NATIONAL INSTITUTE OF PUBLIC HEALTH AND ENVIRONMENTAL PROTECTION
BILTHOVEN, THE NETHERLANDS

RIVM report no 719102 028

Ordering aquatic species by their sensitivity to chemical compounds: a principal component analysis of acute toxicity data.

M.A. Vaal, J.T. Van der Wal and J.A. Hoekstra

March 1994

This research was carried out on behalf of the Directorate-General for Environmental Protection, Directorate of Chemicals, External Safety and Radiation Protection, in the frame of the project no. 719102.

MAILING LIST

1-4 Directoraat-Generaal Milieubeleid, Directie Stoffen, Veiligheid en Straling
5 Plv. Directeur-Generaal Milieubeleid, dr.ir.B.C.J. Zoeteman
6 Ir. J. Hendriks (RIZA)
7 Dr. J. Hermens (RITOX)
8 Drs. H. Schobben (TNO)
9 Prof.dr. S.A.L.M. Kooijman (VU)
10 Depôt Nederlandse publikaties en Nederlandse bibliografie
11 Directie RIVM
12 Sectordirecteur Stoffen en Risico's, dr.ir. G. de Mik
13 Adviescentrum Toxicologie, drs. A.G.A.C. Knaap
14 Laboratorium voor Ecotoxicologie, prof.dr. H.A.M. de Kruijf
15 Centrum voor Wiskundige methoden, drs. A. van der Giessen
16 Dr. W. Slob
17 Drs. J.H. Canton
18 Dr.ir. J. Notenboom
19 Dr. W. Slooff
20 Ir. C.J. Roghair
21 Dr.ir. J. Struijs
22 Drs. D. de Zwart
23-25 Auteurs
26 Hoofd bureau Voorlichting en Public Relations
27-28 Bibliotheek RIVM
29 Bibliotheek RIVM, depôt ECO/LWD
30-31 Archief ECO/AET
32 Bureau Projecten- en Rapportenregistratie
33-45 Reserve-exemplaren

CONTENT

LIST OF TABLES AND FIGURES

SUMMARY

The research described in this report is part of a project directed at the variation in the sensitivity of species to toxicants. The object of this project is to develop Quantitative Species Sensitivity Relationships (QSSRs) which predict the sensitivity of a species to a chemical compound.
This report presents the analysis of a species by chemical compound matrix of acute toxicity values. A statistical pattern recognition techniques (principal component analysis) was used to find structure in the variation of species sensitivity to chemical compounds. Patterns may identify groups of species that have similar toxicological profiles. The data were obtained from the database Aquire and literature surveys.
The data matrix consisted of 26 species and 21 compounds. The species included were eight fishes, two amphibians, twelve arthropods, one mollusc, one annelid, one planarian and one coelenterate. The compounds included in the analyses were three heavy metals and eighteen organic compounds. With respect to their mode of action the organic compounds can be classified as six non-polar narcotics, four polar narcotics, two ester narcotics, one reactive compound and five compounds with a specific mode of action.

The conclusions are as follows:
Major part of the variation in species sensitivity is determined by the toxicity of compounds and not by intrinsic differences between the species. Practically every species can be used to order compounds with respect to their average toxicity, since the *pattern in compound toxicity* could be described by one single component. This component explained about 80% of the among-compounds variation in toxicity, as measured by the 26 species. Table 19 presents the rank order of compounds, as calculated by this component, with respect to their average toxicity. The compounds with the highest overall toxicity also had the largest variation in toxicity for different species. An exception is aniline which had a high among-species variation, mainly due to sensitive daphnids, but was on average not that toxic (log LC_{50} = 4.79 µg/l).
The toxicity of non-polar narcotics correlates well with the log K_{ow} (Figure 6). The observed toxicity scores were less than a factor 10 away from the values predicted by the regression line. Other organic chemicals within a factor 10 away from the regression were the polar and ester narcotics (except allylamine), pentachlorophenol and lindane. More toxic than predicted by their log K_{ow} were: allylamine, dieldrin, malathion, parathion and salicylaldehyde.

The *patterns in the sensitivity of the species* are less unambiguous. However, a three component model of the species-by-compounds matrix explained about 56% of the among-species variation in sensitivity. Graphical analysis of the model and comparison with the original data matrix led to an ordering of species with respect to the nine compounds with highest among-species variation (Table 18). Fishes had largely the same pattern in sensitivity as the other vertebrates, the amphibians. They were more sensitive to dieldrin, lindane and pentachlorophenol than the invertebrates. The Phyllopoda (daphnids) were the most sensitive species to aniline, the heavy metals, malathion and parathion. The insect species were not very sensitive compared to the other groups of species: according to model estimation they are moderate sensitive to parathion, malathion, dieldrin and lindane.

In future work the analysis will be extended to chronic toxicity data.

SAMENVATTING

Het in dit rapport beschreven onderzoek maakt deel uit van het project dat gericht is op de oorzaken van de spreiding in de gevoeligheid van soorten voor chemische stoffen. Het doel van dit project is het ontwikkelen van Quantitative Species Sensitivity Relationships (QSSRs). Deze modellen voorspellen de gevoeligheid van een soort voor een chemische stof.
In dit rapport wordt de analyse beschreven van een soorten-bij-stoffen matrix met toxiciteitswaarden die afkomstig zijn uit kortdurende experimenten. Een statistische patroonherkenningstechniek (principale componenten analyse) is gebruikt om de patronen in de variatie van de gevoeligheid van soorten voor stoffen te analyseren. Zulke patronen kunnen worden gebruikt om soorten te groeperen die eenzelfde toxicologisch profiel hebben.
De data zijn afkomstig van de Aquire database en aanvullende literatuurstudies.
De matrix bevatte toxiciteitsgegevens van 26 soorten en 21 stoffen. De soorten zijn onder te verdelen als acht vissen, twee amfibiën, twaalf arthropoden, een weekdier, een wormachtige, een platworm en een holtedier. De stoffen in de studie waren drie zware metalen en achttien organische stoffen. Deze organische stoffen zijn onder te verdelen op grond van hun toxicologisch werkingsmechanisme als zes apolair narcotische, vier polair narcotische, twee ester narcotische, een reactieve stof en vijf stoffen met een specifiek werkingsmechanisme.

De conclusies zijn als volgt:
Het grootste deel van de variatie in de gevoeligheid van soorten wordt bepaald door de toxiciteit van stoffen en slechts in geringe mate door intrinsieke verschillen in de gevoeligheid tussen soorten. Om de stoffen te rangschikken naar hun gemiddelde toxiciteit voor de 26 soorten kan, met hetzelfde resultaat, vrijwel iedere soort gebruikt worden. De *patronen in de toxiciteit van stoffen* lieten zich namelijk door één component beschrijven. Deze verklaarde ongeveer 80% van de tussen-stoffen spreiding in de toxiciteit. In Tabel 19 staat de rangvolgorde van de stoffen, zoals berekend door deze component, gebaseerd op de gemiddelde toxiciteit. De stoffen die gemiddeld het meest toxisch waren vertoonden tevens de grootste spreiding in toxiciteit voor de verschillende soorten. Een uitzondering was aniline omdat deze stof een grote spreiding in toxiciteit vertoonde, maar niet zo toxisch was (log LC_{50} = 4.79 µg/l). Die grote spreiding was het gevolg van de gevoeligheid van daphnia's voor deze stof.
De toxiciteit van de apolaire narcotische stoffen correleert goed met de log K_{ow} (Figuur 6). De afwijking tussen de waargenomen toxiciteitswaarden en de door de regressielijn voorspelde waarde was niet meer dan een factor 10. Andere organische stoffen die minder dan een factor 10 van de regressielijn lagen waren de polaire en ester narcotische stoffen (behalve allylamine), pentachloorfenol en lindaan. Allylamine, dieldrin, malathion, parathion en salicylaldehyde waren toxischer dan kan worden voorspeld op grond van hun log K_{ow}.

De *patronen in de gevoeligheid van soorten* zijn niet zo eenduidig. Een drie-componenten model van de soorten-bij-stoffen matrix verklaarde ongeveer 56% van de tussen-soorten variatie in gevoeligheid. Soorten konden worden geordend, gebaseerd op hun gevoeligheid voor de negen stoffen met de grootste tussen-soorten variatie, door grafische analyse van het model en vergelijking met de originele gegevens in de matrix (Tabel 18). Vissen hadden een gevoeligheidspatroon dat vergelijkbaar was met dat van de andere vertebraten in de analyse, de amfibiën. Zij waren voor de stoffen dieldrin, lindaan en pentachloorfenol gevoeliger dan de evertebraten. De Phyllopoda (de daphnia's) waren de meest gevoelige organismen voor

aniline, de zware metalen, malathion en parathion. De insecten waren niet veel gevoeliger dan andere groepen van soorten: volgens de schatting van het model zijn zij gematigd gevoelig voor parathion, malathion, dieldrin en lindaan.

In een vervolgonderzoek zullen gegevens afkomstig van langdurend toxiciteitsexperimenten worden geanalyseerd.

1 INTRODUCTION

The research described in this report is part of a project directed at the variation in the sensitivity of species to toxicants. The object of this project is to develop Quantitative Species Sensitivity Relationships (QSSRs) which predict the sensitivity of a species to a chemical compound.
The development of QSSRs may give answer to two questions: first, how can differences between species be explained, and second, how can this information be used for generalization and prediction? More systematic knowledge concerning species differences may help to predict the sensitivity of untested species-compound combinations. In general no or only few toxicity data are available for species that have the particular interest of environmental and nature policy makers. Estimation of the sensitivity of these 'species of concern' to compounds of interest may improve the development and evaluation of environmental management. The development of QSSRs may also help to determine the choice of appropriate species for toxicity tests and strategic choices of input data for the derivation of safe environmental concentrations.
Earlier results of this project are published in a report describing the results of a pilot study (Hoekstra et al., 1992), an article about the concept of QSSRs (Vaal et al., 1993) and an article about a quantitative measure of the variation in species sensitivity, the sensitivity ratio (Hoekstra et al., in press).

QSSRs are developed by a two way approach. The first approach is directed at the development of comparative toxicological models. Such mechanistic models describe the relationship between effect concentrations and biological characteristics of species (see Vaal et al., 1993).
In the second approach statistical pattern recognition techniques are used to find structure in the variation of species sensitivity to chemical compounds. These techniques are applied to a large data matrix of effect concentrations of species-compound combinations. Patterns may identify groups of species that have similar toxicological profiles: such species show a similar pattern in their sensitivity to compounds. This leads to the development of empirical QSSRs. An analysis of the biological similarities in such groups may elaborate relevant biological characteristics that determine species sensitivity.
This report describes the first results of this pattern recognition approach. We used principal component analysis (PCA) to a species-by-compound data matrix to identify patterns in species sensitivity to compounds. A short description of this technique is presented in 2.2. In the presentation of the results (Chapter 3) and the discussion (Chapter 4) those who are not familiar with this technique are taken into account.

2 METHODS

2.1 Selection of data

A search of the database Aquire (US-EPA, Duluth, MN, version April 4th, 1993) was performed to achieve data for this study. A data matrix was filled with aquatic toxicity data of as many species and compounds as possible. A species was included if the resulting matrix contained toxicity data for that particular species for at least 40% of the compounds. Compounds were selected similarly. The Aquire search included LC_{50}'s and also EC_{50}-immobility of arthropods of experiments of 48-96 hours exposure time (24 hours exposure time of *Culex pipiens* and *Aedes aegypti* were also included). Data from abstracts or in foreign languages were not selected. The following criteria were applied to ensure a reasonable quality of the data:

Included were data:

- of compounds with, if stated, a purity or with an active ingredient equal to or larger than 80%
- of experiments with, if stated, a static, renewal or flow-through test system
- from ranges between which the effect concentration would fall, in that case the geometric mean of the upper and lower limit was calculated; results with a lower or higher limit were not included
- lower than two times the water solubility limit
- of pentachlorophenol if the pH, if stated, fell between 6.5 and 8.5. This was used as an alternative criterium to the water solubility criterium above.
- of heavy metals if the hardness, if stated, was 'hard' or larger than 150 mg $CaCO_3$/l.

If more data per species-compound combination were available the geometric mean was taken. However, the resulting standard deviations of some of these geometric means were rather large (> 0.5 on log scale). In those cases additional criteria were used to exclude outlying data (see Appendix 1). Most of the variation per compound-species combination was due to different life stages (insensitive fish eggs, one adult versus many juveniles) or extreme experimental conditions (e.g. temperature, hardness).
Besides the Aquire database other literature sources were searched selectively to fill in gaps of the matrix.

2.2 Principal components analysis

Principal component analysis (PCA) is a multivariate statistical technique to detect patterns in a large data matrix. A data matrix is observed as a matrix of rows of objects and columns of variables. We present a short description of the PCA technique, more detailed information can be found in Kowalski (1983).
In our case, the objects are the species, and the variables the LC_{50}'s for the chemical compounds. This corresponds with an analysis of the *patterns in species sensitivity*. The opposite case, LC_{50}'s of compounds as objects and species as variables, describes the *patterns in toxicity of compounds*. The first case is used in the explanation of the PCA method.

The toxicity of a compound is one axes in a multidimensional space which is shaped by all

the compounds in the analysis, see Figure 1. The toxicity of each of these compounds to one species situates the species in this space. Species that have similar sensitivity to each of the compounds will be situated close to each other. Thus, the data of all species in the matrix are represented by a swarm of points in this space. In the PC analysis a simplification of this matrix and a reduction of data is obtained when the matrix is projected on a few dimensional space (like a line or plane). The reduction is successful when only a few components have to be used to describe most of the variation in the data matrix.

In the PC analysis a representation of the matrix can be decomposed into the average LC_{50} of a compound $\bar{x}$, the product of the loading matrix P and the score matrix T, plus a matrix of deviations, residuals E (Figure 1). This is described as

$$\mathbf{X}=\mathbf{1}\cdot\bar{x}+\mathbf{T}\cdot\mathbf{P}+\mathbf{E}$$

The score matrix T describes the projection of the n species on the hyperplane with as dimensions the components defined by the loading matrix P. The residual matrix E contains what is left over. When these residuals are small compared to the variation in X, the PC model is a good representation of X itself. When the number of components is two or three, the columns in T can be plotted against each other to get a few two dimensional pictures of the species and their relations in the multidimensional space. The loadings of the compounds can also be plotted against each other to show the relation between the compounds and to interpret the components. Comparison of a loading plot with its accompanying score plot indicates the patterns in sensitivity: the position of a species in the score plot indicates its sensitivity to a compound with a corresponding position in the loading plot.
The significance of a component is calculated by cross validation. This is a procedure that during model fitting keeps out data and than predicts these values by the developing model. This is repeated until all data have been left out once an only. The lower the PRESS/SS value in the tables the higher the significance of the component or the contribution of the variable to the model.

PCA provides an efficient way to convert a data matrix into a few informative pictures showing the relation between the species as measured by the compounds. We used the software package SIMCA (Umetri, 1992). SIMCA works well with moderate amounts of missing values.

Data transformation prior to the analysis
Data were ^{10}log transformed. Subsequently they were multiplied by -1 to facilitate the interpretation of the PC analyses (species with high, positive scores on a component are sensitive species).
The principal component model in SIMCA is fitted to the data using a method of least squares. To give each variable equal influence on the determination of a component, the data should be centred and scaled to unit variance. Centring of data corresponds to moving the coordinate system to be centred in the point $\bar{x}$, the averages of the p compounds. This is calculated by subtracting the averages of each compound from the data of that compound. After centring, a compound with a large average log LC_{50} does not have more influence on the determination of a component than a compound with a small average log LC_{50}. Scaling

is performed to prevent compounds with large variation to have more influence than compounds with less variation. It is computed by dividing the data by the standard deviation of the centred data of the compound. After scaling the axis of each compound has the same length.

Analyses performed

An analysis of variance of the toxicity of compounds to species was calculated by a two way ANOVA in the software package Excel (version 4.0a, Microsoft, 1992), using species and compounds as sources of variation.

The following PCA analyses were performed:

1) Patterns in species sensitivity:
   - species as objects and compounds as variables, centred and scaled data.
   - species as objects and compounds as variables, centred and unscaled data.
2) Patterns in toxicity of compounds:
   - compounds as objects and species as variables, centred and scaled data.

The K nearest neighbour analysis of SIMCA was used to classify the three nearest species neighbours in the multidimensional space. This analysis was used to qualitatively indicate the role of the taxonomic groups to the patterns in species sensitivity.

3 RESULTS

3.1 Data, descriptive statistics

The resulting matrix consisted of 26 species and 21 compounds. The species included eight fishes, two amphibians, twelve arthropods, one mollusc, one annelid, one planarian and one coelenterate (Table 1). The compounds included in the analyses were three heavy metals and eighteen organic compounds. According to the classification of Verhaar et al. (1992), the organic compounds can be classified as six narcotics (class I), four polar narcotics (II), three reactive compounds (III) and five compounds with a specific mode of action (IV); the ester narcotics cannot be classified as such (Table 2).
In Table 3 the matrix with the log LC_{50}-value per compound-species combination is presented. Toxicity data are expressed as µg/l, the metals as µg/l metal ion. The final matrix contained toxicity data for nearly 80% of the species-compound combinations. Of this data 15.8% was obtained by additional literature search.
31.5% of the data were averages of several log LC_{50}-values. The standard deviations and the number of data per combination are presented in Appendix 3 and 4, respectively. For 7.6% of these means the application of the criteria mentioned in Appendix 1 did not reduce the standard deviation below 0.5.

Table 3 shows that sensitivity of species, averaged over all compounds, ranged from 3.45 to 5.12 (log LC_{50}-values, column average in Table 3). Toxicity of compounds, averaged over all species, ranged from 1.15 to 6.99 (log LC_{50}-values, row average in Table 3). The standard deviations of species sensitivity, averaged over all compounds, ranged from 0.93 to 2.14 (column standard deviations, Table 3). The standard deviations of the toxicity of compounds, averaged over all species, were smaller: they ranged from 0.18 to 1.39.

Analysis of variance
The results of the two way ANOVA is presented in Table 4. It shows that the major part of the variance in the data matrix was due to differences between the compounds (MS = 45.44, $p < 0.001$). Differences between species contributed less to the total variance (MS = 1.00, $p < 0.001$). To give a further appreciation of the relative magnitude and consistency in the variation between species over compounds, we estimated the variance components of the model:
$^{10}\log LC_{50}$ = overall mean + species contribution + compound contribution + 'remainder'.
The remainder does not only absorb random variation but also interactions between species and compounds. That is, if reversals occur in the sensitivity ordering of species with respect to the compounds, these will increase the remainder component. The method used for variance components estimating was equating the adjusted mean squares to their expectation (Fraybill, 1961).
The squared roots from the estimated variance components were:

between species :	$\sigma_{species}$	= 0.17
between compounds:	$\sigma_{compounds}$	= 1.48
remainder:	$\sigma_{remainder}$	= 0.67

This gives clear evidence that a most sensitive species over all compounds included did not exist, and that variation in toxicity between compounds was the largest source of variation

in the compounds x species table.

3.2 Principal component analysis

3.2.1 Patterns in species sensitivity

Species as objects and compounds as variables, centred and scaled data

The results of the modelling components are presented in Table 5. Five components contributed significantly to the model. Only the first three components explained a considerable amount of the matrix variance (> 10% each), they are discussed in this section. Principal component 1 (PC1) explained 17.9% of the matrix variance.

The contribution of the compounds to the model is presented in Table 6. The importance of a compound in a PC model is indicated by the size of its explained variance. The PRESS/SS (prediction sum of squares, i.e. the squared difference between observed and predicted values by cross validation/residual sum of squares of previous dimension, for explanation see 2.2) of ethylpropionate, heptanol and salicylaldehyde was low. Only these compounds were at all modelled by this component.

PC1 loadings of the compounds are presented in Table 7. The contrast between the loadings (the occurrence of as well positive as negative values) indicated that species could not simply be arranged in order of sensitivity: some were relatively sensitive to some compounds but relatively insensitive to others. The compounds dieldrin, malathion and parathion (class IV) had the largest negative loadings. High positive loadings were calculated for o-cresol, ethylacetate, ethylpropionate, heptanol, phenol and salicylaldehyde (non-polar, polar and ester narcotics) and PCP (IV). Thus, PC1 points to a contrast between the three compounds from class IV (except PCP) and the narcotics.

Table 8 presents the scores, residual standard deviations of the objects and the object leverages. Species with high positive scores are sensitive to compounds with high positive loadings, and vice versa. The residual standard deviation of an object is proportional to the object distance to the PC model.

The fishes *Oncorhynchus mykiss*, *Oryzias latipes*, *Pimelas promelas* and the insect *Nemoura cinerea* had the highest positive scores. They were relatively sensitive for the compounds with a high positive loading and relatively insensitive for dieldrin, malathion and parathion. The opposite applied for the insects *Culex pipiens*, *Aedes aegypti*, *Corixa punctata*, *Ischnura elegans* and the annelid *Erpobdella octoculata* who had high negative scores. All these species had a high leverage and not a large residual standard deviation, which indicates that they reinforce this component.

The second component (PC2) modelled 12.3 % of the matrix variance (Table 5). Compounds modelled by PC2 are dieldrin, lindane, malathion and parathion (Table 6).

Loadings of PC2 are presented in Table 7. The organophosphatic acetylcholinesterase (Ache) inhibitors malathion and parathion, the heavy metals and aniline had a high positive loading. Benzene and pyridine and the neurotoxic compounds lindane and dieldrin had a high negative loading.

Table 8 shows that all three daphnid species had a high positive score for PC2: they were relatively sensitive to parathion, malathion, aniline and the heavy metals. The insects *C. punctata*, *I. elegans* and *N. cinerea* and the fishes *O. mykiss* and *P. promelas* were relatively insensitive to these compounds and relatively sensitive to lindane, dieldrin, benzene and pyridine.

In Figure 3 the species scores of PC1 and PC2 are plotted. The position of species in a given direction in a score plot is influenced by compounds lying in the same direction in the loading plot. The daphnids could be clearly distinguished from all other species especially because of the discriminating effect of PC2. They were relatively sensitive to mercury, aniline and cadmium (compare with loading plot, Figure 2). The fishes were relatively sensitive to the compounds at the right end of the PC1 axes that are salicylaldehyde (class III), some class I and II narcotics, the ester narcotics and PCP. Compared to the other insect species *N. cinerea* had a very deviate position in this plot, mainly caused by its position on PC1.
The loading plot of the compounds of PC1 and PC2 (Figure 2) points that the ester and polar narcotics could be distinguished from the non-polar narcotics because the first group had a high loading for PC1 and the second group a high negative loading for PC2. Exceptions to this observations were heptanol and trichloroethylene. This plot indicates a difference between the specific mode of actions of PCP (uncoupler of oxidation), malathion and parathion (AChE-inhibitors), dieldrin and lindane (neurotoxic). The metals were situated close to each other with aniline nearby.

The third component, PC3, modelled 11.6 % of the matrix variance (Table 5). This was mainly due to allylamine, mercury and parathion (Table 6).
With PC3 included in the model 12 of the 21 compounds had more than 50% of their variance explained (Table 6). Parathion, ethylpropionate and mercury had more than 75% of their variance explained.

Species as objects and compounds as variables, centred and unscaled data
The first three components contributed considerably to the model: together they explain 55.7% the matrix variance (Table 8). The initial matrix variance was 0.48, with scaled data the initial matrix variance is 1.00. The description of PC1-PC3 is presented in the Tables 9-12. Regarding PC1 aniline, malathion, parathion and the heavy metals had the largest positive loadings (Table 11). They were the compounds with the largest standard deviations when averaged over all species (see column standard deviation, Table 3) which gave them more weight in this unscaled analysis compared to the scaled analysis.
Daphnids had the largest positive scores, they were relatively the most sensitive organisms for the above mentioned compounds. The fish species *P. promelas*, *Carassius auratus*, *Poecilia reticulata* and the insect *Ischnura elegans* were the least sensitive species for these compounds.
The PC2 loadings of the compounds (Table 11) showed a contrast between parathion (high negative) and cadmium and PCP (high positive). In general the insects were relatively sensitive to the first, while fishes were relatively sensitive to the other two compounds (Table 12).
Figures 4 and 5 are plots of the loadings of PC1 and PC2 and the scores of PC1 and PC2 respectively. The loading plot shows that compounds that are situated near the end of either axis could be arranged in three groups: malathion and parathion; the heavy metals and aniline; PCP. All other compounds were situated near the centre of both axes. The score plot in Figure 5 demonstrates that the arthropods were relatively sensitive to the compounds with a positive loading on PC1 and a negative loading on PC2: malathion, parathion, the heavy metals and aniline. Exceptions were the insects *N. cinerea* and *Chironomus thummi*.
With the third PC included in the model the following compounds had a considerable percentage of their initial variance (> 50%) explained (Table 10): aniline, dieldrin, lindane,

malathion, parathion, PCP and the heavy metals. These were the compounds that had the highest overall toxicity and the largest standard deviation, when averaged over all species (Table 3). They had more weight in this unscaled analysis compared to the scaled analysis.

K nearest neighbours of species, scaled data
The summarized results of the Knn analysis of species neighbours are presented in Table 13. Vertebrates were mainly situated close to other vertebrates. Insects were mainly situated close to other insects or, in the case of *C. thummi* and *N. cinerea*, close to fish species. All three daphnid species were situated near each other.
These results confirm the observations of the PCA-analyses.

K nearest neighbours of species, unscaled data
The results of this Knn analysis resembled those of the scaled analysis. They were presented in Table 14. *C. thummi* and *N. cinerea* had even a more prominent position among the fishes which caused the fishes to have mainly insects as there nearest neighbours.

3.2.2 Patterns in the toxicity of compounds

Compounds as objects and species as variables, centred and scaled data
The results of this analysis is presented in Table 15. The first component modelled the major part of the matrix variance: 78.6%. PC2 - PC8, although all significant, only explained < 3.5% each of the variance. In Tables 16 and 17 a description of PC1 is presented.
All PC1 loadings were positive, which means that all species have approximately the same weight in arranging the compounds by their toxicity. The compound scores give an overall measure of toxicity. Dieldrin, mercury and parathion were the most toxic compounds and had the highest scores; acetone and propanol were the least toxic compounds and had the lowest scores.

4 DISCUSSION

Selection of data

The database Aquire contains data of a broad diversity of ecotoxicological experiments. This forced us to apply specific selection criteria to ensure a reasonable quality of the data and to reduce standard deviations per species-compound combination. Most variation per species-compound combination was due to differences in sensitivity between life stages. Modifications in body size, metabolism, rate of biotransformation and body lipid content and composition during ontogeny are supposed to be important factors determining differences in sensitivity between life stages. Therefore, it would be better to differentiate between life stages in our analyses. However, it would not be possible then to construct a sufficiently large and filled data matrix.

Due to the toxicokinetic behaviour of compounds, toxicity may increase with increasing exposure time until equilibrium between the internal organism concentration and the external toxicant concentration is reached. In this study the results of 48-96 hours exposure duration were included. In particular for larger species and more hydrophobic compounds this exposure period might have been to short to reach this equilibrium. Usually, in long-term experiments lower test concentrations are applied compared to short-term experiments so that other modes of action can be involved (De Bruijn and Hermens, 1991). Therefore, analysis with data from long-term toxicity experiments may result in different patterns than the ones observed in this report. We will compare these patterns in future research.

Patterns in species sensitivity

As has been stated before (e.g. Cairns, 1986; Slooff, 1983; Suter, 1993) there is not one species that is the most sensitive to all compounds. This is clearly demonstrated in our results. If it would be possible to arrange species with respect to their sensitivity to all compounds, one principal component should have been sufficient to explain most of the variation in the data matrix. As was to be expected, the patterns in species sensitivity are more diffuse, species that are very sensitive to one (group of) compound might be among the least sensitive species for other compounds.

The unscaled analysis was mainly determined by compounds with a high overall toxicity and a large standard deviation when averaged over all species. Our following description of the patterns in species sensitivity is focused on this analysis, because, from an environmental protection point of view, these compounds are the most threatening to species.

We tried to qualitatively describe the observed patterns in species sensitivity in Table 18, by graphical analysis of the model (PC1-PC3) and comparison with the original data matrix. Species were grouped taxonomically, and compounds according to their position on the PC-axes. The sensitivity of each species-by-compounds group combination was derived from interpretation of their scores and loadings in the model. Fishes had largely the same pattern in sensitivity as the other vertebrates, the amphibians. They were more sensitive to dieldrin, lindane and PCP in comparison to the invertebrates. Among the arthropods, the Phyllopoda (daphnids) were the most sensitive species. They were very sensitive to aniline and the heavy metals and malathion and parathion. For the purpose of QSSR development these observations should be interpreted in terms of modes of action of the compounds and relevant biological species characteristics. An explanation for the relative insensitivity of fishes and amphibians to malathion and parathion might be found in their ability to hydrolyse these organophosphates. This enhances their excretion, whereas insects (and possibly other

arthropods) are known to metabolize these compounds into the more toxic metabolites malaoxon and paraoxon. In our ordering daphnids appeared to be more sensitive then the insects.
The occurrence of the polar narcotic aniline among the class IV compounds and the heavy metals is unexpected. Aniline had not such a high overall toxicity compared to these compounds (Table 3). However, the daphnids appeared to be extremely sensitive to aniline compared to the other species, which caused the large overall standard deviation and thus the role of aniline in the analysis. This pattern was similarly to that observed with the heavy metals.
A more thorough search for explanations of the observed patterns will be presented as a result of our future activities together with the results of a study with data from chronic toxicity studies.

Patterns in the toxicity of compounds
In contrast to patterns in *species sensitivity*, the patterns in *compounds toxicity* could be largely modelled by one axis. This axis indicated an arrangement of the compounds by their average toxicity to all the species in this study. The species had a comparable weight in the arrangement of the compounds, i.e. practically every species can be used to order the toxicants with respect to their toxicity. The arrangement of compounds based on their scores in the PC analysis is presented in Table 19. The log LC_{50}-values in the Table were averaged over all species with missing species-compound combinations estimated by the model, using the first component.
For the purpose of the estimation of the toxicity of a non-tested chemical it may suffice to test just one 'average' sensitive species. However, part of the compounds did show a rather large variation between the more and less sensitive species. This is demonstrated by their sensitivity ratio: the ratio between the 95% and 5% percentile of the distribution of species sensitivity for the respective compounds (Hoekstra et al., in press). For aniline, malathion, parathion and the heavy metals the sensitivity ratio was larger than 1000 (Table 19). In these cases testing an 'average' sensitive species may underestimate the sensitivity of groups of very sensitive species.

Figure 6 demonstrates the relationship between the acute toxicity of the compounds (using the estimated averages from Table 19) and their hydrofobicity as expressed by the log K_{ow}. As known, toxicity increases with increasing log K_{ow}. This relationship is especially valid for compounds with a minimum toxicity: non-polar narcotics. The line drawn in the figure is based on the data of the non-polar narcotic compounds and, regarding McKim and Schmieder (1989), can be considered as the baseline toxicity. The toxicity of the non-polar narcotics in this study correlated well with the log K_{ow}. The observed toxicity scores of these compounds were less than a factor 10 away from the values predicted by the regression line. The regression equation for this line is:
$\log LC_{50} = 6.47 - 0.77 (\log K_{ow})$, $r^2 = 0.91$, $p \leq 0.01$.
Compounds with a more than minimum toxicity are expected to lie below this line. Within a factor 10 under the line lay the polar and ester narcotic compounds (except allylamine), lindane and PCP. PCP lay somewhat above the minimum toxicity line, deviate from the other compounds with a specific mode of action. This can be explained by the fact that the pH in the experiments was such that PCP was ionized to the less hydrophobe phenolate form. Compounds that were more than a factor 10 toxic than predicted by the regression line are: allylamine, dieldrin, malathion, parathion and salicylaldehyde.

Using principal component analysis for the development of empiric QSSRs

PC analysis appeared to be a useful method to describe the patterns in a multivariate data matrix. In this study it resulted in a quantitative description of the patterns in species sensitivity and in the toxicity of compounds. However, if more than two components are needed to explain major part of the variance, some qualitative interpretation has to be used to describe the observed patterns in one simple table.

PC analysis can also be used for predictive purposes. The species PC model presented here described only 60% of the matrix variance, which does not give it much value for prediction.

Taxonomy and the choice of test organisms

Ecotoxicological protection levels are derived from observed interspecific variability in sensitivity (EPA, 1984; Van Straalen and Denneman, 1989; Wagner and Løkke, 1991). More insight in the background of differences in species sensitivity may help to improve the choice of test organisms that serve as representants of communities of species in ecosystems. The general feeling is that the test organisms used as input data for the derivation of safe environmental concentrations should not all belong to the same taxonomic entity. The analyses presented in this report indicate roughly that among species of the phylum Chordata (the fishes and amphibians) and the phylum Arthropoda distinct groups at the level of the taxonomic class can be identified.

Some papers have been published dealing with the taxonomic relationship and sensitivity of species (e.g. Slooff et al. 1983; Slooff and Canton, 1983; LeBlanc, 1984; Mayer and Ellersieck, 1986; Suter, 1993). These papers suggest that in general close taxonomic relationships imply some similarity in sensitivity. The use of taxonomic hierarchy is based on the assumption that evolutionary processes that resulted in increasing differences in morphological and biochemical traits used to classify species also resulted in increasing differences in biological factors that determine the sensitivity of an organism. This will only be true if those traits are not evolutionary labile or highly modified by adaptation to the specific environments that distinguish species from the same taxonomic group. A thorough investigation of the explanatory value of phylogeny to species sensitivity to toxicants has not yet been performed. To have a more thorough insight in the role of the taxonomic similarity on lower hierarchical levels than phylum and class much more species and taxa should be included in the analysis. However, the lack of useful toxicity data is an important limiting factor.

5 CONCLUSIONS

The results in this report show that major part of the variation in species sensitivity is determined by the toxicity of compounds and not by intrinsic differences between the species. Practically every species can be used to order compounds with respect to their average toxicity, since the pattern in compound toxicity could be described by one single component. Table 19 presents the calculated rank order of compounds with respect to their average toxicity. The ordering explains about 80% of the among-compounds variation in toxicity, as measured by the 26 species. The compounds with the highest overall toxicity also have the largest variation in toxicity for different species. An exception is aniline which has a high among-species variation but is on average not so toxic (log LC_{50} = 4.82 µg/l).

The toxicity of non-polar narcotics correlates well with the log K_{ow}. The observed toxicity scores are less than a factor 10 away from the values predicted by the regression line. Other organic chemicals within a factor 10 away from the regression are the polar and ester narcotics (except allylamine), lindane and PCP. PCP lies somewhat above the minimum toxicity line, deviate from the other compounds with a specific mode of action. More toxic than predicted by the regression are: allylamine, dieldrin, malathion, parathion and salicylaldehyde.

Patterns in species sensitivity are more diffuse. Only part of the variance (56%) in species sensitivity could be explained by principal components, leaving the remaining variance unexplained. Graphical analysis of the model and comparison with the original data matrix leads to an ordering of species with respect to the nine compounds with highest among-species variation (Table 18). Fishes have largely the same pattern in sensitivity as the other vertebrates, the amphibians. They are more sensitive to dieldrin, lindane and PCP than the invertebrates. Among the arthropod patterns the Phyllopoda (daphnids) are the most sensitive species. They are very sensitive to aniline and the heavy metals and malathion and parathion.

LITERATURE

Cairns Jr, J., 1986.
The myth of the most sensitive species. BioScience 36 (10): 670-672.

De Bruijn, J.L.M and J.H.M. Hermens, 1991.
Qualitative and quantitative modelling of toxic effects of organophosphorus compounds to fish. Sci. Total Environ. 109/110: 441-456.

EPA, 1984.
Estimating concern levels for concentrations of chemical substances in the environment. US-EPA, 43 p.

Fraybill, F.A, 1961.
An introduction to linear statistical models, vol. 1. McGraw Hill, New York. p 299.

Hoekstra, J.A., Vaal, M.A. and J. Notenboom, 1992.
Sensitivity patterns of aquatic species to toxicants: a pilot study. RIVM rapportnr 719102016.

Hoekstra, J., M. Vaal and J. Notenboom, in press.
Variation in the sensitivity of aquatic species to toxicants. Bull. Environm. Cont. Tox.

Kowalski, B.R. 1983
Chemometrics. Mathematics and statistics in chemistry. Reidel Publ. Comp. Dordrecht, The Netherlands. 79 pp.

LeBlanc, G.A., 1984.
Interspecies relationships in acute toxicity of chemicals to aquatic organisms. Envir. Tox. and Chem. 3: 47-60.

Mayer, F.L. M. and M.R. Ellersieck, 1986.
Manual of acute toxicity: interpretation and database for 410 chemicals and 66 species of freshwater animals. US Dept. of the Interior, Fish and Wildlife service, Washington, D.C. Resource publ. 160.

McKim, J.M. and P.K. Schmieder, 1991.
Bioaccumulation: does it reflect toxicity?. In: K. Nagel and R. Loskill (eds.). Bioaccumulation in aquatic systems; contributions to the assessments. Berlin, 1990.

Slooff, W., 1983.
Biological effects of chemical pollutants in the aquatic environment and their indicative value. Thesis, University of Utrecht, The Netherlands. 191 pp.

Slooff, W., J.H. Canton and J.L.M. Hermens, 1983.
Comparison of the susceptibility of 22 freshwater species to 15 chemical compounds. I. (Sub)acute toxicity tests. Aquat. Toxicol. 4: 113-128.

Slooff, W. and Canton, J.H. 1983.
Comparison of the susceptibility of 11 freshwater species to 8 chemicals compounds. II. (Semi)chronic toxicity tests. Aquat. Toxicol. 4: 271-282.

Suter II, G.W., 1993.
Ecological risk assessment. Lewish Publishers, Michigan, USA. 538 pp.

Umetri, 1992.
SIMCA-R Multivariate modelling and analysis, vers. 4.4, user's guide, 92-06-12. UMETRI AB, Umeå, Sweden.

Vaal, M., J. Notenboom en J. Hoekstra, 1993.
Het voorspellen van de gevoeligheid van waterorganismen voor giftige stoffen. H_2O 26 (23): 676-686.

Van Straalen, N.M. and C.A.J. Denneman, 1989.
Ecotoxicological evaluation of soil quality criteria. Ecotox. and Environm. Saf. 18: 241-251.

Verhaar, H.J.M, C.J. Van Leeuwen and J.L.M. Hermens, 1992.
Classifying environmental pollutants. 1: Stucture-activity relationships for prediction of aquatic toxicity. Chemosphere, 25 (4): 471-491.

Wagner, C. and H. Løkke, 1991.
Estimations of ecotoxicological protection levels from NOEC toxicity data. Water Res. 14: 141-150.

TABLES AND FIGURES

TABLE 1 Species and taxonomic groups used in the analyses.

Species name	Common name	Phylum	Class	Abbreviation
Aedes aegypti	Mosquito	Arthropoda	Insecta	Aede aeg
Ambystoma mexicanum	Salamander	Chordata	Amphibia	Amby mex
Asellus aquaticus	Aquatic sowbug	Arthropoda	Malacostraca	Asel aqu
Carassius auratus	Goldfish	Chordata	Osteichtyes	Cara aur
Chironomus thummi	Midge	Arthropoda	Insecta	Chir thu
Cloeon dipterum	Mayfly	Arthropoda	Insecta	Cloe dip
Corixa punctata	Water boatman	Arthropoda	Insecta	Cori pun
Culex pipiens	Mosquito	Arthropoda	Insecta	Cule pip
Daphnia pulex	Water flea	Arthropoda	Phyllopoda	Daph pul
Daphnia cucullata	Water flea	Arthropoda	Phyllopoda	Daph cuc
Daphnia magna	Water flea	Arthropoda	Phyllopoda	Daph mag
Dugesia lugubris	Flatworm	Platyhelminthes	Turbellaria	Duge lug
Erpobdella octoculata	Leech	Annelida	Hirudinea	Erpo oct
Gambusia affinis	Mosquitofish	Chordata	Osteichtyes	Gamb aff
Gammarus pulex	Scud	Arthropoda	Malacostraca	Gamm pul
Hydra oligactis	Hydra	Coelenterata	Hydrozoa	Hydr oli
Ischnura elegans	Dragonfly	Arthropoda	Insecta	Isch ele
Lepomis macrochirus	Bluegill	Chordata	Osteichtyes	Lepo mac
Leuciscus idus	Ide, silver or golden orfe	Chordata	Osteichtyes	Leuc idu
Lymnaea stagnalis	Great pond snail	Mollusca	Gastropoda	Lymn sta
Nemoura cinerea	Stonefly	Arthropoda	Insecta	Nemo cin
Oncorhynchus mykiss	Rainbow trout,donaldson trout	Chordata	Osteichtyes	Onco myk
Oryzias latipes	Medaka	Chordata	Osteichtyes	Oryz lat
Pimephales promelas	Fathead minnow	Chordata	Osteichtyes	Pime pro
Poecilia reticulata	Guppy	Chordata	Osteichtyes	Poec ret
Xenopus laevis	Clawed toad	Chordata	Amphibia	Xeno lae

TABLE 2 Compounds and chemical classes used in the analyses.

Compound name	CAS-nr	Toxicological mechanism	Chem. class [1]	Abbreviation
Acetone	67641	non-polar narcosis	I	Acetone
Allylamine	107119	polar narcosis	II	Allylamine
Aniline	62533	polar narcosis	II	Aniline
Benzene	71432	non-polar narcosis	I	Benzene
$Cadmium^{2+}$	10108642 [2]		M	Cd2+
$Copper^{2+}$	7447394 [2]		M	Cu2+
Cresol, o-	95487	polar narcosis	II	Cresol
Dieldrin	60571	neurotoxic	IV	Dieldrin
Ethylacetate	141786	ester narcosis		Ethylace
Ethylpropionate	105373	ester narcosis		Ethylpro
Heptanol, 1-	111706	non-polar narcosis	I	Heptanol
Lindane	58899	neurotoxic	IV	Lindane
Malathion	121755	AChE-inhibition	IV	Malathion
$Mercury^{2+}$	7487947 [2]		M	Hg2+
Parathion	56382	AChE-inhibition	IV	Parathion
Pentachlorophenol	87865 [2]	uncoupling of oxidation	IV	PCP
Phenol	108952	polar narcosis	II	Phenol
Propanol, 1-	71238	non-polar narcosis	I	Propanol
Pyridine	110861	non-polar narcosis	I	Pyridine
Salicylaldehyde	90028	carbonyl reactivity	III	Salicylald
Trichloroethylene, 1,1,2-	79016	non-polar narcosis	I	TriClEth

[1] Chemical classification of organic compounds according to Verhaar et al. (1992):
I = inert chemicals, II = less inert chemicals, III = reactive chemicals, IV = specifically acting chemicals; M = heavy metal.
Ethylacetate and ethylpropionate cannot be classified as such but appear in figures as class III.

[2] Data for this compound were taken from several salts (See Appendix 2).

TABLE 3 Matrix of data, mean log LC_{50} (µg/l) per species-compound combination, overall mean, standard deviation and percentage of matrix filling per species and per compound.

Species	Acde aeg	Amby mex	Asel aqu	Cara aur	Chir tha	Clae dip	Cori pua	Cule pip	Daph cuc	Daph mag	Daph pul	Duge lug	Erpo oct	Gamb aff	Gamm pul	Hydr oli	Iach ele	Lepo mac	Leuc ida	Lymn sta	Nemo cla	Onco myk	Oryz lat	Pime pro	Poec ret	Xeno lae	Average	St.dev.	n (=21)	% avail. data
Acetone	7.18	7.30	6.88	-	7.11	6.88	6.70	7.23	6.88	7.13	6.94	6.88	6.85	7.11	6.78	7.13	6.81	6.92	6.97	6.85	7.01	6.74	7.16	6.87	6.98	7.38	6.99	0.18	25	96.2%
Allylamine	5.08	3.26	5.00	4.10	4.15	4.48	5.08	5.23	4.45	4.60	4.53	4.82	4.41	-	5.11	4.24	-	-	4.79	3.70	5.00	4.18	4.20	3.32	4.08	3.90	4.42	0.56	23	88.5%
Aniline	5.19	5.64	4.83	5.27	5.24	5.34	5.18	4.97	2.84	2.50	2.00	5.19	5.88	-	5.05	5.61	5.37	4.69	4.76	5.90	4.81	4.38	4.01	4.89	5.00	5.82	4.82	1.00	25	96.2%
Benzene	5.30	5.57	5.08	4.54	5.00	4.53	4.68	4.85	5.57	5.61	5.48	4.87	-	5.59	4.62	4.53	4.00	4.50	4.81	5.36	5.11	3.82	5.40	4.43	4.53	5.28	4.92	0.50	25	96.2%
Cd2+	3.72	1.79	3.51	2.87	-	4.11	-	3.00	1.99	1.50	1.82	-	3.29	3.86	2.16	2.88	-	3.58	3.97	2.88	4.37	0.98	2.48	3.87	4.74	4.37	3.08	1.04	22	84.6%
Cresol	4.90	4.60	4.36	4.37	4.53	4.70	4.90	4.66	4.22	4.17	3.98	4.38	5.13	-	4.32	4.88	4.66	4.32	4.13	5.20	4.00	3.92	4.61	4.19	4.34	4.58	4.48	0.35	25	96.2%
Cu2+	-	-	4.51	3.14	-	0.68	-	-	-	1.48	1.34	-	-	3.62	2.13	-	-	3.53	2.50	-	-	2.16	-	3.50	3.47	2.58	2.67	1.09	13	50.0%
Dieldrin	0.18	-	-	1.61	-	-	-	-	-	-	2.40	-	-	1.61	-	-	-	0.96	1.28	-	-	0.61	0.60	1.12	1.08	1.81	1.15	0.63	12	46.2%
Ethylace	5.54	5.18	6.20	-	5.88	5.68	5.78	6.60	5.22	5.85	5.42	6.48	6.08	-	5.88	6.13	5.78	-	5.46	6.04	5.11	5.58	5.10	5.36	5.32	5.26	5.69	0.43	23	88.5%
Ethylpro	5.54	4.73	5.08	-	4.81	5.29	5.78	-	4.66	5.21	4.84	6.00	5.32	-	5.22	5.53	5.39	-	4.89	5.23	4.30	4.75	4.76	4.85	-	4.75	5.09	0.42	21	80.8%
Heptanol	5.20	4.72	4.85	-	4.90	5.13	5.20	5.09	4.92	4.84	4.69	5.30	5.11	-	4.78	5.20	5.04	-	4.53	4.60	4.56	4.63	4.68	4.54	4.81	4.64	4.87	0.25	23	88.5%
Hg2+	3.62	2.48	2.17	2.54	2.61	1.57	2.21	2.89	0.37	0.50	0.34	1.74	2.37	2.56	0.87	1.75	3.88	2.14	2.63	2.65	2.10	2.07	2.74	1.63	4.02	1.85	2.17	0.95	26	100.0%
Lindane	-	-	2.57	2.17	-	-	-	1.96	-	2.95	3.58	-	-	2.84	1.39	-	-	2.02	1.64	3.71	-	1.48	2.08	1.84	1.97	-	2.30	0.73	14	53.8%
Malathion	1.96	-	-	3.98	-	-	-	1.58	-	0.86	0.26	-	-	0.54	-	-	-	1.95	-	-	-	2.22	2.88	4.17	3.12	-	2.14	1.31	11	42.3%
Parathion	0.68	-	1.27	3.34	-	0.35	-	0.18	-	-0.07	-0.16	-	-	2.57	-	-	-	2.67	2.79	-	-	3.15	3.46	3.29	1.79	-	1.81	1.39	14	53.8%
PCP	4.86	2.48	3.46	2.25	2.04	3.77	4.04	4.53	3.18	2.74	2.64	2.11	2.40	2.45	2.85	2.86	4.62	2.10	2.30	2.75	2.58	1.95	1.91	2.34	2.82	2.41	2.86	0.85	26	100.0%
Phenol	-	-	5.28	4.67	-	4.48	-	-	-	4.31	4.93	5.00	5.66	4.61	4.78	-	-	4.29	4.13	-	-	3.92	-	4.53	4.64	-	4.66	0.46	14	53.8%
Propanol	6.64	6.60	6.40	-	6.37	6.49	6.30	6.68	6.76	6.85	6.48	6.67	6.15	6.12	5.26	6.06	5.61	-	5.34	5.54	5.40	3.66	5.52	5.00	6.14	6.60	6.56	0.23	23	88.5%
Pyridine	5.11	5.98	5.34	-	5.36	5.22	4.48	4.82	6.39	6.11	5.76	6.28	6.38	-	5.26	6.06	5.61	-	5.34	5.54	5.40	3.66	5.52	5.00	6.14	6.29	5.55	0.66	24	92.3%
Salicylald	4.20	3.85	3.60	-	3.97	4.11	4.11	4.73	3.74	3.63	3.73	3.82	3.91	-	3.28	3.85	4.08	-	3.54	3.81	3.11	3.13	3.62	3.36	3.72	3.89	3.77	0.36	23	88.5%
TriClEth	4.68	4.68	4.48	-	4.81	4.62	5.04	4.74	4.76	4.71	4.65	4.62	4.88	-	4.38	4.88	4.69	4.65	5.26	4.75	4.85	4.62	3.28	4.69	5.26	4.65	4.69	0.36	24	92.3%
Average	4.42	4.59	4.47	3.45	4.77	4.30	4.96	4.13	4.40	3.77	3.60	4.94	4.92	3.62	4.16	4.82	5.12	3.45	4.12	4.74	4.57	3.55	3.96	4.02	4.23	4.47	4.26			
St.dev.	1.84	1.58	1.43	1.13	1.33	1.78	1.10	2.02	1.79	2.14	2.08	1.51	1.39	1.93	1.67	1.52	0.93	1.57	1.55	1.37	1.28	1.68	1.67	1.48	1.57	1.62				
n (=26)	18	15	19	13	14	18	14	18	15	20	21	15	15	12	18	15	13	14	20	16	15	21	19	21	20	17			436	
% avail. data	85.7%	71.4%	90.5%	61.9%	66.7%	85.7%	66.7%	85.7%	71.4%	95.2%	100.0%	71.4%	71.4%	57.1%	85.7%	71.4%	61.9%	66.7%	95.2%	76.2%	71.4%	100.0%	90.5%	100.0%	95.2%	81.0%				79.9%

TABLE 4 Analysis of variance, species and compounds as sources of variation.

Source of Variation	SS	df	MS	F	P-value	F crit
Compounds	982.06	20	45.44	100.89	< 0.001	1.59
Species	25.03	25	1.00	2.22	0.001	1.53
Error	175.63	390	0.45			
Total	1182.72	435				

TABLE 5 Analyses of patterns in species sensitivity, description of principal components, scaled data.

Compo -nent	SS expl. (%) [1]	Total SS expl. (%)	Variance expl. (%)	Total Variance expl. (%)	Normalized eigenvalue	Residual matrix variance	PRESS /SS [2]	Limit [3]	Sign.[4]
1	26.8	26.8	17.9	17.9	5.6	0.82	0.99	0.92	S2
2	18.2	45.0	12.3	30.2	3.8	0.70	1.07	0.91	S2
3	14.9	59.9	11.6	41.8	3.1	0.58	0.97	0.91	S2
4	8.7	68.6	5.5	47.3	1.8	0.53	1.31	0.91	S2
5	7.2	75.8	4.9	52.3	1.5	0.48	1.53	0.90	S2
6	5.8	81.6	4.1	56.4	1.2	0.44	1.07	0.90	NS

[1] SS expl. (%): Sum of squares explained.

[2] PRESS/SS: Prediction Sum of Squares / residual Sum of Squares of previous dimension. Prediction Sum of Squares is the squared difference between observed values and values predicted by cross validation.

[3] Limit: significance limit.

[4] Significance of the principal components at the 5% confidence level is based on the following SIMCA cross validation rules (Umetri, 1992).

S1: $PRESS/SS_{(total)} \leq$ LIMIT.

S2: At least $(K)^{1/3}$ variables, have $(PRESS/SS)_K \leq$ LIMIT. K is the number of variables.

NS: Not significant.

TABLE 6 Analyses of patterns in species sensitivity, residuals of compounds as variables, scaled data.

Variable	PC 1		PC 2		PC 3		Total after 3 components	
	PRESS/SS [1]	S2 expl.(%) [2]	PRESS/SS	S2 expl.(%)	PRESS/SS	S2 expl.(%)	PRESS/SS	S2 expl.(%)
Acetone	1.13	-3.4	1.08	8.7	0.88	41.4	0.88	46.7
Allylamine	1.16	19.6	0.95	-2.8	0.75	32.6	0.71	49.4
Aniline	1.01	10.5	1.06	29.9	0.88	26.1	0.88	66.4
Benzene	1.17	-2.2	1.22	52.2	0.99	8.3	0.99	58.3
Cd2+	1.08	-1.7	1.10	16.2	1.10	24.4	1.00	38.9
Cresol	0.81	45.8	1.16	-2.5	0.81	23.4	0.66	66.8
Cu2+	1.10	-8.4	0.93	12.8	0.85	20.6	0.79	25.1
Dieldrin	1.28	6.5	0.70	49.5	1.14	-4.8	0.70	51.2
Ethylace	0.89	51.3	1.02	-1.6	0.98	4.1	0.87	53.9
Ethylpro	0.38	76.1	1.17	-0.5	0.93	2.0	0.36	77.5
Heptanol	0.58	69.7	1.00	1.0	1.12	0.4	0.58	71.0
Hg2+	1.10	7.2	1.03	31.3	0.54	36.6	0.54	75.1
Lindane	1.49	-4.1	0.77	62.9	0.97	4.3	0.75	63.1
Malathion	1.28	-3.3	0.66	57.6	1.37	-3.5	0.66	50.9
Parathion	1.24	41.7	0.53	38.0	0.69	8.3	0.37	87.9
PCP	0.76	52.9	1.07	-1.7	1.02	-1.7	0.76	49.4
Phenol	0.97	46.7	1.39	-4.8	1.24	-5.8	0.97	36.1
Propanol	1.14	-4.7	1.41	14.6	1.27	16.3	1.00	26.2
Pyridine	1.07	-4.4	1.35	32.3	0.96	17.1	0.96	45.0
Salicylald	0.60	61.0	0.85	4.4	0.80	8.9	0.41	74.4
TriClEth	1.06	0.2	1.03	-4.2	1.20	-5.2	1.00	-9.1

[1] PRESS/SS: Prediction Sum of Squares / residual Sum of Squares of previous dimension. Predicton Sum of Squares is the squared difference between observed values and values predicted by cross validation.

[2] S2 expl.(%): Variance explained.

TABLE 7 Analyses of patterns in species sensitivity, loadings of compounds as variables, scaled data.

Variable	Loading PC1	Loading PC2	Loading PC3
Acetone	-0.04	-0.17	0.38
Allylamine	0.20	0.05	-0.34
Aniline	0.16	0.28	0.31
Benzene	-0.06	-0.35	0.18
Cd2+	0.08	0.22	0.29
Cresol	0.29	0.00	0.30
Cu2+	-0.04	0.20	0.24
Dieldrin	-0.15	-0.37	0.04
Ethylace	0.30	-0.04	-0.14
Ethylpro	0.37	-0.04	-0.10
Heptanol	0.34	-0.07	-0.08
Hg2+	0.14	0.28	0.36
Lindane	0.09	-0.36	0.14
Malathion	-0.10	0.32	0.08
Parathion	-0.28	0.29	0.18
PCP	0.31	0.03	-0.04
Phenol	0.38	0.00	0.01
Propanol	-0.01	-0.20	0.25
Pyridine	-0.02	-0.28	0.25
Salicylald	0.32	-0.12	0.18
TriClEth	0.09	0.03	0.01

TABLE 8 Analyses of patterns in species sensitivity, scores of species as objects, scaled data. Species with high positive scores are sensitive to compounds with high positive loadings in Table 7, and vice versa.

Object name	PC 1			PC 2			PC 3		
	SDO [1]	Score	OLEV [2]	SDO [1]	Score	OLEV [2]	SDO [1]	Score	OLEV [2]
Aede aeg	0.76	-3.76	0.32	0.79	-0.56	0.25	0.78	-1.13	0.23
Amby mex	0.97	1.87	0.16	0.98	1.60	0.16	0.68	-2.65	0.21
Asel aqu	0.72	-1.11	0.09	0.74	-0.69	0.08	0.76	0.65	0.08
Cara aur	0.64	1.60	0.14	0.55	-1.36	0.14	0.57	-0.63	0.13
Chir thu	0.57	0.18	0.02	0.60	-0.48	0.03	0.58	-0.89	0.06
Cloe dip	0.72	-1.79	0.15	0.76	0.06	0.12	0.77	0.82	0.11
Cori pun	0.80	-3.76	0.32	0.74	-1.97	0.27	0.61	1.72	0.26
Cule pip	0.75	-4.77	0.40	0.79	-0.02	0.31	0.83	0.18	0.27
Daph cuc	1.01	2.03	0.17	0.69	3.87	0.28	0.64	1.21	0.26
Daph mag	1.11	0.76	0.06	0.53	4.45	0.29	0.40	1.48	0.27
Daph pul	1.25	1.36	0.12	0.77	4.36	0.30	0.56	2.30	0.29
Duge lug	0.90	-2.36	0.20	0.91	1.54	0.18	0.90	1.23	0.18
Erpo oct	0.92	-2.76	0.23	0.97	-0.61	0.18	1.03	-0.30	0.16
Gamb aff	0.86	0.76	0.06	0.81	1.37	0.10	0.52	-2.58	0.17
Gamm pul	1.03	0.44	0.04	1.03	-1.56	0.11	0.66	3.15	0.20
Hydr oli	0.70	-2.15	0.18	0.70	1.22	0.16	0.71	-0.78	0.15
Isch ele	0.84	-3.07	0.26	0.71	-2.52	0.26	0.76	-0.23	0.23
Lepo mac	0.47	1.63	0.14	0.38	-1.18	0.13	0.40	0.17	0.12
Leuc idu	0.66	1.87	0.16	0.63	-1.24	0.15	0.66	0.06	0.13
Lymn sta	1.04	-0.96	0.08	1.07	1.19	0.10	0.99	-1.93	0.14
Nemo cin	0.91	3.14	0.27	0.83	-2.33	0.25	0.87	0.46	0.22
Onco myc	1.15	3.04	0.26	1.05	-2.39	0.25	0.78	3.09	0.28
Oryz lat	1.11	2.62	0.22	1.17	-0.07	0.17	1.19	-1.31	0.17
Pime pro	0.72	2.67	0.23	0.60	-1.92	0.21	0.62	-0.34	0.19
Poec ret	0.90	0.24	0.02	0.87	-1.42	0.09	0.77	-1.94	0.14
Xeno lae	0.91	1.58	0.13	0.93	1.04	0.12	0.58	-2.85	0.20

[1] SDO: Residual standard deviation of object.

[2] OLEV: Object leverage.

TABLE 9 Analyses of patterns in species sensitivity, description of principal components, unscaled data.

Component	SS expl. (%) [1]	Total SS expl. (%)	Variance expl. (%)	Total Variance expl. (%)	Normalized eigenvalue	Residual matrix variance	PRESS/SS [2]	Limit [3]	Sign. [4]
1	37.2	37.2	29.6	29.6	7.8	0.34	1.08	0.92	S2
2	21.2	58.4	17.6	47.2	4.5	0.26	1.05	0.91	S2
3	11.1	69.5	8.5	55.7	2.3	0.21	1.23	0.91	NS

[1] SS expl. (%): Sum of squares explained.

[2] PRESS/SS: Prediction Sum of Squares / residual Sum of Squares of previous dimension. Prediction Sum of Squares is the squared difference between observed values and values predicted by cross validation.

[3] Limit: significance limit.

[4] Significance of the principal components at the 5% confidence level is based on the following SIMCA cross validation rules (Umetri, 1992).

S1: $PRESS/SS_{(total)} \leq$ LIMIT.

S2: At least $(K)^{1/3}$ variables, have $(PRESS/SS)_K \leq$ LIMIT. K is the number of variables.

NS: Not significant.

TABLE 10 Analyses of patterns in species sensitivity, residuals of compounds as variables, unscaled data.

Variable	PC 1		PC 2		PC 3		Total after 3 components	
	PRESS/SS [1]	S2 expl.(%) [2]	PRESS/SS	S2 expl. (%)	PRESS/SS	S2 expl. (%)	PRESS/SS	S2 expl. (%)
Acetone	1.05	-4.3	1.06	-4.5	1.07	-0.4	1.00	-9.2
Allylamine	0.92	4.7	0.91	31.5	1.15	5.4	0.83	41.6
Aniline	0.96	59.3	0.97	3.5	1.20	3.1	0.94	66.0
Benzene	0.81	21.3	1.09	-3.3	1.18	26.4	0.81	44.4
Cd2+	1.08	26.7	1.61	25.7	1.24	-2.6	1.00	49.8
Cresol	0.96	6.7	0.91	19.0	0.83	18.6	0.73	44.3
Cu2+	0.72	44.6	1.19	-5.0	1.32	29.0	0.72	68.5
Dieldrin	1.14	5.9	0.97	7.3	0.67	43.4	0.66	56.5
Ethylace	1.05	-4.7	1.07	10.9	1.28	6.7	1.00	12.8
Ethylpro	1.10	-4.2	0.89	29.5	1.10	1.2	0.89	26.5
Heptanol	1.05	-4.3	0.62	41.2	1.20	0.3	0.62	37.2
Hg2+	0.91	58.7	0.96	12.9	1.08	-1.2	0.87	70.4
Lindane	0.96	14.0	1.12	-6.2	1.02	68.3	0.96	76.2
Malathion	1.24	64.0	0.89	0.7	1.85	2.6	0.89	67.3
Parathion	1.53	59.7	0.61	34.8	1.61	-0.3	0.61	94.2
PCP	1.21	-3.2	0.98	87.3	1.04	3.8	0.98	87.9
Phenol	1.27	-8.1	0.95	13.2	1.11	41.1	0.95	46.2
Propanol	1.21	-4.3	1.03	-5.2	1.14	-2.9	1.00	-12.4
Pyridine	1.04	-1.4	1.25	-4.8	1.56	54.7	1.00	48.5
Salicylald	1.00	-4.8	0.69	49.3	0.99	1.4	0.68	46.0
TriClEth	1.04	-4.1	0.99	4.7	1.07	-3.5	0.99	-2.9

[1] PRESS/SS: Prediction Sum of Squares / residual Sum of Squares of previous dimension. Predicton Sum of Squares is the squared difference between observed values and values predicted by cross validation.

[2] S2 expl.(%): Variance explained.

TABLE 11 Analyses of patterns in species sensitivity, loadings of compounds as variables, unscaled data.

Variable	Loading PC1	Loading PC2	Loading PC3
Acetone	0.00	-0.01	-0.03
Allylamine	-0.08	0.23	0.12
Aniline	0.38	0.15	-0.17
Benzene	-0.12	-0.02	-0.22
Cd2+	0.27	0.40	-0.02
Cresol	0.06	0.11	-0.12
Cu2+	0.34	0.06	-0.51
Dieldrin	-0.12	-0.14	-0.34
Ethylace	0.00	0.11	-0.11
Ethylpro	0.02	0.17	-0.07
Heptanol	-0.01	0.11	-0.03
Hg2+	0.37	0.25	0.02
Lindane	-0.15	0.06	-0.48
Malathion	0.45	-0.15	0.24
Parathion	0.51	-0.47	-0.06
PCP	-0.04	0.55	0.13
Phenol	0.01	0.17	-0.22
Propanol	-0.01	0.00	-0.03
Pyridine	-0.06	-0.01	-0.36
Salicylald	0.00	0.17	-0.06
TriClEth	0.01	0.07	-0.03

TABLE 12 Analyses of patterns in species sensitivity, scores of species as objects, unscaled data. Species with high positive scores are sensitive to compounds with high positive loadings in Table 11, and vice versa.

Object name	PC 1			PC 2			PC 3		
	SDO [1]	Score	OLEV [2]	SDO [1]	Score	OLEV [2]	SDO [1]	Score	OLEV [2]
Aede aeg	0.81	-0.22	0.02	0.38	-2.95	0.24	0.36	-0.88	0.22
Amby mex	0.63	-0.22	0.02	0.57	1.38	0.11	0.53	1.51	0.15
Asel aqu	0.57	-0.45	0.05	0.52	-1.17	0.10	0.48	1.07	0.12
Cara aur	0.51	-2.09	0.21	0.34	1.39	0.20	0.36	-0.02	0.18
Chir thu	0.28	-1.15	0.12	0.23	0.84	0.12	0.23	0.32	0.11
Cloe dip	0.69	1.46	0.15	0.60	-1.63	0.18	0.52	-1.63	0.20
Cori pun	0.59	-0.56	0.06	0.43	-2.05	0.17	0.37	-1.31	0.18
Cule pip	0.77	0.82	0.08	0.45	-2.61	0.22	0.41	-1.05	0.21
Daph cuc	0.27	4.87	0.49	0.28	0.07	0.40	0.29	0.48	0.36
Daph mag	0.31	4.17	0.42	0.27	0.76	0.35	0.28	0.12	0.31
Daph pul	0.38	4.85	0.49	0.35	0.80	0.40	0.34	0.51	0.36
Duge lug	0.51	0.10	0.01	0.54	0.04	0.01	0.47	1.81	0.13
Erpo oct	0.50	-1.53	0.15	0.53	-0.10	0.13	0.31	2.40	0.21
Gamb aff	0.79	-0.13	0.01	0.84	-0.09	0.01	0.62	1.91	0.14
Gamm pul	0.52	1.30	0.13	0.54	0.46	0.11	0.52	-0.88	0.12
Hydr oli	0.39	-0.41	0.04	0.42	-0.05	0.03	0.38	1.22	0.09
Isch ele	0.66	-2.70	0.27	0.40	-2.60	0.31	0.38	-0.99	0.28
Lepo mac	0.42	-0.87	0.09	0.39	0.71	0.09	0.41	0.27	0.08
Leuc idu	0.41	-1.15	0.12	0.40	0.56	0.10	0.38	-0.77	0.11
Lymn sta	0.58	-0.93	0.09	0.61	0.03	0.08	0.40	2.38	0.19
Nemo cin	0.55	-0.61	0.06	0.58	0.08	0.05	0.60	-0.84	0.08
Onco myc	0.90	-0.22	0.02	0.72	2.49	0.20	0.57	-1.91	0.23
Oryz lat	0.68	-1.12	0.11	0.53	1.88	0.18	0.55	-0.44	0.16
Pime pro	0.56	-2.26	0.23	0.46	1.45	0.22	0.47	-0.43	0.20
Poec ret	0.54	-2.01	0.20	0.50	-1.04	0.18	0.53	0.10	0.17
Xeno lae	0.55	-0.94	0.09	0.57	0.34	0.08	0.57	0.94	0.10

[1] SDO: Residual standard deviation of object.

[2] OLEV: Object leverage.

TABLE 13 K nearest neighbour of species, scaled data.
The number of three nearest species neighbour species in the multidimensional space is summed for all n species in the taxonomic group.

Taxonomic group	Neighbour species			
	Osteichtyes and Amphibia	Insecta	Malacostraca and Phyllopoda	Rest
Osteichtyes and Amphibia (n=10)	25	4	-	1
Insecta (n=7)	6	13	1	1
Malacostraca and Phyllopoda (n=5)	4	2	9	-
Rest (n=4)	4	3	1	4

TABLE 14 K nearest neighbour of species, unscaled data.
The number of three nearest neighbour species in the multidimensional space is summed for all n species in the taxonomic group.

Taxonomic group	Neighbour species			
	Osteichtyes and Amphibia	Insecta	Malacostraca and Phyllopoda	Rest
Osteichtyes and Amphiba (n=10)	8	14	1	7
Insecta (n=7)	7	12	1	1
Malacostraca and Phyllopoda (n=5)	-	4	9	2
Rest (n=4)	2	2	-	8

TABLE 15 Analyses of patterns in toxicity of compounds, description of principal components.

Compo-nent	SS expl. (%) [1]	Total SS expl. (%)	Variance expl. (%)	Total Variance expl. (%)	Residual matrix variance	Normalized eigenvalue	PRESS/ SS [2]	Limit [3]	Sign. [4]
1	81.0	81.0	78.6	78.6	0.21	17.0	0.25	0.92	S1
2	5.0	85.9	3.5	82.1	0.18	1.0	1.09	0.91	S2
3	4.2	90.1	3.6	85.6	0.14	0.9	0.93	0.91	S2
4	2.7	92.8	2.2	87.8	0.12	0.6	1.17	0.91	S2
5	1.9	94.7	1.5	89.3	0.11	0.4	1.15	0.90	S2
6	1.4	96.1	1.1	90.5	0.10	0.3	1.19	0.89	S2
7	1.0	97.1	0.8	91.3	0.09	0.2	1.05	0.89	S2
8	0.7	97.8	0.3	91.6	0.08	0.2	1.31	0.88	S2
9	0.5	98.4	-0.3	91.3	0.09	0.1	1.12	0.87	NS

[1] SS expl. (%): Sum of squares explained.

[2] PRESS/SS: Prediction Sum of Squares / residual Sum of Squares of previous dimension . Prediction Sum of Squares is the squared difference between observed values and values predicted by cross validation.

[3] Limit: significance limit.

[4] Significance of the principal components at the 5% confidence level is based on the following SIMCA cross validation rules (Umetri, 1992).

S1: $PRESS/SS_{(total)} \leq$ LIMIT.

S2: At least $(K)^{1/3}$ variables, have $(PRESS/SS)_K \leq$ LIMIT. K is the number of variables.

NS: Not significant.

TABLE 16 Analyses of patterns in toxicity of compounds, residuals and loadings of species as variables.

	PC 1		
Variable	PRESS/SS [1]	S2 expl. (%) [2]	Loading
Aede aeg	0.31	76.8	0.18
Amby mex	0.22	80.9	0.21
Asel aqu	0.14	88.4	0.21
Cara aur	0.62	40.6	0.13
Chir thu	0.16	85.9	0.22
Cloe dip	0.30	77.8	0.21
Cori pun	0.35	73.9	0.20
Cule pip	0.25	77.6	0.17
Daph cuc	0.33	75.4	0.21
Daph mag	0.25	82.4	0.20
Daph pul	0.41	68.1	0.17
Duge lug	0.22	83.1	0.22
Erpo oct	0.16	86.5	0.22
Gamb aff	0.41	61.7	0.13
Gamm pul	0.08	92.0	0.23
Hydr oli	0.10	91.2	0.23
Isch ele	0.39	63.7	0.18
Lepo mac	0.30	72.8	0.16
Leuc idu	0.13	89.2	0.20
Lymn sta	0.16	86.1	0.21
Nemo cin	0.21	82.1	0.22
Onco myc	0.25	80.0	0.19
Oryz lat	0.21	82.9	0.18
Pime pro	0.24	82.1	0.19
Poec ret	0.24	79.4	0.18
Xeno lae	0.13	90.0	0.20

[1] PRESS/SS: Prediction Sum of Squares / residual Sum of Squares of previous dimension . Prediction Sum of Squares is the squared difference between observed values and predicted values by cross validation.

[2] S2 expl. (%): Variance explained.

TABLE 17 Analyses of patterns in toxicity of compounds, scores of compounds as objects. The scores indicate the rank order of the compounds by their average toxicity.

Object name	PC 1		
	SDO [1]	Score	OLEV [2]
Acetone	0.37	-8.35	0.39
Allylamine	0.46	-0.10	0.00
Aniline	0.56	-1.76	0.08
Benzene	0.52	-1.74	0.08
Cd2+	0.60	3.59	0.17
Cresol	0.31	-0.35	0.02
Cu2+	0.59	4.12	0.19
Dieldrin	0.44	9.63	0.44
Ethylace	0.27	-4.01	0.19
Ethylpro	0.28	-2.02	0.09
Heptanol	0.22	-1.31	0.06
Hg2+	0.52	7.25	0.33
Lindane	0.49	5.66	0.26
Malathion	0.70	5.08	0.23
Parathion	0.78	6.83	0.32
PCP	0.55	5.00	0.23
Phenol	0.34	-1.89	0.09
Propanol	0.33	-6.71	0.31
Pyridine	0.43	-3.55	0.16
Salicylald	0.34	2.28	0.11
TriClEth	0.35	-0.88	0.04

[1] SDO: Residual standard deviation of object.
[2] OLEV: Object leverage.

TABLE 18 Ordering of species by their sensitivity to compound groups, qualitative interpretation of the three principal components model.

Taxonomic group	Compounds			
	parathion malathion	aniline heavy metals	dieldrin lindane	pentachlorophenol
Amphibians	(-)	-	(++)	+
Osteichtyes (fishes)	-	-	++	+
Phyllopoda (daphnids)	++	++	-	-
Insecta	(+)	-	(+)	-
Malacostraca (*Asellus* and *Gammarus*)	(-)	-	-	-

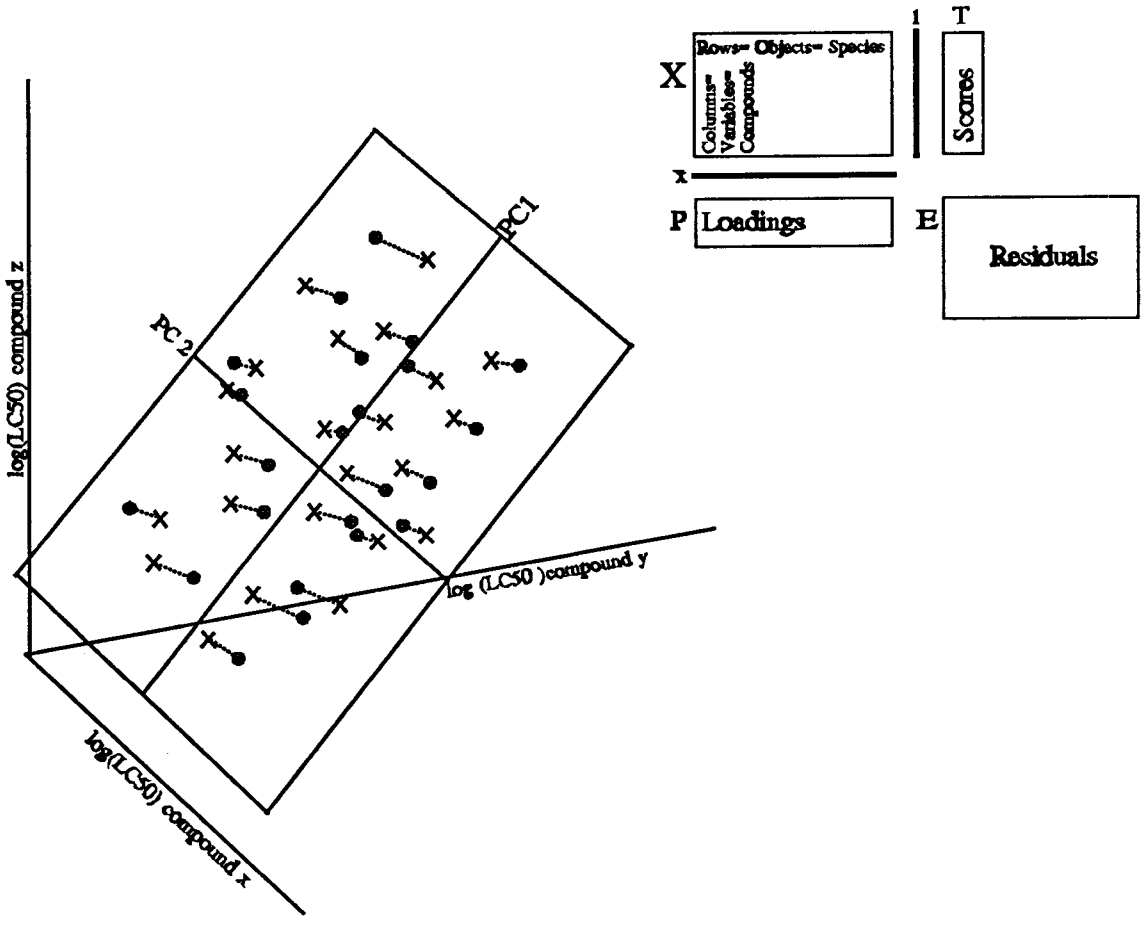
-: relatively insensitive to this compound group compared to other species
+: relatively sensitive to this compound group compared to other species
( ): estimation, based on the PC model, no or few data available

TABLE 19 Ordering of compounds by their toxicity to species and their Sensitivity Ratio (the ratio of the 95% and 5% percentile of the distribution of LC_{50}-values for that compound).

Objectname	log LC50 [1]	$SR_{95:5}$ [2]
Dieldrin	1.37	433
Hg2+	2.17	3240
Parathion	2.19	3429
Lindane	2.60	169
Malathion	2.67	4349
PCP	2.86	468
Cu2+	3.03	1200
Cd2+	3.23	1010
Salicylald	3.69	15
Allylamine	4.38	65
Cresol	4.45	16
TriClEth	4.62	47
Heptanol	4.74	26
Aniline	4.79	2398
Phenol	4.88	22
Benzene	4.94	39
Ethylpro	4.94	66
Pyridine	5.44	252
Ethylace	5.54	106
Propanol	6.38	50
Acetone	6.90	4

[1] Average log LC50, missing species-compounds combinations estimated by first component of Compounds PC model (X= 1 . $\bar{x}$ - T . P . s), data from tables 3, 16 and 17.

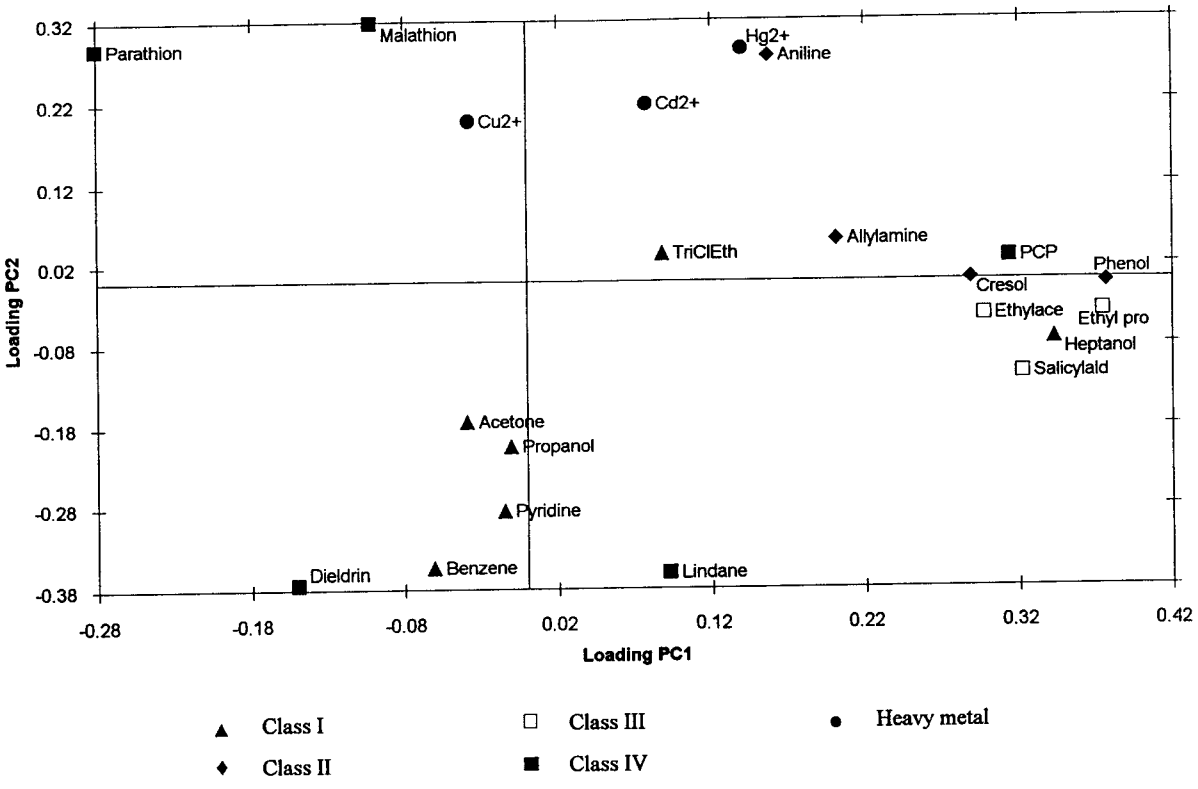
[2] Sensitivity ratio 95% percentile : 5% percentile of the distribution of LC50's over all species (see Hoekstra et al., in press). $SR_{95:5}$ is based on same data as Average log LC50.



●: Observation for species a on compounds x, y and z.

×: Projection of species on Principal Component plane (PC1 vs. PC2).

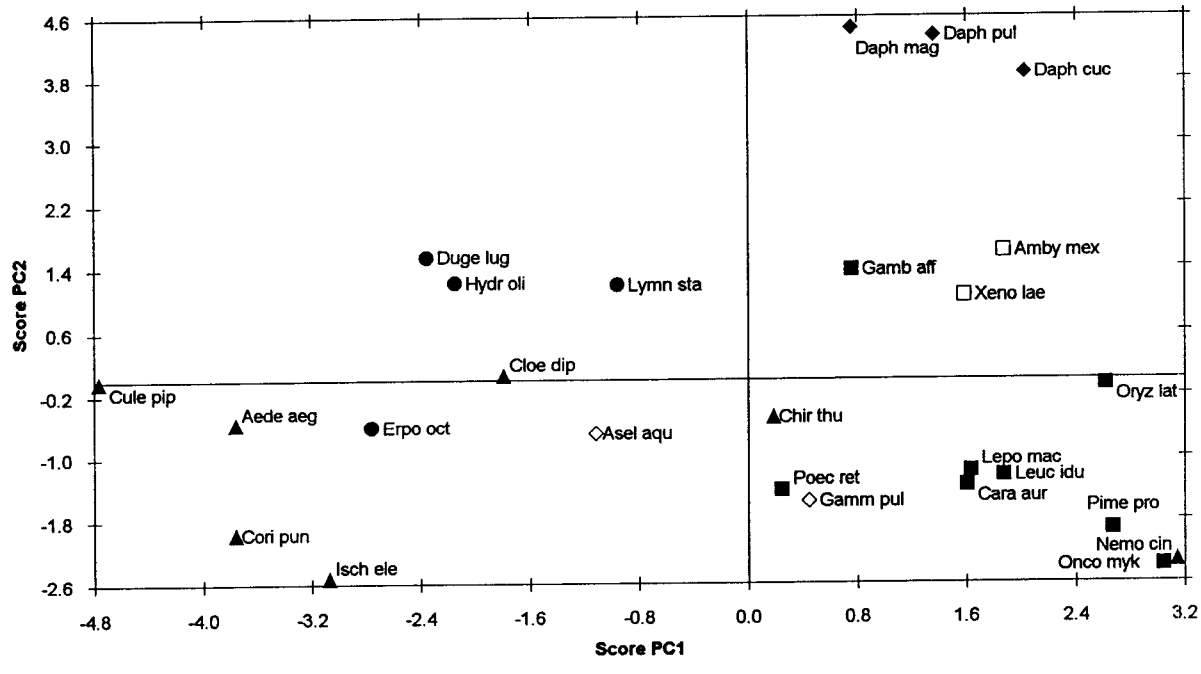
FIGURE 1 Description of PCA method. Reduction of a three-dimensional (p) space to an easier to interpretate two-dimensional picture. This two-dimensional picture is described by two components (PC1 and PC2), the swarm of object points is projected on this plane. The smaller this distance, the better the components describe the swarm of datapoints. The figure in the right corner demonstrates the matrix notation. From Kowalski, 1983.



▲ Class I □ Class III ● Heavy metal

♦ Class II ■ Class IV

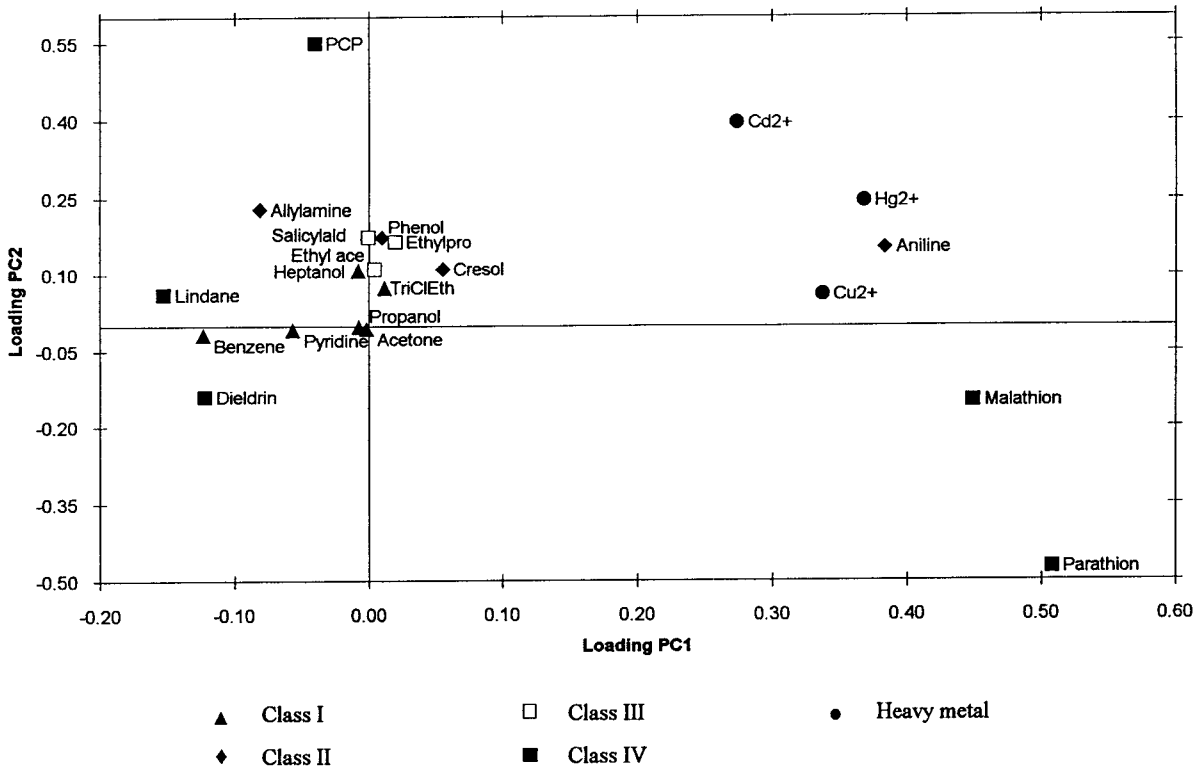
FIGURE 2 Loadings of compounds that determine patterns in species sensitivity (Figure 3), scaled data.



■ Osteichtyes ▲ Insecta ◇ Malacostraca

□ Amphibia ♦ Phyllopoda ● Rest

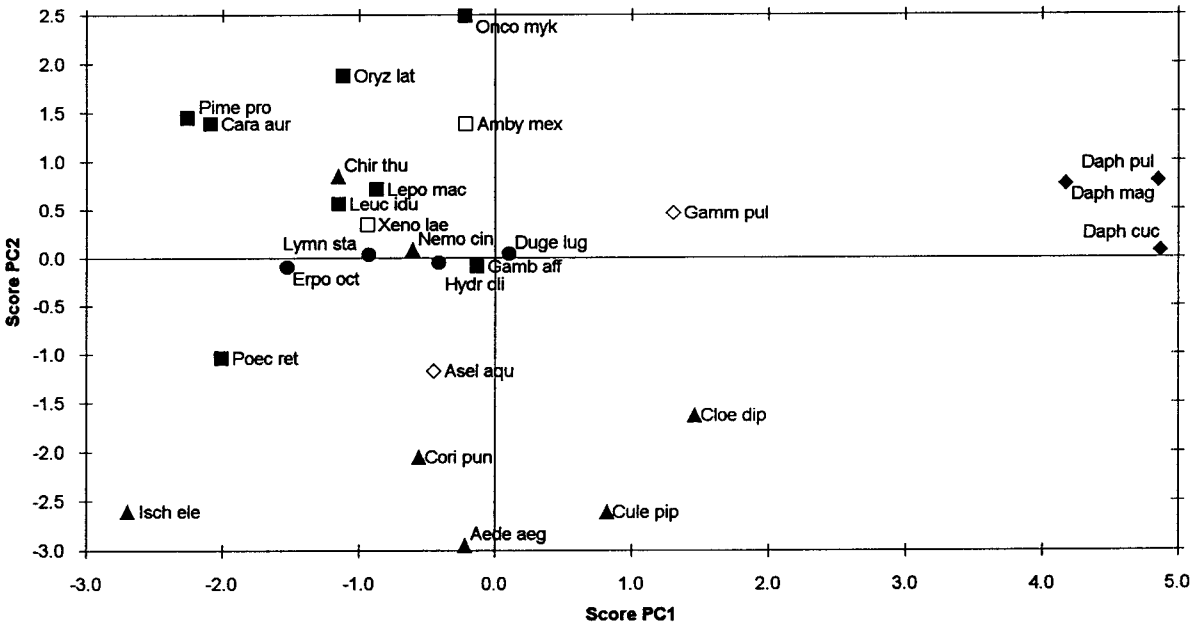
FIGURE 3 Patterns in species sensitivity, plot of scores of species, scaled data. Species has high sensitivity to compound on corresponding position in Figure 2.1



▲ Class I □ Class III ● Heavy metal

♦ Class II ■ Class IV

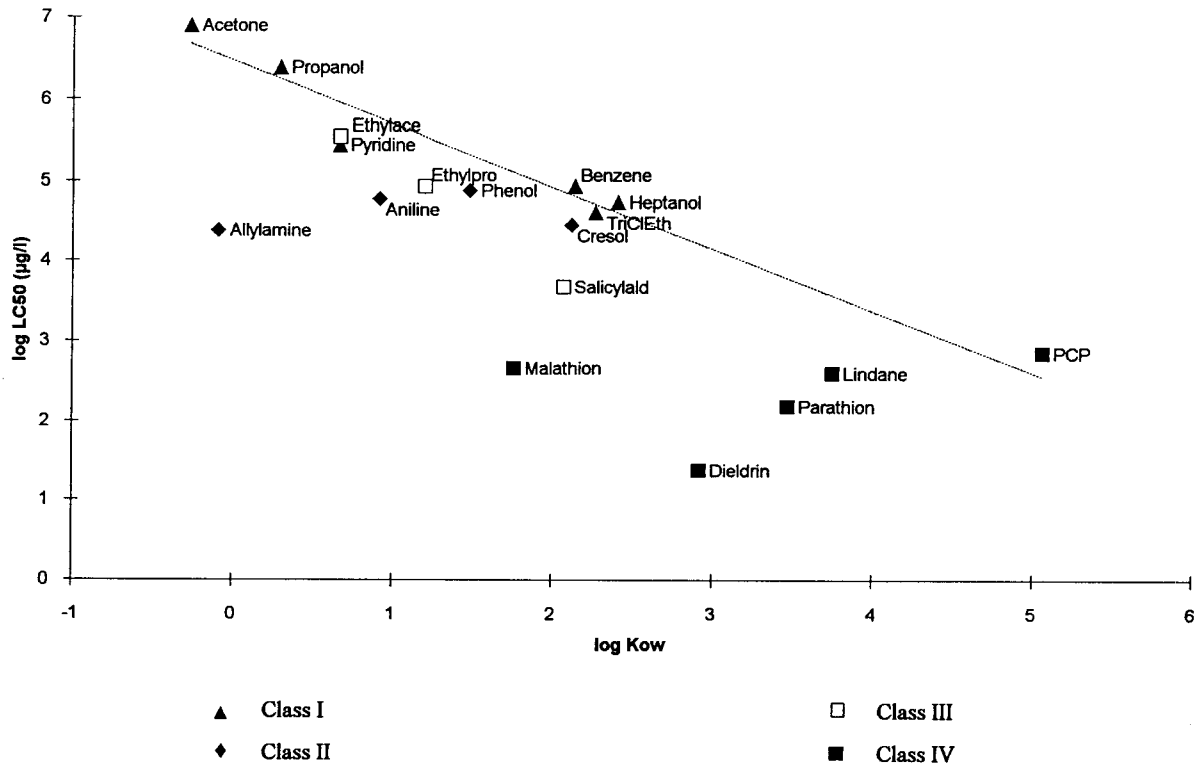
FIGURE 4 Loadings of compounds that determine patterns in species sensitivity (Figure 5), unscaled data.



■ Osteichtyes ▲ Insecta ◇ Malacostraca

□ Amphibia ♦ Phyllopoda ● Rest

FIGURE 5 Patterns in species sensitivity, plot of scores of species, unscaled data. Species has high sensitivity to compound on corresponding position in Figure 4.



▲ Class I
□ Class III
♦ Class II
■ Class IV

FIGURE 6 Average toxicity of compounds in relation to their hydrofobicity (log K_{ow}). Line drawn is the baseline toxicity, based on the class I (non-polar narcotic) compounds: log LC_{50} = 6.47 - 0.77 (log K_{ow}), r^2 = 0.91, $p \leq 0.01$.

APPENDIX

APPENDIX 1 Additional criteria used for the selection of data

Criteria used to reduce standard deviations larger than 0.5 (log scale) of species-compound combinations with more than one effect concentration.
Results of experiments were excluded if:

- compared to the other results relatively extreme experimental conditions were used (e.g. temperature, hardness, salinity, pH)
- compared to the other results relatively very different life stages were used (e.g. an adult versus many juveniles, a heavy juvenile vice light larvae)
- resistant populations were used
- in case of *Oncorhynchus mykiss* data were derived from egg or sac fry lifestages
- they could be identified as outliers by the Grubbs test (F.E. Grubbs and G. Beck, Technometrics, 14(4):847-854, 1972; F.E. Grubbs, Ann. of Math. Stat., 21(10):27-58, 1950; F.E. Grubbs, Technometrics 11(1):1-21, 1969).

APPENDIX 2 Data of salts used in the analyses.

Salts mentioned in the Appendix are used to derive data for the metal-ions and pentachlorophenol. The respective compound name and abbreviation used in the report for these salts are also presented.

CAS-nr	Salts	CAS-nr	Compound name	Abbreviation
543908	Cadmium acetate	10108642	Cadmium chloride	Cd2+
7790809	Cadmium iodide			
10108642	Cadmium chloride			
10124364	Cadmium sulfate			
10325947	Cadmium nitrate			
142712	Cupric acetate	7447394	Cupric chloride	Cu2+
1317380	Cupric oxide			
3251238	Cupric nitrate			
7447394	Cupric chloride			
7758987	Cupric sulfate			
20427592	Cupric hydroxide			
1600277	Mercuric acetate	7487947	Mercuric chloride	Hg2+
7487947	Mercuric chloride			
10045940	Mercuric nitrate			
87865	Pentachlorophenol	87865	Pentachlorophenol	PCP
131522	NaPCP			

APPENDIX 3 Matrix of data, standard deviation of the mean log LC_{50} (µg/l) per species-compound combination.

Compound/ Species	Aede aeg	Ambly mex	Asel aqu	Cara aur	Chir thu	Cloe dip	Cori pun	Cule pip	Daph cuc	Daph mag	Daph pul	Duge lug	Erpo oct	Gamb aff	Gamm pul	Hydr oli	Isch ele	Lepo mac	Leuc idu	Lyms sta	Nemo cin	Onco myk	Oryz lat	Plume pro	Poec ret	Xeno lae
Acetone	-	-	-	-	-	-	-	-	0.01	0.15	-	-	-	0.00	-	-	-	0.00	0.09	-	-	-	-	0.05	-	-
Allylamine	-	-	-	0.46	-	-	-	-	0.01	0.09	0.04	-	-	-	-	-	-	-	0.08	-	-	-	-	-	-	0.28
Aniline	-	-	-	-	-	-	-	-	0.00	0.33	-	-	-	-	-	-	-	-	0.06	-	-	0.20	0.95	0.24	-	0.29
Benzene	-	-	-	0.00	-	-	-	-	0.03	0.16	0.08	-	-	0.01	-	-	-	0.33	0.30	-	-	0.13	-	0.17	0.06	0.00
Cd2+	-	-	0.91	-	-	0.73	-	-	0.04	0.46	0.50	-	0.28	-	0.63	-	-	0.33	0.49	-	-	0.19	-	0.36	-	0.51
Cresol	-	-	-	-	-	-	-	-	0.04	0.17	0.07	-	-	-	-	-	-	0.00	0.18	-	-	-	-	0.11	0.09	-
Cu2+	-	-	-	-	-	-	-	-	-	0.49	0.19	-	-	0.77	0.12	-	-	0.68	0.00	-	-	0.60	-	0.59	0.25	-
Dieldrin	-	-	-	0.00	-	-	-	0.03	-	-	-	-	-	0.17	-	-	-	0.19	-	-	-	0.43	-	0.30	0.37	0.30
Ethylace	-	-	-	-	-	-	-	-	0.04	0.06	0.08	-	-	-	-	-	-	-	0.05	-	-	0.14	-	-	-	-
Ethylpro	-	-	-	-	-	-	-	-	0.02	0.13	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Heptanol	-	-	-	-	-	-	-	-	0.01	0.04	-	-	-	-	-	-	-	-	0.08	-	-	-	-	-	-	-
Hg2+	-	-	-	-	-	-	-	-	0.01	0.34	0.00	-	-	0.30	-	-	-	0.11	0.10	-	-	0.49	0.04	0.09	0.18	0.00
Lindane	-	-	-	0.11	-	-	-	0.58	-	0.25	-	-	-	0.49	0.42	-	-	0.49	0.74	0.22	-	0.11	-	0.09	0.22	-
Malathion	-	-	-	0.08	-	-	-	0.18	-	0.93	0.00	-	-	-	-	-	-	0.16	-	-	-	0.10	-	0.18	0.32	-
Parathion	-	-	0.13	0.11	-	0.10	-	0.74	-	0.15	0.09	-	-	0.33	-	-	-	0.42	0.13	-	-	0.01	-	0.16	0.06	-
PCP	-	-	-	0.13	-	-	-	-	-	0.29	0.18	-	-	0.01	-	-	-	0.27	0.43	-	-	0.32	-	0.30	0.29	-
Phenol	-	-	0.03	0.03	-	-	-	-	-	0.44	0.03	-	-	0.13	0.13	-	-	0.08	0.09	-	-	0.12	-	0.12	0.09	-
Propanol	-	-	-	-	-	-	-	-	-	0.04	0.02	-	-	-	-	-	-	-	0.02	-	-	-	-	0.01	-	-
Pyridine	-	-	-	-	-	-	-	-	0.02	0.11	0.06	-	-	0.01	-	-	-	-	0.04	-	-	0.00	-	0.04	-	0.18
Salicylald	-	-	-	-	-	-	-	-	-	0.13	0.05	-	-	-	-	-	-	-	0.06	-	-	-	-	-	-	-
TriClEth	-	-	-	-	-	-	-	-	0.01	0.25	0.08	-	-	-	-	-	-	-	0.10	-	-	-	-	0.08	0.00	-

APPENDIX 4 Matrix of data, number of data per species-compound combination.

Compound/ Species	Aede aeg	Arnby mex	Asel aqu	Cara aur	Chir tbu	Cloe dip	Cori pun	Cule pip	Daph cuc	Daph mag	Daph pul	Duge lug	Erpo oct	Gamb aff	Gamm pul	Hydr oli	Lepo mac	Leuc idu	Lyms sta	Onco myk	Oryz lat	Plme pro	Poec ret	Xeno lae	n (=21)
Acetone	1	1	1	0	1	1	1	1	2	5	1	1	1	2	1	1	2	3	1	1	1	5	1	1	38
Allylamine	1	1	1	2	1	1	1	1	2	6	2	1	1	0	1	1	0	3	1	1	1	1	1	2	34
Aniline	1	1	1	1	1	1	1	1	2	8	1	1	1	0	1	1	1	3	1	6	2	5	1	10	54
Benzene	1	1	1	2	1	1	1	1	2	8	2	1	0	2	1	1	4	3	1	3	1	6	3	2	51
Cd2+	1	1	4	1	0	5	0	1	2	59	42	0	3	1	3	1	2	3	1	6	1	15	1	3	157
Cresol	1	1	1	1	1	1	1	1	2	7	2	1	1	0	1	1	2	2	1	1	1	6	2	1	41
Cu2+	0	0	1	1	0	1	0	0	0	27	6	0	0	10	3	0	24	2	0	15	0	77	19	1	197
Dieldrin	1	0	0	2	0	0	0	2	0	0	1	0	0	2	0	0	10	1	0	10	1	5	9	4	51
Ethylace	1	1	1	0	1	1	1	1	2	6	2	1	1	0	1	1	0	3	1	3	1	1	1	1	34
Ethylpro	1	1	1	0	1	1	1	0	2	6	2	1	1	0	1	1	0	1	1	1	1	1	0	1	28
Heptanol	1	1	1	0	1	1	1	1	2	5	1	1	1	0	1	1	0	3	1	1	1	1	1	1	30
Hg2+	1	1	1	1	1	1	1	1	2	9	2	1	1	4	1	1	3	3	1	5	2	2	2	2	52
Lindane	0	0	1	5	0	0	0	5	0	3	1	0	0	6	12	0	14	3	2	11	1	6	4	0	83
Malathion	1	0	0	3	0	0	0	6	0	2	2	0	0	1	0	0	8	0	0	13	1	7	3	0	58
Parathion	1	0	4	4	0	3	0	2	0	4	2	0	0	6	0	0	9	3	0	2	1	13	2	0	70
PCP	1	1	1	16	1	1	1	1	1	65	42	1	1	3	1	1	20	2	1	16	1	56	10	1	252
Phenol	0	0	2	2	0	1	0	0	0	24	5	1	1	6	8	0	8	3	0	27	0	27	9	0	126
Propanol	1	1	1	0	1	1	1	1	1	6	2	1	1	0	1	1	0	3	1	1	1	2	1	1	32
Pyridine	1	1	1	0	1	1	1	1	2	6	2	1	1	2	1	1	0	3	1	3	1	2	1	9	47
Salicylald	1	1	1	0	1	1	1	1	1	3	2	1	1	0	1	1	0	3	1	1	1	1	1	1	28
TriClEth	1	1	1	0	1	1	1	1	2	7	2	1	1	0	1	1	1	6	1	1	1	9	2	1	46
n (=26)	18	15	26	41	14	24	14	29	27	266	124	15	17	45	40	15	108	56	17	128	21	248	74	42	1509