855	2.293	2.547	4.943	<3.163
D	1.001	1.39	2.807	2.939	3.636	4.926	6.103
E	1.097	1.747	2.815	3.086	3.17	5.303	6.364
F	0.967	1.666	2.87	3.198	3.358	5.441	6.12
G	1.108	1.614	2.719	2.751	3.724	5.234	5.975
H	0.981	1.706	2.785	2.975	3.609	5.059	6.304
I	0.991	1.562	2.68	2.917	3.486	5.222	6.004
J	1.134	1.69	2.936	2.848	3.319	5.279	6
K	1.164	1.564	2.516	2.978	2.13	4.589	5.948
L	4.24	2.014	2.709	5.046	2.165	3.91	5.471
M	1.912	1.782	3.532	2.429	2.314	4.455	5.425
N	1.582	1.709	2.865	3.672	2.102	4.466	6.522
O	1.567	1.756	2.683	3.125	2.401	4.374	5.425
P	1.554	1.829	2.852	3.125	2.374	4.26	5.384
Q	1.625	1.791	2.896	3.125	2.401	4.542	5.471

Table 4. Data matrix after centering by column.

	1	2	3	4	5	6	7
A	-0.2033	0.1136	0.1714	0.2271	-0.2532	0.3864	0.6189
B	0.0107	-0.1704	-0.2026	-0.2949	0.0008	0.4294	-1.6211
C	-0.7483	0.1826	0.0224	-0.8029	-0.2852	0.0884	-2.4821
D	-0.4263	-0.3204	-0.0256	-0.1569	0.8038	0.0714	0.4579
E	-0.3303	0.0366	-0.0176	-0.0099	0.3378	0.4484	0.7189
F	-0.4603	-0.0444	0.0374	0.1021	0.5258	0.5864	0.4749
G	-0.3193	-0.0964	-0.1136	-0.3449	0.8918	0.3794	0.3299
H	-0.4463	-0.0044	-0.0476	-0.1209	0.7768	0.2044	0.6589
I	-0.4363	-0.1484	-0.1526	-0.1789	0.6538	0.3674	0.3589
J	-0.2933	-0.0204	0.1034	-0.2479	0.4868	0.4244	0.3549
K	-0.2633	-0.1464	-0.3166	-0.1179	-0.7022	-0.2656	0.3029
L	2.8127	0.3036	-0.1236	1.9501	-0.6672	-0.9446	-0.1741
M	0.4847	0.0716	0.6994	-0.6669	-0.5182	-0.3996	-0.2201
N	0.1547	-0.0014	0.0324	0.5761	-0.7302	-0.3886	0.8769
O	0.1397	0.0456	-0.1496	0.0291	-0.4312	-0.4806	-0.2201
P	0.1267	0.1186	0.0194	0.0291	-0.4582	-0.5946	-0.2611
Q	0.1977	0.0806	0.0634	0.0291	-0.4312	-0.3126	-0.1741

Table 5. Data matrix after double centering (rows and columns).

	1	2	3	4	5	6	7
A	-0.3548	-0.0380	0.0199	0.0755	-0.4048	0.2349	0.4673
B	0.2747	0.0936	0.0614	-0.0309	0.2648	0.6934	-1.3571
C	-0.1733	0.7576	0.5974	-0.2279	0.2898	0.6634	-1.9071
D	-0.4840	-0.3781	-0.0833	-0.2146	0.7461	0.0137	0.4002
E	-0.4994	-0.1325	-0.1867	-0.1791	0.1686	0.2793	0.5498
F	-0.6348	-0.2190	-0.1371	-0.0725	0.3512	0.4119	0.3003
G	-0.4231	-0.2002	-0.2174	-0.4488	0.7879	0.2756	0.2261
H	-0.5921	-0.1502	-0.1934	-0.2668	0.6309	0.0586	0.5131
I	-0.5026	-0.2147	-0.2188	-0.2452	0.5875	0.3012	0.2926
J	-0.4087	-0.1358	-0.0120	-0.3633	0.3714	0.3090	0.2395
K	-0.0477	0.0692	-0.1010	0.0977	-0.4866	-0.0500	0.5185
L	2.3617	-0.1474	-0.5746	1.4991	-1.1182	-1.3956	-0.6251
M	0.5632	0.1500	0.7779	-0.5885	-0.4398	-0.3211	-0.1417
N	0.0804	-0.0757	-0.0418	0.5018	-0.8045	-0.4628	0.8026
O	0.2922	0.1980	0.0029	0.1815	-0.2788	-0.3281	-0.0677
P	0.2724	0.2643	0.1652	0.1748	-0.3125	-0.4488	-0.1154
Q	0.2759	0.1588	0.1416	0.1072	-0.3531	-0.2344	-0.0959

Table 6. Residuals between the doubly centered values and the reconstitution of the toxicity matrix by the first principal component.

	1	2	3	4	5	6	7
A	-0.2157	-0.0265	0.0125	0.1567	-0.5015	0.1498	0.4248
B	0.2452	0.0912	0.0630	-0.0481	0.2853	0.7114	-1.3480
C	-0.0431	0.7683	0.5905	-0.1519	0.1993	0.5837	-1.9468
D	0.0745	-0.3321	-0.1131	0.1115	0.3578	-0.3280	0.2294
E	-0.0384	-0.0945	-0.2113	0.0901	-0.1519	-0.0028	0.4088
F	-0.0805	-0.1733	-0.1667	0.2512	-0.0342	0.0728	0.1307
G	0.2180	-0.1474	-0.2516	-0.0745	0.3422	-0.1166	0.0300
H	0.0080	-0.1008	-0.2254	0.0836	0.2137	-0.3086	0.3295
I	0.0769	-0.1670	-0.2497	0.0932	0.1846	-0.0534	0.1153
J	0.0917	-0.0946	-0.0387	-0.0711	0.0235	0.0028	0.0864
K	-0.1479	0.0609	-0.0957	0.0391	-0.4169	0.0113	0.5492
L	0.1560	-0.3292	-0.4570	0.2110	0.4160	-0.0460	0.0498
M	0.2435	0.1237	0.7950	-0.7752	-0.2175	-0.1255	-0.0439
N	-0.3394	-0.1103	-0.0195	0.2566	-0.5126	-0.2059	0.9310
O	-0.0682	0.1683	0.0221	-0.0289	-0.0282	-0.1076	0.0426
P	-0.1250	0.2316	0.1864	-0.0572	-0.0362	-0.2057	0.0062
Q	-0.0552	0.1315	0.1593	-0.0862	-0.1229	-0.0318	0.0054