QSPR MODELING THE AQUEOUS SOLUBILITY OF POLYCHLORINATED BIPHENYLS BY OPTIMIZATION OF CORRELATION WEIGHTS OF LOCAL AND GLOBAL GRAPH INVARIANTS

E. A. Castro[1] *, A. A. Toropov[2], A. I. Nesterova[2], O. M. Nabiev[2]

[1]INIFTA , Departamento de Química, Facultad de Ciencias Exactas, Universidad, Nacional de La Plata, Diag. 113 y 64, Suc.4, C.C. 16, La Plata 1900, Argentina

[2]Uzbekistan Academy of Sciences Algorithm-Engineering Institute,
F. Khodjaev Street 25, 700125 Tashkent, Uzbekistan

Abstract: Aqueous solubilities of polychlorinated biphenyls have been correlated with topological molecular descriptors which are functions of local and global invariants of labeled hydrogen filled graphs. Morgan extended connectivity and nearest neighboring codes have been used as local graph invariants. The number of chlorine atoms in biphenyls has been employed as a global graph invariant. Present results show that taking into account correlation weights of global invariants gives quite reasonable improvement of statistical characteristics for the prediction of aqueuos solubilities of polychlorinated biphenyls.

---

* Corresponding author (castro@quimica.unlp.edu.ar / jubert@arnet.com.ar)

INTRODUCTION

Since the atmosphere is a significant pathway for the transport of organic pollutants, considerable efforts have been expended for the measurement of physicochemical properties that govern the movement of chemicals in the environment. Polychlorinated biphenyls (PCBs), a class of persistent organic chemicals, have attracted the attention of scientists in recent years because they are found at an appreciable concentration in the polar region, presumably as a result of long-range atmospheric transport **[1]**. Although the manufacture and use of PCBs have been banned since 1979 **[2]**, these persistent organic pollutants remain widely distributed in the environment due to their chemical stability. Among the environmental pollutants that may be able to disrupt the endocrine system of human and animals, PCBs have been particularly noteworthy **[3,4]**.

The ability of PCBs to mimic natural hormones may reflect a close relationship between the physicochemical properties encoded in the molecular structure of these compounds and the toxic responses they elicit in biological systems. Due to their remarkable insulating capacity and flame resistant nature, PCBs replaced combustible insulating fluids in capacitors and transformers and reduced the risk of fire in hospitals, schools, and factories. PCBs entered the environment as components of pesticides, plasticizers, and adhesives. The nonflammability and chemical stability of PCBs have contributed to the widespread environmental problems associated with these organohalogen compounds. The lipophilicity of these compounds is responsible for their accumulation in the food chain and the cause of adverse human health effects.

In light of the probable carcinogenic activity of these compounds **[5,6]** and their tendency to sorb and bioaccumulate in aquatic environment, the aqueous solubility of the PCBs has been measured by a variety of investigators **[7-11]**. Recently, Puri *et al* **[12]** have calculated aqueous solubilities of PCBs using Mobile Order and Disorder Theory **[7-8]** for a representative set of 61 molecules. They have obtained a quite good agreement between calculated and experimental values of solubilities of PCBs at 298.15 K (standard deviation = ±0.41 log units, Table 6 in Ref. 12) which demonstrates the utility and capability of CoMF-predicted values of fusion enthalpies to calculate the aqueous solubilities of any PCB.

Since there are other alternatives to predict molecular aqueous solubilities within the frame of the QSPR theory, we have deemed sensible to resort to look for ways to improve these predictions. An option is the approach based upon the correlation weights of local graph invariants **[13-16]**, which has proved to be a quite suitable tool to calculate thermodynamic properties for a wide variety of molecular species **[17-21]**.

The aim of the present study was to develop simple and predictive models that correlate aqueous solubility of PCBs with Morgan extended connectivity and nearest neighboring codes. This study shows that this particular set of molecular topological descriptors makes up a suitable option to predict this physicochemistry property with a greater accuracy than previous approaches, so that these models provide a numerical value that can be used in cases when experimental data are unavailable.

The paper is organized as follows: next section deals with the presentation of the method and the mathematical algorithm applied in this study. Then, we display the set of PCBs together with available experimental data and previous theoretical prediction of aqueous solubility plus the results derived from the present approach and they are discussed in a comparative fashion. Finally, we analyze the main conclusions derived from this study and point out some possible further extensions of this calculation method.

METHOD

The modeling of the aqueous solubility of PCBs under consideration has been based on the optimization of the correlation weights of graph invariants in the **L**abeled **H**ydrogen-**F**illed **G**raph (LHFG) and in the **G**raph of **A**tomic **O**rbitals (GAO) versions. Since the methodological principles and specific formulae have been presented elsewhere **[13-21]**, we do not consider necessary to introduce them here so that potentially interested readers can resort to the pertinent literature cited before.

The molecular descriptors are defined as

$$D(a,LI) = \sum_{k} CW(a_k) + \sum_{k} CW(LI_k) \qquad \textbf{(1)}$$

$$D(ao,LI) = \sum_{k} CW(ao_k) + \sum_{k} CW(LI_k) \qquad \textbf{(2)}$$

where a_k is a chemical element that is image of the k-th vertex in the LHFG; ao_k is the atomic orbital that is image of the k-th vertex in the GAO; LI_k is some numerical local invariant of LHFG or GAO. As local invariants (i.e.LIs) we have chosen the Morgan extended connectivity of zero (^{0}EC) and first (^{1}EC) order in the LHFG and also in the GAO nearest neighboring codes (NNC) in the LHFG.

The aqueous solubilities of 61 PCBs have been reported in the literature **[7,8]**. This molecular set has been chosen by Puri *et al* **[12]** to develop three-dimensional quantitative-structure-property relationship (3D-QSPR) models for prediction of enthalpies of fusion and their application to estimates of enthalpies of sublimation and aqueous solubilities, so that we have selected this set of PCBs to be able to make a direct comparison of our predictions with previous results.

We have chosen two calculation strategies to report results:

a) We have made calculations on the whole molecular set of PCBs (*i.e.* 61 molecules).
b) We have divided the complete molecular set into two partial sets: a training set (31 molecules) and a test set (30 molecules). The regression models were determined resorting to the training set and then true predictions were made for the molecules belonging to the test set.

Since in principle, the partition is arbitrary, we have tried with several choices in order to determine the dependence of final results on such partitions. However, we have found that final results are nearly independent of the chosen partition, so that we report results for a typical choice. Table 8 lists the composition of each partial sets.

RESULTS AND DISCUSSION

As it is usual in this sort of calculations, we have proved more than one numerical probe, so that in order to reach internal consistency we have tried three different probes to test some possible dependency on a given particular one.

In Table 1 we present results of OCWLI based on local graph invariants in the LHFGs and in the GAO. From the results shown Table 1 one can see that results are independent of the probes for each variable, so that them are internally consistent. Besides, statistical characteristics are practically the same in the five cases presented here. Most probably such situation is due to the great similarity among molecular structures under consideration. In other words, descriptors calculated with the CWs often have equal numerical values. Under such circumstances, the taking into account of global graph invariants becomes a reasonable concept of modeling for this molecular set. The optimization of correlation weights of the mentioned local invariants together with the number of chlorine atoms which are present in the LHFG of a PCB (denoted as N_{Cl}) and the number of $3p^5$ orbitals which are present in the GAO of a PCB (denoted as N_{3p5}) may be considered as one of the possible ways to define local and global optimization scheme in the QSPR modeling. In other words, the QSPR analysis of PCBs descriptors are calculated as

$$D(a,LI) = \left\{\sum_k CW(a_k) + \sum_k CW(LI_k)\right\} + CW(N_{Cl}) \qquad \textbf{(3)}$$

$$D(ao,LI) = \left\{\sum_k CW(ao_k) + \sum_k CW(LI_k)\right\} + CW(N_{3p5}) \qquad \textbf{(4)}$$

Results derived from the calculation with Eqs. (3)-(4) are presented in Table 2. One can see that statistical characteristics of models displayed in the Table 2 are better than those ones given in Table 1. Final results are also nearly independent of the chosen probe, as happens with data analyzed in Table 1.

Correlation weights for calculating D(a,LI) of Eqs. (3)-(4) are presented in Tables 3-7. The results derived from the calculation the aqueous solubility of the PCBs with the optimized fitting linear polynomials

$$\log S_B = 2.599\ D(a,{}^0EC) - 314.8 \qquad \textbf{(5)}$$
$$\log S_B = 1.870\ D(a,{}^1EC) - 127.5 \qquad \textbf{(6)}$$
$$\log S_B = 0.5388\ D(ao,{}^0EC) - 79.15 \qquad \textbf{(7)}$$
$$\log S_B = 0.3547\ D(ao,{}^1EC) - 78.96 \qquad \textbf{(8)}$$
$$\log S_B = 1.659\ D(a, NNC) - 97.58 \qquad \textbf{(9)}$$

are shown in Tables 8-12.

The analysis of data shows a quite satisfactory agreement among experimental and theoretical predictions of aqueous solubility. Particularly noticeable are the lower absolute average deviations for the test set, save the predictions derived from Eq. (11) (0.33 vs 0.40). In fact, although differences are not spectacular (i.e. 0.38 vs 0.35 (twice); 0.38 vs 0.33; and 0.39 vs 0.37) they are significant. Large deviations are scarce and they amount to around 16% (for example, molecules 14 and 15 in Tables 8, 9 and 12 (training set), molecule 12 in Table 11 (test set)). Once again we found a quite good predictive capability in the fitting equations since just in one case (molecule 12, Table 11) the deviation is rather large for a member of the test set. In order to judge suitably these findings we must take into account that results for test sets are true predictions and not the outcome of numerical fittings.

The comparison with previous results **[12]** for this molecular set shows the relative merits of the present approach. In fact, the average absolute deviation obtained by Puri *et al* was 0.32 (see Table 6 in Ref. 12) which is smaller than present results. However, in the precedent analysis has not resorted to true predictions as in our method, so that final results are biased since authors have not divided the total molecular set into a training one and a test set.

CONCLUSIONS

The optimization of correlation weights of local and global graph invariants in the LHFG and/or GAO approaches may be considered as a reasonable well fitted tool to predict the aqueous solubility of the PCBs for the molecular set under consideration in the present study. The relative deviations are bounded and true predictions are satisfactory. In fact, from a general viewpoint average absolute deviations are rather small and in particular just in only one case we have found a relatively large deviation for a member of the test set. We conclude that present approach based on the optimization of correlation weights of local and global graph invariants is a suitable way to predict aqueous solubility of PCBs. This finding is in line with our previous findings about this special kind of molecular descriptors. Global descriptors show to be better variables than local ones in order to calculate aqueous solubility of this particular set of PCBs.

REFERENCES

[1] F. Wania, D. Mackay, Global Fractionation and Cold Condensation of Low Volatility Organochlorine Compounds in Polar Regions, *AMBIO* **1993**, *22*, 10-18.

[2] Agency for Toxic Substances and Disease Registry (ATSDR). Toxicological Profile for Selected PCBs. U.S. Public Health Service. U.S. Department of Health and Human Services: Atlanta, GA, 1993.

[3] J. D. McKinney, C. L. Waller, Polychlorinated Biphenyls as Hormonally Active Structural Analogues, *Environ. Health Perspect.* **1994**, *102*, 290-297.

[4] A. Brouwer, M. P. Longnecker, L. S. Birnbaum, J. Cogliano, P. Kostyniak, J. Moore, S. Schantz, G. Winneke, Characterization of Potential Endocrine-Related Health Effects at Low-Dose Levels of Expopsure to PCBs, *Environ. Health Perspect.* **1999**, *107*, 639-649.

[5] U.S. Environmental Protection Agency. *Integrated Risk Information System (IRIS) on PCBs*, Environmental Criteria and Assesment Office. Office of Health and Environmental Assessment. Office of Research and Development, Cincinnati, OH, 1994.

[6] U.S. Environmental Protection Agency. *PCBs: Cancer Dose-Response Assessment and Application to Environmental Mixtures*, Report # EPA 600P-96002F, 1996.

[7] P. Ruelle, U. W. Kesserling, Aqueous Solubility Prediction of Environmentally Important Chemicals from the Mobile Order Thermodynamic, *Chemosphere* **1997**, *34*, 275-298.

[8] P. Ruelle, A. Cuendet, U. W. Kesselring, Hydrophobic and Solvation Effects on the Solubility of Hydroxysteroids in Various Solvents: Quantitative and Qualitative Assessment by Application of the Mobile Order and Disorder Theory, *Perspective Drug Discovery Design*, **2000**, *18*, 61-112, and references herein.

[9] W. J. Doucette, A. W. Andren, Estimation of Octanol/Water Partition Coefficients: Evaluation of Six Methods for Highly Hydrophobic Aromatic Hydrocarbons, *Chemosphere* **1988**, *17*, 345-359.

[10] R. Kuhne, R. –U. Ebert, F. Kleint, G. Schmidt, G. Schuurmann, Group Contribution Methods to Estimate Water Solubility of Organic Chemicals, *Chemosphere* **1995**, *30*, 2061-2077.

[11] T. T. Blair, E. Gifford, W. E. Acree, Jr., C. –C. Tsai, Quantitative Structure-Property Relationships for Aqueous Solubilities of Halogenated Aromatic Compounds, *Phys. Chem. Liq.* **1992**, *24*, 137-160.

[12] S. Puri, J.S. Chickos, W. J. Welsh Three-Dimensional Quantitative Structure – Property Relationship (3D-QSPR) Models for Prediction of Thermodynamic Properties of Polychlorinated Biphenyls (PCBs): Enthalpies of Fusion and Their Application to Estimates of Enthalpies of Sublimation and Aqueous Solubilities, *J. Chem. Inf. Comput. Sci.* **2003**, *43(1)*, 55-62.

[13] P. J. Peruzzo, D. J. G. Marino, E. A. Castro, A. A. Toropov, Calculation of pK Values of Flavylium Salts from the Optimization of Correlation Weights of Local Graph Invariants, *J. Mol. Struct. THEOCHEM* **2001,** *572*, 53-60.

[14] D. J. G. Marino, P. J. Peruzzo, E. A. Castro, A. A. Toropov, QSAR Carcinogenic Study of Methylated Polycyclic Aromatic Hydrocarbons Based on Topological Descriptors Derived from Distance Matrices and Correlation Weights of Local Graph Invariants, *Internet Electron. J. Mol. Des.* **2002**, *1*, 108-133, http://www.biochempress.com.

[15] A. A. Toropov and A. P. Toropova Optimization of Correlation Weights of the Local Graph Invariants: Use of the Enthalpies of Formation of Complex Compounds for the QSPR Modeling, *Russ. J. Coord. Chem.* **1998**, *24*, 81-85

[16] A. A. Toropov and A. P. Toropova, Modeling of Acyclic Compounds Normal Boiling Points by Correlation Weighting of Nearest Neighboring Codes, *J. Mol. Struct. THEOCHEM* **2002**, *581*, 11-15

[17] A. A. Toropov, O. M. Nabiev, P. R. Duchowicz, E. A. Castro, F. Torrens, QSPR Modeling of Hydrocarbon Dipole Moments by Means of Correlation Weighting of Local Graph Invariants, *J. Theor. Comp. Chem.* **2003**, *2(2)*, 139-146.

[18] D. J. G. Marino, P. J. Peruzzo, E. A. Castro, A. A. Toropov, QSPR Modeling of Lipophilicity by Means of Correlation Weights of Local Graph Invariants, *Internet Elect. J. Molec. Design* **2003**, *2*, 334-347.

[19] A. A. Toropov, P. R. Duchowicz, E. A. Castro, Structure-Toxicity Relationships for Aliphatic Compounds Based on Correlation Weighting of Local Graph Invariants, *Int. J. Mol. Sci.* **2003**, *4*, 272-283.

[20] P. R. Duchowicz, E. A. Castro, A. A. Toropov, QSPR Modeling of Normal Boiling Points of Aldehydes, Ketones and Esters by Means of Nearest Neighboring Codes Correlation Weighting, *J. Arg. Chem. Soc.* **2002**, *90(1-3)*, 91-107.

[21] P. R. Duchowicz, E. A. Castro, A. A. Toropov, Improved QSPR Analysis of Standard Entropy of Acyclic and Aromatic Compounds Using Optimized Correlation Weights of Linear Graph Invariants, *Comp. Chem.* **2002**, *26(4)*, 327-332.

Table 1. Statistical characteristics of PCB solubility models based on local graph invariants.

	Training Set (n = 31)			**Test Set (n = 30)**			**Complete set (n = 61)**		
Descriptor	R	s	F	R	s	F	R	s	F
Probe 1 D(a,^{0}EC)	0.9143	0.609	148	0.9623	0.404	351	0.9320	0.514	390
Probe 2 D(a,^{0}EC)	0.9143	0.609	148	0.9623	0.404	351	0.9320	0.514	390
Probe 3 D(a,^{0}EC)	0.9143	0.609	148	0.9623	0.404	351	0.9320	0.514	390
Probe 1 D(a,^{1}EC)	0.9143	0.609	148	0.9623	0.404	351	0.9320	0.514	390
Probe 2 D(a,^{1}EC)	0.9143	0.609	148	0.9623	0.404	351	0.9320	0.514	390
Probe 3 D(a,^{1}EC)	0.9143	0.609	148	0.9623	0.404	351	0.9320	0.514	390
Probe 1 D(ao,^{0}EC)	0.9140	0.609	148	0.9620	0.404	351	0.9320	0.514	390
Probe 2 D(ao,^{0}EC)	0.9140	0.609	148	0.9620	0.404	351	0.9320	0.514	390
Probe 3 D(ao,^{0}EC)	0.9140	0.609	148	0.9620	0.404	351	0.9320	0.514	390
Probe 1 D(ao,^{1}EC)	0.9301	0.552	186	0.9328	0.477	187	0.9306	0.512	382
Probe 2 D(ao,^{1}EC)	0.9301	0.552	186	0.9330	0.477	188	0.9307	0.512	382
Probe 3 D(ao,^{1}EC)	0.9301	0.552	186	0.9331	0.476	189	0.9308	0.512	383
Probe 1 D(a,NNC)	0.9143	0.609	148	0.9623	0.404	351	0.9320	0.514	390
Probe 2 D(a,NNC)	0.9143	0.609	148	0.9623	0.404	351	0.9320	0.514	390
Probe 3 D(a,NNC)	0.9143	0.609	148	0.9623	0.404	351	0.9320	0.514	390

Table 2. Statistical characteristics of PCB solubility models based on local and global graph invariants.

	Training Set (n = 31)			**Test Set (n = 30)**			**Complete set (n = 61)**		
Descriptor	R	s	F	R	S	F	R	s	F
Probe 1 D(a,^{0}EC)	0.9372	0.525	209	0.9620	0.437	347	0.9429	0.479	473
Probe 2 D(a,^{0}EC)	0.9371	0.525	209	0.9626	0.437	353	0.9430	0.479	474
Probe 3 D(a,^{0}EC)	0.9372	0.525	209	0.9627	0.436	354	0.9431	0.479	474
Probe 1 D(a,^{1}EC)	0.9372	0.524	209	0.9629	0.434	356	0.9432	0.478	476
Probe 2 D(a,^{1}EC)	0.9372	0.525	209	0.9619	0.434	347	0.9430	0.478	474
Probe 3 D(a,^{1}EC)	0.9372	0.525	209	0.9631	0.434	358	0.9432	0.478	476
Probe 1 D(ao,^{0}EC)	0.9372	0.524	209	0.9627	0.434	354	0.9432	0.478	475
Probe 2 D(ao,^{0}EC)	0.9372	0.524	209	0.9625	0.435	353	0.9431	0.479	474
Probe 3 D(ao,^{0}EC)	0.9372	0.524	209	0.9626	0.435	354	0.9431	0.478	475
Probe 1 D(ao,^{1}EC)	0.9533	0.454	289	0.9141	0.524	142	0.9374	0.486	428
Probe 2 D(ao,^{1}EC)	0.9534	0.454	290	0.9140	0.524	142	0.9375	0.485	428
Probe 3 D(ao,^{1}EC)	0.9536	0.453	291	0.9136	0.524	141	0.9375	0.485	428
Probe 1 D(a,NNC)	0.9372	0.524	209	0.9627	0.433	355	0.9433	0.478	476
Probe 2 D(a,NNC)	0.9371	0.525	209	0.9621	0.432	349	0.9432	0.477	475
Probe 3 D(a,NNC)	0.9371	0.525	209	0.9619	0.435	347	0.9430	0.479	473

Table 3. Correlation weights for the calculation of D(a,^{0}EC)

	Probe 1	**Probe 2**	**Probe 3**

Correlation weights of the a_k values

H	1.140	1.037	1.140
C	3.583	1.517	1.191
Cl	0.898	0.889	0.934

Correlation weights of the ^{0}EC values

0001	2.986	1.505	1.235
0003	2.837	1.948	1.384

Correlation weights of the N_{Cl} values

H000	1.173	1.287	1.221
H001	1.070	1.199	1.113
H002	1.002	1.128	1.045
H003	1.080	1.150	1.099
H004	0.963	1.046	0.988
H005	1.210	1.200	1.198
H006	1.128	1.113	1.105
H007	1.287	1.195	1.230
H008	1.244	1.150	1.188
H009	1.002	0.950	0.965
H010	0.777	0.769	0.750

Table 4. Correlation weights for the calculation of $D(a,{}^{1}EC)$.

	Probe 1	**Probe 2**	**Probe 3**

Correlation weights of the a_k values

H	1.353	1.100	1.200
C	1.517	1.026	0.866
Cl	1.015	0.923	0.834

Correlation weights of the ${}^{1}EC$ values

0003	1.655	1.128	1.339
0007	1.315	1.200	1.298
0009	1.635	0.900	1.187

Correlation weights of the N_{Cl} values

H000	1.187	1.098	1.304
H001	1.040	1.019	1.150
H002	0.950	0.963	1.052
H003	1.058	1.019	1.156
H004	0.909	0.923	0.998
H005	1.244	1.100	1.353
H006	1.126	1.037	1.240
H007	1.344	1.141	1.458
H008	1.295	1.120	1.421
H009	0.968	0.923	1.049
H010	0.656	0.750	0.710

Table 5. Correlation weights for the calculation of D(ao,^{0}EC).

	Probe 1	**Probe 2**	**Probe 3**

Correlation weights of the ao_k values

$1s^1$	2.022	2.048	2.457
$1s^2$	0.616	0.650	0.834
$2p^2$	1.562	1.307	2.416
$2s^2$	0.649	0.720	0.569
$2p^6$	0.686	0.804	0.615
$3s^2$	0.723	0.703	0.689
$3p^5$	0.620	0.745	0.554

Correlation weights of the ^{0}EC values

0003	0.656	0.673	0.590
0007	2.281	2.488	2.125
0009	1.380	2.092	1.501
0011	0.568	0.731	0.598

Correlation weights of the N_{3p5} values

H000	1.497	1.517	1.615
H001	1.102	1.149	1.163
H002	0.836	0.920	0.860
H003	1.278	1.262	1.360
H004	0.818	0.857	0.849
H005	2.053	1.863	2.252
H006	1.736	1.588	1.912
H007	2.548	2.238	2.838
H008	2.454	2.139	2.737
H009	1.372	1.230	1.497
H010	0.368	0.363	0.368

Table 6. Correlation weights for the calculation of D(ao,^{1}EC).

	Probe 1	**Probe 2**	**Probe 3**

Correlation weights of the ao_k values

$1s^1$	3.038	2.194	3.452
$1s^2$	1.170	1.192	0.600
$2p^2$	4.861	7.942	15.895
$2s^2$	0.781	1.073	1.171
$2p^6$	0.773	0.839	0.714
$3s^2$	0.711	0.703	0.692
$3p^5$	0.813	0.579	0.885

Correlation weights of the ^{1}EC values

0021	2.790	2.828	2.757
0033	0.604	0.651	0.769
0045	2.372	2.893	2.561
0051	1.596	1.945	1.755
0057	1.893	2.188	1.944
0063	1.302	1.467	1.381
0069	1.360	1.410	1.275
0075	0.946	0.909	0.912
0081	1.069	0.951	0.930
0093	0.804	0.530	0.617

Correlation weights of the N_{3p5} values

H000	0.459	0.318	0.424
H001	0.812	0.723	1.001
H002	0.785	0.750	1.032
H003	1.469	1.696	1.976
H004	1.349	1.561	1.836
H005	2.940	3.660	3.834
H006	3.006	3.823	4.008
H007	5.072	6.495	6.688
H008	4.633	5.961	6.215
H009	2.246	2.880	3.039
H010	0.068	0.018	0.028

Table 7. Correlation weights for the calculation of D(a,NNC).

	Probe 1	**Probe 2**	**Probe 3**

Correlation weights of the a_k values

H	1.440	1.230	0.975
C	1.076	0.974	1.113
Cl	0.953	1.128	1.100

Correlation weights of the NNC values

0110	1.353	0.926	0.929
0320	1.188	1.025	0.988
0321	1.067	1.051	1.200
0330	1.663	0.760	1.051

Correlation weights of the N_{Cl} values

H000	1.368	0.963	0.997
H001	1.197	0.926	0.974
H002	1.073	0.900	0.950
H003	1.180	0.950	0.975
H004	0.996	0.900	0.938
H005	1.351	1.025	1.025
H006	1.218	0.999	1.000
H007	1.443	1.073	1.050
H008	1.379	1.071	1.037
H009	0.980	0.964	0.950
H010	0.623	0.855	0.878

Table 8. Modeling of the PCBs solubility with Eq. (5) based on D(a,^{0}EC).

Training Set

n	Molecule	D(a,^{0}EC)	log S_B	Eq. (5)	Delta
1	Biphenyl	119.473	-4.31	-4.29	-0.02
2	4-monochlorobiphenyl	119.128	-5.20	-5.19	-0.01
3	2,4'-dichlorobiphenyl	118.818	-5.28	-5.99	0.71
4	3,3'-dichlorobiphenyl	118.818	-5.80	-5.99	0.19
5	3,4-dichlorobiphenyl	118.818	-6.39	-5.99	-0.40
6	4,4'-dichlorobiphenyl	118.818	-6.56	-5.99	-0.57
7	2,2',5-trichlorobiphenyl	118.654	-6.02	-6.42	0.40
8	2,4',5-trichlorobiphenyl	118.654	-6.25	-6.42	0.17
9	3,4,4'-trichlorobiphenyl	118.654	-7.06	-6.42	-0.64
10	2,2',3,5'-tetrachlorobiphenyl	118.295	-6.47	-7.35	0.88
11	2,2',5,5'-tetrachlorobiphenyl	118.295	-7.00	-7.35	0.35
12	2,3',4,4'-tetrachlorobiphenyl	118.295	-6.68	-7.35	0.67
13	2,4,4',6-tetrachlorobiphenyl	118.295	-6.94	-7.35	0.41
14	3,3',4,4'-tetrachlorobiphenyl	118.295	-8.53	-7.35	-1.18
15	3,3',5,5'-tetrachlorobiphenyl	118.295	-8.54	-7.35	-1.19
16	2,2',3,3',5-pentachlorobiphenyl	118.300	-6.96	-7.34	0.38
17	2,2',3,4,5-pentachlorobiphenyl	118.300	-7.21	-7.34	0.13
18	2,2',3,4,5'-pentachlorobiphenyl	118.300	-7.91	-7.34	-0.57
19	2,2',3,4,6-pentachlorobiphenyl	118.300	-7.43	-7.34	-0.09
20	2,2',4,5,5'-pentachlorobiphenyl	118.300	-7.33	-7.34	0.01
21	2,2',4,6,6'-pentachlorobiphenyl	118.300	-7.32	-7.34	0.02
22	2,3',4,4',5-pentachlorobiphenyl	118.300	-7.39	-7.34	-0.05
23	2,2',3,3',4,4'-hexachlorobiphenyl	117.976	-9.01	-8.18	-0.83
24	2,2',3,3',4,5-hexachlorobiphenyl	117.976	-8.07	-8.18	0.11
25	2,2',3,4,4',5'-hexachlorobiphenyl	117.976	-8.32	-8.18	-0.14
26	2,2',3,5,5',6-hexachlorobiphenyl	117.976	-7.42	-8.18	0.76
27	2,2',3,4,4',5',6-heptachlorobiphenyl	117.893	-7.92	-8.40	0.48
28	2,2',3,4',5,5',6-heptachlorobiphenyl	117.893	-8.94	-8.40	-0.54
29	2,2',3,3',5,5',6,6'-octachlorobiphenyl	117.608	-9.15	-9.14	-0.01
30	2,2',3,3',4,5,5',6,6'-nonachlorobiphenyl	117.124	-10.41	-10.40	-0.01
31	decachlorobiphenyl	116.657	-11.62	-11.61	-0.01

Average absolute deviation = 0.38

Test Set

n	Molecule	D(a,^{0}EC)	log S_B	Eq. (5)	Delta
1	2-monochlorobiphenyl	119.128	-4.54	-5.19	0.65
2	2,2'-dichlorobiphenyl	118.818	-5.27	-5.99	0.72
3	2,6-dichlorobiphenyl	118.818	-5.21	-5.99	0.78
4	2,3,4'-trichlorobiphenyl	118.654	-6.26	-6.42	0.16
5	2,3,6-trichlorobiphenyl	118.654	-6.29	-6.42	0.13
6	2,3',5-trichlorobiphenyl	118.654	-6.01	-6.42	0.41
7	2,4,4'-trichlorobiphenyl	118.654	-6.21	-6.42	0.21
8	2,4,5-trichlorobiphenyl	118.654	-6.27	-6.42	0.15
9	2,4,6-trichlorobiphenyl	118.654	-6.14	-6.42	0.28
10	2',3,4-trichlorobiphenyl	118.654	-6.29	-6.42	0.13
11	2,2',3,3'-tetrachlorobiphenyl	118.295	-7.28	-7.35	0.07
12	2,2',4,4'-tetrachlorobiphenyl	118.295	-6.51	-7.35	0.84

13	2,2',4,5'-tetrachlorobiphenyl	118.295	-6.57	-7.35	0.78
14	2,2',5,6'-tetrachlorobiphenyl	118.295	-7.08	-7.35	0.27
15	2,2',6,6'-tetrachlorobiphenyl	118.295	-7.21	-7.35	0.14
16	2,3,4,5-tetrachlorobiphenyl	118.295	-7.16	-7.35	0.19
17	2,3',4',5-tetrachlorobiphenyl	118.295	-7.25	-7.35	0.10
18	2,2',3,3',4-pentachlorobipheny	118.300	-7.05	-7.34	0.29
19	2,3,4,5,6-pentachlorobiphenyl	118.300	-7.92	-7.34	-0.58
20	2,2',3,3',5,6-hexachlorobiphenyl	117.976	-8.60	-8.18	-0.42
21	2,2',3,3',6,6'-hexachlorobiphenyl	117.976	-8.65	-8.18	-0.47
22	2,2',3,4,5,5'-hexachlorobiphenyl	117.976	-7.68	-8.18	0.50
23	2,2',4,4',5,5'-hexachlorobiphenyl	117.976	-8.56	-8.18	-0.38
24	2,2',4,4',6,6'-hexachlorobiphenyl	117.976	-8.71	-8.18	-0.53
25	2,3,3',4,4',5-hexachlorobiphenyl	117.976	-7.82	-8.18	0.36
26	2,3,3',4,4',6-hexachlorobiphenyl	117.976	-7.66	-8.18	0.52
27	2,2',3,3',4,4',6-heptachlorobiphenyl	117.893	-8.30	-8.40	0.10
28	2,2',3,4,5,5',6-heptachlorobiphenyl	117.893	-8.46	-8.40	-0.06
29	2,2',3,3',4,4',5,5'-octachlorobiphenyl	117.608	-9.16	-9.14	-0.02
30	2,2',3,3',4,4',5,5',6-nonachlorobiphenyl	117.124	-10.26	-10.40	0.14

Average absolute deviation = 0.35

Table 9. Modeling of the PCBs solubility with Eq. (6) based on D(a,^{1}EC).

Training Set

n	**Molecule**	**D(a,^{1}EC)**	**log S_B**	**Eq. (6)**	**Delta**
1	Biphenyl	65.891	-4.31	-4.28	-0.03
2	4-monochlorobiphenyl	65.406	-5.20	-5.19	-0.01
3	2,4'-dichlorobiphenyl	64.978	-5.28	-5.99	0.71
4	3,3'-dichlorobiphenyl	64.978	-5.80	-5.99	0.19
5	3,4-dichlorobiphenyl	64.978	-6.39	-5.99	-0.40
6	4,4'-dichlorobiphenyl	64.978	-6.56	-5.99	-0.43
7	2,2',5-trichlorobiphenyl	64.748	-6.02	-6.42	0.40
8	2,4',5-trichlorobiphenyl	64.748	-6.25	-6.42	0.17
9	3,4,4'-trichlorobiphenyl	64.748	-7.06	-6.42	-0.64
10	2,2',3,5'-tetrachlorobiphenyl	64.261	-6.47	-7.33	0.86
11	2,2',5,5'-tetrachlorobiphenyl	64.261	-7.00	-7.33	0.33
12	2,3',4,4'-tetrachlorobiphenyl	64.261	-6.68	-7.33	0.65
13	2,4,4',6-tetrachlorobiphenyl	64.261	-6.94	-7.33	0.39
14	3,3',4,4'-tetrachlorobiphenyl	64.261	-8.53	-7.33	-1.20
15	3,3',5,5'-tetrachlorobiphenyl	64.261	-8.54	-7.33	-1.21
16	2,2',3,3',5-pentachlorobiphenyl	64.258	-6.96	-7.34	0.38
17	2,2',3,4,5-pentachlorobiphenyl	64.258	-7.21	-7.34	0.13
18	2,2',3,4,5'-pentachlorobiphenyl	64.258	-7.91	-7.34	-0.57
19	2,2',3,4,6-pentachlorobiphenyl	64.258	-7.43	-7.34	-0.09
20	2,2',4,5,5'-pentachlorobiphenyl	64.258	-7.33	-7.34	0.01
21	2,2',4,6,6'-pentachlorobiphenyl	64.258	-7.32	-7.34	0.02
22	2,3',4,4',5-pentachlorobiphenyl	64.258	-7.39	-7.34	-0.05
23	2,2',3,3',4,4'-hexachlorobiphenyl	63.802	-9.01	-8.19	-0.82
24	2,2',3,3',4,5-hexachlorobiphenyl	63.802	-8.07	-8.19	0.12
25	2,2',3,4,4',5'-hexachlorobiphenyl	63.802	-8.32	-8.19	-0.13
26	2,2',3,5,5',6-hexachlorobiphenyl	63.802	-7.42	-8.19	0.77
27	2,2',3,4,4',5',6-heptachlorobiphenyl	63.682	-7.92	-8.42	0.50
28	2,2',3,4',5,5',6-heptachlorobiphenyl	63.682	-8.94	-8.42	-0.52
29	2,2',3,3',5,5',6,6'-octachlorobiphenyl	63.295	-9.15	-9.14	-0.01
30	2,2',3,3',4,5,5',6,6'-nonachlorobiphenyl	62.630	-10.41	-10.38	-0.03
31	Decachlorobiphenyl	61.980	-11.62	-11.60	-0.02

Average absolute deviation = 0.38

Test Set

n	**Molecule**	**D(a,^{1}EC)**	**log S_B**	**Eq. (6)**	**Delta**
1	2-monochlorobiphenyl	65.406	-4.54	-5.19	0.65
2	2,2'-dichlorobiphenyl	64.978	-5.27	-5.99	0.72
3	2,6-dichlorobiphenyl	64.978	-5.21	-5.99	0.78
4	2,3,4'-trichlorobiphenyl	64.748	-6.26	-6.42	0.16
5	2,3,6-trichlorobiphenyl	64.748	-6.29	-6.42	0.13
6	2,3',5-trichlorobiphenyl	64.748	-6.01	-6.42	0.41
7	2,4,4'-trichlorobiphenyl	64.748	-6.21	-6.42	0.21
8	2,4,5-trichlorobiphenyl	64.748	-6.27	-6.42	0.15
9	2,4,6-trichlorobiphenyl	64.748	-6.14	-6.42	0.28
10	2',3,4-trichlorobiphenyl	64.748	-6.29	-6.42	0.13
11	2,2',3,3'-tetrachlorobiphenyl	64.261	-7.28	-7.33	0.05

12	2,2',4,4'-tetrachlorobiphenyl	64.261	-6.51	-7.33	0.82
13	2,2',4,5'-tetrachlorobiphenyl	64.261	-6.57	-7.33	0.76
14	2,2',5,6'-tetrachlorobiphenyl	64.261	-7.08	-7.33	0.25
15	2,2',6,6'-tetrachlorobiphenyl	64.261	-7.21	-7.33	0.12
16	2,3,4,5-tetrachlorobiphenyl	64.261	-7.16	-7.33	0.17
17	2,3',4',5-tetrachlorobiphenyl	64.261	-7.25	-7.34	0.09
18	2,2',3,3',4-pentachlorobiphenyl	64.258	-7.05	-7.34	0.29
19	2,3,4,5,6-pentachlorobiphenyl	64.258	-7.92	-7.34	-0.58
20	2,2',3,3',5,6-hexachlorobiphenyl	63.802	-8.60	-8.19	-0.41
21	2,2',3,3',6,6'-hexachlorobiphenyl	63.802	-8.65	-8.19	-0.46
22	2,2',3,4,5,5'-hexachlorobiphenyl	63.802	-7.68	-8.19	0.51
23	2,2',4,4',5,5'-hexachlorobiphenyl	63.802	-8.56	-8.19	-0.37
24	2,2',4,4',6,6'-hexachlorobiphenyl	63.802	-8.71	-8.19	-0.52
25	2,3,3',4,4',5-hexachlorobiphenyl	63.802	-7.82	-8.19	0.37
26	2,3,3',4,4',6-hexachlorobiphenyl	63.802	-7.66	-8.19	0.53
27	2,2',3,3',4,4',6-heptachlorobiphenyl	63.682	-8.30	-8.42	0.12
28	2,2',3,4,5,5',6-heptachlorobiphenyl	63.682	-8.46	-8.42	-0.04
29	2,2',3,3',4,4',5,5'-octachlorobiphenyl	63.295	-9.16	-9.14	-0.02
30	2,2',3,3',4,4',5,5',6-nonachlorobiphenyl	62.630	-10.26	-10.38	0.12

Average absolute deviation = 0.33

Table 10. Modeling of the PCBs solubility with Eq. (7) based on D(ao,^{0}EC).

Training Set

n	Molecule	D(ao,^{0}EC)	log S_B	Eq. (7)	Delta
1	Biphenyl	138.911	-4.31	-4.31	0.00
2	4-monochlorobiphenyl	137.273	-5.20	-5.19	-0.01
3	2,4'-dichlorobiphenyl	135.764	-5.28	-6.00	0.72
4	3,3'-dichlorobiphenyl	135.764	-5.80	-6.00	0.20
5	3,4-dichlorobiphenyl	135.764	-6.39	-6.00	-0.39
6	4,4'-dichlorobiphenyl	135.764	-6.56	-6.00	-0.56
7	2,2',5-trichlorobiphenyl	134.963	-6.02	-6.43	0.41
8	2,4',5-trichlorobiphenyl	134.963	-6.25	-6.43	0.18
9	3,4,4'-trichlorobiphenyl	134.963	-7.06	-6.43	-0.63
10	2,2',3,5'-tetrachlorobiphenyl	133.260	-6.47	-7.35	0.88
11	2,2',5,5'-tetrachlorobiphenyl	133.260	-7.00	-7.35	0.35
12	2,3',4,4'-tetrachlorobiphenyl	133.260	-6.68	-7.35	0.67
13	2,4,4',6-tetrachlorobiphenyl	133.260	-6.94	-7.35	0.41
14	3,3',4,4'-tetrachlorobiphenyl	133.260	-8.53	-7.35	-1.18
15	3,3',5,5'-tetrachlorobiphenyl	133.260	-8.54	-7.35	-1.19
16	2,2',3,3',5-pentachlorobiphenyl	133.252	-6.96	-7.35	0.39
17	2,2',3,4,5-pentachlorobiphenyl	133.252	-7.21	-7.35	0.14
18	2,2',3,4,5'-pentachlorobiphenyl	133.252	-7.91	-7.35	-0.56
19	2,2',3,4,6-pentachlorobiphenyl	133.252	-7.43	-7.35	-0.08
20	2,2',4,5,5'-pentachlorobiphenyl	133.252	-7.33	-7.35	0.02
21	2,2',4,6,6'-pentachlorobiphenyl	133.252	-7.32	-7.35	0.03
22	2,3',4,4',5-pentachlorobiphenyl	133.252	-7.39	-7.35	-0.04
23	2,2',3,3',4,4'-hexachlorobiphenyl	131.692	-9.01	-8.19	-0.82
24	2,2',3,3',4,5-hexachlorobiphenyl	131.692	-8.07	-8.19	0.12
25	2,2',3,4,4',5'-hexachlorobiphenyl	131.692	-8.32	-8.19	-0.13
26	2,2',3,5,5',6-hexachlorobiphenyl	131.692	-7.42	-8.19	0.77
27	2,2',3,4,4',5',6-heptachlorobiphenyl	131.261	-7.92	-8.43	0.51
28	2,2',3,4',5,5',6-heptachlorobiphenyl	131.261	-8.94	-8.43	-0.51
29	2,2',3,3',5,5',6,6'-octachlorobiphenyl	129.924	-9.15	-9.15	0.00
30	2,2',3,3',4,5,5',6,6'-nonachlorobiphenyl	127.599	-10.41	-10.40	-0.01
31	Decachlorobiphenyl	125.352	-11.62	-11.61	-0.01

Average absolute deviation = 0.38

Test Set

n	Molecule	D(ao,^{0}EC)	log S_b	Eq.(7)	Delta
1	2-monochlorobiphenyl	137.273	-4.54	-5.19	0.65
2	2,2'-dichlorobiphenyl	135.764	-5.27	-6.00	0.73
3	2,6-dichlorobiphenyl	135.764	-5.21	-6.00	0.79
4	2,3,4'-trichlorobiphenyl	134.963	-6.26	-6.43	0.17
5	2,3,6-trichlorobiphenyl	134.963	-6.29	-6.43	0.14
6	2,3',5-trichlorobiphenyl	134.963	-6.01	-6.43	0.42
7	2,4,4'-trichlorobiphenyl	134.963	-6.21	-6.43	0.22
8	2,4,5-trichlorobiphenyl	134.963	-6.27	-6.43	0.16
9	2,4,6-trichlorobiphenyl	134.963	-6.14	-6.43	0.29
10	2',3,4-trichlorobiphenyl	134.963	-6.29	-6.43	0.14
11	2,2',3,3'-tetrachlorobiphenyl	133.260	-7.28	-7.35	0.07

12	2,2',4,4'-tetrachlorobiphenyl	133.260	-6.51	-7.35	0.84
13	2,2',4,5'-tetrachlorobiphenyl	133.260	-6.57	-7.35	0.78
14	2,2',5,6'-tetrachlorobiphenyl	133.260	-7.08	-7.35	0.27
15	2,2',6,6'-tetrachlorobiphenyl	133.260	-7.21	-7.35	0.14
16	2,3,4,5-tetrachlorobiphenyl	133.260	-7.16	-7.35	0.19
17	2,3',4',5-tetrachlorobiphenyl	133.260	-7.25	-7.35	0.10
18	2,2',3,3',4-pentachlorobiphenyl	133.252	-7.05	-7.35	0.30
19	2,3,4,5,6-pentachlorobiphenyl	133.252	-7.92	-7.35	-0.57
20	2,2',3,3',5,6-hexachlorobiphenyl	131.692	-8.60	-8.19	-0.41
21	2,2',3,3',6,6'-hexachlorobiphenyl	131.692	-8.65	-8.19	-0.46
22	2,2',3,4,5,5'-hexachlorobiphenyl	131.692	-7.68	-8.19	0.51
23	2,2',4,4',5,5'-hexachlorobiphenyl	131.692	-8.56	-8.19	-0.37
24	2,2',4,4',6,6'-hexachlorobiphenyl	131.692	-8.71	-8.19	-0.52
25	2,3,3',4,4',5-hexachlorobiphenyl	131.692	-7.82	-8.19	0.37
26	2,3,3',4,4',6-hexachlorobiphenyl	131.692	-7.66	-8.19	0.53
27	2,2',3,3',4,4',6-heptachlorobiphenyl	131.261	-8.30	-8.43	0.13
28	2,2',3,4,5,5',6-heptachlorobiphenyl	131.261	-8.46	-8.43	-0.03
29	2,2',3,3',4,4',5,5'-octachlorobiphenyl	129.924	-9.16	-9.15	-0.01
30	2,2',3,3',4,4',5,5',6-nonachlorobiphenyl	127.599	-10.26	-10.40	0.14

Average absolute deviation = 0.35

Table 11. Modeling of the PCBs solubility with Eq. (8) based on D(ao,^{1}EC).

Training Set

n	**Molecule**	**D(ao,^{1}EC)**	**log S_B**	**Eq.(8)**	**Delta**
1	Biphenyl	210.491	-4.31	-4.30	-0.01
2	4-monochlorobiphenyl	207.973	-5.20	-5.19	-0.01
3	2,4'-dichlorobiphenyl	206.194	-5.28	-5.82	0.54
4	3,3'-dichlorobiphenyl	206.185	-5.80	-5.83	0.03
5	3,4-dichlorobiphenyl	205.306	-6.39	-6.14	-0.25
6	4,4'-dichlorobiphenyl	205.075	-6.56	-6.22	-0.34
7	2,2',5-trichlorobiphenyl	205.681	-6.02	-6.01	-0.01
8	2,4',5-trichlorobiphenyl	204.562	-6.25	-6.40	0.15
9	3,4,4'-trichlorobiphenyl	203.119	-7.06	-6.91	-0.15
10	2,2',3,5'-tetrachlorobiphenyl	202.897	-6.47	-6.99	0.52
11	2,2',5,5'-tetrachlorobiphenyl	203.245	-7.00	-6.87	-0.13
12	2,3',4,4'-tetrachlorobiphenyl	201.085	-6.68	-7.64	0.96
13	2,4,4',6-tetrachlorobiphenyl	201.889	-6.94	-7.35	0.41
14	3,3',4,4'-tetrachlorobiphenyl	200.359	-8.53	-7.89	-0.64
15	3,3',5,5'-tetrachlorobiphenyl	201.793	-8.54	-7.38	-1.16
16	2,2',3,3',5-pentachlorobiphenyl	201.662	-6.96	-7.43	0.47
17	2,2',3,4,5-pentachlorobiphenyl	202.421	-7.21	-7.16	-0.05
18	2,2',3,4,5'-pentachlorobiphenyl	202.019	-7.91	-7.30	-0.61
19	2,2',3,4,6-pentachlorobiphenyl	202.499	-7.43	-7.13	-0.30
20	2,2',4,5,5'-pentachlorobiphenyl	201.479	-7.33	-7.50	0.17
21	2,2',4,6,6'-pentachlorobiphenyl	202.925	-7.32	-6.98	-0.34
22	2,3',4,4',5-pentachlorobiphenyl	200.036	-7.39	-8.01	0.62
23	2,2',3,3',4,4'-hexachlorobiphenyl	199.268	-9.01	-8.28	-0.73
24	2,2',3,3',4,5-hexachlorobiphenyl	199.823	-8.07	-8.08	0.01
25	2,2',3,4,4',5'-hexachlorobiphenyl	198.728	-8.32	-8.47	0.15
26	2,2',3,5,5',6-hexachlorobiphenyl	200.054	-7.42	-8.00	0.58
27	2,2',3,4,4',5',6-heptachlorobiphenyl	198.958	-7.92	-8.39	0.47
28	2,2',3,4',5,5',6-heptachlorobiphenyl	198.763	-8.94	-8.46	-0.48
29	2,2',3,3',5,5',6,6'-octachlorobiphenyl	196.833	-9.15	-9.14	-0.01
30	2,2',3,3',4,5,5',6,6'-nonachlorobiphenyl	193.267	-10.41	-10.41	0.00
31	Decachlorobiphenyl	189.910	-11.62	-11.60	-0.02

Average absolute deviation = 0.33

Test Set

n	**Molecule**	**D(ao,^{1}EC)**	**log S_B**	**Eq.(8)**	**Delta**
1	2-monochlorobiphenyl	209.092	-4.54	-4.80	0.26
2	2,2'-dichlorobiphenyl	207.313	-5.27	-5.43	0.16
3	2,6-dichlorobiphenyl	207.391	-5.21	-5.40	0.19
4	2,3,4'-trichlorobiphenyl	204.214	-6.26	-6.53	0.27
5	2,3,6-trichlorobiphenyl	205.411	-6.29	-6.10	-0.19
6	2,3',5-trichlorobiphenyl	205.117	-6.01	-6.21	0.20
7	2,4,4'-trichlorobiphenyl	203.845	-6.21	-6.66	0.45
8	2,4,5-trichlorobiphenyl	204.076	-6.27	-6.57	0.30
9	2,4,6-trichlorobiphenyl	204.880	-6.14	-6.29	0.15
10	2',3,4-trichlorobiphenyl	204.238	-6.29	-6.52	0.23
11	2,2',3,3'-tetrachlorobiphenyl	202.549	-7.28	-7.12	-0.16

12	2,2',4,4'-tetrachlorobiphenyl	200.359	-6.51	-7.89	1.38
13	2,2',4,5'-tetrachlorobiphenyl	202.528	-6.57	-7.12	0.55
14	2,2',5,6'-tetrachlorobiphenyl	203.887	-7.08	-6.64	-0.44
15	2,2',6,6'-tetrachlorobiphenyl	204.529	-7.21	-6.41	-0.80
16	2,3,4,5-tetrachlorobiphenyl	202.582	-7.16	-7.10	-0.06
17	2,3',4',5-tetrachlorobiphenyl	201.802	-7.25	-7.38	0.13
18	2,2',3,3',4-pentachlorobiphenyl	201.671	-7.05	-7.43	0.38
19	2,3,4,5,6-pentachlorobiphenyl	202.877	-7.92	-7.00	-0.92
20	2,2',3,3',5,6-hexachlorobiphenyl	199.706	-8.60	-8.12	-0.48
21	2,2',3,3',6,6'-hexachlorobiphenyl	200.858	-8.65	-7.72	-0.93
22	2,2',3,4,5,5'-hexachlorobiphenyl	200.171	-7.68	-7.96	0.28
23	2,2',4,4',5,5'-hexachlorobiphenyl	198.188	-8.56	-8.66	0.10
24	2,2',4,4',6,6'-hexachlorobiphenyl	199.796	-8.71	-8.09	-0.62
25	2,3,3',4,4',5-hexachlorobiphenyl	198.728	-7.82	-8.47	0.65
26	2,3,3',4,4',6-hexachlorobiphenyl	198.806	-7.66	-8.44	0.78
27	2,2',3,3',4,4',6-heptachlorobiphenyl	199.498	-8.30	-8.20	-0.10
28	2,2',3,4,5,5',6-heptachlorobiphenyl	200.941	-8.46	-7.69	-0.77
29	2,2',3,3',4,4',5,5'-octachlorobiphenyl	197.067	-9.16	-9.06	-0.10
30	2,2',3,3',4,4',5,5',6-nonachlorobiphenyl	193.384	-10.26	-10.37	0.11

Average absolute deviation = 0.40

Table 12. Model of the PCBs solubility with Eq. (9) based on D(a,NNC).

Training Set

n	Molecule	D(a,NNC)	log S_B	Eq.(9)	Delta
1	Biphenyl	56.206	-4.31	-4.33	0.02
2	4-monochlorobiphenyl	55.669	-5.20	-5.23	0.03
3	2,4'-dichlorobiphenyl	55.179	-5.28	-6.04	0.76
4	3,3'-dichlorobiphenyl	55.179	-5.80	-6.04	0.24
5	3,4-dichlorobiphenyl	55.179	-6.39	-6.04	-0.35
6	4,4'-dichlorobiphenyl	55.179	-6.56	-6.04	-0.52
7	2,2',5-trichlorobiphenyl	54.920	-6.02	-6.47	0.45
8	2,4',5-trichlorobiphenyl	54.920	-6.25	-6.47	0.22
9	3,4,4'-trichlorobiphenyl	54.920	-7.06	-6.47	-0.59
10	2,2',3,5'-tetrachlorobiphenyl	54.370	-6.47	-7.38	0.91
11	2,2',5,5'-tetrachlorobiphenyl	54.370	-7.00	-7.38	0.38
12	2,3',4,4'-tetrachlorobiphenyl	54.370	-6.68	-7.38	0.70
13	2,4,4',6-tetrachlorobiphenyl	54.370	-6.94	-7.38	0.44
14	3,3',4,4'-tetrachlorobiphenyl	54.370	-8.53	-7.38	-1.15
15	3,3',5,5'-tetrachlorobiphenyl	54.370	-8.54	-7.38	-1.16
16	2,2',3,3',5-pentachlorobiphenyl	54.359	-6.96	-7.40	0.44
17	2,2',3,4,5-pentachlorobiphenyl	54.359	-7.21	-7.40	0.19
18	2,2',3,4,5'-pentachlorobiphenyl	54.359	-7.91	-7.40	-0.51
19	2,2',3,4,6-pentachlorobiphenyl	54.359	-7.43	-7.40	-0.03
20	2,2',4,5,5'-pentachlorobiphenyl	54.359	-7.33	-7.40	0.07
21	2,2',4,6,6'-pentachlorobiphenyl	54.359	-7.32	-7.40	0.08
22	2,3',4,4',5-pentachlorobiphenyl	54.359	-7.39	-7.40	0.01
23	2,2',3,3',4,4'-hexachlorobiphenyl	53.860	-9.01	-8.23	-0.78
24	2,2',3,3',4,5-hexachlorobiphenyl	53.860	-8.07	-8.23	0.16
25	2,2',3,4,4',5'-hexachlorobiphenyl	53.860	-8.32	-8.23	-0.09
26	2,2',3,5,5',6-hexachlorobiphenyl	53.860	-7.42	-8.23	0.81
27	2,2',3,4,4',5',6-heptachlorobiphenyl	53.719	-7.92	-8.46	0.54
28	2,2',3,4',5,5',6-heptachlorobiphenyl	53.719	-8.94	-8.46	-0.48
29	2,2',3,3',5,5',6,6'-octachlorobiphenyl	53.289	-9.15	-9.17	0.02
30	2,2',3,3',4,5,5',6,6'-nonachlorobiphenyl	52.524	-10.41	-10.44	0.03
31	Decachlorobiphenyl	51.801	-11.62	-11.64	0.02

Average absolute deviation = 0.39

Test Set

n	Molecule	D(a,NNC)	log S_B	Eq.(9)	Delta
1	2-monochlorobiphenyl	55.669	-4.54	-5.23	0.69
2	2,2'-dichlorobiphenyl	55.179	-5.27	-6.04	0.77
3	2,6-dichlorobiphenyl	55.179	-5.21	-6.04	0.83
4	2,3,4'-trichlorobiphenyl	54.920	-6.26	-6.47	0.21
5	2,3,6-trichlorobiphenyl	54.920	-6.29	-6.47	0.18
6	2,3',5-trichlorobiphenyl	54.920	-6.01	-6.47	0.46
7	2,4,4'-trichlorobiphenyl	54.920	-6.21	-6.47	0.26
8	2,4,5-trichlorobiphenyl	54.920	-6.27	-6.47	0.20
9	2,4,6-trichlorobiphenyl	54.920	-6.14	-6.47	0.33
10	2',3,4-trichlorobiphenyl	54.920	-6.29	-6.47	0.18
11	2,2',3,3'-tetrachlorobiphenyl	54.370	-7.28	-7.38	0.10

12	2,2',4,4'-tetrachlorobiphenyl	54.370	-6.51	-7.38	0.87
13	2,2',4,5'-tetrachlorobiphenyl	54.370	-6.57	-7.38	0.81
14	2,2',5,6'-tetrachlorobiphenyl	54.370	-7.08	-7.38	0.30
15	2,2',6,6'-tetrachlorobiphenyl	54.370	-7.21	-7.38	0.17
16	2,3,4,5-tetrachlorobiphenyl	54.370	-7.16	-7.38	0.22
17	2,3',4',5-tetrachlorobiphenyl	54.370	-7.25	-7.38	0.13
18	2,2',3,3',4-pentachlorobiphenyl	54.359	-7.05	-7.40	0.35
19	2,3,4,5,6-pentachlorobiphenyl	54.359	-7.92	-7.40	-0.52
20	2,2',3,3',5,6-hexachlorobiphenyl	53.860	-8.60	-8.23	-0.37
21	2,2',3,3',6,6'-hexachlorobiphenyl	53.860	-8.65	-8.23	-0.42
22	2,2',3,4,5,5'-hexachlorobiphenyl	53.860	-7.68	-8.23	0.55
23	2,2',4,4',5,5'-hexachlorobiphenyl	53.860	-8.56	-8.23	-0.33
24	2,2',4,4',6,6'-hexachlorobiphenyl	53.860	-8.71	-8.23	-0.48
25	2,3,3',4,4',5-hexachlorobiphenyl	53.860	-7.82	-8.23	0.41
26	2,3,3',4,4',6-hexachlorobiphenyl	53.860	-7.66	-8.23	0.57
27	2,2',3,3',4,4',6-heptachlorobiphenyl	53.719	-8.30	-8.46	0.16
28	2,2',3,4,5,5',6-heptachlorobiphenyl	53.719	-8.46	-8.46	0.00
29	2,2',3,3',4,4',5,5'-octachlorobiphenyl	53.289	-9.16	-9.17	0.01
30	2,2',3,3',4,4',5,5',6-nonachlorobiphenyl	52.524	-10.26	-10.44	0.18

Average absolute deviation = 0.37