Abstract: The paper presents a unitary approach of the use of a Molecular Descriptors Family in structure-property/activity relationships, particularly in modelling the chromatographic retention times of polychlorinated biphenyls. Starting from molecular structure, viewed as a graph, and considering the bonds and bond types, atom types and often the 3D geometry of the molecule, a huge family of molecular descriptors called MDF was calculated. A preliminary selection of MDF members was done by simple linear regression (LR) against the measured property. The best fitted MDF subset is then submitted to multivariate linear regression (MLR) analysis in order to find the best pairs of MDF members that produce a reliable QSPR (Quantitative Structure-Property Relationship) model. The predictive capability was finally tested by randomly splitting of data into training and test sets. The best obtained models are presented and the results are discussed.
Keywords: Quantitative Structure-Property Relationship (QSPR), Molecular Descriptors Family (MDF), Polychlorinated Biphenyls (PCBs), Chromatographic Retention Time.

1. Introduction

Polychlorinated biphenyls (PCBs), organic compounds with 1 to 10 chlorine atoms attached to biphenyl, have the general chemical formula C_{12}H_{10-x}Cl_{x}. First manufactured by Monsanto in 1929, the PCBs production was banned in the 1970th due to the high toxicity of most PCBs (209) and mixtures [1]. PCBs were used as insulating fluids for industrial transformers and capacitors, and are known as persistent organic pollutants. Even if the production of the PCBs was stopped, they still have an influence on the human [2-4] and animal [5] health due to their accumulation in the environment. Moreover, the toxicity and carcinogenicity of PCBs could be related to mechanistic studies of their truncated analogue vinyl chloride [6]. Ecological and toxicological aspects of polychlorinated biphenyls (PCBs) in the environment are under investigation due to their worldwide distribution [7-10].

Starting with the 20th century, several mathematical approaches, that link chemical structure and property/activity in a quantitative manner, have been introduced [11]. Nowadays, quantitative structure-property/activity relationships (QSPRs/QSARs) are currently used in pharmaceutical chemistry, toxicology and other related fields [12].

A series of properties and activities of PCBs have been investigated by QSPR/QSAR modelling: aqueous solubility [13], gas/particle partitioning in the atmosphere [14], photo degradation half-life in n-hexane solution under UV irradiation [15], n-octanol/water partition coefficients [16,17], vaporization [18,19], and sublimation enthalpy [20]. The retention time of PCB congeners has also been previously investigated and reported [21-25]. Some of the reported results are: • Hasan and Jurs [22] - five-variable regression equation with $R^2 = 0.997$ and standard deviation of 0.017; • Liu et. al [24] - five-variable regression equation with the correlation coefficient of 0.9964 ($R^2 = 0.9928$); • Ren et. al [25] - four descriptors regression model with a correlation coefficient of 0.988 ($R^2 = 0.9761$) for the test set and an average absolute relative deviation of 3.08%.

The family of molecular descriptors MDF, designed by treating the interactions among fragments of a molecular structure with the formalism of electrostatic fields and potentials, and molecular topology as well, was developed and tested in QSPR/QSAR studies [26-29].

The aim of the present study was to investigate the ability of our MDF in modelling the retention times of 209 polychlorinated biphenyls.
2. Materials and Methods

2.1 Polychlorinated Biphenyls (PCBs)

The relative response times of all PCBs obtained by using temperature-programmed, high-resolution gas chromatography on a capillary column of SE-54, reported by Mullin et al. [30] served as experimental data in this study.

Molecular structure of PCBs was drawn by using HyperChem software [31] and their 3D geometry optimised at the Extended Hückel level of theory. These calculations also provided partial charges of atoms inside the molecules. The output files *.hin files, which store the information about topology, geometry and charge distribution of the PCBs, represented the primary data for the generation of the molecular descriptors family.

2.2 Methodology of using Molecular Descriptors Family in QSPR/QSAR

Our MDF implements three criteria of fragmentation, related to pairs of atoms, in order to generate molecular fragments. Let \(i\) and \(j\) be the atoms forming a pair. The criteria are as follows:

(a) A minimal fragment is that one containing only the atom \(i\), while a maximal fragment will contain all the atoms connected to \(i\), excluding the atom \(j\).

(b) A Szeged fragment is the set of vertices located closer to \(i\) than \(j\) (a distance-based criterion), the distance \(d(i, k)\) being lesser than \(d(k, j)\), and

(c) A Cluj fragment is generated by excluding the path from \(i\) to \(j\) (except its terminal points) and then applying the above Szeged criterion.

Every MDF member is named with seven ordered case sensitive letters: \(lMfOIpdp\), every letter encoding an operator, as follows.

The 7th letter (\(d\)) encodes the distance metric and is either \('g'\) (geometric) or \('t'\) (topological). The 6th letter (\(p\)) encodes the atomic property and can be \('M'\) (mass), \('Q'\) (charge), \('C'\) (cardinality), \('E'\) (electronegativity), \('G'\) (group electronegativity), or \('H'\) (number of attached hydrogens). The 5th letter (\(l\)) encodes the interaction descriptor (involving two participants): \('D(d)'\), \('d(1/d)'\), \('O(p_1)'\), \('o(1/p_1)'\), \('P(p_1p_2)'\), \('p(1/p_1p_2)'\), \('Q(\sqrt{p_1p_2})'\), \('q(\sqrt{1/p_1p_2})'\), \('J(p_1d)'\), \('j(1/p_1d)'\), \('K(p_1p_2d)'\), \('k(1/p_1p_2d)'\), \('L(T(p_1p_2d)^2)'\), \('T(p_1p_2d)'\), \('W(p_1d)'\), \('W(p_1p_2d)'\), \('F(p_1d^2)'\), \('f(p_1p_2d^2)'\), \('S(p_1p_2d^3)'\), \('s(p_1p_2d^3)'\), \('T(p_1^2/d^2)'\), \('t(p_1p_2/d^2)'\). The 4th letter (\(O\)) encodes the type of overlapping interactions, which are either scalar ('\(R'\), 'r', 'M', 'm') or vectorial ('D', 'd'). The 3rd letter (\(f\)) encodes the fragmentation algorithm and can be: \('m'\) (minimal), \('M'\) (maximal), \('D'\) (Szeged, distance based), and \('P'\) (Cluj, shortest paths based). The 2nd letter (\(M\)) encodes overlapping fragmental descriptors, which are of the type: sized group (\('m'\), smallest; \('M'\), largest; \('n'\), smallest absolute; \('N'\), largest absolute); averaged group (\('S'\), sum; \('A'\), average over all values; \('a'\), \(S\) divided by the number of all fragments; \('B'\), average first by atom group and then by the whole molecule; \('b'\), by bond); geometric group (\('P'\), multiplication; \('G'\), geometric mean, by fragments; \('g'\), adjusted \(G\); \('F'\), geometric mean by atom group and then by the whole molecule; \('f'\), by bond); harmonic group (\('s'\), harmonic mean, \('H'\) harmonic mean, by fragments, and similarly to above \('h'\), \(I\), and \('i'\).
MDF values enter in QSPR/QSAR modelling after a transformation (linearization procedure), one of: 'I' (identity), 'i' (inverse), 'A' (absolute), 'a' (inverse of absolute), 'L' (logarithm of absolute), 'l' (logarithm), which are encoded by the 1st letter.

MDF use a genetic algorithm for QSPR/QSAR modelling (genetic algorithms are a particular class of evolutionary algorithms, being categorized as global search heuristics [32]). The peculiarities of the genetic algorithm used are:

- Step 1 (implies inheritance and mutation). To the solution domain (2x6x24x6x4x19 MDF members) having the genetic representation with six letters words) are applied the linearization procedure from above, when every descendent is obtained from a parent (inheritance) through a transformation (mutation). Six times more (than parents) descendants are obtained. In this step, the fitness function is defined as “have real and distinct values”. A number of 490030 descendants dye due to mutation on PCB data set (remaining 297938 descendants, having genetic representation with seven letters words now).

- Step 2 (implies selection). To the solution domain (MDF descendants from Step 1) a bias procedure (selection) is applied. In this step, the fitness function is defined as “have distinct first nine digits of determination coefficient with measured property”. For PCBs data set, only 99806 members pass selection. From this solution domain another selection is made: best descriptor (which correlates the best with measured property (for PCBs result being presented in Eq(1)).

- Step 3 (implies crossover). Pairs of MDF members are crossover in order to obtain models with two descriptors. Two fitness functions are used here: “have better determination coefficient” and “have better cross-validation leave-one-out score”. The result for PCBs data set is given in Eq(2).

2.3 Computational Details

The MDF is calculated by a set of original programs written in PHP (Pre Hypertext Processor, [33]) and stored into a MySQL database [34] under a FreeBSD server [35]. This set of programs completes the MDF generation task. The programs create tables, insert, drop, delete, and select requests on `MDF` database (Figure 1). All programs run in a directory with the name of the set of selected compounds (actually, PCB).

![Figure 1. ‘MDF’ database for PCBs.](image)

The first program, `a_mdf_prepare.php`, orders the molecules, contained as *.hin files in a `data` subdirectory, in the same ordering as the measured property, contained in a `property.txt` file. The names of *.hin files and corresponding property are used to create a temporary `PCB_tmpx` table and finally the `PCB_data` table. The second program, `b_mdf_generate.php` (the most time consuming
The third program, `c_mdf_linearize.php`, completes the `PCB_data`, `PCB_xval`, and `PCB_yval` tables with linearized MDF members and statistical parameters. Note that, only real and distinct values are stored into the database. The fourth program, `d_mdf_bias.php` applied a bias procedure for data reduction. Finally, the fifth program, `e_mdf_order.php`, re-arrange the data from the `PCB_xval`, and `PCB_yval` tables in descending order of the squared correlation coefficient. When the task is complete, the fifth program writes in the `ready` table a record with the set name (Figure 2).

**Figure 2.** Preparing data for Multiple Linear Regression analysis.

The QSPR/QSAR finding procedure is made by a client programs built in Delphi programming language [36]. Bivariate correlations are performed, one with any other MDF members. A client program (Figure 3) connects the `MDF` database, query the ready tables all together, for the ready set (now PCB set is ready), and runs for finding the best QSAR/QSPR model. Every new better QSPR/QSAR is stored into a table called `qspr_qsar`, within the same `MDF` database.

**Figure 3.** MLR MDF QSPR client-server.

This program, called `i_mdf_query.php`, provides complete statistical analysis of models. The user of MDF can modify, by means of this `i_mdf_query.php` program, the criteria for the best QSPR/QSAR models.
3. Results and Discussion

The above described procedure was used for finding the best QSPR model of the PCBs relative chromatographic retention times.

In monovariate correlation, the best MDF QSPR model was provided by the $iIDRwHg$ MDF member, Eq(1):

$$\hat{Y}_{1d} = -0.16 + 0.09 \cdot iIDRwHg$$

$$R^2 = 0.9840; \text{95\%CI}_R [0.9894 - 0.9939]; \text{StErr} = 0.02; F = 12806; p < 0.0001$$

$$Q^2_{cv-loo} = 0.9838$$

(1)

where $\hat{Y}_{1d}$ = estimated retention time by MDF-SAR equation with one descriptor; $iIDRwHg$ = molecular descriptor; $R^2$ = square correlation coefficient; 95\%CI$_R$ = 95\% confidence interval for correlation coefficient; $Q^2_{cv-loo}$ = cross-validation leave-one-out score.

The quality of statistics is given by $R^2$ (the square correlation coefficient), StErr (standard error of estimate), F (Fisher parameter) and p (type I error, or $\alpha$ error). The cross-validation leave-one-out score is given as $Q^2$. Clearly, the model shows a good predictability. The type I error of the model from Eq(1) is very small, showing a very small error of rejecting the null hypothesis when it is actually true.

About ninety-eight percents of variation in PCBs chromatographic retention time can be explained by its linear relation with a single MDF member, $iIDRwHg$, which accounts for the actual geometry (by the geometric distance operator (‘$g$’)) and the number of directly bonded hydrogen atoms (‘$H$’).

The best model with two descriptors was:

$$\hat{Y}_{2d} = -5.96 + 0.024 \cdot ISDmsHt - 1.02 \cdot lADrtHg$$

$$R^2 = 0.9967; \text{95\% CI} [0.9977-0.9987]; \text{StdErr} = 0.01; F = 30752; p < 0.0001$$

$$Q^2_{cv-loo} = 0.9962$$

(2)

where $\hat{Y}_{2d}$ = estimated retention time by MDF-SAR equation with two descriptors.

The multi-colinearity analysis shown that the two descriptors used by Eq(2) rather inter-related ($R^2(ISDmsHt, lADrtHg) = 0.944$) and each of them ($R^2(Y, ISDmsHt) = 0.907; \text{rank} = 12614; R^2(Y, lADrtHg) = 0.973; \text{rank} = 277$) are not the best descriptor in monovariate regression model (see Eq(1)). The ISDmsHt descriptor is built by a topological distance operator (‘t’) while lADrtHg takes into account the genuine distance (‘g’). Both of them consider the directly bonded hydrogen atom (‘$H$’). The topological description explains more than 90\% of the variance, the remaining 9.7\% being completed by the information on molecular geometry.

The plot corresponding to Eq(2) is illustrated in Figure 4.

The values of the best descriptors in uni and bivariate regressions (Eq(1)&(2)), the experimental and estimated chromatographic retention time, and residuals for the PCBs set are listed in Table 1.

The accuracy of description is extremely high, even as the set of molecules is quite large. The excellent model (Eq(2)), derived for such a large set, is by itself a test of predictive ability. Indeed, if various ratios training/testing selections were considered, the quality of statistics remained very high (Table 2).
As it can be observed from Table 2, the lowest $R^2$ is about 0.996 in both training and test sets, which demonstrates the ability of ($ISDmsHt, lADrtHg$) MDF pair to described the PCBs relative retention time (Eq(2)). Note that, $R^2$ exceeds the upper bond of the confidence interval of Eq(2) in almost 20% of cases and is less then the lower bond in other 20% of cases. In the test set, in four cases the values of $Q^2$ were greater than the upper confidence boundary.

By analysing of the obtained models (Eq(1) and Eq(2)) in the light of the previously reported models, it can be observed that with a single exception ([25], $p = 0.3528$) out of three the model with one descriptor - Eq(1) - did not obtains a greater squared correlation coefficient compared with models reported in the references [22] and [24] (the differences are of -0.0064 [22], and of -0.0043 [24] respectively).

Analyzing the model with two molecular descriptors it was identified a statistical significant differences between correlation coefficient of this model and of the model reported by [24] ($p < 0.0001$) or by [25] ($p < 0.0001$). There was not identified a statistical difference between the Eq(2) and the model reported by [22] ($p = 0.7263$). The following remarks can be revealed by summarizing the above results:

- The MDF model obtained by Eq(1) is a better model comparing with previously reported ones [22, 24,25] in terms of number of variables used (one descriptor for the model from Eq(1), five descriptors for the model reported in [22] and [24], four descriptors for the model reported in [25]).
- The MDF model obtained by Eq(2) is significantly better models comparing with models reported in [24] and [25] in terms of correlation coefficients. Moreover, it is a better model comparing with model reported in [22] in terms of number of variables used (two descriptors used by the Eq(2), and five descriptors used by the model reported in [22]).
| Mol   | PCB structure | Y   | iIDRwHg | Ŷ_1d | Y - Ŷ_1d | ISDmsHt | lADrtHg | Ŷ_2d | Y - Ŷ_2d |
|-------|---------------|-----|---------|------|----------|----------|---------|------|----------|
| PCB001 | Cl            | 0.0997 | 10.02 | -0.0122 | -0.5363 | 133.20  | -3.42  | 0.1119 | -0.0122 |
| PCB002 | Cl            | 0.1544 | 10.60 | 0.0041  | 0.1800  | 134.27  | -3.47  | 0.1503 | 0.0041  |
| PCB003 | Cl            | 0.1937 | 9.96  | 0.0376  | 1.6541  | 135.23  | -3.40  | 0.1561 | 0.0376  |
| PCB004 | ClCl          | 0.2245 | 10.14 | 0.0054  | 0.2377  | 134.89  | -3.42  | 0.2191 | 0.0054  |
| PCB005 | ClCl          | 0.2785 | 9.75  | 0.0251  | 1.1035  | 133.36  | -3.41  | 0.2534 | 0.0251  |
| PCB006 | ClCl          | 0.2709 | 10.15 | -0.0193 | -0.8496 | 136.72  | -3.38  | 0.2902 | -0.0193 |
| PCB007 | ClCl          | 0.2566 | 10.72 | 0.0028  | 0.1234  | 134.60  | -3.48  | 0.2538 | 0.0028  |
| PCB008 | ClCl          | 0.2783 | 10.27 | -0.0048 | -0.2094 | 133.35  | -3.43  | 0.2831 | -0.0048 |
| PCB009 | ClCl          | 0.2570 | 11.12 | 0.0348  | 1.5315  | 134.95  | -3.55  | 0.2222 | 0.0348  |
| PCB010 | ClCl          | 0.2243 | 11.75 | 0.0333  | 1.4623  | 133.57  | -3.57  | 0.1910 | 0.0333  |
| PCB011 | ClCl          | 0.3238 | 11.26 | 0.0168  | 0.7378  | 133.12  | -3.52  | 0.3070 | 0.0168  |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1) & Eq(2); Ŷ_1d, 2d = relative retention time: estimated by Eq(1) and Eq(2), respectively; Y = relative retention time: experimental [30]; Y - Ŷ_1d, 2d = residuals.
| Mol   | PCB structure | Y   | iDRwHg $\hat{Y}_{1d}$ | Y - $\hat{Y}_{1d}$ | ISDmsHt | lADrtHg $\hat{Y}_{2d}$ | Y - $\hat{Y}_{2d}$ |
|-------|---------------|-----|------------------------|---------------------|----------|------------------------|---------------------|
| PCB012 | ![PCB012 Structure](image) | 0.3298 | 11.52 | 0.0442 | 1.9425 | 132.24 | -3.57 | 0.2856 | 0.0442 |
| PCB013 | ![PCB013 Structure](image) | 0.3315 | 11.09 | 0.0065 | 0.2857 | 134.24 | -3.49 | 0.3250 | 0.0065 |
| PCB014 | ![PCB014 Structure](image) | 0.2373 | 10.98 | -0.0393 | -1.7268 | 133.10 | -3.50 | 0.2766 | -0.0393 |
| PCB015 | ![PCB015 Structure](image) | 0.3387 | 10.98 | 0.0036 | 0.1567 | 131.58 | -3.51 | 0.3351 | 0.0036 |
| PCB016 | ![PCB016 Structure](image) | 0.3625 | 10.45 | 0.0193 | 0.8481 | 132.74 | -3.45 | 0.3432 | 0.0193 |
| PCB017 | ![PCB017 Structure](image) | 0.3398 | 10.97 | -0.0184 | -0.8086 | 133.06 | -3.51 | 0.3582 | -0.0184 |
| PCB018 | ![PCB018 Structure](image) | 0.3378 | 10.72 | -0.0125 | -0.5510 | 131.64 | -3.50 | 0.3503 | -0.0125 |
| PCB019 | ![PCB019 Structure](image) | 0.3045 | 10.16 | 0.0042 | 0.1849 | 132.62 | -3.44 | 0.3003 | 0.0042 |
| PCB020 | ![PCB020 Structure](image) | 0.4170 | 11.56 | 0.0015 | 0.0644 | 132.66 | -3.57 | 0.4155 | 0.0015 |
| PCB021 | ![PCB021 Structure](image) | 0.4135 | 11.09 | -0.0179 | -0.7855 | 133.32 | -3.50 | 0.4314 | -0.0179 |
| PCB022 | ![PCB022 Structure](image) | 0.4267 | 11.05 | 0.0005 | 0.0212 | 131.66 | -3.52 | 0.4262 | 0.0005 |

$iDRwHg$, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1)& Eq(2); $\hat{Y}_{1d,2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; $Y = \text{relative retention time: experimental [30]}$; $Y - \hat{Y}_{1d,2d}$ = residuals.
### Table 1. (Continued)

| Mol      | PCB structure | Y   | iIDRwHg $\hat{\gamma}_{1d}$ | Y-$\hat{\gamma}_{1d}$ | ISDmsHt | lADrtHg $\hat{\gamma}_{2d}$ | Y-$\hat{\gamma}_{2d}$ |
|----------|---------------|-----|------------------------------|------------------------|----------|------------------------------|------------------------|
| PCB023   | ![Cl-Cl-Cl](image) | 0.3770 | 11.05 | -0.0239 | -1.0517 | 132.19 | -3.51 | 0.4009 | -0.0239 |
| PCB024   | ![Cl-Cl-Cl](image) | 0.3508 | 10.52 | 0.0042 | 0.1838 | 131.49 | -3.46 | 0.3466 | 0.0042 |
| PCB025   | ![Cl-Cl-Cl](image) | 0.3937 | 10.80 | -0.0283 | -1.2445 | 133.28 | -3.48 | 0.4220 | -0.0283 |
| PCB026   | ![Cl-Cl-Cl](image) | 0.3911 | 10.24 | -0.0015 | -0.0653 | 133.94 | -3.42 | 0.3926 | -0.0015 |
| PCB027   | ![Cl-Cl-Cl](image) | 0.3521 | 10.75 | 0.0056 | 0.2482 | 132.13 | -3.50 | 0.3465 | 0.0056 |
| PCB028   | ![Cl-Cl-Cl](image) | 0.4031 | 10.23 | -0.0294 | -1.2916 | 131.25 | -3.45 | 0.4325 | -0.0294 |
| PCB029   | ![Cl-Cl-Cl](image) | 0.3820 | 12.03 | -0.0161 | -0.7060 | 132.78 | -3.65 | 0.3981 | -0.0161 |
| PCB030   | ![Cl-Cl-Cl](image) | 0.3165 | 11.48 | -0.0323 | -1.4195 | 133.66 | -3.57 | 0.3488 | -0.0323 |
| PCB031   | ![Cl-Cl-Cl](image) | 0.4094 | 11.55 | 0.0036 | 0.3793 | 132.00 | -3.60 | 0.4008 | 0.0086 |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1) & Eq(2); $\hat{\gamma}_{1d,2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; $Y$ = relative retention time: experimental [30]; $Y-\hat{\gamma}_{1d,2d}$ = residuals.


| Mol  | PCB structure | Y   | iIDRwHg $\hat{Y}_{1d}$ | Y- $\hat{Y}_{1d}$ | ISDmsHt | lADrtHg $\hat{Y}_{2d}$ | Y- $\hat{Y}_{2d}$ |
|------|---------------|-----|------------------------|--------------------|----------|------------------------|--------------------|
| PCB032 | ![PCB032 Structure](image) | 0.3636 | 11.22 | 0.0089 | 0.3932 | 131.75 | -3.57 | 0.3547 | 0.0089 |
| PCB033 | ![PCB033 Structure](image) | 0.4163 | 11.25 | 0.0057 | 0.2490 | 132.12 | -3.58 | 0.4106 | 0.0057 |
| PCB034 | ![PCB034 Structure](image) | 0.3782 | 12.89 | -0.0103 | -0.4521 | 130.80 | -3.73 | 0.3885 | -0.0103 |
| PCB035 | ![PCB035 Structure](image) | 0.4738 | 12.32 | 0.0138 | 0.6063 | 131.52 | -3.67 | 0.4600 | 0.0138 |
| PCB036 | ![PCB036 Structure](image) | 0.4375 | 12.42 | 0.0027 | 0.1167 | 130.24 | -3.68 | 0.4348 | 0.0027 |
| PCB037 | ![PCB037 Structure](image) | 0.4858 | 11.87 | 0.0184 | 0.8096 | 129.97 | -3.61 | 0.4674 | 0.0184 |
| PCB038 | ![PCB038 Structure](image) | 0.5102 | 12.09 | 0.0635 | 2.7897 | 130.07 | -3.66 | 0.4467 | 0.0635 |
| PCB039 | ![PCB039 Structure](image) | 0.4488 | 11.53 | 0.0041 | 0.1782 | 131.14 | -3.59 | 0.4447 | 0.0041 |
| PCB040 | ![PCB040 Structure](image) | 0.5102 | 11.58 | 0.0012 | 0.0545 | 129.81 | -3.60 | 0.5090 | 0.0012 |
| PCB041 | ![PCB041 Structure](image) | 0.4990 | 12.15 | -0.0127 | -0.5568 | 130.26 | -3.67 | 0.5117 | -0.0127 |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1)& Eq(2); $\hat{Y}_{1d,2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; $Y$ = relative retention time: experimental [30]; Y- $\hat{Y}_{1d,2d}$ = residuals.
Table 1. (Continued)

| Mol  | PCB structure | Y | iIDRwHg $\hat{Y}_{1d}$ | Y- $\hat{Y}_{1d}$ | ISDmsHt | IADrtHg $\hat{Y}_{2d}$ | Y- $\hat{Y}_{2d}$ |
|------|--------------|---|-----------------------|-------------------|---------|-----------------------|-------------------|
| PCB042 | | 0.4870 | 11.30 | -0.0324 | -1.4222 | 129.77 | -3.59 | 0.5194 | -0.0324 |
| PCB043 | | 0.4587 | 12.11 | -0.0267 | -1.1744 | 130.59 | -3.66 | 0.4854 | -0.0267 |
| PCB044 | | 0.4832 | 11.60 | -0.0088 | -0.3869 | 129.80 | -3.61 | 0.4920 | -0.0088 |
| PCB045 | | 0.4334 | 12.57 | 0.0004 | 0.0168 | 130.50 | -3.74 | 0.4330 | 0.0004 |
| PCB046 | | 0.4450 | 13.43 | 0.0088 | 0.3881 | 128.67 | -3.82 | 0.4362 | 0.0088 |
| PCB047 | | 0.4639 | 12.87 | -0.0562 | -2.4723 | 128.32 | -3.76 | 0.5201 | -0.0562 |
| PCB048 | | 0.4651 | 12.62 | -0.0098 | -0.4320 | 128.24 | -3.75 | 0.4749 | -0.0098 |
| PCB049 | | 0.4610 | 14.04 | -0.0314 | -1.3821 | 126.70 | -3.90 | 0.4924 | -0.0314 |
| PCB050 | | 0.4007 | 10.02 | -0.0122 | -0.5363 | 133.20 | -3.42 | 0.1119 | -0.0122 |
| PCB051 | | 0.4242 | 10.60 | 0.0041 | 0.1800 | 134.27 | -3.47 | 0.1503 | 0.0041 |

iIDRwHg, ISDmsHt, and IADrtHg = the value of the descriptor - Eq(1)& Eq(2); $\hat{Y}_{1d,2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; $Y$ = relative retention time: experimental [30]; $Y-\hat{Y}_{1d,2d}$ = residuals.
| Mol     | PCB structure | Y   | iDRwHg | Ŷ₁,₁d | Y - Ŷ₁,₁d | ISDmsHt | lADrtHg | Ŷ₂,₁d | Y - Ŷ₂,₁d |
|---------|---------------|-----|--------|-------|-----------|----------|----------|-------|-----------|
| PCB052  |               | 0.4557 | 9.96   | 0.0376 | 1.6541    | 135.23   | -3.40    | 0.1561 | 0.0376    |
| PCB053  |               | 0.4187 | 10.14  | 0.0054 | 0.2377    | 134.89   | -3.42    | 0.2191 | 0.0054    |
| PCB054  |               | 0.3800 | 9.75   | 0.0251 | 1.1035    | 133.36   | -3.41    | 0.2534 | 0.0251    |
| PCB055  |               | 0.5562 | 10.15  | -0.0193 | -0.8496  | 136.72   | -3.38    | 0.2902 | -0.0193   |
| PCB056  |               | 0.5676 | 10.72  | 0.0028 | 0.1234    | 134.60   | -3.48    | 0.2538 | 0.0028    |
| PCB057  |               | 0.5515 | 10.27  | -0.0048 | -0.2094  | 133.35   | -3.43    | 0.2831 | -0.0048   |
| PCB058  |               | 0.5267 | 11.12  | 0.0348 | 1.5315    | 132.24   | -3.57    | 0.2856 | 0.0348    |
| PCB059  |               | 0.4860 | 11.75  | 0.0333 | 1.4623    | 133.57   | -3.57    | 0.1910 | 0.0333    |
| PCB060  |               | 0.5676 | 11.26  | 0.0168 | 0.7378    | 133.12   | -3.52    | 0.3070 | 0.0168    |
| PCB061  |               | 0.5331 | 11.52  | 0.0442 | 1.9425    | 132.24   | -3.57    | 0.2856 | 0.0442    |

iDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1)& Eq(2); Ŷ₁,₁d = relative retention time: estimated by Eq(1) and Eq(2), respectively; Y = relative retention time: experimental [30]; Y - Ŷ₁,₁d = residuals.
| Mol   | PCB structure | Y  | iIDRwHg $\hat{Y}_{1d}$ | Y-$\hat{Y}_{1d}$ | ISDmsHt | lADrtHg $\hat{Y}_{2d}$ | Y-$\hat{Y}_{2d}$ |
|-------|--------------|----|------------------------|------------------|---------|------------------------|------------------|
| PCB062 | ![Cl4](attachment:chemical.png) | 0.4685 11.09 | 0.0065 0.2857 | 134.24 | -3.49 | 0.3250 0.0065 |
| PCB063 | ![Cl4](attachment:chemical.png) | 0.5290 10.98 | -0.0393 1.7268 | 133.10 | -3.50 | 0.2766 -0.0393 |
| PCB064 | ![Cl4](attachment:chemical.png) | 0.4999 10.98 | 0.0036 0.1567 | 131.58 | -3.51 | 0.3351 0.0036 |
| PCB065 | ![Cl4](attachment:chemical.png) | 0.4671 10.45 | 0.0193 0.8481 | 132.74 | -3.45 | 0.3432 0.0193 |
| PCB066 | ![Cl4](attachment:chemical.png) | 0.5447 10.97 | -0.0184 0.8086 | 133.06 | -3.51 | 0.3582 -0.0184 |
| PCB067 | ![Cl4](attachment:chemical.png) | 0.5214 10.72 | -0.0125 0.5510 | 131.64 | -3.50 | 0.3503 -0.0125 |
| PCB068 | ![Cl4](attachment:chemical.png) | 0.5040 10.16 | 0.0042 0.1849 | 132.62 | -3.44 | 0.3003 0.0042 |
| PCB069 | ![Cl4](attachment:chemical.png) | 0.5410 11.56 | 0.0015 0.0644 | 132.66 | -3.57 | 0.4155 0.0015 |
| PCB070 | ![Cl4](attachment:chemical.png) | 0.5407 11.09 | -0.0179 0.7855 | 133.32 | -3.50 | 0.4314 -0.0179 |
| PCB071 | ![Cl4](attachment:chemical.png) | 0.4989 11.05 | 0.0005 0.0212 | 131.66 | -3.52 | 0.4262 0.0005 |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1) & Eq(2); \( \hat{Y}_{1d,2d} \) = relative retention time: estimated by Eq(1) and Eq(2), respectively; \( Y \) = relative retention time: experimental \([30]\); \( Y-\hat{Y}_{1d,2d} \) = residuals.
Table 1. (Continued)

| Mol | PCB structure | Y | iIDRwHg $\hat{Y}_{1d}$ | Y - $\hat{Y}_{1d}$ | ISDmsHt lADrtHg $\hat{Y}_{2d}$ | Y - $\hat{Y}_{2d}$ |
|-----|---------------|---|-------------------------|-------------------|-------------------------------|-------------------|
| PCB072 | ![Image](image1.png) | 0.4984 11.05 | -0.0239 | -1.0517 | 132.19 | -3.51 | 0.4009 | -0.0239 |
| PCB073 | ![Image](image2.png) | 0.4554 10.52 | 0.0042 | 0.1838 | 131.49 | -3.46 | 0.3466 | 0.0042 |
| PCB074 | ![Image](image3.png) | 0.5341 10.80 | -0.0283 | -1.2445 | 133.28 | -3.48 | 0.4220 | -0.0283 |
| PCB075 | ![Image](image4.png) | 0.4643 10.24 | -0.0015 | -0.0653 | 133.94 | -3.42 | 0.3926 | -0.0015 |
| PCB076 | ![Image](image5.png) | 0.5408 10.75 | 0.0056 | 0.2482 | 132.13 | -3.50 | 0.3465 | 0.0056 |
| PCB077 | ![Image](image6.png) | 0.6295 10.23 | -0.0294 | -1.2916 | 131.25 | -3.45 | 0.4325 | -0.0294 |
| PCB078 | ![Image](image7.png) | 0.6024 12.03 | -0.0161 | -0.7060 | 132.78 | -3.65 | 0.3981 | -0.0161 |
| PCB079 | ![Image](image8.png) | 0.5894 11.48 | -0.0323 | -1.4195 | 133.66 | -3.57 | 0.3488 | -0.0323 |
| PCB080 | ![Image](image9.png) | 0.5464 11.55 | 0.0086 | 0.3793 | 132.00 | -3.60 | 0.4008 | 0.0086 |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1)& Eq(2); $\hat{Y}_{1d,2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; Y = relative retention time: experimental [30]; Y - $\hat{Y}_{1d,2d}$ = residuals.
### Table 1. (Continued)

| Mol | PCB structure | Y   | iIDRwHg $\hat{Y}_{1d}$ | Y- $\hat{Y}_{1d}$ | ISDmsHt lADrtHg $\hat{Y}_{2d}$ | Y- $\hat{Y}_{2d}$ |
|-----|---------------|-----|------------------------|------------------|--------------------------------|-------------------|
| PCB081 | ![Cl-Cl-Cl-Cl](image1) | 0.6149 11.22 | 0.0089 0.3932 | 131.75 | -3.57 | 0.3547 0.0089 |
| PCB082 | ![Cl-Cl-Cl-Cl](image2) | 0.6453 11.25 | 0.0057 0.2490 | 132.12 | -3.58 | 0.4106 0.0057 |
| PCB083 | ![Cl-Cl-Cl-Cl](image3) | 0.6029 12.89 | -0.0103 -0.4521 | 130.80 | -3.73 | 0.3885 -0.0103 |
| PCB084 | ![Cl-Cl-Cl-Cl](image4) | 0.5744 12.32 | 0.0138 0.6063 | 131.52 | -3.67 | 0.4600 0.0138 |
| PCB085 | ![Cl-Cl-Cl-Cl](image5) | 0.6224 12.42 | 0.0027 0.1167 | 130.24 | -3.68 | 0.4348 0.0027 |
| PCB086 | ![Cl-Cl-Cl-Cl](image6) | 0.6105 11.87 | 0.0184 0.8096 | 129.97 | -3.61 | 0.4674 0.0184 |
| PCB087 | ![Cl-Cl-Cl-Cl](image7) | 0.6175 12.09 | 0.0635 2.7897 | 130.07 | -3.66 | 0.4467 0.0635 |
| PCB088 | ![Cl-Cl-Cl-Cl](image8) | 0.5486 11.53 | 0.0041 0.1782 | 131.14 | -3.59 | 0.4447 0.0041 |
| PCB089 | ![Cl-Cl-Cl-Cl](image9) | 0.5779 11.58 | 0.0012 0.0545 | 129.81 | -3.60 | 0.5090 0.0012 |

*iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1)& Eq(2); $\hat{Y}_{1d,2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; $Y$ = relative retention time: experimental [30]; $Y$- $\hat{Y}_{1d,2d}$ = residuals.*
Table 1. (Continued)

| Mol    | PCB structure | Y    | iIDRwHg $\hat{Y}_{1d}$ | Y - $\hat{Y}_{1d}$ | ISDmsHt | lADrtHg $\hat{Y}_{2d}$ | Y - $\hat{Y}_{2d}$ |
|--------|---------------|------|------------------------|---------------------|---------|------------------------|---------------------|
| PCB090 | Cl Cl Cl Cl   | 0.5814 12.15 | -0.0127 -0.5568 130.26 | -3.67 | 0.5117 -0.0127 |
| PCB091 | Cl Cl Cl Cl Cl Cl | 0.5549 11.30 | -0.0324 -1.4222 129.77 | -3.59 | 0.5194 -0.0324 |
| PCB092 | Cl Cl Cl Cl Cl Cl Cl | 0.5742 12.11 | -0.0267 -1.1744 130.59 | -3.66 | 0.4854 -0.0267 |
| PCB093 | Cl Cl Cl Cl Cl Cl Cl Cl | 0.5437 11.60 | -0.0088 -0.3869 129.80 | -3.61 | 0.4920 -0.0088 |
| PCB094 | Cl Cl Cl Cl Cl Cl Cl Cl Cl | 0.5331 12.57 | 0.0004 0.0168 130.50 | -3.74 | 0.4330 0.0004 |
| PCB095 | Cl Cl Cl Cl Cl Cl Cl Cl Cl Cl | 0.5464 13.43 | 0.0088 0.3881 128.67 | -3.82 | 0.4362 0.0088 |
| PCB096 | Cl Cl Cl Cl Cl Cl Cl Cl Cl Cl Cl | 0.5057 12.87 | -0.0562 -2.4723 128.32 | -3.76 | 0.5201 -0.0562 |
| PCB097 | Cl Cl Cl Cl Cl Cl Cl Cl Cl Cl Cl Cl | 0.6100 12.62 | -0.0098 -0.4320 128.24 | -3.75 | 0.4749 -0.0098 |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1)& Eq(2); $\hat{Y}_{1d,2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; $Y$ = relative retention time: experimental [30]; $Y - \hat{Y}_{1d,2d}$ = residuals.
### Table 1. (Continued)

| Mol    | PCB structure | Y    | iIDRwHg $\hat{Y}_{1d}$ | Y - $\hat{Y}_{1d}$ | ISDmsHt  | lADrtHg $\hat{Y}_{2d}$ | Y - $\hat{Y}_{2d}$ |
|--------|---------------|------|------------------------|-------------------|----------|------------------------|-------------------|
| PCB098 | ![PCB098](image) | 0.5415 | 14.04 | -0.0314 | -1.3821 | 126.70 | -3.90 | 0.4924 | -0.0314 |
| PCB099 | ![PCB099](image) | 0.5880 | 10.02 | -0.0122 | -0.5363 | 133.20 | -3.42 | 0.1119 | -0.0122 |
| PCB100 | ![PCB100](image) | 0.5212 | 10.60 | 0.0041 | 0.1800 | 134.27 | -3.47 | 0.1503 | 0.0041 |
| PCB101 | ![PCB101](image) | 0.5816 | 9.96  | 0.0376 | 1.6541 | 135.23 | -3.40 | 0.1561 | 0.0376 |
| PCB102 | ![PCB102](image) | 0.5431 | 10.14 | 0.0054 | 0.2377 | 134.89 | -3.42 | 0.2191 | 0.0054 |
| PCB103 | ![PCB103](image) | 0.5142 | 9.75  | 0.0251 | 1.1035 | 133.36 | -3.41 | 0.2534 | 0.0251 |
| PCB104 | ![PCB104](image) | 0.4757 | 10.15 | -0.0193 | -0.8496 | 136.72 | -3.38 | 0.2902 | -0.0193 |
| PCB105 | ![PCB105](image) | 0.7049 | 10.72 | 0.0028 | 0.1234 | 134.60 | -3.48 | 0.2538 | 0.0028 |
| PCB106 | ![PCB106](image) | 0.6680 | 10.27 | -0.0048 | -0.2094 | 133.35 | -3.43 | 0.2831 | -0.0048 |

$iIDRwHg$, $ISDmsHt$, and $lADrtHg = $ the value of the descriptor - Eq(1) & Eq(2); $\hat{Y}_{1d, 2d} = $ relative retention time: estimated by Eq(1) and Eq(2), respectively; $Y = $ relative retention time: experimental [30]; $Y - \hat{Y}_{1d, 2d} = $ residuals.
Table 1. (Continued)

| Mol | PCB structure | Y  | iIDRwHg $\hat{Y}_{1d}$ | Y- $\hat{Y}_{1d}$ | ISDmsHt $lADrtHg$ $\hat{Y}_{2d}$ | Y- $\hat{Y}_{2d}$ |
|-----|---------------|----|------------------------|------------------|-----------------------------------|-------------------|
| PCB107 | ![PCB107 structure](image) | 0.6628 11.12 | 0.0348 | 1.5315 | 134.95 | -3.55 | 0.2222 0.0348 |
| PCB108 | ![PCB108 structure](image) | 0.6626 11.75 | 0.0333 | 1.4623 | 133.57 | -3.57 | 0.1910 0.0333 |
| PCB109 | ![PCB109 structure](image) | 0.6016 11.26 | 0.0168 | 0.7378 | 133.12 | -3.52 | 0.3070 0.0168 |
| PCB110 | ![PCB110 structure](image) | 0.6314 11.52 | 0.0442 | 1.9425 | 132.24 | -3.57 | 0.2856 0.0442 |
| PCB111 | ![PCB111 structure](image) | 0.6183 11.09 | 0.0065 | 0.2857 | 134.24 | -3.49 | 0.3250 0.0065 |
| PCB112 | ![PCB112 structure](image) | 0.5986 10.98 | -0.0393 | -1.7268 | 133.10 | -3.50 | 0.2766 -0.0393 |
| PCB113 | ![PCB113 structure](image) | 0.5862 10.98 | 0.0036 | 0.1567 | 131.58 | -3.51 | 0.3351 0.0036 |
| PCB114 | ![PCB114 structure](image) | 0.6828 10.45 | 0.0193 | 0.8481 | 132.74 | -3.45 | 0.3432 0.0193 |
| PCB115 | ![PCB115 structure](image) | 0.6171 10.97 | -0.0184 | -0.8086 | 133.06 | -3.51 | 0.3582 -0.0184 |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1) & Eq(2); $\hat{Y}_{1d,2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; Y = relative retention time: experimental [30]; Y- $\hat{Y}_{1d,2d}$ = residuals.
Table 1. (Continued)

| Mol   | PCB structure | Y       | IDRwHg  | Y-Ŷ₁<sub>d</sub> | ISDmsHt | lADrtHg  | Y-Ŷ₂<sub>d</sub> |
|-------|---------------|---------|---------|------------------|---------|-----------|------------------|
| PCB116 |               | 0.6132  | 10.72   | -0.0125  | -0.5510 | 131.64    | -3.50  | 0.3503  | -0.0125 |
| PCB117 |               | 0.6150  | 10.16   | 0.0042   | 0.1849  | 132.62    | -3.44  | 0.3003  | 0.0042  |
| PCB118 |               | 0.6693  | 11.56   | 0.0015   | 0.0644  | 132.66    | -3.57  | 0.4155  | 0.0015  |
| PCB119 |               | 0.5968  | 11.09   | -0.0179  | -0.7855 | 133.32    | -3.50  | 0.4314  | -0.0179 |
| PCB120 |               | 0.6256  | 11.05   | 0.0005   | 0.0212  | 131.66    | -3.52  | 0.4262  | 0.0005  |
| PCB121 |               | 0.5518  | 11.05   | -0.0239  | -1.0517 | 132.19    | -3.51  | 0.4009  | -0.0239 |
| PCB122 |               | 0.6871  | 10.52   | 0.0042   | 0.1838  | 131.49    | -3.46  | 0.3466  | 0.0042  |
| PCB123 |               | 0.6658  | 10.80   | -0.0283  | -1.2445 | 133.28    | -3.48  | 0.4220  | -0.0283 |
| PCB124 |               | 0.6584  | 10.24   | -0.0015  | -0.0653 | 133.94    | -3.42  | 0.3926  | -0.0015 |

iDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1) & Eq(2); Ŷ<sub>1d, 2d</sub> = relative retention time: estimated by Eq(1) and Eq(2), respectively; Y = relative retention time: experimental [30]; Y-Ŷ<sub>1d, 2d</sub> = residuals.
Table 1. (Continued)

| Mol  | PCB structure | \(Y\)  | IDRwHg \(\hat{Y}_{1d}\) | \(Y - \hat{Y}_{1d}\) | ISDmsHt \(\hat{Y}_{2d}\) | lADrtHg \(\hat{Y}_{2d}\) | \(Y - \hat{Y}_{2d}\) |
|------|---------------|-------|-----------------|----------------|------------------|------------------|----------------|
| PCB125 | ![PCB125](image) | 0.6142 | 10.75 | 0.0056 | 0.2482 | 132.13 | -3.50 | 0.3465 | 0.0056 |
| PCB126 | ![PCB126](image) | 0.7512 | 10.23 | -0.0294 | 1.2916 | 131.25 | -3.45 | 0.4325 | -0.0294 |
| PCB127 | ![PCB127](image) | 0.7078 | 12.03 | -0.0161 | 0.7060 | 132.78 | -3.65 | 0.3981 | -0.0161 |
| PCB128 | ![PCB128](image) | 0.7761 | 11.48 | -0.0323 | 1.4195 | 133.66 | -3.57 | 0.3488 | -0.0323 |
| PCB129 | ![PCB129](image) | 0.7501 | 11.55 | 0.0086 | 0.3793 | 132.00 | -3.60 | 0.4008 | 0.0086 |
| PCB130 | ![PCB130](image) | 0.7184 | 11.22 | 0.0089 | 0.3932 | 131.75 | -3.57 | 0.3547 | 0.0089 |
| PCB131 | ![PCB131](image) | 0.6853 | 11.25 | 0.0057 | 0.2490 | 132.12 | -3.58 | 0.4106 | 0.0057 |
| PCB132 | ![PCB132](image) | 0.7035 | 12.89 | -0.0103 | 0.4521 | 130.80 | -3.73 | 0.3885 | -0.0103 |
| PCB133 | ![PCB133](image) | 0.6871 | 12.32 | 0.0138 | 0.6063 | 131.52 | -3.67 | 0.4600 | 0.0138 |

iDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1) & Eq(2); \(\hat{Y}_{1d,2d}\) = relative retention time: estimated by Eq(1) and Eq(2), respectively; \(Y\) = relative retention time: experimental [30]; \(Y - \hat{Y}_{1d,2d}\) = residuals.
| Mol    | PCB structure | Y | iIDRwHg $\hat{Y}_{1d}$ | Y - $\hat{Y}_{1d}$ | ISDmsHt $\hat{Y}_{2d}$ | Y - $\hat{Y}_{2d}$ |
|--------|---------------|---|------------------------|---------------------|------------------------|---------------------|
| PCB134 | Cl Cl Cl Cl   | 0.6796 | 12.42 | 0.0027 | 0.1167 | 130.24 | -3.68 | 0.4348 | 0.0027 |
| PCB135 | Cl Cl Cl Cl   | 0.6563 | 11.87 | 0.0184 | 0.8096 | 129.97 | -3.61 | 0.4674 | 0.0184 |
| PCB136 | Cl Cl Cl Cl   | 0.6257 | 12.09 | 0.0635 | 2.7897 | 130.07 | -3.66 | 0.4467 | 0.0635 |
| PCB137 | Cl Cl Cl Cl   | 0.7329 | 11.53 | 0.0041 | 0.1782 | 131.14 | -3.59 | 0.4447 | 0.0041 |
| PCB138 | Cl Cl Cl Cl   | 0.7403 | 11.58 | 0.0012 | 0.0545 | 129.81 | -3.60 | 0.5090 | 0.0012 |
| PCB139 | Cl Cl Cl Cl   | 0.6707 | 12.15 | -0.0127 | -0.5568 | 130.26 | -3.67 | 0.5117 | -0.0127 |
| PCB140 | Cl Cl Cl Cl   | 0.6707 | 11.30 | -0.0324 | -1.4222 | 129.77 | -3.59 | 0.5194 | -0.0324 |
| PCB141 | Cl Cl Cl Cl   | 0.7200 | 12.11 | -0.0267 | -1.1744 | 130.59 | -3.66 | 0.4854 | -0.0267 |
| PCB142 | Cl Cl Cl Cl   | 0.6848 | 11.60 | -0.0088 | -0.3869 | 129.80 | -3.61 | 0.4920 | -0.0088 |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1)& Eq(2); $\hat{Y}_{1d, 2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; Y = relative retention time: experimental [30]; Y - $\hat{Y}_{1d, 2d}$ = residuals.
| Mol       | PCB structure | Y     | iIDRwHg | \( \hat{Y}_{1d} \) | Y- \( \hat{Y}_{1d} \) | ISDmsHt | lADrtHg | \( \hat{Y}_{2d} \) | Y- \( \hat{Y}_{2d} \) |
|-----------|---------------|-------|---------|----------------|---------------------|---------|---------|----------------|----------------|
| PCB143    | Cl Cl Cl Cl   | 0.6789| 12.57   | 0.0004 | 0.0168 | 130.50 | -3.74   | 0.4330 | 0.0004 |
| PCB144    | Cl Cl Cl Cl   | 0.6563| 13.43   | 0.0088 | 0.3881 | 128.67 | -3.82   | 0.4362 | 0.0088 |
| PCB145    | Cl Cl Cl Cl   | 0.6149| 12.87   | -0.0562| -2.4723| 128.32 | -3.76   | 0.5201 | -0.0562|
| PCB146    | Cl Cl Cl Cl   | 0.6955| 12.62   | -0.0098| -0.4320| 128.24 | -3.75   | 0.4749 | -0.0098|
| PCB147    | Cl Cl Cl Cl   | 0.6608| 14.04   | -0.0314| -1.3821| 126.70 | -3.90   | 0.4924 | -0.0314|
| PCB148    | Cl Cl Cl Cl   | 0.6243| 10.02   | -0.0122| -0.5363| 133.20 | -3.42   | 0.1119 | -0.0122|
| PCB149    | Cl Cl Cl Cl   | 0.6672| 10.60   | 0.0041 | 0.1800 | 134.27 | -3.47   | 0.1503 | 0.0041 |
| PCB150    | Cl Cl Cl Cl   | 0.5969| 9.96    | 0.0376 | 1.6541 | 135.23 | -3.40   | 0.1561 | 0.0376 |
| PCB151    | Cl Cl Cl Cl   | 0.6499| 10.14   | 0.0054 | 0.2377 | 134.89 | -3.42   | 0.2191 | 0.0054 |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1)& Eq(2); \( \hat{Y}_{1d, 2d} \) = relative retention time: estimated by Eq(1) and Eq(2), respectively; Y = relative retention time: experimental [30]; Y- \( \hat{Y}_{1d, 2d} \) = residuals.
Table 1. (Continued)

| Mol   | PCB structure                   | Y     | iIDRwHg $\hat{Y}_{1d}$ | Y - $\hat{Y}_{1d}$ | ISDmsHt lADrtHg $\hat{Y}_{2d}$ | Y - $\hat{Y}_{2d}$ |
|-------|---------------------------------|-------|------------------------|---------------------|---------------------------------|---------------------|
| PCB152| ![PCB152 structure](image)      | 0.6062| 9.75                   | 0.0251              | 1.1035                         | -3.41              |
| PCB153| ![PCB153 structure](image)      | 0.7036| 10.15                  | -0.0193             | -0.8496                        | 136.72             |
| PCB154| ![PCB154 structure](image)      | 0.6349| 10.72                  | 0.0028              | 0.1234                         | -3.48              |
| PCB155| ![PCB155 structure](image)      | 0.5666| 10.27                  | -0.0048             | -0.2094                        | 133.35             |
| PCB156| ![PCB156 structure](image)      | 0.8105| 11.12                  | 0.0348              | 1.5315                         | 134.95             |
| PCB157| ![PCB157 structure](image)      | 0.8184| 11.75                  | 0.0333              | 1.4623                         | 133.57             |
| PCB158| ![PCB158 structure](image)      | 0.7429| 11.26                  | 0.0168              | 0.7378                         | 133.12             |
| PCB159| ![PCB159 structure](image)      | 0.7655| 11.52                  | 0.0442              | 1.9425                         | 132.24             |
| PCB160| ![PCB160 structure](image)      | 0.7396| 11.09                  | 0.0065              | 0.2857                         | 134.24             |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1) & Eq(2); $\hat{Y}_{1d,2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; $Y = \hat{Y}_{1d}$, $\hat{Y}_{2d}$ = relative retention time: experimental [30]; $Y - \hat{Y}_{1d,2d}$ = residuals.
Table 1. (Continued)

| Mol   | PCB structure | Y   | iIDRwHg $\hat{Y}_{ld}$ | Y - $\hat{Y}_{ld}$ | ISDmsHt lADrtHg $\hat{Y}_{2d}$ | Y - $\hat{Y}_{2d}$ |
|-------|---------------|-----|------------------------|-------------------|-------------------------------|-------------------|
| PCB161 |               | 0.6968 | 10.98 | -0.0393 | 1.7268 | 133.10 | -3.50 | 0.2766 | -0.0393 |
| PCB162 |               | 0.7737 | 10.98 | 0.0036 | 0.1567 | 131.58 | -3.51 | 0.3351 | 0.0036 |
| PCB163 |               | 0.7396 | 10.45 | 0.0193 | 0.8481 | 132.74 | -3.45 | 0.3432 | 0.0193 |
| PCB164 |               | 0.7399 | 10.97 | -0.0184 | -0.8086 | 133.06 | -3.51 | 0.3582 | -0.0184 |
| PCB165 |               | 0.6920 | 10.72 | -0.0125 | -0.5510 | 131.64 | -3.50 | 0.3503 | -0.0125 |
| PCB166 |               | 0.7572 | 10.16 | 0.0042 | 0.1849 | 132.62 | -3.44 | 0.3003 | 0.0042 |
| PCB167 |               | 0.7814 | 11.56 | 0.0015 | 0.0644 | 132.66 | -3.57 | 0.4155 | 0.0015 |
| PCB168 |               | 0.7068 | 11.09 | -0.0179 | -0.7855 | 133.32 | -3.50 | 0.4314 | -0.0179 |
| PCB169 |               | 0.8625 | 11.05 | 0.0005 | 0.0212 | 131.66 | -3.52 | 0.4262 | 0.0005 |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1) & Eq(2); $\hat{Y}_{ld,2d} = $ relative retention time: estimated by Eq(1) and Eq(2), respectively; Y = relative retention time: experimental [30]; Y - $\hat{Y}_{ld,2d} = $ residuals.
| Mol   | PCB structure | Y     | iIDRwHg $\hat{Y}_{1d}$ | Y - $\hat{Y}_{1d}$ | ISDmsHt $\hat{Y}_{2d}$ | lADrtHg $\hat{Y}_{2d}$ | Y - $\hat{Y}_{2d}$ |
|-------|---------------|-------|-------------------------|---------------------|------------------------|------------------------|---------------------|
| PCB170|               | 0.8740 11.05 | -0.0239 | -1.0517 | 132.19 | -3.51 | 0.4009 | -0.0239 |
| PCB171|               | 0.8089 10.52 | 0.0042 | 0.1838 | 131.49 | -3.46 | 0.3466 | 0.0042 |
| PCB172|               | 0.8278 10.80 | -0.0283 | -1.2445 | 133.28 | -3.48 | 0.4220 | -0.0283 |
| PCB173|               | 0.8152 10.24 | -0.0015 | -0.0653 | 133.94 | -3.42 | 0.3926 | -0.0015 |
| PCB174|               | 0.7965 10.75 | 0.0056 | 0.2482 | 132.13 | -3.50 | 0.3465 | 0.0056 |
| PCB175|               | 0.7611 10.23 | -0.0294 | -1.2916 | 131.25 | -3.45 | 0.4325 | -0.0294 |
| PCB176|               | 0.7305 12.03 | -0.0161 | -0.7060 | 132.78 | -3.65 | 0.3981 | -0.0161 |
| PCB177|               | 0.8031 11.48 | -0.0323 | -1.4195 | 133.66 | -3.57 | 0.3488 | -0.0323 |
| PCB178|               | 0.7537 11.55 | 0.0086 | 0.3793 | 132.00 | -3.60 | 0.4008 | 0.0086 |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1) & Eq(2); $\hat{Y}_{1d, 2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; $Y$ = relative retention time: experimental [30]; $Y - \hat{Y}_{1d, 2d}$ = residuals.
Table 1. (Continued)

| Mol   | PCB structure | Y   | iIDRwHg $\hat{Y}_{1d}$ | Y - $\hat{Y}_{1d}$ | ISDmsHt $\hat{Y}_{2d}$ | lADrtHg $\hat{Y}_{2d}$ | Y - $\hat{Y}_{2d}$ |
|-------|---------------|-----|------------------------|--------------------|------------------------|------------------------|--------------------|
| PCB179| ![PCB structure](image) | 0.7205 | 11.22 | 0.0089 | 0.3932 | 131.75 | -3.57 | 0.3547 | 0.0089 |
| PCB180| ![PCB structure](image) | 0.8362 | 11.25 | 0.0057 | 0.2490 | 132.12 | -3.58 | 0.4106 | 0.0057 |
| PCB181| ![PCB structure](image) | 0.7968 | 12.89 | -0.0103 | -0.4521 | 130.80 | -3.73 | 0.3885 | -0.0103 |
| PCB182| ![PCB structure](image) | 0.7653 | 12.32 | 0.0138 | 0.6063 | 131.52 | -3.67 | 0.4600 | 0.0138 |
| PCB183| ![PCB structure](image) | 0.7720 | 12.42 | 0.0027 | 0.1167 | 130.24 | -3.68 | 0.4348 | 0.0027 |
| PCB184| ![PCB structure](image) | 0.7016 | 11.87 | 0.0184 | 0.8096 | 129.97 | -3.61 | 0.4674 | 0.0184 |
| PCB185| ![PCB structure](image) | 0.7848 | 12.09 | 0.0635 | 2.7897 | 130.07 | -3.66 | 0.4467 | 0.0635 |
| PCB186| ![PCB structure](image) | 0.7416 | 11.53 | 0.0041 | 0.1782 | 131.14 | -3.59 | 0.4447 | 0.0041 |
| PCB187| ![PCB structure](image) | 0.7654 | 11.58 | 0.0012 | 0.0545 | 129.81 | -3.60 | 0.5090 | 0.0012 |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1)& Eq(2); $\hat{Y}_{1d,2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; $Y$ = relative retention time: experimental [30]; $Y - \hat{Y}_{1d,2d}$ = residuals.
| Mol     | PCB structure | Y     | iIDRwHg Ŷ\textsuperscript{1d} | Y - Ŷ\textsuperscript{1d} | ISDmsHt lADrtHg Ŷ\textsuperscript{2d} | Y - Ŷ\textsuperscript{2d} |
|---------|---------------|-------|-------------------------------|--------------------------|---------------------------------|--------------------------|
| PCB188  |               | 0.6920 12.15 | -0.0127 -0.5568 130.26 | -3.67 | 0.5117 -0.0127 |
| PCB189  |               | 0.9142 11.30 | -0.0324 -1.4222 129.77 | -3.59 | 0.5194 -0.0324 |
| PCB190  |               | 0.8740 12.11 | -0.0267 -1.1744 130.59 | -3.66 | 0.4854 -0.0267 |
| PCB191  |               | 0.8447 11.60 | -0.0088 -0.3869 129.80 | -3.61 | 0.4920 -0.0088 |
| PCB192  |               | 0.8269 12.57 | 0.0004 0.0168 130.50 | -3.74 | 0.4330 0.0004 |
| PCB193  |               | 0.8397 13.43 | 0.0088 0.3881 128.67 | -3.82 | 0.4362 0.0088 |
| PCB194  |               | 0.9620 12.87 | -0.0562 -2.4723 128.32 | -3.76 | 0.5201 -0.0562 |
| PCB195  |               | 0.9321 12.62 | -0.0098 -0.4320 128.24 | -3.75 | 0.4749 -0.0098 |
| PCB196  |               | 0.8938 14.04 | -0.0314 -1.3821 126.70 | -3.90 | 0.4924 -0.0314 |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1)& Eq(2); Ŷ\textsuperscript{1d,2d} = relative retention time: estimated by Eq(1) and Eq(2), respectively; Y = relative retention time: experimental [30]; Y - Ŷ\textsuperscript{1d,2d} = residuals.
| Mol   | PCB structure | Y   | iIDRwHg $\bar{Y}_{1d}$ | Y - $\bar{Y}_{1d}$ | ISDmsHt lADrtHg $\bar{Y}_{2d}$ | Y - $\bar{Y}_{2d}$ |
|-------|---------------|-----|------------------------|---------------------|-------------------------------|-------------------|
| PCB197|               | 0.8293 10.02 | -0.0122 0.5363 133.20 | -3.42 0.1119 0.0122 |                               |
| PCB198|               | 0.8845 10.60 | 0.0041 1.8000 134.27 | -3.47 0.1503 0.0041 |                               |
| PCB199|               | 0.8494 9.96 | 0.0376 1.6541 135.23 | -3.40 0.1561 0.0376 |                               |
| PCB200|               | 0.8197 10.14 | 0.0054 0.2377 134.89 | -3.42 0.2191 0.0054 |                               |
| PCB201|               | 0.8875 9.75 | 0.0251 1.1035 133.36 | -3.41 0.2534 0.0251 |                               |
| PCB202|               | 0.8089 10.15 | -0.0193 -0.8496 136.72 | -3.38 0.2902 -0.0193 |                               |
| PCB203|               | 0.8938 10.72 | 0.0028 0.1234 134.60 | -3.48 0.2538 0.0028 |                               |
| PCB204|               | 0.8217 10.27 | -0.0048 -0.2094 133.35 | -3.43 0.2831 -0.0048 |                               |
| PCB205|               | 0.9678 11.12 | 0.0348 1.5315 134.95 | -3.55 0.2222 0.0348 |                               |

iIDRwHg, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1) & Eq(2); $\bar{Y}_{1d, 2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; Y = relative retention time: experimental [30]; Y - $\bar{Y}_{1d, 2d}$ = residuals.
Table 1. (Continued)

| Mol   | PCB structure | Y    | iIDRwHg $\hat{\gamma}_{1d}$ | Y - $\hat{\gamma}_{1d}$ | ISDmsHt | lADrtHg $\hat{\gamma}_{2d}$ | Y - $\hat{\gamma}_{2d}$ |
|-------|---------------|------|----------------------------|--------------------------|---------|-----------------------------|--------------------------|
| PCB206 | ![](image) | 1.0103 11.75 | 0.0333 1.4623 133.57 | -3.57 | 0.1910 | 0.0333 |
| PCB207 | ![](image) | 0.9423 11.26 | 0.0168 0.7378 133.12 | -3.52 | 0.3070 | 0.0168 |
| PCB208 | ![](image) | 0.9320 11.52 | 0.0442 1.9425 132.24 | -3.49 | 0.2856 | 0.0065 |
| PCB209 | ![](image) | 1.0496 11.09 | 0.0065 0.2857 134.24 | -3.49 | 0.3250 | 0.0065 |

$iIDRwHg$, ISDmsHt, and lADrtHg = the value of the descriptor - Eq(1)& Eq(2); $\hat{\gamma}_{1d,2d}$ = relative retention time: estimated by Eq(1) and Eq(2), respectively; Y = relative retention time: experimental [30]; Y - $\hat{\gamma}_{1d,2d}$ = residuals.

Table 2. Training vs Test Experiments: Results.

| No PCBs | Training set | Test set |
|---------|--------------|----------|
|         | Coefficients | Statistics | No PCBs | Statistics |         |         |
|         | Intercept   | ISDmsHt | lADrtHg | $R^2$ | F | $Q^2$ | F |
| 9       | -6.06       | 0.0243  | -1.0294 | 0.999 | 2640$^\dagger$ | 200 | 0.997 | 32807$^\dagger$ |
| 19      | -6.18       | 0.0248  | -1.0442 | 0.998 | 4827$^\dagger$ | 190 | 0.997 | 28567$^\dagger$ |
| 29      | -5.94       | 0.0237  | -1.0195 | 0.996 | 3342$^\dagger$ | 180 | 0.997 | 32047$^\dagger$ |
| 39      | -5.80       | 0.0230  | -1.0040 | 0.998 | 8406$^\dagger$ | 170 | 0.997 | 27678$^\dagger$ |
| 49      | -5.89       | 0.0234  | -1.0172 | 0.998 | 9608$^\dagger$ | 160 | 0.997 | 25667$^\dagger$ |
| 59      | -6.30       | 0.0257  | -1.0448 | 0.996 | 6924$^\dagger$ | 150 | 0.998 | 27578$^\dagger$ |
| 69      | -6.21       | 0.0251  | -1.0440 | 0.996 | 7641$^\dagger$ | 140 | 0.998 | 29667$^\dagger$ |
| 79      | -5.95       | 0.0238  | -1.0170 | 0.996 | 9186$^\dagger$ | 130 | 0.998 | 29915$^\dagger$ |
| 89      | -6.12       | 0.0246  | -1.0348 | 0.997 | 16315$^\dagger$ | 120 | 0.997 | 19049$^\dagger$ |
| 99      | -6.06       | 0.0244  | -1.0278 | 0.997 | 15763$^\dagger$ | 110 | 0.997 | 20314$^\dagger$ |
| 109     | -6.07       | 0.0244  | -1.0304 | 0.996 | 13489$^\dagger$ | 100 | 0.998 | 26764$^\dagger$ |
| 119     | -6.10       | 0.0245  | -1.0361 | 0.997 | 19333$^\dagger$ | 90  | 0.997 | 14990$^\dagger$ |
| 129     | -6.07       | 0.0245  | -1.0284 | 0.997 | 19823$^\dagger$ | 80  | 0.998 | 17306$^\dagger$ |
| 139     | -5.98       | 0.0240  | -1.0219 | 0.997 | 21316$^\dagger$ | 70  | 0.997 | 11610$^\dagger$ |
| 149     | -6.07       | 0.0244  | -1.0297 | 0.997 | 25972$^\dagger$ | 60  | 0.997 | 10077$^\dagger$ |
| 159     | -6.03       | 0.0241  | -1.0287 | 0.997 | 31071$^\dagger$ | 50  | 0.997 | 5692$^\dagger$ |
| 169     | -6.03       | 0.0242  | -1.0258 | 0.997 | 25723$^\dagger$ | 40  | 0.998 | 12671$^\dagger$ |
| 179     | -5.97       | 0.0239  | -1.0203 | 0.997 | 30942$^\dagger$ | 30  | 0.997 | 4938$^\dagger$ |
| 189     | -6.02       | 0.0242  | -1.0247 | 0.997 | 31570$^\dagger$ | 20  | 0.998 | 3383$^\dagger$ |
| 199     | -6.01       | 0.0241  | -1.0243 | 0.997 | 34566$^\dagger$ | 10  | 0.998 | 1450$^\dagger$ |

$\dagger$ p < 0.0001
4. Conclusions

The MDF methodology provides excellent QSPR models, with good stability and predictive ability. It has the disadvantage to be time consuming (it calculates a huge pool of molecular descriptors and provides exhaustive mono- and bivariate regressions) but this is compensated by the high quality of the QSPR models.

Thus, the variance of chromatographic retention time of PCBs is 99.7% explained by two molecular descriptors, showing us that the property is related with geometry and topology, as well as with directly bounded hydrogen’s of PCBs.

The selection of the MDF members from a huge family offers not only a QSPR model, but also a strong instrument to investigate the structural causality of a measured property. Thus, the chromatographic property of PCBs is determined by the molecular topology, geometry and the non-chlorinated (i.e., the remained hydrogenated) positions on the PCB structure.

Acknowledgements

The research was partly supported by UEFISCSU Romania through research projects.

Notes

Virtual library of QSPR/QSAR models:
- http://l.academicdirect.org/Chemistry/SARs/MDF_SARs/sar/
  - Training and test analysis:
    - http://l.academicdirect.org/Chemistry/SARs/MDF_SARs/qsar_qspr_s/

References

1. National Research Council (U.S.). Committee on the Assessment of Polychlorinated Biphenyls in the Environment. Polychlorinated biphenyls: a report; National Academy of Sciences: Washington, 1979.
2. Angulo Lucena, R.; Farouk Allam, M.; Serrano Jiménez, S.; Luisa Jodral Villarejo, M. A review of environmental exposure to persistent organochlorine residuals during the last fifty years. *Curr. Drug Safety* **2007**, *2*(2), 163-172.
3. Roveda, A. M.; Veronesi, L.; Zoni, R.; Colucci, M. E.; Sansebastiano, G. Exposure to polychlorinated biphenyls (PCBs) in food and cancer risk: recent advances. *Igiene e sanità pubblica* **2006**, *62*(6), 677-696.
4. Lundqvist, C.; Zuurbier, M.; Leijis, M.; Johansson, C.; Ceccatelli, S.; Saunders, M.; Schoeters, G.; Ten Tusscher, G.; Koppe, J. G. The effects of PCBs and dioxins on child health. *Acta. Paediatr.* **2006**, *95*(453), 55-64.
5. Poppenga, R. H. Current environmental threats to animal health and productivity. *Vet. Clin. N. Am.-Food A.* **2000**, *16*(3), 545-558.
6. Bren, U.; Zupan, M.; Guengerich, F. P.; Mavri, J. Chemical Reactivity as a Tool to Study Carcinogenicity: Reaction between Chloroethylene Oxide and Guanine. *J. Org. Chem.* **2006**, *71*(11), 4078-4084.
7. Lebeuf, M.; Noël, M.; Trottier, S.; Measures, L. Temporal trends (1987-2002) of persistent, bioaccumulative and toxic (PBT) chemicals in beluga whales (Delphinapterus leucas) from the St. Lawrence Estuary, Canada. Sci. Total Environ. 2007, 383(1-3), 216-231.

8. Tan, J.; Cheng, S. M.; Loganath, A.; Chong, Y. S.; Obbard, J. P. Selected organochlorine pesticide and polychlorinated biphenyl residues in house dust in Singapore. Chemosphere 2007, 68(9), 1675-1682.

9. Borrell, A.; Cantos, G.; Aguilar, A.; Androukaki, E.; Dendrinos, P. Concentrations and patterns of organochlorine pesticides and PCBs in Mediterranean monk seals (Monachus monachus) from Western Sahara and Greece. Sci. Total Environ. 2007, 381(1-3), 316-325.

10. Klánová, J.; Kohoutek, J.; Kostrhounová, R.; Holoubek, I. Are the residents of former Yugoslavia still exposed to elevated PCB levels due to the Balkan wars? Part 1: air sampling in Croatia, Serbia, Bosnia and Herzegovina. Environ. Int. 2007, 33(6), 719-726.

11. Hansch, C. Quantitative approach to biochemical structure-activity relationships. Acc. Chem. Res. 1969, 2(8), 232-239.

12. Hansch, C.; Leo, A. Substituent Constants for Correlation Analysis in Chemistry and Biology. John Wiley & Sons: New York, 1979.

13. Castro, E. A.; Toropov, A. A.; Nesterova A. I.; Nabiev, O. M. QSAR modeling aqueous solubility of polychlorinated biphenyls by optimization of correlation weights of local and global graph invariants. Central European Journal of Chemistry 2004, 2(3), 500-523.

14. Wei, B.; Xie, S.; Yu, M.; Wu, L. QSAR-based prediction of gas/particle partitioning of polychlorinated biphenyls in the atmosphere. Chemosphere 2007, 66(10), 1807-1820.

15. Niu, J. F.; Yang, Z. F.; Shen, Z. Y.; Wang L. L. QSARs for the prediction of photodegradation half-life of PCBs in n-hexane. SAR QSAR Environ. Res. 2006, 17(2), 173-182.

16. Padmanabhan, J.; Parthasarathi, R.; Subramanian, V.; Chattaraj, P. K. QSAR models for polychlorinated biphenyls: n-Octanol/water partition coefficient. Bioorg. Med. Chem. Lett. 2006, 14(4), 1021-1028.

17. Jäntschi, L.; Bolboacă, S. Molecular Descriptors Family on Structure Activity Relationships 6. Octanol-Water Partition Coefficient of Polychlorinated Biphenyls. Leonardo El. J. Pract. Technol. 2006, 8, 71-86.

18. Puri, S.; Chickos, J. S.; Welsh, W. J. Three-dimensional quantitative structure - Property relationship (3D-QSPR) models for prediction of thermodynamic properties of polychlorinated biphenyls (PCBs): Enthalpy of vaporization. J. Chem. Inf. Comp. Sci. 2002, 42(2), 299-304.

19. Padmanabhan, J.; Parthasarathi, R.; Subramanian, V.; Chattaraj, P. K. Using QSAR models to predict the enthalpy of vaporization of 209 polychlorinated biphenyl congeners. QSAR Comb. Sci. 2007, 26(2), 227-237.

20. Puri, S.; Chickos, J. S.; Welsh, W. J. Three-dimensional quantitative structure - Property relationship (3D-QSPR) models for prediction of thermodynamic properties of polychlorinated biphenyls (PCBs): Enthalpy of sublimation. J. Chem. Inf. Comp. Sci. 2002, 42(1), 109-116.

21. Devillers J. A simple method for the prediction of the GLC retention times of all the 209 PCB congeners. Fresenius Z. Anal. Chem. 1988, 332 (1), 61-62.

22. Hasan, M.N.; Jurs, P.C. Computer-assisted prediction of gas chromatographic retention times of polychlorinated biphenyls. Anal. Chem. 1988, 60(10), 978-982.
23. Makino, M. Novel classification to predict relative gas chromatographic retention times and noctanol/water partition coefficients of polychlorinated biphenyls. *Chemosphere* 1999, 39(6), 893-903.

24. Liu, S.-S.; Liu, Y.; Yin, D.-Q.; Wang, X.-D.; Wang, L.-S. Prediction of chromatographic relative retention time of polychlorinated biphenyls from the molecular electronegativity distance vector. *J. Sep. Sci.* 2006, 29(2), 296-301.

25. Ren, Y.; Liu, H.; Yao, X.; Liu, M. An accurate QSRR model for the prediction of the GC\times GCTOFMS retention time of polychlorinated biphenyl (PCB) congeners. *Anal. Bioanal. Chem.* 2007, 388(1), 165-172.

26. Jäntschi, L.; Katona, G.; Diudea, M. Modeling Molecular Properties by Cluj Indices. *MATCH* Commun. Math. Comput. Chem. 2000, 41, 151-188.

27. Jäntschi, L., MDF - A New QSPR/QSAR Molecular Descriptors Family. *Leonardo J. Sci.* 2004, 4, 68-85.

28. Jäntschi, L. Molecular Descriptors Family on Structure Activity Relationships 1. Review of the Methodology. *Leonardo El. J. Pract. Technol.* 2005, 6, 76-98.

29. Jäntschi, L.; Bolboacă, S. Results from the Use of Molecular Descriptors Family on Structure Property/Activity Relationships. *Int. J. Mol. Sci.* 2007, 8(3), 189-203.

30. Mullin, M. D.; Pochini, C. M.; McCrindle, S.; Romkes, M.; Safe, S. H.; Safe, L. M. High resolution PCB analysis: synthesis and chromatographic properties of all 209 PCB congeners. *Environ. Sci. Technol.* 1984, 18, 468-476.

31. HyperChem, Molecular Modelling System [software]; ©2003, Hypercube [cited 2007 June]. Available from: URL: http://hyper.com/products/

32. Chambers, D.L. The practical handbook of genetic algorithms. Chapman & Hall: Boca Raton, 2001.

33. The PHP Group [online]; ©2001-2007, The PHP Group [cited 2007 June]. Available from: URL: http://php.net

34. MySQL AB [online]; ©1995-2007 MySQL AB [cited 2007 June]. Available from: URL: http://mysql.com

35. The FreeBSD Project [online]; ©1995-2007 The FreeBSD Project [cited 2007 June]. Available from: URL: http://freebsd.org

36. Borland Software Corporation [online]; ©1994 - 2007 Borland Software Corporation [cited 2007 June]. Available from: URL: http://borland.com

© 2007 by MDPI (http://www.mdpi.org). Reproduction is permitted for noncommercial purposes.