Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 14 (2006) 2779–2788
Bioorganic & Medicinal Chemistry

QSAR models for Daphnia toxicity of pesticides based on combinations of topological parameters of molecular structures

Andrey A. Toropov* and Emilio Benfenati

Laboratory of Environmental Chemistry and Toxicology, Istituto di Ricerche Farmacologiche 'Mario Negri,' Via Eritrea 62, 20157 Milan, Italy

Received 14 November 2005; revised 29 November 2005; accepted 29 November 2005
Available online 27 December 2005

Abstract—A topological parameter is defined as an integer value of a given local or global invariant of a molecular graph. We examined three types of local graph invariants, the vertex degrees (^{0}EC), the extended connectivity of first order (^{1}EC), and the numbers of paths of length two (P2), as elementary invariants for construction of quantitative structure–activity relationships (QSAR). We also examined combined invariants, obtained by multiplying one of these three elementary types with another (i.e., [^{0}EC · ^{1}EC], [^{0}EC · P2], and [^{1}EC · P2]), as graph invariants. Finally, global invariants were used in the QSAR analyses, codifying the presence and nature of cycles in the molecular structures under consideration. We used the correlation weights of these invariants to obtain optimal descriptors. These descriptors have been used in one-variable models to predict toxicity toward *Daphnia magna* for a set of pesticides. Statistical characteristics of the best model, based on the correlation weight of local topological parameters (the [^{0}EC · P2]) together with the global topological parameters, are the following: $n = 220$, $r^2 = 0.7822$, $s = 0.849$, $F = 783$ (training set); $n = 42$, $r^2 = 0.7388$, $s = 0.941$, $F = 113$ (test set). The role of these topological parameters is discussed.

1. Introduction

Quantitative structure–property/activity relationships (QSPR/QSAR) are useful for predicting many physical, chemical, and biological characteristics of substances.[1–14] The reasons for modeling biochemical endpoints in general, and toxicity in particular, are well known. First, the experimental measurement of biochemical parameters involves considerable amount of time and money, whereas prediction using a computer is, as a rule, faster and cheaper. Second, a good or at least satisfactory result using a given molecular descriptor in modeling the endpoint is important in view of developments in theoretical chemistry and biochemistry, allowing hypotheses about the mechanisms of biological actions.

In a first approximation, toxicity can be examined as a mathematical function of different physical and chemical phenomena.[15–20] In other words, complex biochemical phenomena affecting toxicity can be described in terms of mathematics and chemistry. In the present study we examined the representation of pesticide molecules as the hydrogen-filled molecular graphs (HFG). The HFG contain information on atom composition (physical reality) taking account of the presence of double, aromatic, and triple covalent bonds. The saturation/unsaturation of chemical bonds in HFG can be detected from differences in vertex degree. For instance, the HFG vertex degree of carbon in ethane, ethene, and ethyne equals four, three, and two, respectively.

Toxicity is affected by very complex conditions, such as the hydrogen-bond interactions, possibility of several stereo-chemical configurations, possibility of multistage chemical transformations of membranes, and mediums of living systems.

The socalled optimization of correlation weights of local and global graph invariants (OCWLGI) has been suggested as a tool for QSPR/QSAR analyses.[12–14,18–21] Results with the OCWLGI depend on the choice of graph invariants.

The present study assessed the OCWLGI technique for predicting toxicity, based on correlation weights of special invariants, which are the results of multiplication of elementary graph invariants.

Keywords: QSAR; Toxicity; Pesticide.
* Corresponding author. E-mail: aatoropov@yahoo.com

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.11.060

2. Method

2.1. Data set

We used 262 pesticides, randomly split into a training ($n = 220$) and a test ($n = 42$) set. QSAR models were developed using only chemicals in the training set. Results were validated using the test set. Numerical data on the acute toxic effects towards *Daphnia magna* were used.

2.2. Toxicity data

As endpoint we used the decimal logarithm $\log(1/C)$, where C is the concentration, in mmol/L, expressed as LC_{50}-96 h, which is the dose that kills 50% of daphnias in 96 h. The data set contains heterogeneous pesticides which is a complex case. The toxicity data from the EPA-OPP data set were kindly provided by Brian Montague (Washington, DC). Toxicity values were selected according to the criteria described in Ref. 22.

2.3. Topological parameters and QSAR models

OCWLGI in the molecular graph is basically an attempt to estimate how local and global topological parameters of the molecular structure contribute to the toxicological phenomenon under consideration. The local topological parameter can be defined as an integer that characterizes a given position of the vertex in a graph (or position of an atom in a molecule), for instance, the number of neighbors connected to a given kth vertex. This invariant is the valence or degree of a given vertex (^0EC_k).[1–14] The sum of the numerical valences over all neighbors of the kth vertex is the extended connectivity of first order (^1EC_k), described in Ref. 3. Finally, the number of paths of length 2, which starts from a given kth vertex, can also be examined as local topological parameter of a given molecular structure.[5,6]

Table 1. Examples of A, B, and C, and their interpretations

Digit of A, B, or C	Global topological parameters in molecular structure
0	Absence of cycles
1	Presence of cycles
3	Presence of heteroatoms in the cycle
5	Presence of vertex with valence equal to 4
7	Presence of heteroatoms together with vertices of valence equal to 4

The recently described nearest neighboring codes (NNC), used in the OCWLGI models, have shown quite good potential as local invariants.[12,13,19,20] We defined NNC by the following integers, according to the topological parameter of a given vertex Eq. 1:

$$^{(O,N)}NNC_k = \begin{cases} 1000, \text{ if at least one vertex connected to vertex } k \text{ is an oxygen atom;} \\ 2000, \text{ if at least one vertex connected to vertex } k \text{ is a nitrogen atom;} \\ 3000, \text{ if vertices connected to vertex } k \text{ contain both oxygen and nitrogen atoms;} \\ 0, \text{ in all other cases.} \end{cases} \quad (1)$$

Thus, the final version of the local invariants in a generalized form can be represented as Eq. 2

$$V_k = {}^{(O,N)}NNC_k + LI_k, \quad (2)$$

where LI_k is ^0EC_k, ^1EC_k, $P2_k$, $[^0EC \cdot {}^1EC]$, $[^0EC \cdot P2]$, and $[^1EC \cdot P2]$.[2,5,12,13,19,20]

Table 2. Statistical characteristics of OCWLGI models of toxicity on three probes

Probe No.	Flexible descriptor	Number of optimized parameters	Training set, $n = 220$			Test set, $n = 42$		
			r^2	s	F	r^2	s	F
1	$^0X_{CW}$ (a_k, ^0EC_k, CC)	49	0.6728	1.041	448	0.2881	1.548	16
2			0.6750	1.038	453	0.2801	1.555	16
3			0.6764	1.035	456	0.2870	1.547	16
1	$^0X_{CW}$ (a_k, ^1EC_k, CC)	82	0.7518	0.907	660	0.6416	1.068	72
2			0.7536	0.903	667	0.6279	1.092	68
3			0.7538	0.903	667	0.6306	1.082	68
1	$^0X_{CW}$ (a_k, $P2_k$, CC)	74	0.7140	0.973	544	0.5121	1.240	42
2			0.7059	0.987	523	0.5488	1.194	49
3			0.7106	0.979	535	0.5376	1.207	47
1	$^0X_{CW}$ (a_k, $[^0EC_k \cdot {}^1EC_k]$, CC)	111	0.8023	0.809	885	0.6569	1.046	77
2			0.8042	0.805	895	0.6428	1.063	72
3			0.8029	0.808	888	0.6170	1.100	64
1	$\mathbf{^0X_{CW}}$ **(a_k, $[^0EC_k \cdot P2_k]$, CC)**	**100**	**0.7822**	**0.849**	**783**	**0.7388**	**0.941**	**113**
2			**0.7850**	**0.844**	**796**	**0.7229**	**0.959**	**104**
3			**0.7882**	**0.838**	**811**	**0.7280**	**0.953**	**107**
1	$^0X_{CW}$ (a_k, $[^1EC_k \cdot P2_k]$, CC)	115	0.8030	0.808	888	0.6688	1.032	81
2			0.8056	0.802	904	0.6450	1.067	73
3			0.8070	0.800	912	0.6627	1.037	79

Table 3. Correlation weights for calculation of the ${}^{0}X_{CW}$ (a_k, [${}^{0}EC_k \cdot P2_k$], CC), Eq. 8

Graph invariant	Correlation weights on three OCWLGI probes		
	Probe 1	Probe 2	Probe 3
Chemical elements, a_k			
H	0.9998340	0.9995351	1.0001855
B	0.9991865	1.0001772	1.0015599
C↑	1.0000410	1.0012682	1.0014873
N↓	0.9986569	0.9976272	0.9984693
O↓	0.9990589	0.9979545	0.9980054
F	0.9999842	0.9998251	1.0004956
P↑	1.0071452	1.0115583	1.0119361
S↑	1.0015137	1.0020779	1.0029424
Cl↑	1.0009335	1.0010469	1.0016312
As	0.9988792	1.0000591	0.9995905
Br↑	1.0025396	1.0032833	1.0040118
Sn↑	1.0110815	1.0175736	1.0177644
Vertex degree, V_k; $V_k = {}^{(O,\ N)}NNC_k + [{}^{0}EC_k \cdot P2_k]$			
0000↓	0.9976885	0.9963117	0.9956264
0001↑	1.0012995	1.0008622	1.0003258
0002↓	0.9991405	0.9985137	0.9986249
0003↓	0.9999238	0.9999297	0.9994806
0004↓	0.9997487	0.9998733	0.9993658
0006↑	1.0022984	1.0030836	1.0020457
0008↑	1.0012346	1.0011016	1.0004140
0009↑	1.0050735	1.0070186	1.0055096
0010	1.0008595	1.0001990	0.9995228
0012↑	1.0014052	1.0010766	1.0001283
0015	1.0009044	1.0003884	0.9994818
0016↓	0.9974633	0.9964616	0.9960375
0018	1.0007257	0.9997334	0.9994272
0020	1.0000775	0.9995645	0.9990928
0021	1.0002728	0.9990545	0.9986009
0024	1.0010078	1.0005467	0.9998885
0027↑	1.0022504	1.0001859	1.0015291
0028	0.9997687	1.0000610	0.9996549
0032↑	1.0033152	1.0032452	1.0030452
0036↓	0.9999993	0.9990000	0.9987338
0040	0.9964538	0.9955673	0.9956906
0044↑	1.0024235	1.0024519	1.0031876
0048↓	0.9976243	0.9961204	0.9957531
1001↓	0.9963491	0.9949948	0.9948703
1002↑	1.0112949	1.0145131	1.0136740
1004↑	1.0004901	1.0011229	1.0010269
1006↑	1.0027511	1.0039415	1.0035489
1008↑	1.0025957	1.0041884	1.0040253
1009↑	1.0009720	1.0014842	1.0010609
1012↑	1.0015342	1.0022930	1.0019043
1015↑	1.0014771	1.0016453	1.0018110
1016↓	0.9995780	0.9996553	0.9996786
1018↓	0.9990600	0.9982326	0.9981233
1020↓	0.9979906	0.9977518	0.9988808
1024↓	0.9979694	0.9981254	0.9979998
1028↓	0.9982349	0.9980939	0.9983123
1032↑	1.0023755	1.0037498	1.0037656
1036↓	0.9986516	0.9992800	0.9979997
1040↑	1.0063597	1.0089712	1.0093494
2001↑	1.0016881	1.0033469	1.0018755
2002↑	1.0003980	1.0010274	1.0008357
2004↓	0.9990305	0.9985007	0.9983726
2006	1.0004532	1.0006748	0.9996687
2008↑	1.0019059	1.0025843	1.0018330
2009↑	1.0012961	1.0015293	1.0007398
2010↑	1.0068617	1.0082848	1.0075990
2012↓	0.9996479	0.9993119	0.9983751
2015	1.0001005	0.9999291	0.9994017
2016	1.0005927	1.0023917	0.9998568
2018	1.0005284	1.0002412	0.9997648
2020	1.0002175	0.9999562	0.9996233
2021	0.9985933	0.9964266	0.9967526
2024↑	1.0037228	1.0044433	1.0041288
2028↓	0.9914080	0.9880660	0.9885603
2032↓	0.9998259	0.9997321	0.9995136
2036↓	0.9981864	0.9989968	0.9994344
040↑	1.0057307	1.0077729	1.0085289
2044	0.9996237	1.0004030	0.9999970
3003	0.9954763	0.9955652	0.9945065
3006↑	1.0077740	1.0148444	1.0096637
3009↑	1.0020170	1.0038488	1.0031898
3012↑	1.0009684	1.0019065	1.0014588
3015↓	0.9948782	0.9932389	0.9930215
3016↓	0.9960057	0.9967296	0.9963785
3020↓	0.9926485	0.9917179	0.9917178
Cycle code, CC			
9000↑	1.0016570	1.0030643	1.0033934
9030↓	0.9957570	0.9950150	0.9958475
9050↓	0.9947585	0.9917798	0.9920252
9055↑	1.0060757	1.0082072	1.0087171
9070	0.9996192	1.0008696	1.0002415
9100↑	1.0010785	1.0028965	1.0032642
9105↑	1.0069560	1.0120634	1.0123149
9107↑	1.0005926	1.0006442	1.0028704
9130	0.9995461	1.0006853	1.0010635
9135	0.9986245	0.9997472	1.0008700
9150↑	1.0002457	1.0031867	1.0032167
9170	0.9994672	1.0005909	1.0008095
9300	0.9988432	1.0001158	1.0008111
9305↑	1.0023651	1.0053815	1.0057704
9330	0.9983743	0.9986196	1.0006838
9370↑	1.0009969	1.0015447	1.0027625
9500↑	1.0020145	1.0030309	1.0032426
9550↓	0.9954142	0.9961897	0.9954962
9570↑	1.0018114	1.0040032	1.0030620
9575↑	1.0095385	1.0166236	1.0161125
9700↓	0.9970648	0.9980941	0.9970628
9730↑	1.0059454	1.0079552	1.0086505
9770↑	1.0077712	1.0084790	1.0063486

'↑,' increases the toxicity; '↓,' reduces the toxicity.

Global topological (or topochemical) parameters can be useful in toxicity modeling. The presence of cycles differing in size and atom composition can be expressed by an integer. As global invariant integers we use the code '9ABC,' where digits A, B, and C are six-member, five-member, and three-member cycle characteristics, respectively. Table 1 shows the meanings of these digits. For instance, '9170' can be interpreted as the presence in the molecule of an aromatic six-member cycle without heteroatoms (9170), together with a five-member cycle that contains at least one saturated atom and at least one heteroatom (9170); finally, this code indicates the absence of three-member cycles (9170). This global invariant can be named the code of cycles (CC). The symbol '9' excludes obtaining identical values for V_k and CC. All (global and local) topological parameters

of molecular structures can be calculated with the adjacency matrix.

Flexible descriptors used in QSAR modeling of toxicity have the following generalized form as in Eq. 3

$$ {}^{0}X_{\mathrm{CW}}(a_k, \mathrm{LI}_k, \mathrm{CC}) = \mathrm{CW}(\mathrm{CC}) \times \prod_{k=1}^{n} \mathrm{CW}(a_k) \times \mathrm{CW}(V_k), \quad (3) $$

where a_k is the chemical element presented by the kth vertex, LI_k is an integer calculated with Eq. 2, and CC is the cycle code; $\mathrm{CW}(a_k)$, $\mathrm{CW}(\mathrm{LI}_k)$, and CW(CC) are correlation weights of the graph invariants. The correlation coefficient between ${}^{0}X_{\mathrm{CW}}$ (a_k, LI_k, CC) and toxicity is a mathematical function of the correlation weights: R (${}^{0}X_{\mathrm{CW}}$, $\log(1/C)$) = F(CW(CC), $\mathrm{CW}(a_k)$, and $\mathrm{CW}(V_k)$). From these values, one can obtain OCWLGI, that is, the values maximizing the correlation coefficient for the training set. With these numerical data one can calculate the values for the descriptors on all compounds of the training set. The least-squares method can be used to obtain the one-variable model of $\log(1/C) = C0 + C1 \times {}^{0}X_{\mathrm{CW}}$. The predictive ability of this model can be validated with compounds of an internal test set.

The algorithm of optimization used in the present study is described in Ref. 21. This software was implemented by one of the authors (A.T.) using Borland Turbo Pascal 7.0.

QSAR analysis was carried out in OCWLGI using the following six flexible descriptors:

$$ {}^{0}X_{\mathrm{CW}}(a_k, {}^{0}\mathrm{EC}_k, \mathrm{CC}) = \mathrm{CW}(\mathrm{CC}) \times \prod_{k=1}^{n} \mathrm{CW}(a_k) \times \mathrm{CW}({}^{(\mathrm{O,N})}\mathrm{NNC}_k + {}^{0}\mathrm{EC}_k) \quad (4) $$

$$ {}^{0}X_{\mathrm{CW}}(a_k, {}^{1}\mathrm{EC}_k, \mathrm{CC}) = \mathrm{CW}(\mathrm{CC}) \times \prod_{k=1}^{n} \mathrm{CW}(a_k) \times \mathrm{CW}({}^{(\mathrm{O,N})}\mathrm{NNC}_k + {}^{1}\mathrm{EC}_k) \quad (5) $$

$$ {}^{0}X_{\mathrm{CW}}(a_k, \mathrm{P2}_k, \mathrm{CC}) = \mathrm{CW}(\mathrm{CC}) \times \prod_{k=1}^{n} \mathrm{CW}(a_k) \times \mathrm{CW}({}^{(\mathrm{O,N})}\mathrm{NNC}_k + \mathrm{P2}_k) \quad (6) $$

$$ {}^{0}X_{\mathrm{CW}}(a_k, [{}^{0}\mathrm{EC}_k \cdot {}^{1}\mathrm{EC}_k], \mathrm{CC}) = \mathrm{CW}(\mathrm{CC}) \times \prod_{k=1}^{n} \mathrm{CW}(a_k) \times \mathrm{CW}({}^{(\mathrm{O,N})}\mathrm{NNC}_k + [{}^{0}\mathrm{EC}_k \cdot {}^{1}\mathrm{EC}_k]) \quad (7) $$

$$ {}^{0}X_{\mathrm{CW}}(a_k, [{}^{0}\mathrm{EC}_k \cdot \mathrm{P2}_k], \mathrm{CC}) = \mathrm{CW}(\mathrm{CC}) \times \prod_{k=1}^{n} \mathrm{CW}(a_k) \times \mathrm{CW}({}^{(\mathrm{O,N})}\mathrm{NNC}_k + [{}^{0}\mathrm{EC}_k \cdot \mathrm{P2}_k]) \quad (8) $$

$$ {}^{0}X_{\mathrm{CW}}(a_k, [{}^{1}\mathrm{EC}_k \cdot \mathrm{P2}_k], \mathrm{CC}) = \mathrm{CW}(\mathrm{CC}) \times \prod_{k=1}^{n} \mathrm{CW}(a_k) \times \mathrm{CW}({}^{(\mathrm{O,N})}\mathrm{NNC}_k + [{}^{1}\mathrm{EC}_k \cdot \mathrm{P2}_k]) \quad (9) $$

3. Results and discussion

Table 2 shows the results of the OCWLGI modeling based on Eqs. 4–9: with Eq. 8 it gives the best model. The toxicity model based on correlation weights of probe 1, obtained by the least-squares method, is the following:

$$ \log(1/C) = -226.250 + 227.538 \times {}^{0}X_{\mathrm{CW}}(a_k, [{}^{0}\mathrm{EC}_k \cdot \mathrm{P2}_k], \mathrm{CC}) $$
$$ n = 220, \quad r^2 = 0.7822, \quad s = 0.849, $$
$$ F = 783(\text{training set}); $$
$$ n = 42, \quad r^2 = 0.7388, \quad s = 0.941, \quad F = 113(\text{test set}) \quad (10) $$

Table 4. Demonstration of calculation of the ${}^{0}X_{\mathrm{CW}}$ (a_k, [${}^{0}\mathrm{EC}_k \cdot \mathrm{P2}_k$], CC) on the structure of DCDMH (Gluchlor)

Chemical element	No.	[${}^{0}\mathrm{EC}_k \cdot \mathrm{P2}_k$]	CW (a_k)	CW (V_k)
N	1	12	0.9986569	1.0014052
C	2	3015	1.0000410	0.9948782
C	3	2040	1.0000410	1.0057307
N	4	15	0.9986569	1.0009044
C	5	3012	1.0000410	1.0009684
Cl	6	2002	1.0009335	1.0003980
O	7	2	0.9990589	0.9991405
Cl	8	2002	1.0009335	1.0003980
O	9	2	0.9990589	0.9991405
C	10	12	1.0000410	1.0014052
H	11	3	0.9998340	0.9999238
H	12	3	0.9998340	0.9999238
H	13	3	0.9998340	0.9999238
C	14	12	1.0000410	1.0014052
H	15	3	0.9998340	0.9999238
H	16	3	0.9998340	0.9999238
H	17	3	0.9998340	0.9999238

The CAS number is 118-52-5, CC = 9070, CW(CC) = 0.9996192. Numbering of atoms is shown in Figure 1. ${}^{0}X_{\mathrm{CW}}$ (a_k, [${}^{0}\mathrm{EC}_k \cdot \mathrm{P2}_k$], CC) = 1.00141.

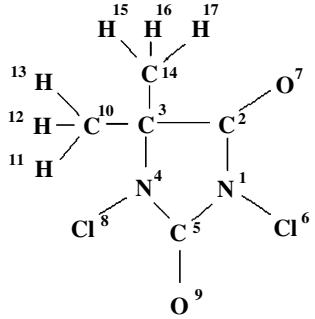

Figure 1. Numbering of atoms in the molecular graph of DCDMH (Gluchlor).

Table 5. Pesticides considered in the training and test sets, experimental and calculated with Eq. 10 toxicity towards *Daphnia magna*

No.	EPA ID	Pesticides	CAS	$^{0}X_{CW}$	Experimental	Calculated	Experimental − calculated
Training set							
1	002	*Z*-11-Hexadecanol	53939-28-9	1.00438	2.37700	2.28462	0.09238
2	003	DCDMH (Gluchlor formulation)	118-52-5	1.00141	2.59500	1.60883	0.98617
3	004	Dichloropropene	542-75-6	1.00462	3.09100	2.33923	0.75177
4	005	Naphthaleneacetic acid	8687-3	0.99764	0.01500	0.75101	−0.73601
5	006	2,4-D-Acid	94-75-7	1.00266	0.94600	1.89325	−0.94725
6	010	2,4-D-Butoxyethanol ester	1929-73-3	1.00733	1.64900	2.95585	−1.30685
7	013	2-Benzyl-4-chlorophenol	120-32-1	1.00010	2.56900	1.31075	1.25825
8	014	3-Chloro-*p*-toluidine hydrochloride	7745-89-3	1.00672	3.40500	2.81706	0.58794
9	017	4-Chloro-3,5-xylenol	88-04-0	1.00166	1.30800	1.66571	−0.35771
10	020	Acetochlor	34256-82-1	0.99895	1.51700	1.04909	0.46791
11	021	Alachlor	15972-60-8	0.99871	1.10900	0.99448	0.11452
12	022	Aldicarb	116-06-03	1.00597	2.66600	2.64640	0.01960
13	023	Aldoxycarb	1646-88-4	1.00457	2.90000	2.32785	0.57215
14	025	Ametryn	834-12-8	1.00180	0.90900	1.69757	−0.78857
15	026	Amitraz	33089-61-1	1.00790	3.92300	3.08555	0.83745
16	027	Anilazine	101-05-3	1.00340	2.39900	2.06163	0.33737
17	030	Bendiocarb	22781-23-3	1.01057	3.88300	3.69308	0.18992
18	032	Bensulide	741-58-2	1.00827	2.83600	3.16974	−0.33374
19	033	Cyfluthrin	68359-37-5	1.02197	6.48800	6.28701	0.20099
20	035	Bromacil	314-40-9	1.00041	0.33400	1.38129	−1.04729
21	036	Bromoxynil (phenol)	1689-84-5	1.00464	1.15800	2.34378	−1.18578
22	038	Bromoxynil (buturate)	3861-41-4	1.01069	3.23900	3.72038	−0.48138
23	039	Bronopol	52-51-7	0.99586	2.09700	0.34599	1.75101
24	043	Captan	133-06-2	1.00362	1.55370	2.11169	−0.55799
25	044	Carbaryl	63-25-2	1.00866	4.55500	3.25848	1.29652
26	048	Oxythioquinox	2439-01-2	1.00643	3.29100	2.75107	0.53993
27	049	Kepone	143-50-0	1.00931	3.27570	3.40638	−0.13068
28	051	Chlorhexidine diacetate	55-56-1	1.01150	3.99700	3.90469	0.09231
29	054	Chlorophacinone	3691-35-8	1.00775	2.94400	3.05142	−0.10742
30	057	Chlorpropham	101-21-3	1.00479	1.76100	2.37791	−0.61691
31	063	Clomazone	81777-89-1	1.00288	1.66370	1.94331	−0.27961
32	067	Cycloate	1134-23-2	1.00499	0.95300	2.42341	−1.47041
33	068	Cyhexatin	13121-70-5	1.01387	6.35500	4.44395	1.91105
34	069	Cymoxanil	57966-95-7	1.00153	0.85000	1.63613	−0.78613
35	072	Cyprodinil	121552-61-2	1.01119	3.84700	3.83415	0.01285
36	073	Daminozide	1596-84-5	0.99216	0.21100	−0.4959	0.70690
37	075	DBNPA	10222-01-2	1.00299	2.44900	1.96834	0.48066
38	076	Dichloro-diphenyl-trichloroethane	50-29-3	1.01175	4.87700	3.96157	0.91543
39	081	Dibromodicyanobutane	35691-65-7	1.00584	2.08200	2.61682	−0.53482
40	087	Dichloran (DCNA)	99-30-9	1.00785	2.00000	3.07417	−1.07417
41	090	Dienochlor	2227-17-0	1.00624	2.59700	2.70784	−0.11084
42	091	Difenoconazole	119446-68-3	1.00537	2.72200	2.50988	0.21212
43	094	Diflufenzopyr-sodium	109293-97-2	0.99931	1.37600	1.13100	0.24500
44	095	Dimethenamid	87674-68-8	0.99874	1.36100	1.00130	0.35970
45	104	Dowicil	4080-31-3	0.99690	0.79790	0.58263	0.21527
46	110	Ethalfluralin	55283-68-6	1.01128	3.74500	3.85463	−0.10963
47	111	Ethion	563-12-2	1.03017	6.83700	8.15282	−1.31582
48	116	Farnesol	4602-84-0	1.00391	2.00500	2.17767	−0.17267
49	119	Fenarimol	60168-88-9	1.00185	1.68700	1.70895	−0.02195
50	122	Fenoxaprop-ethyl	66441-23-4	1.00921	2.05600	3.38362	−1.32762
51	124	Fenpropathrin	39515-41-8	1.01528	5.81900	4.76478	1.05422
52	125	Fendridazone-sodium	68254-10-4	0.99564	0.18300	0.29593	−0.11293
53	126	Fenthion	55-38-9	1.01779	4.72900	5.33590	−0.60690
54	127	Fentin hydroxide	76-87-9	1.01521	4.56500	4.74885	−0.18385
55	131	Fludioxonil (Maxim4FS)	131341-86-1	1.00518	2.44000	2.46665	−0.02665
56	133	Fluridone	59756-60-4	1.00425	1.87400	2.25504	−0.38104
57	134	Flurprimidol	56425-91-3	1.00389	1.42300	2.17312	−0.75012
58	137	Folpet	133-07-3	1.01151	4.17100	3.90696	0.26404
59	138	Fonofos	944-22-9	1.01474	5.09000	4.64191	0.44809
60	139	Formethanate hydrochloride	22259-30-9	1.00831	3.47300	3.17884	0.29416
61	140	Fosamineammonium	59682-52-9	0.99970	−0.9520	1.21974	−2.17174
62	143	Glyphosate (8OWDG)	1071-83-6	0.99511	0.10100	0.17534	−0.07434
63	145	Hexadecadienol acetate	53042-79-8	1.00335	2.60300	2.05025	0.55275
64	147	Imazethabenz	81405-85-8	0.99530	0.11700	0.21857	−0.10157
65	148	Imazaquin	81335-37-7	0.99226	0.04600	−0.4731	0.51914

(*continued on next page*)

Table 5 (*continued*)

No.	EPA ID	Pesticides	CAS	$^{0}X_{CW}$	Experimental	Calculated	Experimental − calculated
66	152	Isofenphos	25311-71-1	1.00914	4.94700	3.36770	1.57930
67	157	Lithium per-fluorooctane sulfonate	29457-72-5	1.00548	0.87800	2.53491	−1.65691
68	158	Malathion	121-75-5	1.01483	5.51900	4.66239	0.85661
69	163	Methidathion	950-37-8	1.01310	4.67400	4.26875	0.40525
70	165	Methomyl	16752-77-5	1.00598	4.26600	2.64868	1.61732
71	168	Methylchloroform	71-55-6	1.00641	1.07600	2.74652	−1.67052
72	169	Methylene bis(thiocyanate)	6317-18-6	1.00876	3.52300	3.28123	0.24177
73	171	Metolachlor	87392-12-9	0.99786	1.08200	0.80107	0.28093
74	172	Metribuzin	21087-64-9	1.00134	1.71000	1.59290	0.11710
75	174	MTI	82633-79-2	1.00376	2.07700	2.14354	−0.06654
76	176	DEET	134-62-3	1.00425	0.40700	2.25504	−1.84804
77	177	Naled	300-76-5	1.01790	5.97900	5.36093	0.61807
78	178	Napropamide	15299-99-7	0.99922	1.27800	1.11052	0.16748
79	179	Naptalam	132-66-1	1.00302	0.42200	1.97516	−1.55316
80	182	Nitrapyrin	1929-82-4	1.00583	2.02100	2.61455	−0.59355
81	183	*N*-Methylneodecanamide	105726-67-8	0.99748	0.15400	0.71460	−0.56060
82	185	Imidacloprid	105827-78-9	0.99665	0.47700	0.52575	−0.04875
83	186	OBPA	58-36-6	1.01640	5.02000	5.01962	0.00038
84	187	Octhilinone	26530-20-1	1.00614	3.07400	2.68508	0.38892
85	188	Oryzalin	19044-88-3	1.00602	2.36300	2.65778	−0.29478
86	189	Oxadiazon	19666-30-9	1.00429	2.20000	2.26414	−0.06414
87	191	Oxazolidine E	7747-35-5	1.00171	0.53300	1.67709	−1.14409
88	192	Oxydemeton-methyl	301-12-2	1.00855	3.01100	3.23345	−0.22245
89	197	Methylparathion	298-00-0	1.01582	6.27400	4.88765	1.38635
90	198	Pentachlorophenol	87-86-5	1.00518	3.04500	2.46665	0.57835
91	199	Pebulate	1114-71-2	1.00437	1.47400	2.28234	−0.80834
92	200	Pendimethalin	40487-42-1	1.00757	3.00200	3.01046	−0.00846
93	205	Phosmet	732-11-6	1.01293	4.75300	4.23007	0.52293
94	209	Piperonylbutoxide	51-03-6	1.00689	2.82200	2.85574	−0.03374
95	213	Profenofos	41198-08-7	1.01773	5.60400	5.32225	0.28175
96	214	Prometon	1610-18-0	1.00138	0.94300	1.60200	−0.65900
97	216	Propachlor	1918-16-7	0.99849	1.48700	0.94442	0.54258
98	217	Propanil	709-98-8	0.99969	1.51200	1.21746	0.29454
99	224	Rotenone	83-79-4	1.01448	5.02800	4.58275	0.44525
100	227	Simazine	122-34-9	0.99903	2.26300	1.06729	1.19571
101	233	TBTF	1983-10-4	1.02001	6.09200	5.84104	0.25096
102	236	Phostebupirin	96182-53-5	1.02298	6.61000	6.51682	0.09318
103	237	Tebuthiuron	34014-18-1	1.00153	−0.1140	1.63613	−1.75013
104	238	Themephos	3383-96-8	1.03374	7.62700	8.96513	−1.33813
105	239	Terbacil	5902-51-2	0.99900	0.52300	1.06046	−0.53746
106	240	Terbufos	13071-79-9	1.01780	5.96900	5.33818	0.63082
107	241	Terbutylazine	5915-41-3	1.00035	1.03500	1.36764	−0.33264
108	244	Triazopyr	117718-60-2	1.00427	1.81300	2.25959	−0.44659
109	248	Tralomethrin	66841-25-6	1.02835	7.23200	7.73870	−0.50670
110	254	Trichloromelamine	7673-09-8	1.00514	2.28100	2.45755	−0.17655
111	256	Trichlopyr acid	55335-06-3	1.00037	0.28500	1.37219	−1.08719
112	257	Triclosan	3380-34-5	1.00801	2.87100	3.11058	−0.23958
113	258	Tridiphane	58138-08-2	1.00810	3.10800	3.13106	−0.02306
114	262	Trimethacarb	2686-99-9	1.00761	4.03100	3.01956	1.01144
115	263	Tris(hydroxymethyl)nitromethane	126-11-4	0.98991	0.27600	−1.0079	1.28386
116	269	Propionic acid	79-09-4	0.99677	0.51400	0.55305	−0.03905
117	272	Acetaldehyde	26532-25-2	1.00294	1.51000	1.95696	−0.44696
118	273	3,5-Dimethyl-(1-hydroxymethyl)pyrazole	85264-33-1	0.99428	0.60000	−0.0135	0.61352
119	275	2,4-D-Isopropyl ester	94-11-1	1.00708	2.00500	2.89897	−0.89397
120	276	2-Hydroxyethyl octyl sulfide	3547-33-9	0.99937	2.71100	1.14465	1.56635
121	278	4,5-Dichloro-1,2-dithio-3-one	1192-52-5	1.00931	4.23100	3.40638	0.82462
122	279	Acibenzolar-*s*-methyl	135158-54-2	1.00207	1.86000	1.75900	0.10100
123	280	Azoxystrobin	131860-33-8	1.00768	3.19200	3.03549	0.15651
124	285	Bifenazate	149877-41-8	1.00450	2.77900	2.31192	0.46708
125	286	Bifenthrin	82657-04-3	1.02125	5.42200	6.12318	−0.70118
126	288	Biobor	2665-13-6	0.99226	−0.4070	−0.4731	0.06614
127	290	Bromonitrostyrene	7166-19-0	1.00606	3.97800	2.66688	1.31112
128	297	DDAC	7173-51-5	1.01294	4.30300	4.23234	0.07066
129	300	Diflubenzuron	35367-38-5	1.00703	4.92400	2.88759	2.03641
130	302	Dimethepin	55290-64-7	0.99632	0.99400	0.45066	0.54334
131	303	Diphenylamine	122-39-4	1.00672	2.14900	2.81706	−0.66806

Table 5 (*continued*)

No.	EPA ID	Pesticides	CAS	$^{0}X_{CW}$	Experimental	Calculated	Experimental − calculated
132	304	Dipropyl-iso-cinchomeronate	136-45-8	0.99988	1.14500	1.26070	−0.11570
133	305	DTEA	36362-09-1	1.00782	3.84800	3.06735	0.78065
134	306	Ethoxyquin	91-53-2	1.00875	2.03600	3.27896	−1.24296
135	307	Etridiazole	2593-15-9	1.00438	1.70300	2.28462	−0.58162
136	308	Fenbuconazole	114369-43-6	1.00535	2.16600	2.50533	−0.33933
137	310	Fenpyroximate	134098-61-6	1.01551	5.06800	4.81711	0.25089
138	311	Fluazinam	79622-59-6	1.00835	3.41200	3.18794	0.22406
139	315	Flumioxazin(V-53482)	103361-09-7	1.00265	1.80900	1.89098	−0.08198
140	317	Imazalil	35554-44-0	1.00155	1.97300	1.64068	0.33232
141	318	Kresoxim methyl	143390-89-0	1.01020	2.97500	3.60889	−0.63389
142	319	Lambda Cynalothrin	91465-08-6	1.02134	6.29100	6.14366	0.14734
143	320	Methylisothiocyanate	556-61-6	1.00314	3.12400	2.00247	1.12153
144	321	MethylNonylketone	112-12-9	1.00381	2.49900	2.15492	0.34408
145	323	Naphthalene	91-20-3	1.00598	1.90400	2.64868	−0.74468
146	328	Parachloromethacresol	59-50-7	1.00122	1.79200	1.56560	0.22640
147	329	PCNB	82-68-8	1.00960	2.58300	3.47236	−0.88936
148	331	Pirimiphos-methyl	29332-93-7	1.01212	6.16300	4.04576	2.11724
149	332	Potassium salt oleic acid	143-18-0	1.00208	2.75000	1.76128	0.98872
150	333	Prallethrin	23031-36-9	1.01481	4.68500	4.65784	0.02716
151	334	Pyridaben	96489-71-3	1.01272	5.83800	4.18228	1.65572
152	337	Strychnine	57-24-9	1.00359	1.52400	2.10486	−0.58086
153	338	Tabuconazole	107534-96-3	1.00619	1.88600	2.69646	−0.81046
154	339	Tefluthrin	79538-32-2	1.01859	6.77700	5.51793	1.25907
155	340	Thiodicarb	59669-26-0	1.01294	4.11800	4.23234	−0.11434
156	342	Trifloxystrobin	141517-21-7	1.01350	4.20800	4.35976	−0.15176
157	345	Clopyralid-(2-hydroxyethyl)ammonium	57754-85-5	0.99684	0.05100	0.56898	−0.51798
158	346	Sodium dichloro-*s*-triazinethrione	2893-78-9	0.99954	3.30100	1.18333	2.11767
159	348	2,4-**D**-Isooctyl ester	25168-26-7	1.01031	4.24900	3.63392	0.61508
160	349	2,4-DB	94-82-6	1.00090	4.82000	1.49278	3.32722
161	350	Asulam	3337-71-1	1.00030	0.93100	1.35626	−0.42526
162	351	Azinphos-methyl	86-50-0	1.01084	5.46000	3.75451	1.70549
163	352	Bromadiolone	28772-56-7	1.00530	2.42100	2.49395	−0.07295
164	353	Bromethalin	63333-35-7	1.01388	5.46500	4.44623	1.01877
165	354	BCDMH	16079-88-2	1.00301	2.50800	1.97289	0.53511
166	355	2,6-Dibromo-4-cyanophenylheptanoate	56634-95-8	1.01240	4.09900	4.10947	−0.01047
167	356	Captafol	2425-06-1	1.00339	2.00600	2.05935	−0.05335
168	357	Carbendazim	10605-21-7	1.00253	2.73700	1.86367	0.87333
169	358	Chlormequat	7003-89-6	0.99991	0.97600	1.26752	−0.29152
170	359	Chlorothalonil	1897-45-6	1.00712	3.58000	2.90807	0.67193
171	360	Chlorsulfuron	64902-72-3	0.99894	−0.0140	1.04681	−1.06081
172	362	Cyanazine	21725-46-2	0.99865	0.69100	0.98082	−0.28982
173	363	Cyclanide	113136-77-9	1.00675	1.73900	2.82388	−1.08488
174	364	Cyproconazole	94361-06-5	0.99899	1.05000	1.05819	−0.00819
175	365	Decanol	112-30-1	1.00356	1.38700	2.09804	−0.71104
176	366	Diazinon	333-41-5	1.01663	5.50100	5.07196	0.42904
177	368	Dicamba (acid)	1918-00-9	1.00115	0.30000	1.54967	−1.24967
178	369	Dichlofop-methyl	51338-27-3	1.00984	3.17100	3.52697	−0.35597
179	370	Diphenamid	957-51-7	1.00191	0.61600	1.72260	−1.10660
180	371	Endosulfan	115-29-7	1.00881	3.38900	3.29261	0.09639
181	372	Endothall	145-73-3	0.99521	0.75800	0.19809	0.55991
182	373	Ethephon	16672-87-0	1.00319	0.65900	2.01385	−1.35485
183	375	Ethoprop	13194-48-4	1.01211	3.41600	4.04349	−0.62749
184	376	Fenamiphos	22224-92-6	1.01015	5.20300	3.59751	1.60549
185	377	Fenitrothion	122-14-5	1.01627	4.40100	4.99004	−0.58904
186	378	Fluchloralin	33245-39-5	1.00791	2.80300	3.08783	−0.28483
187	380	1Flumeturon	2164-17-2	1.00626	1.37000	2.71239	−1.34239
188	381	4Famesafen	72178-02-0	1.00322	0.17400	2.02067	−1.84667
189	382	Glufosinate-ammonium	51276-47-2	0.99548	−0.5270	0.25953	−0.78653
190	385	Hexazinone	51235-04-2	0.99996	0.22100	1.27890	−1.05790
191	386	Hymexazol	10004-44-1	0.99585	0.50700	0.34372	0.16328
192	387	Indole-3-butyric acid	133-32-4	0.99642	0.55200	0.47341	0.07859
193	388	Kinoprene	42588-37-4	1.00926	3.44200	3.39500	0.04700
194	390	Maleic hydrazide	123-33-1	0.99518	0.01800	0.19127	−0.17327
195	392	Methoprene	40596-69-8	1.01187	2.54300	3.98888	−1.44588
196	393	Methoxyfenazide	161050-58-4	1.00498	1.99800	2.42114	−0.42314
197	394	Methylbromide	74-83-9	1.00112	1.56200	1.54284	0.01916

(*continued on next page*)

Table 5 (*continued*)

No.	EPA ID	Pesticides	CAS	$^{0}X_{CW}$	Experimental	Calculated	Experimental − calculated
198	395	N6-Benzuladenine	1214-39-7	1.00089	1.04100	1.49051	−0.44951
199	396	Oxadixyl	77732-09-3	0.99302	−0.2800	−0.3002	0.02022
200	397	Phosphamidon	13171-21-6	1.01082	4.37300	3.74996	0.62304
201	398	Picloram (acid)	1918-02-1	0.99950	0.54800	1.17423	−0.62623
202	399	Piperalin	3478-94-2	1.00344	2.24200	2.07073	0.17127
203	400	Propiconazole	60207-90-1	1.00148	1.85300	1.62476	0.22824
204	401	Propoxur	114-26-1	1.00676	4.27900	2.82616	1.45284
205	402	Pymetrozine	123312-89-0	0.99293	0.39700	−0.3207	0.71769
206	403	Pyraclostrobin	175013-18-0	1.01217	4.39200	4.05714	0.33486
207	405	Pyriproxyfen	95737-68-1	1.00768	2.90500	3.03549	−0.13049
208	407	Rimsulfuron	122931-48-0	0.99602	−0.3650	0.38240	−0.74740
209	410	Sulfentrazone	122836-35-5	0.99726	0.80700	0.66455	0.14245
210	411	Sulfluramid	4151-50-2	1.00108	3.15400	1.53374	1.62026
211	412	TCMTB	21564-17-0	1.01266	4.01500	4.16863	−0.15363
212	413	Tetrachlorvinphos	22248-79-9	1.01839	5.28500	5.47242	−0.18742
213	414	Tetramethrin	7696-12-0	1.01076	3.86700	3.73631	0.13069
214	415	Thiabendazole	148-79-8	1.00862	2.81200	3.24938	−0.43738
215	416	Thidiazuron	51707-55-2	1.00122	1.34300	1.56560	−0.22260
216	417	Thiophanate-methyl	23564-05-8	1.00571	1.80200	2.58724	−0.78524
217	418	Tricosene	27519-02-4	1.01185	2.47500	3.98433	−1.50933
218	419	Triforine	26644-46-2	0.99784	1.19100	0.79652	0.39448
219	420	Triticonazole	131983-72-7	1.00611	1.62100	2.67826	−1.05726
220	421	Vinclozolin	50471-44-8	1.00033	1.89400	1.36309	0.53091
Test set							
1	001	Nonanoic acid	112-05-0	1.00015	0.21700	1.32213	−1.10513
2	008	2,4-D-2-Butoxymethylethyl ester	1320-18-9	1.00694	2.44600	2.86711	−0.42111
3	016	Dichloro-2*N*-octyl-3(2*H*)iso-thiozolone	64359-81-5	1.00836	4.73300	3.19022	1.54278
4	034	Bis(trichloromethyl)sulfone	3064-70-8	1.00716	3.24000	2.91717	0.32283
5	037	Bromoxynil octanoate	1689-99-2	1.01297	3.56400	4.23917	−0.67517
6	046	Carbofuran	1563-66-2	1.01308	3.88200	4.26420	−0.38220
7	047	Carboxin	5234-68-4	0.99826	0.44500	0.89208	−0.44708
8	053	Chloroneb	2675-77-6	1.00746	1.52400	2.98543	−1.46143
9	059	Chlorpyrifos-methyl	5598-13-0	1.01526	5.46300	4.76023	0.70277
10	083	Dichlobenil	1194-65-6	1.00536	1.44300	2.50760	−1.06460
11	088	Dicrotophos	141-66-2	1.01118	4.27100	3.83187	0.43913
12	092	Difenzoquat methylsulfate	49866-87-7	1.00377	2.15370	2.14582	0.00788
13	093	Difethialone	104653-34-1	1.01280	5.08800	4.20049	0.88751
14	098	Diphacinone	82-66-6	1.00664	2.27700	2.79885	−0.52185
15	101	Dithiopyr	97886-45-8	1.00574	1.37300	2.59407	−1.22107
16	123	Fenoxycarb	79127-80-3	1.00837	2.87700	3.19249	−0.31549
17	146	Hydramethylnon	67485-29-4	1.00299	2.63700	1.96834	0.66866
18	151	Igrarol	28159-98-0	1.00636	1.67900	2.73514	−1.05614
19	153	Isopropalin	33820-53-0	1.01095	3.05900	3.77954	−0.72054
20	180	Neurolidol	7212-44-4	1.00263	2.00500	1.88642	0.11858
21	193	Paclobutrazol	76738-62-0	0.99711	0.94700	0.63042	0.31658
22	215	Prometryn	7287-19-6	1.00237	1.11300	1.82727	−0.71427
23	225	Sethoxydim	74051-80-2	0.99754	0.62300	0.72826	−0.10526
24	228	Sodium 2-mercaptobenzothiolate	2492-26-4	1.00780	1.81500	3.06280	−1.24780
25	230	Sodium fluoroacetate	62-74-8	0.99711	−0.5440	0.63042	−1.17442
26	245	Thiobencarb	28249-77-6	1.00886	3.40700	3.30399	0.10301
27	271	1,3,5-Triethylhexahydroxy-*s*-triazine	7779-27-3	1.00170	1.04900	1.67481	−0.62581
28	274	1,2-Benzenedicarboxyaldehyde	643-79-8	1.00422	3.17300	2.24821	0.92479
29	289	Brodifacoum	56073-10-0	1.01090	2.72800	3.76816	−1.04016
30	322	MGK-264	113-48-4	1.00305	2.07800	1.98199	0.09601
31	330	Pirimicarb	23103-98-2	1.00606	4.56400	2.66688	1.89712
32	347	2-(Hydroxymethylamino)ethanol	34275-28-5	0.99577	0.55800	0.32551	0.23249
33	361	Coumaphos	56-72-4	1.01656	6.27600	5.05603	1.21997
34	374	Ethofenprox	80844-07-1	1.01649	5.82000	5.04010	0.77990
35	379	Flumetsulam	98967-40-9	0.99574	0.10700	0.31869	−0.21169
36	383	Halofenoxide	112226-61-6	1.00393	1.96300	2.18222	−0.21922
37	384	Hexaflumuron	86479-06-3	1.00962	6.61900	3.47692	3.14208
38	391	Methiocarb	2032-65-7	1.00914	4.07400	3.36770	0.70630
39	404	Pyridate	55512-33-9	1.00821	2.54500	3.15609	−0.61109
40	406	Quinolorac	84087-01-4	0.99921	0.91000	1.10824	−0.19824
41	408	Sodium 2-phenylphenoate	90-43-7	1.00289	1.70400	1.94558	−0.24158
42	422	Warfarin	81-81-2	0.99858	0.37500	0.96490	−0.58990

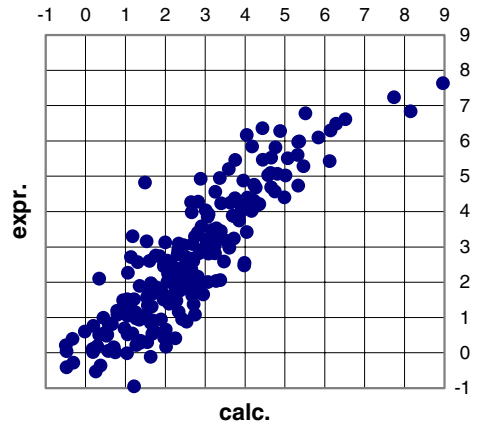

Figure 2. Plot of Daphnia toxicity of pesticides calculated with Eq. 10 against the experimental values of log(1/*C*) on the training set.

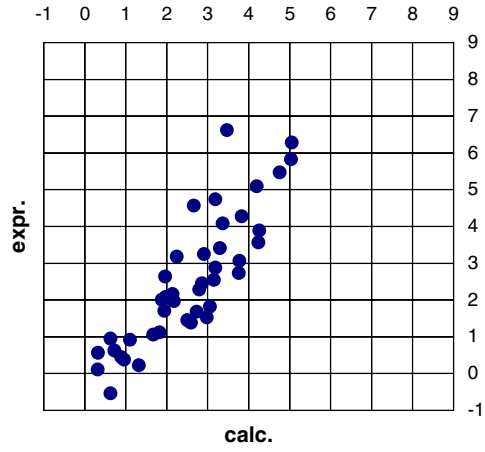

Figure 3. Plot of Daphnia toxicity of pesticides calculated with Eq. 10 against the experimental values of log(1/*C*) on the test set.

The statistical characteristics of OCWLGI models obtained without correlation weights of the CC are unsatisfactory, so cycles have an important role in biochemical processes, like in physical and chemical phenomena.[23] The influence of cycles on different kinds of toxicity is well known.[24–32]

Table 3 lists the correlation weights for calculation of ${}^{0}X_{CW}$ (a_k, $[{}^{0}EC_k \cdot P2_k]$, CC). These were obtained in three Monte Carlo method optimization processes. The numerical data on the weights are not identical over the three probes, but in spite of this the statistical characteristics of models with the same list of invariants are almost identical. There are three classes of topological parameters:

Class 1: Topological parameters that have stable correlation weights > 1.0 over all three probes; this class of topological parameters is related to an increase of toxicity toward *D. magna*.

Class 2: Topological parameters that have stable correlation weights < 1.0 over all three probes, meaning that this class is related to less toxicity.

Class 3: A topological parameter that has correlation weights both >1.0 and <1.0 in the three probes of optimization. Most probably this class has no significant influence on the toxicity under consideration.

This interpretation of the graph invariants related to the chemical element is easily understandable and in several cases reasonable. For instance, the relationship between the presence of Cl atom and toxicity can be explained by the toxicity of chlorinated pesticides. The relationship with P atom is likely to be due to the toxicity of organophosphates. In the case of descriptors related to a ring, the presence of a three-member ring is like to be related to increased toxicity. Less clear is the meaning of the vertex-based invariants. Classes 1 and 2 are shown in Table 3, by characters ↑ and ↓, respectively.

Table 4 shows the calculation of the OCWLGI descriptor with Eq. 8 for Gluchlor, whose structure is shown in Figure 1. Table 5 presents experimental toxicity values and those calculated with Eq. 10 as well as CAS numbers. Figures 2 and 3 represent graphically the correlations between experimental and calculated toxicity values on the training and test sets.

Table 6 lists some models of Daphnia toxicity from the literature. Comparison of these statistical characteristics indicates that the present model is good.

4. Conclusions

A QSAR model based on topological descriptors has been obtained and validated on an independent test set of compounds. The model is powerful considering the widely heterogeneous data set, containing pesticides which are active compounds with many functional groups, predicting toxic effects on Daphnia through a number of different modes of action.

Table 6. Statistical characteristics on some Daphnia toxicity models from the literature

N (training)	*N* (test)	R^2 (training)	R^2 (test)	Method	Reference
20	—	0.92	—	Quantum chemical descriptors	29
24	5	0.87	—	2D,3D descriptors	30
700	76	0.76	—	Probabilistic neural networks	31
217	—	0.97	—	Group contribution method	32
62	31	0.57	0.60	Physical and chemical parameters, quantum chemistry	33
220	42	0.78	0.73	OCWLGI	Present study

The topological parameters described in the present paper can be related to toxicity. The local invariant of the molecular graph, that is equal to the multiplication of vertex degree by number of paths of length 2, is the most appropriate for constructing the optimal descriptors for modeling toxicity toward *D. magna*.

Acknowledgment

This work was funded by EU contract QLK5-CT-2002-00691, Demetra.

References and notes

1. Ren, B. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1121.
2. Bonchev, D.; Kier, L. B. *J. Math. Chem.* **1992**, *9*, 75.
3. Basak, S. C.; Niemi, G. J.; Veith, G. D. *J. Math. Chem.* **1990**, *4*, 185.
4. Basak, S. C.; Niemi, G. J.; Veith, G. D. *J. Math. Chem.* **1991**, *7*, 243.
5. Randic, M. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 627.
6. Randic, M.; Basak, S. C. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 261.
7. Dearden, J. C.; Ghafourian, T. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 231.
8. Balaban, A. T.; Basak, S. C.; Mills, D. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 769.
9. Gutman, I.; Das, K. Ch. *MATCH Commun. Math. Comput. Chem.* **2004**, *50*, 83.
10. Gutman, I.; Toropov, A. A.; Toropova, A. P. *MATCH Commun. Math. Comput. Chem.* **2005**, *53*, 215.
11. Gutman, I.; Furtula, B.; Toropov, A. A.; Toropova, A. P. *MATCH Commun. Math. Comput. Chem.* **2005**, *53*, 225.
12. Toropov, A. A.; Toropova, A. P. *J. Mol. Struct. (THEOCHEM)* **2002**, *581*, 11.
13. Toropov, A. A.; Toropova, A. P. *J. Mol. Struct. (THEOCHEM)* **2004**, *711*, 173.
14. Toropov, A. A.; Toropova, A. P. *J. Mol. Struct. (THEOCHEM)* **2001**, *538*, 287.
15. Seward, J. R.; Cronin, M. T. D.; Schultz, T. W. *SAR QSAR Environ. Res.* **2001**, *11*, 489.
16. Schultz, T. W. *Toxicol. Methods* **1997**, *7*, 289.
17. Dimitrov, S. D.; Mekenyan, O. G.; Schultz, T. W. *Bull. Environ. Contam. Toxicol.* **2000**, *65*, 399.
18. Toropov, A. A.; Toropova, A. P. *J. Mol. Struct. (THEOCHEM)* **2002**, *578*, 129.
19. Toropov, A. A.; Benfenati, E. *J. Mol. Struct. (THEOCHEM)* **2004**, *676*, 165.
20. Toropov, A. A.; Benfenati, E. *J. Mol. Struct. (THEOCHEM)* **2004**, *679*, 225.
21. Toropov, A. A.; Schultz, T. W. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 560.
22. Roncaglioni, A.; Benfenati, E.; Borani, E.; Clook, M. *J. Environ. Sci. Health., Part B* **2004**, *4*, 641.
23. Gutman, I.; Petrovic, V.; Mohar, B. *Chem. Phys. Lett.* **1993**, *203*, 378.
24. Tremolada, P.; Finizio, A.; Villa, S.; Gaggi, C.; Vighi, M. *Aquat. Toxicol.* **2004**, *67*, 87.
25. Cleuvers, M. *Chemosphere* **2005**, *59*, 199.
26. Cleuvers, M. *Toxicol. Lett.* **2003**, *142*, 185.
27. Devillers, J.; Flatin, J. *SAR QSAR Environ. Res.* **2000**, *11*, 25.
28. Devillers, J. *SAR QSAR Environ. Res.* **2001**, *11*, 397.
29. Liu, X.; Wang, B.; Huang, Z.; Han, S.; Wang, L. *Chemosphere* **2003**, *50*, 403.
30. Papa, E.; Battaini, F.; Grammatica, P. *Chemosphere* **2005**, *58*, 559.
31. Kaiser, K. L. E. *J. Mol. Struct. (THEOCHEM)* **2003**, *622*, 85.
32. Tao, Sh.; Xi, X.; Xu, F.; Li, B.; Cao, J.; Dawson, R. *Environ. Pollut.* **2002**, *116*, 57.
33. Faucon, J. C.; Bureau, R.; Faisant, J.; Briens, F.; Rault, F. *Chemosphere* **2001**, *44*, 407.