Bulletin of Insectology 56 (1): 103-109, 2003
ISSN 1721-8861

Modeling the acute toxicity of pesticides to *Apis mellifera*

James DEVILLERS[1], Minh Hà PHAM-DELÈGUE[2], Axel DECOURTYE[2], Hélène BUDZINSKI[3], Sophie CLUZEAU[4], Gilbert MAURIN[4]
[1]*CTIS, Rillieux La Pape, France*
[2]*Laboratoire de Neurobiologie Comparée des Invertébrés, INRA, Bures-sur-Yvette, France*
[3]*LPTC, UMR 5472 CNRS, Université de Bordeaux I, Talence, France*
[4]*ACTA, Association de Coordination Technique Agricole, PARIS, France*

Abstract

A Quantitative Structure-Activity Relationship (QSAR) model was proposed for estimating the acute toxicity of pesticides to *Apis mellifera*. Chemicals were described from the autocorrelation method yielding the computation of descriptors encoding different physicochemical properties. A three-layer feedforward neural network trained by the back-propagation algorithm was used as statistical tool for deriving the model. The root mean square residual values for the training set (89 chemicals) and external testing set (11 chemicals) were 0.430 and 0.386, respectively. The usefulness of this type of modeling approach was discussed.

Key words: acute toxicity, *Apis mellifera*, artificial neural network, autocorrelation method, pesticides, QSAR model.

Introduction

Investigations into the development and use of Quantitative Structure-Activity Relationship (QSAR) models (Karcher and Devillers, 1990; Devillers, 1998) to rapidly predict the ecotoxicity of xenobiotics from their molecular structure and/or physicochemical properties have increased dramatically over the past decades in order to save time and money in the design of safer chemicals (Devillers, 2003). These QSAR models are now integrated in most of the methodological frameworks designed for estimating the environmental hazard and risk of organic chemicals (ECETOC, 1998, Walker and Carlsen, 2002). Surprisingly, while numerous QSAR models are available for estimating the toxicity of chemicals to various species of Arthropods, the number of models derived on the honey bee is very scarce. Indeed, only Vighi *et al.* (1991) proposed a structure-toxicity model for estimating the acute toxicity of pesticides to *Apis mellifera*. Even if, from a historical point of view, their model is interesting, its usefulness is limited because it was only designed for simulating the toxicity of the organophosphorus pesticides. Furthermore, its statistical validity is highly questionable because the regression equation was derived from only 15 pesticides and included six molecular descriptors.

Consequently, the aim of our study was to propose a more powerful model based on an artificial neural network (Devillers, 1996a) to simulate the acute toxicity of all the families of pesticides to the honey bee.

Materials and methods

Toxicity data

The acute toxicity data (LC_{50} or LD_{50}) used in this study (table 1) were retrieved from literature (Atkins *et al.*, 1981; Tomlin, 1994; EXTONET, 2002). The laboratory LD_{50} values published by Atkins *et al.* (1981) were selected in priority because they were obtained under well-defined and controlled conditions. The two other sources of information often provided data without explanation on the experimental conditions in which they were raised. Consequently, the Atkins data were exclusively used to constitute the training set of the model while those from Tomlin (1994) and EXTONET (2002) were used to complete the training set and to form the external testing set. In that case, the allocation of the pesticides between both types of sets was randomly performed. The QSAR model was derived from a training set of 89 chemicals and its simulation performances were then tested from an external testing set of 11 pesticides. However, it is worth mentioning that, during the design of the model, four pesticides were randomly selected from this testing set to constitute a cross-validation set allowing to monitor the artificial neural network and avoid overtraining (Devillers, 1996a).

In the literature, the acute toxicity data were reported in μg/bee or mg/bee. For modeling purposes, they were first converted in μmol/bee and then converted into their negative logarithms.

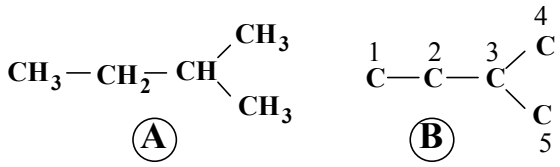
Molecular descriptors

In order to encode all the structural characteristics of the pesticides, the autocorrelation method (Moreau and Broto, 1980a, b) was used. Briefly, the autocorrelation descriptors are simple 2-D molecular descriptors designed from the hydrogen-suppressed graphs of the molecules. Autocorrelation vectors (AVs) can be derived for all physicochemical properties which can be calculated from atomic contributions. They consist of autocorrelation components (ACs) corresponding to the different interatomic distances which can be computed within the studied molecule. The calculation procedure of an AV for isopentane is illustrated in figure 1.

After their calculation, the AVs have to be truncated to obtain strings of descriptors of same dimensionality with a reduced number of null values. This procedure is very important when the model is designed from sets of structurally diverse compounds, as it is the case here. In

the classical algorithm proposed by Moreau and Broto (1980a, b), the different ACs for a property are obtained by summations of products (see step 4 in figure 1). This is annoying when negative atomic contributions have to be used such as for encoding lipophilicity of some functional groups. Indeed, in that case the physicochemical meaning of the ACs is not straightforward. To overcome this problem, a slightly different algorithm was employed (Devillers *et al.*, 1992). With this modified algorithm, the first component of an AV is simply obtained by the sum of the positive and negative contributions attributed to the atoms and functional groups constituting a molecule. Moreover, the calculation procedure takes into account the signs of the different contributions in order to increase the physicochemical significance of the descriptors. Additional information on the algorithms and the potentialities of the autocorrelation method in environmental QSARs can be found in a recent publication (Devillers, 1999).

Step 1. Definition of the molecular graph (B) from the formula (A).



Step 2. Calculation of the shortest interatomic distances d(i,j) (expressed as number of bonds) between each couple of atoms i and j.

d(1, 1) = 0	d(1, 2) = 1	d(1, 3) = 2	d(1, 4) = 3	d(1, 5) = 3
	d(2, 2) = 0	d(2, 3) = 1	d(2, 4) = 2	d(2, 5) = 2
		d(3, 3) = 0	d(3, 4) = 1	d(3, 5) = 1
			d(4, 4) = 0	d(4, 5) = 2
				d(5, 5) = 0

Step 3. Selection of an atomic property G and calculation of the autocorrelation vector AV = (C_0, C_1, C_2, ..., C_n) corresponding to a given property and for which the components C of order 0, 1, 2, ..., n are calculated by means of the following equation:

$$C_n = \Sigma\ g(i) \times g(j)$$

where g(i) and g(j) are the contributions attributed to atoms i and j.

If connectivity (i.e.; number of neighbors of each atom) is chosen as atomic property then we obtain:

$$g(1) = 1,\ g(2) = 2,\ g(3) = 3,\ g(4) = 1,\ \text{and}\ g(5) = 1$$

Step 4. Calculation of the components Ti of the vector T encoding the connectivity.

$T_0 = (1 \times 1) + (2 \times 2) + (3 \times 3) + (1 \times 1) + (1 \times 1) = 16$
$T_1 = (1 \times 2) + (2 \times 3) + (3 \times 1) + (3 \times 1) = 14$
$T_2 = (1 \times 3) + (2 \times 1) + (2 \times 1) + (1 \times 1) = 8$
$T_3 = (1 \times 1) + (1 \times 1) = 2$

For isopentane, higher order components equal zero.

$$T = (16, 14, 8, 2)$$

Figure 1. Principle of the autocorrelation method.

The 100 pesticides listed in table 1 were described by means of four different AVs.

First, from the fragmental constants of Rekker and Mannhold (1992), for each molecule, an AV representing lipophilicity (H) was derived. Second, an AV encoding molar refractivity (MR) was designed from the fragmental constants of Hansch and Leo (1979) or directly from the Lorentz-Lorenz equation (Eq. 1).

$$MR = \frac{n^2 - 1}{n^2 + 2} \times \frac{MW}{d} \quad (1)$$

In Eq. 1, n is the refraction index, d the density, and MW the molecular weight of the molecule.

Last, AVs encoding the H-bonding acceptor ability (HBA) and H-bonding donor ability (HBD) of the molecules were also calculated from Boolean contributions (i.e., 0/1).

These AVs were calculated by means of AUTOCOR™ 2.4. For the AVs H, MR, and HBA, a truncature was performed in order to obtain five ACs (distances 0 to 4). For the HBD AV, only the first component was selected.

Statistical tool

A multi-layer feedforward neural network trained by the back-propagation algorithm (BNN) (Rumelhart *et al.*, 1986, Devillers, 1996a) was used to find nonlinear relationships between the autocorrelation descriptors and the toxicity data. Topologically a BNN presents three types of layers (figure 2):

- one input layer (with a number of neurons corresponding to the number of molecular descriptors),
- one (or more) hidden layer(s) with adjustable numbers of neurons,
- one output layer with a number of neurons depending on the modeled activity or property. This layer generates the calculated outputs (i.e., acute toxicity data).

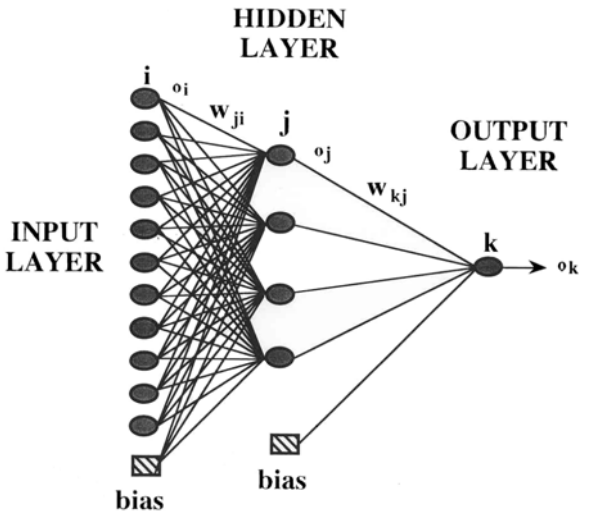


Figure 2. A three-layer feedforward neural network trained by the back-propagation algorithm (BNN).

The neurons of each layer are connected in the forward direction (i.e., input to output). Each of the input and hidden layers can have an additional unit called bias connected as shown in figure 2. Biases allow a more rapid convergence of the training process (Wasserman, 1989). Before starting the training (learning) process, all the weights associated with the connections between the neurons within the network must be initialized to small random numbers (e.g., [-0.3, 0.3]). This ensures that the network is not saturated by large values of the weights, and prevents some training pathologies (Wasserman, 1989). During the training phase, each input pattern of the training set is presented to the network which generates a calculated output. At this stage, the network has performed the feedforward step. An error (Eq. (2)) is computed from the calculated (o_{pk}) and target (t_{pk}) outputs for a pattern p.

$$E_p = \tfrac{1}{2} \sum_k (t_{pk} - o_{pk})^2 \quad (2)$$

For all patterns, we obtain:

$$E = \tfrac{1}{2} \sum_p \sum_k (t_{pk} - o_{pk})^2 \quad (3)$$

Weights are adjusted by backpropagating the error from the output to the input layer. This is performed after presentation of each training pattern (on-line or single pattern training). However, note that calculation times can be reduced by adjusting the weights once all patterns have been presented (batch or epoch training) (Eberhart and Dobbins, 1990). After presentation of a pattern p, the adjustment of the weights located between the output layer k and the hidden layer j is performed by means of the following equation:

$$\Delta_p w_{kj} = \eta \, \delta_{pk} o_{pj} \quad (4)$$

where η is the learning rate and $\delta_{pk} = (t_{pk} - o_{pk})\, o_{pk}\, (1 - o_{pk})$ for a sigmoid function. Note that a large η value corresponds to a rapid learning but might also result in oscillations. If η is set too low, the convergence is difficult and the risk of falling into and remaining in local minima is high (Pao, 1989, p. 128; de Saint Laumer *et al.*, 1991).

The adjustment of the weights located between the hidden layer j and the input layer i are calculated in a similar manner. Rumelhart and coworkers (1986) have modified Eq. (4) by including a momentum term (α) which prevents oscillations (Pao, 1989, p. 128; Wasserman, 1989, p. 54). The adjustment of the weights between the (n)th and the (n + 1)th steps then becomes:

$$\Delta_p w_{ji}(n+1) = \eta \, \delta_{pj} o_{pi} + \alpha \, \Delta_p w_{ji}(n) \quad (5)$$

Preprocessing the data often plays a key role in the modeling performances of a BNN (Devillers, 1996a). Consequently, a classical min/max transformation was used. All the BNN calculations were performed with the STATQSAR™ package.

Results and discussion

Different modeling exercises were performed in order to determine the number of autocorrelation descriptors allowing to correctly describe the pesticides and the optimal architecture of the BNN yielding good simulation results with the training and external testing sets.

The AVs of five components encoding lipophilicity (H), molar refractivity (MR), and the H-bonding acceptor ability (HBA) of the molecules were first introduced separately as inputs in the neural network. They were not able to produce acceptable root mean square residual (RMSR) values (Hair *et al.*, 1992) for both sets. Addition of HBD_0 did not increase the quality of the modeling results.

Conversely, the use of the 15 autocorrelation descriptors (i.e., H_0 to H_4, MR_0 to MR_4, HBA_0 to HBA_4) as inputs in the BNN allowed to obtain satisfying calculated toxicity values for both sets with three neurons on the hidden layer and with about 2000 cycles. These results were interesting but attempts were made to reduce the number of connections in the BNNs. Indeed, it is always necessary to have a BNN model with the smallest number of adjustable parameters and the lowest RMSR errors for the training set and more important, for the external testing set (Devillers, 1996a).

To reduce the number of ACs, the feature selection option of STATISTICA™ based on a genetic algorithm (GA) was used. Indeed, GAs, which are rooted in Darwin's theory of natural selection and evolution, provide an alternative to traditional optimization methods by using powerful search techniques to rapidly locate optimal solutions in complex landscapes (Devillers, 1996b). Different top-ranked solutions of ACs were obtained by modifying the basic parameters of the GA. These solutions were all introduced in the BNN to test their ability to derive an acceptable model.

Because a BNN with H_1 to H_4, MR_0 to MR_4, and HBA_2 to HBA_4 as inputs and 3 neurons on the hidden layer regularly yielded interesting results, this architecture was selected and refined to find an optimal configuration. This refinement was focused on the number of cycles, the learning rate (η), and momentum (α) of the BNN. The acceptable number of cycles was determined by means of a cross-validation set of 4 pesticides randomly selected from the external testing set. During this phase, the BNN was monitored by the cross-validation set and its simulation performances were only tested on the remaining 7 chemicals of the external testing set. After 140 runs with different numbers of cycles and by randomly changing the composition of the cross-validation and external testing sets, the optimal number of runs was set to about 2500 cycles. Thus, this number of cycles was used to optimize the learning rate (η), and momentum (α) from the training set of 79 pesticides and the original external testing set of 11 pesticides. At this stage, the problem of overtraining was solved and it was necessary to secure the design of a BNN model with a predictive power as high as possible.

From 46 runs, the best modeling results were obtained from a 12/3/1 BNN with $\eta = 0.5$, $\alpha = 0.9$, and 2492 cycles. The RMSR values for the training and testing sets were 0.430 and 0.386, respectively. The calculated acute toxicity data are listed in table 1 for both the training set (pesticides no. 1 to 89) and the external testing set (pesticides no. 90 to 100). The distribution of the residual values is given in table 2. Table 2 clearly shows the high simulation performances of the selected BNN model. The number of chemicals not correctly predicted by the model is very limited. Thus, only four pesticides have their residual value greater than 0.9 (in absolute value). In addition, it is important to note that these chemicals only belong to the training set (table 2). Thus, EPN (chemical no. 15), profenofos (chemical no. 53), fenazaquin (chemical no. 79), and metconazole (chemical no. 84) present residual values of 1.10, 0.91, -0.93, and –0.98, respectively. It is impossible to find an explanation for these bad predictions because for each of these chemicals, only one acute toxicity data was available and hence, an experimental error cannot be excluded. Nevertheless, inspection of the residuals clearly shows that the selected model presents a good predictive power.

Table 1. Observed and calculated acute toxicity data (log 1/C) of pesticides tested on *Apis mellifera*.

No.*	Pesticide	Obs.	Cal.
1	TEPP	5.16	5.16
2	Bioethanomethrin	4.02	3.45
3	Resmethrin	3.74	3.29
4	Pay-off (Flucythrinate)	3.76	3.70
5	Deltamethrin	3.88	3.69
6	Chlorpyrifos	3.50	2.68
7	Parathion-methyl	3.38	3.55
8	Dieldrin	3.46	3.12
9	Carbofuran	3.17	2.80
10	Permethrin	3.39	3.42
11	Parathion	3.22	2.87
12	Fenitrothion	3.20	2.96
13	Dimethoate	3.08	2.86
14	Methidathion	3.11	2.70

15	EPN	3.13	2.03
16	Etrimfos	3.04	3.07
17	Aldicarb	2.84	3.15
18	Mexacarbate	2.87	2.51
19	Dicrotophos	2.89	3.15
20	Mevinphos	2.87	2.80
21	Fenthion	2.94	2.50
22	Fensulfothion	2.96	2.98
23	Aldrin	3.02	3.03
24	Monocrotophos	2.80	2.78
25	Diazinon	2.91	2.70
26	Methiocarb	2.78	2.61
27	Fenvalerate	3.01	3.75
28	Famphur	2.90	3.16
29	Azinphos-methyl	2.87	2.47
30	Bendiocarb	2.72	2.58
31	Naled	2.89	2.84
32	Dichlorvos	2.64	2.69
33	Heptachlor	2.85	2.25
34	Isofenphos	2.76	2.71
35	Carbosulfan	2.75	3.43
36	Malathion	2.66	2.74
37	Azinphos-ethyl	2.56	2.61
38	Aminocarb	2.27	2.48
39	Phosmet	2.45	2.27
40	Acephate	2.18	2.76
41	Methomyl	2.10	2.06
42	Propoxur	2.19	2.08
43	Methamidophos	2.01	2.02
44	Stirofos (Tetrachlorvinphos)	2.42	3.22
45	Fenamiphos	2.33	1.89
46	Phosphamidon	2.32	2.27
47	Carbaryl	2.12	1.89
48	Pyrazophos	2.30	2.06
49	Temephos	2.52	2.29
50	Trichloronate	2.22	1.86
51	Crotoxyphos	2.13	2.16
52	Oxydemeton-methyl	1.94	1.87
53	Profenofos	2.03	1.12
54	Terbufos	1.85	1.53
55	Ethoprophos	1.64	1.56
56	Ronnel	1.76	2.42
57	Disulfoton	1.65	1.73
58	DDT	1.76	2.35
59	Ethiofencarb	1.52	1.54
60	Thiodicarb	1.70	2.12
61	Sulprofos	1.65	2.05
62	Fonofos	1.45	0.72
63	Chlordane	1.67	2.15
64	Phosalone	1.61	2.23
65	Phorate	1.40	1.82
66	Oxamyl	1.33	1.77
67	Carbophenothion	1.42	2.13
68	AC 303,630	3.31	2.87
69	Alanycarb	2.70	2.37
70	Chlorpyrifos-methyl	2.93	2.97
71	Bensultap	1.22	1.13
72	Azamethiphos	3.51	3.86
73	EPTC	1.24	1.23
74	Napropamide	0.35	0.49
75	Dicloran	0.06	0.40

76	Diethofencarb	1.13	1.01
77	Dithiopyr	0.70	0.47
78	Orbencarb	0.40	0.79
79	Fenazaquin	1.57	2.50
80	Quinalphos	3.63	3.62
81	Pyridaben	2.82	2.80
82	Mephosfolan	1.89	1.94
83	Imazalil	0.87	1.59
84	Metconazole	0.52	1.50
85	Pyridaphenthion	3.63	3.25
86	Pyrifenox	0.70	0.90
87	Quizalofop	0.84	1.09
88	Tebutam	0.37	0.40
89	Thiometon	2.64	2.15
90	Cypermethrin	4.08	3.66
91	Diafenthiuron	2.26	2.79
92	Tralomethrin	3.74	3.74
93	Propargite	1.37	1.82
94	Diclomezine	1.41	1.35
95	Phenthoate	3.43	2.91
96	Piperophos	1.07	1.25
97	Silafluofen	2.91	2.76
98	Propachlor	-0.17	0.39
99	Tralkoxydim	0.79	0.65
100	Propisochlor	0.45	1.01

*Pesticides no. 1 to 89 have been used as training set and pesticides no. 90 to 100 as external testing set for estimating the simulation performances of the model.

Table 2. Distribution of residuals (absolute values), differences between the experimental and the calculated toxicity data computed from the model.

Range	Training set	Testing set
<0.3	48	5
0.3 to 0.6*	27	6
0.6 to 0.9*	10	0
0.9 to 1.2*	4	0
≥1.2	0	0

*Excluded value

Other BNN models were designed from different training and testing sets of 89 and 11 pesticides, respectively. However, in all cases, as previously indicated in the experimental section, the toxicity data coming from Atkins *et al.* (1981), due to their coherency, were only included in the training set. No significant differences were found in the simulation results.

For comparison purposes, attempts have also been made to derive a model from the partial least squares (PLS) regression method (Devillers *et al*., 2002), which is widely used in QSAR studies due to its ability to work well with large sets of descriptors (Devillers and Doré, 2002). Different procedures proposed by STATISTICA™ were experienced. Unfortunately, no satisfying results were obtained. Indeed, in all cases, very large outliers were obtained for a huge number of pesticides.

These results are not surprising. Indeed, the relationships between the structure of the pesticides and their acute toxicity to bees are not straightforward. Consequently, only an artificial neural network appears suited to derive complex relationships between the autocorrelation descriptors and the toxicity data. In addition, it is worth mentioning that, even if standardized protocols are used to test the acute toxicity of pesticides to honey bees, the obtained data generally suffer from a high degree of variability due to the experimental conditions, the endpoint recorded, the biological material, and so on. Because, it is well known that some degree of fuzziness in the data increases the modeling performances of a BNN (Devillers, 1996a), this also explains the good simulation results obtained with this nonlinear statistical tool.

As long as pesticides are used to protect crops, beekeepers complain about their potential adverse effects on bees (Devillers and Pham-Delègue, 2002). Useful testing guidelines have been developed for estimating the acute toxicity of pesticides to bees prior their commercial use. Even if they provide invaluable information, they are costly, time expensive, and require trained people. The proposed QSAR model does not suffer from these limitations. In addition, it can be used to es-

timate the toxicity of a pesticide not yet synthesized. This could be very useful when attempts are made to design safer pesticides. Consequently, the proposed QSAR model must be seen as a complementary tool to the existing laboratory tests to better estimate the adverse effects of pesticides to honey bees.

Last, it is obvious that during their foraging activity, the honey bees are not only contaminated by pesticides but also by all kind of xenobiotics for which no experimental toxicity data exist. It is expected that this QSAR model should be easily refined to also simulate the acute toxicity of these pollutants.

Acknowledgement

This study was granted by the French Ministry of the Environment (MATE-01133).

References

ATKINS E. L., KELLUM D., ATKINS K. W., 1981.- *Reducing pesticide hazards to honey bees: mortality prediction techniques and integrated management strategies*.- Division of Agricultural Sciences, University of California, Leaflet 2883.

AUTOCOR™ (Version 2.40). CTIS, 3 Chemin de la Gravière, 69140 Rillieux La Pape, France.

DE SAINT LAUMER J. Y., CHASTRETTE M., DEVILLERS J., 1991.- Multilayer neural networks applied to structure-activity relationships.- In: *Applied Multivariate Analysis in SAR and Environmental Studies* (DEVILLERS J., KARCHER W., Eds), Kluwer Academic Publishers, Dordrecht, The Netherlands, 479-521.

DEVILLERS J., 1996a.- Strengths and weaknesses of the back-propagation neural network in QSAR and QSPR studies.- In: *Neural Networks in QSAR and Drug Design* (DEVILLERS J., Ed.), Academic Press, London, 1-46.

DEVILLERS J., 1996b.- *Genetic Algorithms in Molecular Modeling*.- Academic Press, London.

DEVILLERS J., 1998.- Environmental chemistry: QSAR.- In: *The Encyclopedia of Computational Chemistry* (SCHLEYER P. V. P., ALLINGER N. L., CLARK T., GASTEIGER J., KOLLMAN P. A., SCHAEFER H. F., SCHREINER P. R., Eds), John Wiley & Sons, Chichester, 2: 930-941.

DEVILLERS J., 1999.- Autocorrelation descriptors for modeling (eco)toxicological endpoints.- In: *Topological Indices and Related Descriptors in QSAR and QSPR* (DEVILLERS J., BALABAN T. A., Eds).- Gordon and Breach, The Netherlands, 595-612.

DEVILLERS J., 2003.- A decade of research in environmental QSAR.- *SAR and QSAR in Environmental Research*, 14 (1): 1-6.

DEVILLERS J., CHEZEAU A., THYBAUD E., 2002.- PLS-QSAR of the adult and developmental toxicity of chemicals to *Hydra attenuata*.- *SAR and QSAR in Environmental Research*, 13: 705-712.

DEVILLERS J., DOMINE D., CHASTRETTE M., 1992.- A new method of computing the octanol/water partition coefficient.- In: *Proceedings of QSAR92*, July 19-23, 1992, Duluth, USA, 12.

DEVILLERS J., DORE J.C. 2002.- e-Statistics for deriving QSAR models.- *SAR and QSAR in Environmental Research*, 13: 409-416.

DEVILLERS J., PHAM-DELÈGUE M. H. 2002.- *Honey Bees: Estimating the Environmental Impact of Chemicals*.- Taylor & Francis, London.

EBERHART R. C., DOBBINS R. W., 1990.- Implementations.- In: *Neural Network PC Tools. A Practical Guide* (EBERHART R. C., DOBBINS R. W., Eds), Academic Press, San Diego, 35-58.

ECETOC, 1998.- *QSARs in the Assessment of the Environmental Fate and Effects of Chemicals*.- European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium, Technical Report no. 74.

EXTONET, 2002.- http://ace.orst.edu/info/extonet/

HAIR J. F., ANDERSON R. E., TATHAM R. L., BLACK W. C., 1992.- *Mutivariate Data Analysis with Readings*, Third Edition.- Macmillan Publishing Company, New York.

HANSCH C., LEO A., 1979.- *Substituent Constants for Correlation Analysis in Chemistry and Biology*.- J. Wiley & Sons, New York.

KARCHER W., DEVILLERS J., 1990.- *Practical Applications of Quantitative Structure-Activity Relationships (QSAR) in Environmental Chemistry and Toxicology*.- Kluwer Academic Publishers, Dordrecht, The Netherlands.

MOREAU G., BROTO P., 1980a.- The autocorrelation of a topological structure: A new molecular descriptor.- *Nouveau Journal de Chimie*, 4: 359-360.

MOREAU G., BROTO P., 1980b.- Autocorrelation of molecular structures, application to SAR studies.- *Nouveau Journal de Chimie*, 4: 757-764.

PAO Y. H., 1989.- *Adaptive Pattern Recognition and Neural Networks.* Addison-Wesley Publishing Company, Reading.

REKKER R. F., MANNHOLD R. 1992.- *Calculation of Drug Lipophilicity. The Hydrophobic Fragmental Constant Approach*.- VCH, Weinheim.

RUMELHART D. E., HINTON G. E., WILLIAMS R. J., 1986.- Learning representations by back-propagating errors.- *Nature* 323: 533-536.

STATISTICA™, Version 6, StatSoft, 31 Cours des Juilliottes, 94700 Maisons-Alfort, France.

STATQSAR™ Package. CTIS, 3 Chemin de la Gravière, 69140 Rillieux La Pape, France.

TOMLIN C., 1994.- *The Pesticide Manual*, Tenth Edition.- Crop Protection Publications.

VIGHI M., GARLANDA M.M., CALAMARI D., 1991.- QSARs for toxicity of organophosphorus pesticides to *Daphnia* and honeybee.- In*: QSAR in Environmental Toxicology-IV* (HERMENS M. J. L., OPPERHUIZEN A. , Eds), Elsevier, Amsterdam, 605-622.

WALKER J. D., CARLSEN L., 2002.- QSARs for indentifying and prioritizing substances with persistence and bioconcentration potential.- *SAR and QSAR in Environmental Research*, 13 (7-8): 713-725.

WASSERMAN P. D., 1989.- *Neural Computing: Theory and Practice*.- Van Nostrand Reinhold, New York.

Corresponding author: James DEVILLERS, CTIS, 3 Chemin de la Gravière, 69140 Rillieux La Pape, France.
E-mail: j.devillers@ctis.fr