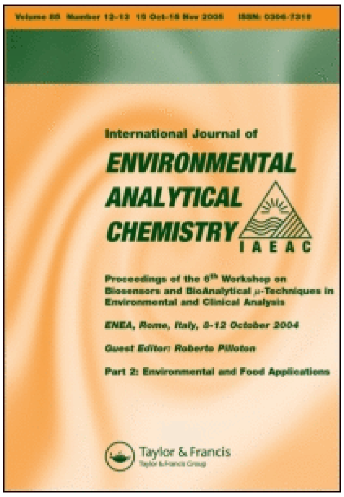

International Journal of Environmental Analytical Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713640455

Ranking and classification of non-ionic organic pesticides for environmental distribution: a qsar approach

Paola Gramatica[a]; Ester Papa[a]; Battaini Francesca[a]
[a] Department of Structural and Functional Biology, QSAR and Environmental Chemistry Research Unit, University of Insubria, 21100 Varese, Italy

To cite this Article Gramatica, Paola , Papa, Ester and Francesca, Battaini(2004) 'Ranking and classification of non-ionic organic pesticides for environmental distribution: a qsar approach', International Journal of Environmental Analytical Chemistry, 84: 1, 65 — 74

To link to this Article: DOI: 10.1080/0306731031000149732

URL: http://dx.doi.org/10.1080/0306731031000149732

Intern. J. Environ. Anal. Chem.
Vol. 84, No. 1–3, pp. 65–74, January–March 2004

Taylor & Francis
Taylor & Francis Group

RANKING AND CLASSIFICATION OF NON-IONIC ORGANIC PESTICIDES FOR ENVIRONMENTAL DISTRIBUTION: A QSAR APPROACH

PAOLA GRAMATICA*, ESTER PAPA and FRANCESCA BATTAINI

Department of Structural and Functional Biology,
QSAR and Environmental Chemistry Research Unit, University of Insubria,
via Dunant 3, 21100 Varese, Italy

(Received 29 September 2002; In final form 3 April 2003)

The partitioning of pesticides into different environmental compartments depends mainly on the physico-chemical properties of the studied chemical. To rank pesticides according to their distribution tendency in various media, we propose a combination of two multivariate approaches: Principal Component Analysis and Hierarchical Cluster Analysis. In such explorative methods we take into account physico-chemical properties like the organic carbon partition coefficient (K_{oc}), the *n*-octanol/water partition coefficient (K_{ow}), water solubility (S_w), vapour pressure (VP) and Henry's law constant (H), being the more relevant to the determination of environmental partitioning.

The 54 studied pesticides of various chemical categories are thus ranked in 4 *a priori* classes according to their environmental behaviour (sorbed, soluble, volatile and non-volatile/medium class) and finally assigned to the defined four classes by different classification methods (CART, K-NN, RDA) using theoretical molecular descriptors. This QSPR approach allows a rapid indication of the environmental distribution of pesticides starting only from their molecular structure.

Keywords: Pesticides; PCA; Clustering; Classification; CART; QSPR

INTRODUCTION

Pesticides have provided mankind with powerful weapons against insect pests, disease and weeds, and this has resulted in great economic and health benefits to society. The widespread use and distribution of pesticide chemicals is of great relevance and their potential for adverse consequences, such as their environmental impact on the quality of water and wildlife habitats, has led to the development of detailed analyses for potential environmental hazards associated with their use [1–3]. Better knowledge of pesticide environmental behaviour, i.e., absorption in soil, possibility of leaching in groundwater

---

*Corresponding author. Fax: +39-0332-421554. E-mail: paola.gramatica@uninsubria.it

ISSN 0306-7319 print: ISSN 1029-0397 online
DOI: 10.1080/0306731031000149732

and volatility in the atmosphere, is a primary goal for an accurate environmental and human risk assessment. Environmental behaviour is strongly influenced by the properties inherent in the compounds themselves, particularly physico-chemical properties such as solubility in water, vapour pressure and partitioning coefficients between organic matter in soil or biota and water.

When assessing pesticide environmental distribution it is most important not only to know, but also to have an understanding of, these properties. Such knowledge can be used to develop a simple and preliminary approach to rank specific compounds according to their intrinsic partitioning tendency, categorising them in different environmental compartments [4–7]. The main goal of this work has been to develop a simple procedure based on a QSAR/QSPR (Quantitative Structure–Activity or Property Relationships) approach for a preliminary screening, ranking and classification of organic pesticides (including those not yet synthesised) according to their environmental partitioning, using only the knowledge of their chemical structure. This initial classification could be invaluable in selecting (at the chemical development level, or later) substances for which further investigations should be made, for instance for dangerous potential leaching in groundwater. Such a preliminary selection could lead to better investment in research efforts for ‘‘environmentally friendly’’ chemicals.

This article deals with a heterogeneous and representative data set of non-ionic pesticides of different organic classes (acetanilides, carbamates, dinitroanilines, organochlorines, organophosphates, phenylureas, triazines); the data set was selected from a bigger data set studied by our team in recent years. These pesticides have already been the subject of QSPR studies, using theoretical molecular descriptors in modeling the K_{oc} and leaching and volatility indexes (LIN and VIN), recently published by Gramatica *et al.* [8, 9].

METHODS

Experimental Data

A data set of 54 non-ionic organic pesticides, representative of different chemical classes, was selected. In order to obtain a reliable and homogeneous data set of the considered partition properties, soil adsorption coefficients were taken from the work of Sabljic *et al.* [10], where the K_{oc} values measured in soils with a low carbon content had been excluded as the pesticide interaction with the inorganic matrix of the soil could have become important; the predicted values were calculated by our published QSPR models [8]. Also the *n*-octanol/water partition coefficients are those of Sabljic *et al.* Water solubility and vapour pressure data are from *The Pesticide Manual* [11]. The Henry's law constants were taken from a critical review made by Suntio *et al.* [12]. The data, all measured at 25°C, were always transformed into logarithmic units and are reported in Table I.

Molecular Descriptors

The theoretical molecular descriptors were computed using the *DRAGON* package of Todeschini and Consonni [13] downloadable (gratis) from the web. The input

TABLE I Experimental data and classes of 54 pesticides

ID	*Compounds*	*CAS*	*Log* K_{oc}	*Log* S_w	*Log* K_{ow}	*Log VP*	*Log H*	*Class*
1	Alachlor	15972-60-8	2.28	2.38	3.52	0.27	−2.21	4
2	Propachlor	1918-16-7	2.42	2.79	2.18	1.49	−1.96	2
3	Aldicarb	116-03-3	1.50	3.78	1.13	0.60	−3.49	1
4	Butylate	2008-41-5	2.11	1.66	4.15	3.24	−0.25	2
5	Carbaryl	63-25-2	2.40	2.08	2.36	−0.80	−2.89	4
6	Carbofuran	1563-66-2	1.75	2.55	2.32	−1.10	−3.29	4
7	Chlorpropham	101-21-3	2.53	1.95	3.51	0.03	−2.68	4
8	Diallate (*cis*)	2303-16-4	3.28	1.15	3.67	1.30	−0.60	2
9	Diallate (*trans*)	2303-16-4	3.28	1.15	3.67	1.30	−0.60	2
10	EPTC	759-94-4	2.38	2.54	3.21	3.66	0.01	2
11	Methomyl	16752-77-5	1.30	4.76	0.60	0.82	−4.19	1
12	Oxamyl	23135-22-0	1.00	5.45	−0.47	1.49	−3.59	1
13	Propoxur	114-26-1	1.67	3.26	1.52	0.11	−0.89	2
14	Triallate	2303-17-5	3.35	0.60	4.53	1.17	0.01	2
15	Vernolate	1929-77-7	2.33	2.03	3.84	3.11	0.31	2
16	Pirimicarb	23103-98-2	1.90	3.43	1.70	0.60	−3.49	1
17	Benfluralin	1861-40-1	3.99	−1.00	5.29	0.94	0.13	3
18	Dinitramine	29091-05-2	3.63	0.04	3.89	−0.32	−0.80	3
19	Profluralin	26399-36-0	4.01	−1.00	5.08	0.92	1.59	3
20	Trifluralin	1582-09-8	3.93	−0.52	5.34	1.17	0.60	3
21	Aldrin	309-00-2	4.69	−1.57	6.50	−0.06	1.96	3
22	Chlordane	12709-03-6	5.15	−1.22	5.80	0.12	0.96	3
23	*p,p*-DDT	50-29-3	5.31	−2.26	6.91	−1.60	0.37	3
24	*p,p*-DDE	72-55-9	4.82	−1.00	6.96	−0.06	0.90	3
25	Dieldrin	60-57-1	4.55	−0.74	5.20	−0.40	0.05	3
26	Endosulfan	115-29-7	4.13	−0.49	3.83	−1.65	0.47	3
27	Lindane	58-89-9	3.00	0.85	3.76	0.64	−0.89	2
28	Azinphos methyl	86-50-0	2.28	1.46	2.75	−1.58	−2.49	4
29	Carbophenothion	786-19-6	4.66	0.00	5.66	0.04	−1.34	3
30	Chlorpyrifos	2921-88-2	3.70	−0.40	5.27	0.35	0.24	3
31	Chlorpyrifos methyl	5598-13-0	3.52	0.60	4.31	0.75	−0.47	2
32	Diazinon	333-41-5	2.75	1.78	3.81	0.90	−1.17	2
33	Dimethoate	60-51-5	1.20	4.60	0.78	0.52	−3.96	1
34	Disulfoton	298-04-4	3.22	1.40	4.02	1.30	−0.66	2
35	Ethion	563-12-2	4.06	0.04	5.07	−0.46	−1.49	3
36	Fenitrothion	122-14-5	2.63	1.48	3.30	−0.88	−2.44	4
37	Malathion	121-75-5	3.07	2.11	2.36	0.03	−2.63	4
38	Parathion ethyl	56-38-2	3.20	1.38	3.83	−0.18	−1.92	4
39	Parathion methyl	298-00-0	3.00	1.78	2.86	0.30	−1.68	4
40	Phorate	298-02-2	2.70	1.34	3.56	1.93	−0.19	2
41	Trichlorfon	52-68-6	1.90	5.08	0.51	−0.58	−5.77	1
42	Dichlorvos	62-73-7	1.67	4.00	1.43	2.42	−0.72	2
43	Fenuron	101-42-8	1.40	3.59	0.98	1.42	−3.57	1
44	Monuron	150-68-5	1.95	2.36	1.94	−1.18	−2.52	4
45	Diuron	330-54-1	2.40	1.62	2.68	−2.04	−2.92	4
46	Linuron	330-55-2	2.70	1.88	3.20	0.35	−2.27	4
47	Ametryn	834-12-8	2.59	2.27	2.98	−0.44	−3.92	4
48	Atrazine	1912-24-9	2.24	1.52	2.61	−1.41	−3.54	4
49	Prometon	1610-18-0	2.60	2.86	2.99	0.01	−4.04	4
50	Prometryn	7287-19-6	2.85	1.52	3.51	−0.78	−3.30	4
51	Propazine	139-40-2	2.40	0.93	2.93	−1.76	−4.00	4
52	Simazine	122-34-9	2.10	0.79	2.18	−2.53	−3.47	4
53	Secbumeton	26259-45-0	2.78	2.78	3.20	−0.01	−3.46	4
54	Terbutryn	886-50-0	2.85	1.34	3.74	−0.55	−2.89	4

S_w mg/L; VP mm Hg; H atm m^3/mol. All at 25°C.

files for descriptor calculation, containing information on atom and bond types, connectivity, partial charge and atomic spatial coordinates relative to the minimum energy conformation of the molecule, were obtained by the molecular mechanics method of Allinger (MM+), using the *HYPERCHEM* package [14]. An initial set of 236 descriptors was used to describe compound structural diversity and to select those useful for the studied classification. We calculated the following: (a) constitutional descriptors (0D and 1D-descriptors, i.e., counting of atoms, bonds and fragments, MW and sum of atomic properties); (b) topological descriptors (2D-descriptors from molecular graphs); (c) WHIM (*Weighted Holistic Invariant Molecular*) descriptors [15] that contain information on the whole 3D-molecular structure in terms of size, symmetry and atom distribution and (d) empirical descriptors: unsaturation index (Ui) and hydrophilic factor (Hy). We also added the number of hydrogen atoms for hydrogen bonds (nHDon) and the number of atom acceptors of hydrogen in the same type of bonds (nHAcc). The meaning of the calculated descriptors for the studied compounds is reported in the cited software and in the literature [16].

Chemometric Methods

Data exploration and multivariate analysis of physico-chemical properties by *Principal Component Analysis* and *Hierarchical Cluster Analysis* was performed on autoscaled data by the SCAN program [17] for the definition of *a priori* classes. In the *Cluster Analysis* the complete linkage and the Euclidean distance among the chemicals on the autoscaled variables (the five studied physico-chemical properties) were applied. The classification models were also obtained using the SCAN package. *Classification And Regression Tree* (CART) is the nonparametric classification strategy [18] that makes an automatic stepwise variable selection (among the 236 molecular descriptors used as input), and displays, as the final result, a binary classification tree that is applicable immediately. The proportional class prior and the splitting criterion of Gini were applied to autoscaled variables. Another classification method used on the descriptors selected by CART was the K-Nearest Neighbor (KNN), a classification method that searches for the k nearest neighbours of each object in the data set, performing the classification of the considered object by considering the majority of the classes to which the k-th nearest objects belong. This method was applied to autoscaled data with *a priori* probability proportional to the size of the classes; the predictive power of the model was checked for K values between 1 and 10.

The third classification method applied was *Regularised Discriminant Analysis* (RDA) (optimal $\lambda = 0.25$ and $\gamma = 0.00$).

To check the classification model prediction ability, Misclassification Risk (MR%) and Misclassification Risk in prediction (MR_{cv}%), calculated by the *leave-one-out* procedure (one chemical out of the training set and predicted by the model at each step), were used. Comparison with the No-Model Misclassification Risk (NOModMR%) allowed the evaluation of model performance: it is a reference measure for classification without any classification model, i.e., all the objects are considered as belonging to the most numerous class and the misclassification risk is calculated as the ratio between the number of these objects and the total number of objects.

RESULTS AND DISCUSSION

The principal aim of this work is the proposal of a simple approach that uses only structural information for pesticide screening/ranking according to the distribution tendency in different environmental media. The first step in this screening/ranking procedure is to combine, by Principal Component Analysis and Hierarchical Cluster Analysis, the most relevant partitioning properties for the grouping of chemicals with "similar" behaviour and the definition of *a priori* classes. These classes are then modelled by different classification methods using theoretical molecular descriptors. This QSPR approach can be used to calculate the partitioning tendency of a new chemical on the basis of only structural information in a pre-screening phase.

Pesticide Ranking According to Environmental Partitioning

Pesticide environmental behaviour is clearly controlled by a variety of physical and chemical processes that are influenced, simultaneously, by several physico-chemical properties of the compounds, properties that are particularly relevant in determining pesticide distribution throughout the different environmental compartments. These basic physico-chemical properties of environmental concern are vapour pressure (VP), water solubility (S_w), and various partition coefficients among different compartments: Henry's law constant (H), octanol/water partition coefficient (K_{OW}), organic carbon sorption coefficient (K_{OC}). For instance, leaching in water is much more evident for chemicals that have, simultaneously, relatively high solubility and low sorption capacity, and this calls for a multivariate approach.

Principal Component Analysis (PCA) and Hierarchical Cluster Analysis are explorative multivariate techniques that, applied to the above-mentioned physico-chemical properties, allow the fast ranking and grouping of pesticides according to their similar environmental behaviour and distribution tendency in different media, contemporaneously taking into account the more environmentally involved physico-chemical properties.

Figure 1 shows the bi-plot of the PCA for 54 pesticides described by the studied physico-chemical properties, the chemicals (dots) being ranked in the space defined by the first two principal components according to their relative partitioning tendency in the different media. These principal components give most of the information in the data: in fact, the cumulative explained variance of the first two principal components is 94.6% and the first component alone provides most of the information (PC1 = 70.1% and PC2 = 24.5%). The loading plot (lines in Fig. 1) reveals the relevance of each variable in each of the first two principal components. It is interesting to note that PC1, where the solubility and sorption coefficients play opposite roles, tends to discriminate between the relatively more-sorbed/less-soluble and the more-soluble/less-sorbed pesticides, while PC2, where volatilisation parameters (vapour pressure and Henry's law constant) dominate, appears to differentiate between volatile and non-volatile compounds. The Henry's law constant has a similar influence on both PC1 and PC2, as can be expected from its calculation (vapour pressure/water solubility).

Some chemicals show extreme behaviour, lying along the edges of Fig. 1: for instance aldrin (21), chlordane (22), *p,p*-DDT (23) and *p,p*-DDE (24) appear to be the most-sorbed/less-soluble compounds in the studied data set, being to the right in the graph along PC1, while methomyl (11), oxamyl (12), dimethoate (33), and trichlorfon (41)

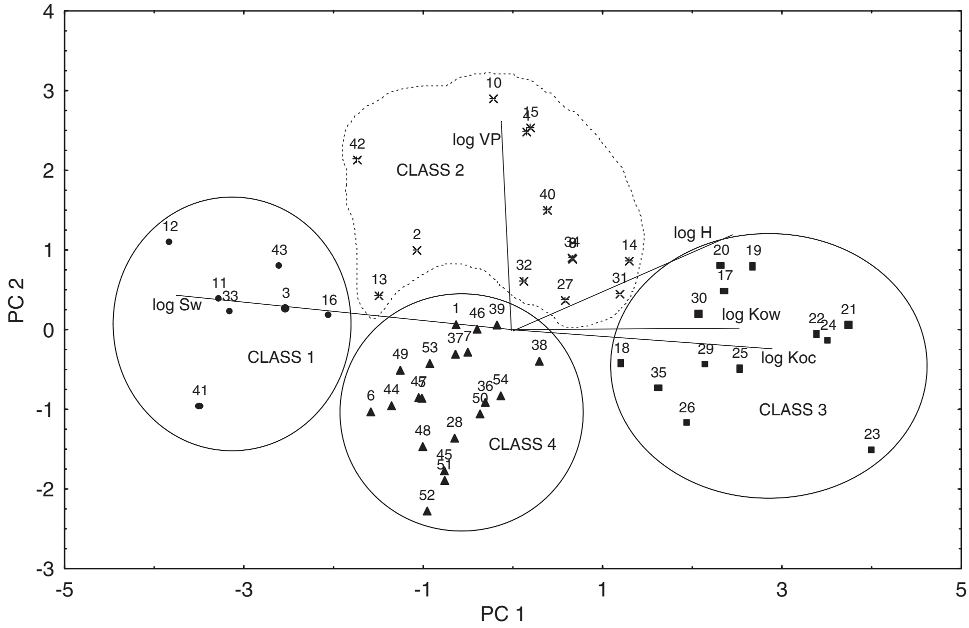

FIGURE 1 Score plot and loading plot of the two first principal components of PCA of five physico-chemical properties (K_{oc}, K_{ow}, S_{w}, VP, Henry's law constant) for 54 pesticides. Cumulative explained variance: 94.6%; expl. var. of PC1: 70.1%.

appear the most-soluble/less-sorbed compounds, being to the left in the graph. The number in parentheses are taken from Table I. Along PC2, the compounds diuron (45), propazine (51) and simazine (52) appear the least volatile, having the lowest score. On the other hand the compounds butylate (4), EPTC (10), vernolate (15) and dichlorvos (42) appear the most volatile because of their high score. The results of the PCA-based screening are similar to those of other approaches, e.g., the GUS index [6] or the Hasse diagram [7], but differ in a way that offers a multimedia and simultaneous picture of the possible partition behaviour, showing possible slight differences among the chemicals. This translates into a fast screening step of the studied chemicals for their partitioning tendency in a multimedia environment, revealing PCA to be a convenient explorative method for an initial assessment.

As the pesticide ranking obtained by PCA alone does not allow the grouping of the pesticides into defined classes of environmental behaviour (in practice groups that are not well separated are identifiable in the PCA-graph of Fig. 1), the same chemicals are then grouped again by combining the same studied physico-chemical properties by Hierarchical Cluster Analysis (using the complete linkage and the Euclidean distance metric); this results in four clusters. Figure 2 shows the dendrogram of this cluster analysis. Note that four clusters are evident: the most-soluble/least-sorbed pesticides are grouped in Cluster 1, and the most-sorbed/least-soluble pesticides are collected in Cluster 3, the most and the least volatile pesticides (together with the compounds of medium behaviour) are divided in two central clusters (Clusters 2 and 4, respectively).

At this stage, the combination of the two multivariate explorative analyses allow the grouping of the studied pesticides into four classes, corresponding to the above

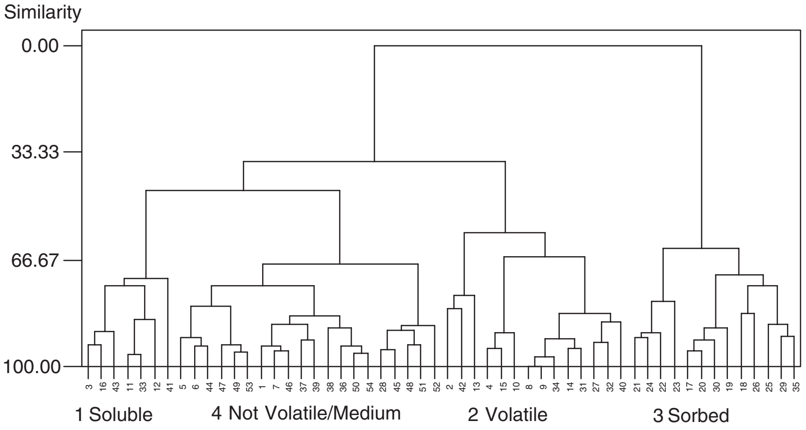

FIGURE 2 Dendrogram of cluster analysis: cluster 1, soluble pesticides; cluster 2, volatile pesticides; cluster 3, sorbed pesticides; cluster 4, non-volatile/medium pesticides.

clusters: soluble, sorbed, volatile and non-volatile (or medium behaviour). These four *a priori* classes of pesticides are highlighted in the PCA-graph (Fig. 1) by the circles and reported in Table I.

Classification of Pesticide Environmental Distribution

The final step in this work is the proposal of a fast and simple tool for classifying non-ionic organic pesticides for their environmental distribution starting only from molecular structure using the QSAR/QSPR approach.

QSAR (Quantitative Structure–Activity Relationships) or QSPR (Quantitative Structure–Property Relationships) studies are based on the fundamental assumptions of Corvin Hansch [19], considered the father of this approach. He demonstrated that "the molecular structure of a chemical influences its physico-chemical properties and biological activity" and that "structurally similar compounds behave similarly".

Classification models are quantitative models based on relationships between one or more independent variables (here the theoretical structural descriptors) and a categorical response variable of integer numerical values, each representing the class of the corresponding sample (here the *a priori* defined classes). For new compounds, where class is obviously unknown, the classification model can be useful to predict an assignment to a defined class. The great advantage of the proposed classification model is that it is possible to assign each compound to a class using only a few molecular descriptors, and the same holds also for heterogeneous chemicals. Different chemometric classification procedures like CART (*Classification And Regression Tree*), RDA (*Regularised Discriminant Analysis*) and K-NN (*K-Nearest Neighbours*) have been applied here.

A first classification model was calculated by CART (Classification And Regression Tree) by using as input all the calculated molecular descriptors, as defined in the Methods section. CART is the non-parametric classification strategy that makes an automatic stepwise variable selection (among the 236 input molecular descriptors) of the descriptors most useful in the discrimination of the chemicals in the *a priori* classes,

FIGURE 3 Classification tree by CART (Classification and Regression Tree). Misclassification risk: 11.11%; misclassification risk in prediction: 18.53%; NoModMR: 62.96%.

and displays as the final result a binary classification tree (Fig. 3), that can be interpreted easily and is simple to use for further predictions.

The performance of the obtained model is quite good, having a misclassification risk (MR%) in fitting of 11.11% and a cross-validated misclassification risk in prediction (MR_{CV}%) of 18.53%. These results are very satisfactory compared with the corresponding misclassification risk of 62.96% in absence of the model. The obtained CART model is based on three very simple molecular descriptors that hierarchically assign each compound to a predefined class (the integer number on the graph baseline in Fig. 3). At each knot of the tree the chemicals with a selected descriptor value higher than the reported cut-off value are assigned to the right class. The first discriminating descriptor is the molecular weight (MW), related to molecular size in terms of the number of atoms, and also to the kind of atoms, in the molecule; the cut-off value, reported in the knot of the tree, allows a first separation of the compounds: pesticides with MW higher than 317.69 are assigned to Class 3 of the most-sorbed/least-soluble chemicals. The topological Balaban index (named J in the software DRAGON, calculating the mean squared distance between the atoms in the chemical structure) is a further discriminator: a value of 2.53 separating all the chemicals of the non-volatile/medium class (Class 4) from the other pesticides. Finally, the number of hydrogen donor groups for hydrogen bonds (nHDon) (for instance OH, NH, etc) allows the separation of the pesticides into the two classes of most-soluble/least-sorbed (Class 1) and most volatile (Class 2). All the selected molecular descriptors are quite easily interpretable: (a) the MW encodes information on the molecule's dimensions and it is well known that big molecules have the greatest tendency to bind, by van der Waals forces, to the organic component of the soil, becoming the most sorbed in organic soils but the least soluble in water (Class 3); (b) the possibility of a chemical forming hydrogen

TABLE II Classification model results

Method	*N. Obj.*	*No-model MR%*	*MR%*	*MR_{cv} %*	*Molec. descriptors*
CART	54	62.96	11.11	18.53	MW - J - nHDon
K-NN	54	62.96	/	16.67	MW - J - nHDon
RDA	54	62.96	14.81	16.67	MW - J - nHDon

CART: Classification and Regression Tree; K-NN: K-Nearest Neighbour; RDA: Regularized Discriminant Analysis. MR: Misclassification Risk; MRcv: Misclassification Risk in prediction; No-Model MR: Misclassification Risk without a model. MW: Molecular Weight, J: topological Balaban index; nHDon: number of donor atom in hydrogen bonds.

bonds with water molecules (encoded in the molecular descriptor nHDon) results in the higher solubility of the Class 1 pesticides; furthermore the chemical with fewer intramolecular hydrogen bonds are the most volatile (Class 2); (c) the last topological descriptor J, that discriminates Class 4 of the medium-behaviour pesticides, is not easily interpretable. The same three molecular descriptors, selected by the CART procedure as the most discriminating, are then applied in two other classification methods: K-Nearest Neighbor (K-NN) and Regularised Discriminant Analysis (RDA). The discriminant power of the variables is also highlighted by these classification methods, the performance being even slightly better than that obtained by CART (see Table II).

All the methods assign more than 44 chemicals correctly, CART makes 10 errors in prediction and K-NN and RDA 9. A comparison of the assignments in the classification predictions of all the applied methods highlights the pesticides classified wrongly by all the methodologies: alachlor (1) from Class 4 in Class 2, propoxur (13) from Class 2 in Class 4, pirimicarb (16) from Class 1 in Class 3, azinphos methyl (28) from Class 4 in Class 3, fenuron (43) from Class 1 in Class 4, while malathion (37), at the centre of the Fig. 1 graph, from Class 4 is wrongly assigned to different classes by all the methods (clearly it is the most problematic chemical to be classified by molecular descriptors). These errors can be considered serious only for pirimicarb (16), azinphos methyl (28) and fenuron (43), while the other chemicals lie at the border of the defined classes, and it is important to note that these borderlines among the classes (derived from previous PC and Cluster Analyses) can be considered as rather arbitrary.

CONCLUSIONS

The experimental data of the physico-chemical properties, considered relevant to the environmental partitioning of 54 pesticides of different chemical classes, are combined in Principal Component Analysis and Hierarchical Cluster Analysis, obtaining a ranking of these pesticides into four *a priori* classes according to environmental behaviour (sorbed, soluble, volatile and non-volatile/medium classes). Classification by different classification methods (CART, K-NN and RDA) using only theoretical molecular structure descriptors allows a fast screening of both existing and new pesticides belonging to the studied chemical domain; such classification is done with regard to behaviour in the environment, and is based simply on the knowledge of molecular structure. In practice, with only three descriptors of molecular structure, the proposed classification models lead to an immediate assessment of pesticide tendency to partitioning into different environmental compartments.

The classification models presented here can be useful as a first step in the preliminary screening and rapid pre-determination of the environmental distribution of pesticides, allowing an easy *a priori* prediction of the environmental compartment in which existing and new pesticides can be found; the starting point needed is only the knowledge of a few molecular structure descriptors, without any *a priori* knowledge of physico-chemical properties or analytical monitoring.

Acknowledgements

We wish to thank Prof. R. Todeschini and Dr. V. Consonni for helpful discussions and the software for molecular descriptor calculation and modelling, and Dr. A. Di Guardo for the scientific reviewing of the article. This work was financially supported by BEAM Program of the Commission of the European Communities (EVK1-CT-1999-00012).

References

[1] A. Sabljic and W.T. Piver, *Environ. Toxicol. Chem.*, **11**, 961–972 (1992).
[2] R. Altenburger, W. Boedeker, M. Faust and L.H. Grimme, *Sci. Tot. Environ. Suppl.*, 1633–1654 (1993).
[3] J.V. Tarazona, A. Fresno, S. Aycard, C. Ramos, M.M. Vega and G. Carbonell, *Sci. Tot. Environ.*, **247**, 151–164 (2000).
[4] J. Kanazawa, *Environ. Toxicol. Chem.*, **8**, 477–484 (1989).
[5] F.B. Taub and L.A. Burns, In: H. Frehse (Ed.), *Pesticide Chemistry*, pp. 423–432. VHC, Weinheim (1991).
[6] D.I. Gustafson, *Environ. Toxicol. Chem.*, **8**, 339–357 (1989).
[7] E. Halfon, S. Galassi, R. Bruggemann and A. Provini, *Chemosphere*, **33**, 1543–1562 (1996).
[8] P. Gramatica, M. Corradi and V. Consonni, *Chemosphere*, **41**, 763–777 (2000).
[9] P. Gramatica and A. Di Guardo, *Chemosphere*, **47**, 947–956 (2002).
[10] A. Sabljic, H. Güsten, H. Verhaar and J. Hermens, *Chemosphere*, **31**, 4489–4514 (1995).
[11] C.R. Worthing, *The Pesticide Manual*, 9th Edn., 1141 pp. The British Crop Protection Council (1991).
[12] L.R. Suntio, W.Y. Shiu, D. MacKay, J.N. Seiber and D. Glotfelty, *Rev. Environ. Contam. Toxicol.*, **103**, 1–59 (1988).
[13] R. Todeschini and V. Consonni, *DRAGON – Software for the Calculation of Molecular Descriptors*, version 1.0 for Windows, (2000). Free download available at: http://www.disat.unimib.it/chm.
[14] HYPERCHE*M*, rel. 4 for Windows, Autodesk, Inc., Sausalito, CA (USA) (1995).
[15] R. Todeschini and P. Gramatica, *SAR QSAR Environ. Res.*, **7**, 89–115 (1997).
[16] R. Todeschini and V. Consonni, *Handbook of Molecular Descriptors*, 667 pp. Wiley-VCH, Weinheim (2000).
[17] SCAN – Software for Chemometric Analysis rel. 1.1 for Windows, Minitab (USA) (1995).
[18] J.E. Frank and J.H. Friedman, *J. Chemom.*, **3**, 463–475 (1989).
[19] C. Hansch and A. Leo, *Exploring QSAR*, 557 pp. ACS Press, Washington (1995).