Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 79 (8) 965–975 (2014)
JSCS–4640
UDC 54–13–72:546.131+547.621:
514.742+511.13:544.341.3
Original scientific paper

QSPR study on the gas/particle partition coefficient of polychlorinated biphenyls using the molecular distance-edge vector index

LONG JIAO[1,2]*, XIAOFEI WANG[1], HUA LI[2] and YUNXIA WANG[2]

[1]*College of Chemistry and Chemical Engineering, Xi'an Shiyou University, Xi'an 710065, P. R. China and* [2]*College of Chemistry and Materials Science, Northwest University, Xi'an 710069, P. R. China*

(Received 27 November, accepted 12 December 2013)

Abstract: The quantitative structure property relationship (QSPR) for the gas/particle partition coefficient, K_p, of polychlorinated biphenyls (PCBs) was investigated. Molecular distance-edge vector (MDEV) index was used as the structural descriptor of the PCBs. The quantitative relationship between the MDEV index and log K_p was modeled by multivariate linear regression (MLR) and artificial neural network (ANN), respectively. The leave-one-out cross validation and external validation were performed to assess the prediction ability of the developed models. When the MLR method was used, the root mean square relative error (*RMSRE*) of the prediction for the leave-one-out cross validation and external validation were 4.72 and 8.62, respectively. When the ANN method was employed, the prediction *RMSRE* of the leave-one-out cross validation and the external validation were 3.87 and 7.47, respectively. It was demonstrated that the developed models are practicable for predicting the K_p of PCBs. The MDEV index was shown to be quantitatively related to the K_p of PCBs.

Keywords: QSPR; PCBs; gas/particle partition coefficient; molecular distance-edge vector index; artificial neural network.

INTRODUCTION

Polychlorinated biphenyls (PCBs) are a group of poly halogenated organic compounds that are persistent organic pollutants. They have gained much attention due to their environmental persistence, high toxicity, tendency to bioaccumulate through the food chain, and risk to human health.[1–3] PCBs are used in various industrial applications as organic diluents, plasticizers, and paint additives, *etc*. It is estimated that hundreds of million of kilograms of PCBs have

* Corresponding author. E-mail: mop@xsyu.edu.cn
doi: 10.2298/JSC130611152J

been released into the environment due to inappropriate disposal practices and accidental releases.[4] Although the production and use of PCBs have been banned in most countries, PCBs have already become ubiquitous pollutants in all environmental compartments, including water, air, vegetation, soil and animals, worldwide.[5] A total of 209 theoretical PCBs exist and approximately 150 have been found in the environment.[6] Being semi-volatile organic compounds, PCBs in the atmosphere are simultaneously present in the gas and particle phase. The gas/particle partition behavior of PCBs in the atmosphere may affect their fate, transport and transformation in the atmosphere. Thus, a quantitative study on gas/particle partition coefficient, K_p, which is a parameter for describing the gas/ /particle partition behavior of PCBs, is useful for an understanding of the environment fate, transport, and transformation of PCBs.[7–9] Many studies focusing on the gas/particle partition coefficient of PCBs have been reported and several methods for determining the K_p of PCBs were proposed.[7–12] However, the experimental determination of K_p is still difficult due to the complexity of the analytical methods and the high cost of experiments.[13] Therefore, attention has been paid to the quantitative structure–property relationship (QSPR) method, which is fast, easy-to-use and cost-effective, for a preliminary assessment and the estimation of K_p of PCBs. Thus, a QSPR model for the K_p of PCBs based on quantum chemical descriptors was recently studied by Wei *et al.*[13] However, it is also time-consuming and complicated to develop a QSPR model based on quantum chemical descriptors, because the calculation and selection of the structural descriptors, which are always time-consuming and complicated, are still necessary to establish the model. Obviously, it is still worthwhile to develop an easy-to-use QSPR model for the K_p of PCBs. The topological index is a structural descriptor that has frequently been used in the field of QSPR research. It can efficiently describe the structure of molecules without the need for detailed molecular orbital calculations. This is useful because, despite its mathematical simplicity, it is able to differentiate molecules with different structures.[14] Thus, the aim of the present study was to develop a QSPR model for the K_p of PCBs based on the topological index. The molecular distance-edge vector (MDEV) index, which has not gained sufficient attention in QSPR research although having been used in a few QSPR studies,[15–17] was introduced and employed as the structural descriptor of PCBs. Multivariate linear regression (MLR) and artificial neural network (ANN) methods were applied in the modeling of the quantitative relationship between the MDEV index and log K_p of PCBs.

CALCULATION AND EXPERIMENTAL

Data set

The MDEV index was calculated according to the approach presented later. The calculated MDEV indexes for the PCBs are listed in Table I. The experimental log K_p values of the 41 PCBs listed in Table II were taken from the literature.[13]

TABLE I. Values of the MDEV index of the investigated PCBs

Group	No.	PCB	μ_1	μ_2
I	1	2,4'	0.0204	2.1510
	2	2,3,4'	0.1471	3.2140
	3	2,2',3,3'	0.3382	4.3470
	4	2,2',3,5'	0.2671	4.3470
	5	2,2',4,5'	0.2063	4.3250
	6	2,3',4,4'	0.2498	4.2540
	7	2,3',4',5	0.2353	4.2760
	8	2,4,4',5	0.2498	4.2540
	9	2,2',3,4,5'	0.6143	5.3870
	10	2,2',4,4',5	0.4127	5.3650
	11	2,2',4',5,5'	0.4056	5.3870
	12	2,3,3',4',6	0.4571	5.3650
	13	2,2',3,3',4,4'	0.7698	6.4270
	14	2,2',3,4,4',5'	0.6987	6.4270
	15	2,2',3,3',4,4',5	1.0470	7.4900
	16	2,2',3,4,4',5,5'	0.9761	7.4900
II	17	2,3'	0.0278	2.1740
	18	2,2',4	0.1229	3.2620
	19	2,2',5	0.1078	3.2850
	20	2,3',4'	0.1059	3.2140
	21	2,4,4'	0.0953	3.1910
	22	2,4',5	0.0760	3.2140
	23	3,4,4'	0.1391	3.1430
	24	2,2',5,5'	0.1960	4.3470
	25	3,3',4,4'	0.2862	4.2050
	26	2,2',3,4',5'	0.4767	5.3870
	27	2,2',3,5',6	0.4374	5.4580
	28	2,3,3',4,4'	0.5080	5.3160
	29	2,3',4,4',5	0.4369	5.3160
	30	2,3',4',4',5'	0.4668	5.3160
	31	2,2',3,4',5',6	0.6674	6.4980
	32	2,2',4,4',5,5'	0.5790	6.4270
	33	2,2',3,5,5',6	0.6991	6.5210
	34	2,3,3',4,4',6	0.7212	6.3650
	35	2,2',3,3',4,5,6'	0.7806	6.4980
	36	2,2',3,3',4,5',6'	1.0160	7.5610
	37	2,2',3,4,4',5',6	0.9519	7.5380
	38	2,2',3,4',5,5',6	0.9375	7.5610
	39	2,3,3′,4,4',5,5'	1.0150	7.3790
	40	2,2',3,3',4,4',5,5'	1.3450	8.5520
	41	2,2',3,3',4,4',6	1.0230	7.5383

The root mean square relative error (*RMSRE*) was used to indicate the prediction performance of the obtained QSPR model. The *RMSRE* is defined as follows:

$$RMSRE = \sqrt{\frac{\sum (RE_i)^2}{n}} \quad (1)$$

where RE_i is the relative error of the i^{th} sample and n is the number of samples.

TABLE II. Experimental and predicted log K_p values of the investigated PCBs

Group	No.	Experimental log K_p	Predicted log K_p	Relative error, %
I	1	–4.60	–4.44	–3.39
	2	–4.03	–4.13	2.53
	3	–3.72	–3.64	–2.03
	4	–3.80	–3.67	–3.39
	5	–3.89	–3.68	–5.41
	6	–3.58	–3.73	4.23
	7	–3.54	–3.72	5.16
	8	–3.63	–3.71	2.21
	9	–3.26	–3.18	–2.49
	10	–3.35	–3.30	–1.38
	11	–3.39	–3.28	–3.28
	12	–3.19	–3.27	2.43
	13	–2.67	–2.79	4.37
	14	–2.79	–2.80	0.36
	15	–2.29	–2.28	–0.49
	16	–2.36	–2.32	–1.69
II	17	–4.45	–4.43	–0.55
	18	–4.27	–4.10	–3.91
	19	–4.35	–4.10	–5.66
	20	–4.03	–4.08	1.27
	21	–4.09	–4.09	0.10
	22	–4.10	–4.10	–0.03
	23	–3.84	–4.07	6.02
	24	–3.88	–3.74	–3.72
	25	–3.10	–3.69	18.90
	26	–3.29	–3.25	–1.17
	27	–3.74	–3.27	–12.69
	28	–3.16	–3.24	2.62
	29	–3.05	–3.28	7.45
	30	–3.25	–3.27	0.51
	31	–2.85	–2.81	–1.40
	32	–2.88	–2.88	0.08
	33	–2.97	–2.79	–6.23
	34	–2.84	–2.82	–0.82
	35	–2.47	–2.74	10.90
	36	–2.57	–2.30	–10.58
	37	–2.50	–2.34	–6.21
	38	–2.63	–2.35	–10.77
	39	–2.04	–2.31	13.31
	40	–1.85	–1.82	–1.42
	41	–2.06	–2.26	9.91

MDEV index

For calculating the MDEV index of a molecule, each non-hydrogen atom is regarded as a point and each chemical bond is regarded as an edge. The whole molecule is regarded as a geometric graph. When dealing with PCBs, the molecular structure of PCBs can be encoded by the MDEV index of chlorine atoms and benzene rings. If the relative electronegativity of each chlorine atom and benzene ring is defined as 1, the *MDEV* index can be defined follows:

$$M_{kl} = \sum_{j \geq i} \left(\frac{1}{d_{ik,jl}^2} \right) \quad k, l = 1, 2 \text{ and } l \geq k) \tag{2}$$

where k and l denote the type of an atom (k = 1 or l = 1 denotes a chlorine atom, and k = 2 or l = 2 denotes a benzene ring). In addition, i and j are the coding number of a chlorine atom or benzene ring in the molecular skeleton graph. Moreover, i and j belong to the k^{th} and l^{th} type, respectively. The term $d_{ik,jl}$ expresses the shortest relative distance between the i^{th} and j^{th} atom. For example, $d_{i1,j1}$ is the nearest relative distance between the i^{th} and j^{th} chlorine atom. The relative bond length between the two adjacent non-hydrogen atoms is defined as d = 1. According to Eq. (2), there are three elements, M_{11}, M_{12} and M_{22} in the MDEV index for a PCB molecule. The three elements are usually noted as μ_1, μ_2 and μ_3, respectively. For example, the MDEV index of 2,2′,3,5′-PCB should be calculated as follows:

$$\begin{aligned} \mu_1 = M_{11} &= \left(\frac{1}{3}\right)^2 + \left(\frac{1}{5}\right)^2 + \left(\frac{1}{6}\right)^2 + \left(\frac{1}{6}\right)^2 + \left(\frac{1}{7}\right)^2 + \left(\frac{1}{5}\right)^2 = 0.6271 \\ \mu_2 = M_{12} &= \left(\frac{1}{1}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{1}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{1}\right)^2 + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{1}\right)^2 + \left(\frac{1}{4}\right)^2 = 4.3472 \\ \mu_3 = M_{22} &= \left(\frac{1}{1}\right)^2 = 1 \end{aligned} \tag{3}$$

Obviously, the M_{22} value of every PCB is 1. Thus, μ_1 and μ_2 are used to describe the structure of PCBs.

Artificial neural networks

As the theory of ANN has been described in detail by many introductory articles,[18-28] only a brief outline of an ANN is presented here.

An ANN is a multivariate calibration method capable of modeling complex functions. Its basic processing unit is the neuron (node). An ANN comprises a number of neurons organized in different layers. A linear artificial neural network (L-ANN),[24-28] which is usually used to develop a linear model, is a neural network having no hidden layers, but an output layer with fully linear neurons (that is, linear neurons with a linear activation function). In an L-ANN, the neurons between the input and output layers fully connect, while the neurons in the same layer do not. The basic architecture of an L-ANN is presented in Fig. 1.

In Fig. 1, x_i (i = 1,2,…,n), y_j (j = 1,2,…,m) and w_{ij} are the input variables, output variables and the element of connection weight matrix **W**, respectively. b_j is the bias vector, which corresponds to the thresholds. The symbol $f_{act}(¤)$ denotes the activation function. Prior to the training procedure, the input and output variables require normalization. When the network is executed, it effectively multiplies the input variables by the weights matrix, and then adds the bias vector. Hence, the postsynaptic potential (PSP) function of the neuron could be described as follows:

$$v_j = \sum_{i=1}^{n} x_i \mathbf{W}_{ij} + b_j \tag{4}$$

Routinely, the activation function used in an L-ANN is a linear function described as:

$$y_j = v_j \tag{5}$$

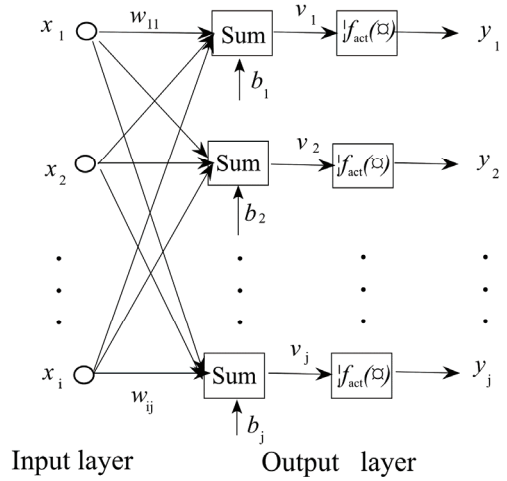

Fig. 1. The architecture of the linear artificial neural network.

As there are no non-linear functions and hidden neurons in the network, an L-ANN is ideal for dealing with linear problems. Actually, training a linear network means finding the optimal setting for the weight matrix **W** to minimize the root mean squared error of the calibration set. In order to achieve this aim, the known samples are generally divided into two parts: a training set and a verification set. The network is trained using the training set, but is also tested after each epoch using the verification set. The training should be stopped once deterioration in the verification error is observed. Overfitting and overlearning can be effectively avoided in this way. Although the verification set is used to identify the best network, actually, training algorithms do not use the verification set to adjust network weights. A standard pseudo-inverse linear optimization algorithm[24] is usually employed to train the network. This algorithm uses the singular value decomposition technique to calculate the pseudo--inverse of the matrix required to set the weights in a linear output layer, to find the least mean squared solution. Essentially, it guarantees the optimal setting for the weights in a linear layer is found.

The main difference between MLR and L-ANN is the optimization algorithm. In MLR, the aim of least square algorithm is to minimize the sum of squared residuals of the training set. As for L-ANN, the aim of training algorithm is to minimize the root mean squared error of verification set.[24] Thus, the prediction ability of L-ANN is usually superior to that of MLR.

In this study, a 2-1 L-ANN (*i.e.*, a network with 2 input variables and 1 output variable) was used to establish the calibration model. The used activation function is presented in Eq. (5). The MDEV index and log K_p were used as the input and output variables, respectively. Prior to the training procedure, the input and output values were normalized.

Leave-one-out cross validation

The leave-one-out cross validation[29] is a commonly used algorithm for estimating the predictive performance of a multivariable calibration model. Usually, practical calibration experiments must be based on a limited set of available samples. The idea behind the leave-one-out cross validation algorithm is to predict the property value of each sample in turn with the calibration model that was developed with the other samples. When applying the algorithm to a dataset with n samples, the calibration modeling is performed n times, each time using (n–1) samples for modeling and one sample for testing. Thus, the procedure of leave-one-out cross validation can be divided into n segment. In each segment i (i = 1,…,n), there are three steps: 1) taking sample i out as temporary "test set", which was not used to develop the calibration model, 2) developing the calibration model with the remaining (n–1) samples and 3) testing the developed model with sample i, calculating and storing the prediction error of the sample.

External validation

External validation[25,30,31] is a conventional approach that is usually used to assess the predictive ability of a calibration model. When using the approach, all the available samples are split into two subsets: a calibration set, which is used to establish the prediction model, and a test set, which is employed to verify the reliability of the established model. Usually, the samples in calibration set and test set are randomly selected from the working dataset.

Software

All data processing was realized with subroutines developed under Matlab (Ver. 7.0). The computation was performed on a personal computer equipped with a Core2 T9400 processor.

RESULTS AND DISCUSSION

First, the values of the MDEV index for the PCBs were calculated according to the method presented in the Calculation and Experimental section and the results are listed in Table I. The MDEV index of different PCB molecules are different, which confirms that the MDEV index can describe the structural differences among these molecules. Thus, it is reasonable to employ MDEV index as structural descriptor for developing QSPR models of PCBs.

Then the two QSPR models (the MLR and L-ANN models) were developed and investigated.

MLR model

Generally, a simple model should always be chosen in preference to a complex model, unless the latter fits the data better. Thus, the MLR model was first investigated to determine whether it could describe adequately the quantitative relationship between the MDEV indexes and log K_p values of the PCBs. The MDEV index was used as the independent variable and the log K_p value as the dependent variable to develop the model. In order to assess the predictive ability of the developed QSPR model, two validation methods, the leave-one-out cross validation and external validation, were performed. The 41 PCBs for which their

log K_p value is known were randomly divided into two groups: Group I, which comprised 20 PCBs and Group II, which comprised 21 PCBs.

Then, the leave-one-out cross validation was applied to Group I and the results are presented in Table II, from which it could be seen that the predicted log K_p were in agreement with the experimental values. For all the 20 compounds, the obtained *RMSRE* was 4.72. Moreover, the predicted log K_p values were plotted *vs*. the corresponding experimental ones (shown in Fig. 2a) and the plot exhibited a linear relationship ($y = 0.9592x - 0.1314$ with $R^2 = 0.9519$). Subsequently, external validation was performed to assess further the predictive ability of the MLR model. In this procedure, the calibration model was developed by using all the 20 compounds in Group I. The obtained regression equation is $\log K_p = 0.883\mu_1 + 0.244\mu_2 - 5.056$. The R^2, standard error of the estimate and F values of the regression model were 0.9636, 0.1501 and 225.0, respectively. Then, the log K_p values of the PCBs in Group II were predicted by the obtained model. The prediction results are also given in Table II, from which it can be seen that the predicted log K_p values were still in agreement with the experimental log K_p. For the 21 compounds, the prediction *RMSRE* was 8.62. The plot of the predicted log K_p *vs*. the experimental values (Fig. 2a) exhibited a linear relationship ($y = 0.9649x - 0.2291$ with $R^2 = 0.9038$).

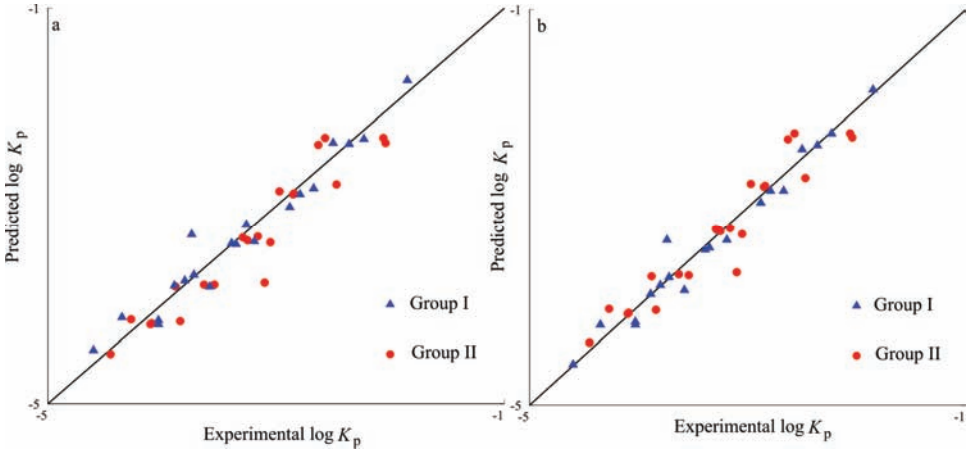

Fig. 2. Experimental log K_p *vs*. predicted log K_p values, a) MLR and b) L-ANN.

The result of the leave-one-out cross validation and external validation demonstrates that the values of MDEV index of the investigated PCBs are quantitatively related to their log K_p values. MLR was shown to be practicable for modeling the quantitative relationship between the MDEV index and log K_p value of PCBs. Obviously, a linear QSPR model based on MDEV index could be used to predict the log K_p of other PCBs.

Thus, an MLR model was developed using the 41 samples listed in Table II. The obtained regression equation was log K_p = 1.046μ_1 + 0.196μ_2 – 4.825. The R^2, standard error of the estimate and F values of the regression model were 0.9282, 0.1958 and 245.5, respectively.

L-ANN model

Besides MLR, an L-ANN, which is the simplest artificial neural network, is a commonly used linear calibration method in QSPR studies. Thus, we investigate whether a better model can be established by using L-ANN. Leave-one-out cross validation and external validation were carried out to assess the predictive ability of the developed model.

Group I was still used to accomplish leave-one-out cross validation. In each run of the ANN, the verification set comprised four randomly selected samples, and the learning rate and momentum were set as 0.6 and 0.3, respectively. The results are presented in Table II, from which it could be seen that the predicted log K_p values were in good agreement with the corresponding experimental ones. For all the 20 compounds, the obtained *RMSRE* was 3.87. Moreover, the predicted log K_p values were plotted *vs*. the experimental ones (Fig. 2b) and the plot exhibited a linear relationship (y = 0.9511x – 0.1613 with R^2 = 0.9725). Then, all the 41 samples listed in Table II were used to accomplish an external validation. The L-ANN model was trained using the 20 samples of Group I. In the training procedure, the verification set comprises four randomly selected samples, and the learning rate and momentum was set as 0.6 and 0.3, respectively. The results of the external validation are given in Table II, from which it could be seen that the predicted log K_p values were in good agreement with the experimental ones. For the 21 compounds, the prediction *RMSRE* was 7.47. The plot of the predicted log K_p values *vs*. the experimental ones (Fig.2b) showed a linear relationship (y = = 0.9574x – 0.1435 with R^2 = 0.9136). Obviously, the prediction accuracy of L-ANN model was slightly higher than that of the MLR model. Hence, using the L-ANN model would be slightly better than the MLR model for the prediction of the log K_p values of PCBs. Thus, it was demonstrated that the L-ANN technique is a practicable and promising method for modeling the quantitative relationship between the values of the MDEV index and the log K_p values of PCBs.

Thus, an L-ANN model was developed using the 41 samples listed in Table II as the calibration set. In the training procedure, the verification set comprised eight randomly selected samples, and the learning rate and momentum was set as 0.6 and 0.3, respectively. Wei *et al.*[13] predicted the log K_p values of the PCBs using a quantum chemistry QSPR model. A comparison of the present result with theirs showed no significant difference between the two sets of predicted results. Thus, the models developed within the present study should provide a reasonable method for the prediction of the log K_p value of PCBs. Compared with quantum chemical descriptors, the MDEV index is easier to generate. Obviously, the

development of QSPR models based on the MDEV index is a simple and easy-to-use method for predicting the log K_p values of PCBs.

CONCLUSIONS

QSPR models for predicting the gas/particle partition coefficient of PCBs were investigated. The MDEV index was used as the structural descriptor of the PCBs. Both a MLR and an L-ANN model were investigated. The predictive ability of the developed models was assessed by the leave-one-out cross validation and by external validation. The results indicate that both the MLR and the L-ANN models are practicable for predicting the log K_p values of PCBs. It is demonstrated that the values of the MDEV index of PCBs are quantitatively related to the log K_p values of PCBs. MDEV index can be generated easier than quantum chemical descriptors. Accordingly, using MDEV index as structural descriptor is more convenient than using a quantum chemical descriptor when developing a QSPR model for the K_p of PCBs. In addition, the result demonstrated that both the MLR and L-ANN models are practicable for modeling the quantitative relationship between the MDEV index and log K_p of PCBs. Therefore, the proposed method should be easy-to-use and practicable for the prediction of the log K_p values of PCBs.

Acknowledgments. The work was supported by the National Natural Science Foundation of China, Nos. 21305108 and 20975081, the project supported by Natural Science Basic Research Plan in Shaanxi Province of China (Program No. 2014JM2039) and the Innovative Research Team of Xi'an Shiyou University.

ИЗВОД

QSPR СТУДИЈА ПОДЕОНОГ КОЕФИЦИЈЕНТА ГАС/ЧЕСТИЦА ПОЛИХЛОРОВАНИХ БИФЕНИЛА ПРИМЕНОМ ВЕКТОРСКОГ ИНДЕКСА РАСТОЈАЊЕ-ГРАНА

LONG JIAO[1,2], XIAOFEI WANG[1], HUA LI[2] и YUNXIA WANG[2]

[1]*Xi'an Shiyou University, Xi'an, China u* [2]*Northwest University, Xi'an, China*

Истраживане су QSPR релације за подеони коефицијент гас/честица, K_p, за полихлороване бифениле (PCB). Као структурни дескриптор коришћен је векторски индекс растојање-грана (MDEV индекс). Квантитативна релација између MDEV индекса и log K_p моделована је методом мултиваријантне линеарне регресије и неуронским мрежама. Показано је да се модел који је развијен може употребити за предвиђање K_p вредности код PCBs. Постоји квантитативна веза између MDEV индекса и K_p за PCB.

(Примљено 27. Новембра, прихваћено 12.децембра 2013)

REFERENCES

1. Environmental protection agency of United States, *Persistent organic pollutants*, http://www.epa.gov/international/toxics/pop.pdf (accessed 26 January 2013)
2. M. S. El-Shahawi, A. Hamza, A. S. Bashammakh, W. T. Al-Saggaf, *Talanta* **80** (2010) 1587
3. S. Jensen, A. G. Johnels, M. Olsson, G. Otterlind, *Nature* **224** (1969) 247

4. X. Y. Han, Z. Y. Wang, Z. C. Zhai, L. S. Wang, *QSAR Comb. Sci*. **25** (2006) 333
5. M. Mandalakis, M. Tsapakis, A. Tsoga, E. G. Stephanou, *Atmos. Environ.* **36** (2002) 4023
6. B. G. Hansen, A. B. Paya-Perez, M. Rahman, B. R. Larsen, *Chemosphere* **39** (1999) 2209
7. J. He, R. Balasubramanian, *Atmos. Environ.* **43** (2009) 4375
8. R. L. Falconer, T. Harner, *Atmos. Environ*. **34** (2000) 4043
9. M. Simick, T. Franz, H. X. Zhang, S. J. Eisenreich, *Environ. Sci. Technol*. **32** (1998) 251
10. T. Harner, T. F. Bidleman, *Environ. Sci. Technol*. **32** (1998) 1494
11. H. G. Yeo, M. Choi, M. Y. Chun, Y. Sunwoo, *Atmos. Environ*. **37** (2003) 3561
12. K. S. Kim, S. Masunaga, *Environ. Pollut*. **138** (2005) 290
13. B. N. Wei, S. D. Xie, M. Yu, L. Wu, *Chemosphere* **66** (2007) 1807
14. I. Gutman, J. Tosovic, *J. Serb. Chem. Soc*. **78** (2013) 805
15. H. H. Liu, X. Xiao, J. Qin, Y. M. Liu, *J. Chongqing Inst. Tech*. **19** (2005) 67 (in Chinese)
16. S. S. Liu, H. L. Liu, Z. N. Xia, C. Z. Cao, Z. L. Li, *J. Chem. Inf. Comput. Sci*. **39** (1999) 951
17. C. S. Yin, W. M. Guo, T. Lin, S. S. Liu, R. Q. Fu, Z. X. Pan, L. S. Wang, *J. Chin. Chem. Soc*. **48** (2001) 739
18. M. Jalali-Heravi, M. Asadollahi-Baboli, P. Shahbazikhah, *Euro. J. Med. Chem*. **43** (2008) 548
19. L. Jiao, *Chemosphere* **80** (2010) 671
20. M. H. Fatemi, Z. G. Chahi, *SAR QSAR Environ. Res*. **23** (2012) 155
21. R. Smiljanić, D. Lazić, M. Gligorić, M. Jotanović, Z. Živković, I. Mihajlović, *J. Serb. Chem. Soc*. **76** (2011) 1163
22. X. J. Yao, A. Panaye, J. P. Doucet, H. F. Chen, R. S. Zhang, B. T. Fan, M. C. Liu, Z. D. Hu, *Anal. Chim. Acta* **535** (2005) 259
23. M. Fatemi, Z. Ghorbannezhad, *J. Serb. Chem. Soc*. **76** (2011) 1003
24. Statsoft, *Model extremely complex functions neural networks*, http://www. statsoft.com/ /textbook/neural-networks (accessed 26 January 2013)
25. L. Jiao, H. Li, *Chemometr. Intell. Lab. Syst*. **103** (2010) 90
26. Y. X. Zhang, H. Li, A. X. Hou, J. Havel, *Chemometr. Intell. Lab. Syst*. **82** (2006) 165
27. C. S. Yin, Y. Shen, S. S. Liu, Q. S. Yin, W. M. Guo, Z. X. Pan, *Comput. Chem*. **25** (2001) 239
28. W. J. Zhang, X. Q. Zhong, G. H. Liu, *Stoch. Environ. Res. Risk, A* **22** (2008) 207
29. H. A. Martens, P. Dardenne, *Chemometr. Intell. Lab. Syst*. **44** (1998) 99
30. E. Kolossov, R. Stanforth, *SAR QSAR Environ. Res*. **18** (2007) 89
31. P. Gramatica, *QSAR Comb. Sci*. **26** (2007) 694.