Quantitative Structure–Property Relationship Study of Retention Time of Some Pesticides in Gas Chromatography

M.R. Hadjmohammadi, M.H.Fatemi, and **K. Kamel***
Department of Chemistry, University of Mazandaran, P.O. Box 453, Babolsar, Iran

Abstract

A quantitative structure–property relationship (QSPR) study based on multiple linear regression (MLR) and artificial neural network (ANN) techniques is carried out to investigate the retention time behavior of some pesticides on the DB-5ms fused-silica column in gas chromatography. Five descriptors selected in the MLR model are: first component WHIM index (E1v), highest eigenvalue n.7 of burden matrix / weighted by atomic van der waals volume (BEHv7); average connectivity index Chi-2 (X2a), 3D-MoRSE signal 23 weighted by atomic Sanderson electronegativity (MoR23m); and principal moments of inertia B (PMIB). A 5-5-1 ANN is also generated to investigate the retention behavior of described pesticides using the same descriptors MLR model as inputs. The statistical parameters derived from MLR and ANN for all molecules are: correlation coefficient $(R)_{MLR}$ = 0.929, standard errors $(SE)_{MLR}$ = 3.452, R_{ANN} = 0.943, and SE_{ANN} = 3.112. The mean of relative errors between the MLR and ANN calculated and the experimental values of the retention times for the prediction set are 13.8% and 9.04%, respectively. The correlation coefficient and standard error of ANN model compared with MLR models showed the superiority of ANNs over regression models. This is partly due to the fact that ANN considers the interaction between different parameters as well as nonlinear relation.

Introduction

Pesticide is a term used for a broad range of chemicals (synthetic or natural) that serve to control insects, fungi, bacteria, and other pests. Monitoring of pesticide residue is one of the most important aspects in minimizing potential hazards to human health. Numerous analytical methods for determination pesticide residue have been published (1–3). The one most frequently used is gas chromatography (GC).

Quantitative structure–property relationships (QSPRs) have been demonstrated to be a powerful tool for the investigation of the chromatographic parameters. QSPRs have been used to obtain simple models to explain and predict the chromatographic behavior of various classes of compounds. QSPRs have been used extensively to explain separation mechanisms, predict retention behavior, and characterize the physicochemical properties of solutes in thin-layer chromatography (4), GC (5,6), and high-performance liquid chromatography (7,8). Also, there are some reports on QSPR studies in capillary electrophoresis (9–11). QSPR study cannot only develop a method for the prediction of the property of interests but also can identify and describe important structural features of molecules that are responsible for variations in molecular properties. The advantage of this approach over other methods lies in the fact that the descriptors used can be calculated from structure alone and are not dependent on any experiment properties. This method has become very useful in the prediction of physicochemical properties.

An artificial neural network (ANN) consists of many pathways and nodes organized into a sequence of layers. The first layer is an input layer with one node for each variable. The last layer is an output layer consisting of one node for the variable to be investigated. In between layers, there is a series of one or more hidden layer(s) consisting of a number of nodes, which are responsible for learning. Nodes of one layer are connected to the nodes of other layer. Each connection is represented by a number called weight. Initially, a learning phase is defined in which each of the input parameter is applied to a processing element. The weight between these parameters is adjusted until the output is correct. Then the system can be applied to unknowns (12). ANNs have been applied to a wide range chemical problems such as: simulation of mass spectra (13), prediction of carbon-13 NMR shift (14), ion interaction (15), GC (16,17), and liquid chromatography (12,18). It also can be used in classification and pattern recognition.

The goal of the present work was to generate a QSPR model between the molecular based structural parameters and observed retention time of some pesticides on the fused silica column. Then, this ANN was employed for the prediction of retention time of some pesticides.

*Author to whom correspondence should be addressed: email k.kamel@umz.ac.ir.

Table I. Data Set and Corresponding Observed and Predicted Values of the Retention Time

Pesticide name	Retention time	MLR	ANN	%Error (MLR)	%Error (ANN)
Training set					
Ametryn	22.65	19.96	23.19	–11.86	2.38
Bromopropylate	38.57	35.83	35.97	–7.1	–6.75
Buprofezin	31.54	27.33	27.94	–13.33	–11.42
Carbaryl	22.23	21.79	24.54	–1.96	10.40
Carbofuran	17.37	18.48	20.36	6.39	17.24
Chinomethionate	28.04	26.78	25.51	–4.5	–9.04
Atrazine	17.54	12.52	20.55	–28.65	17.17
Coumaphos	43.26	41.75	38.84	–3.48	–10.23
Deltamethrin	47.87	47.29	44.57	–1.21	–6.89
Demeton-s-methyl sulfone	23.20	23.40	25.07	0.85	8.05
Dimethoate	16.81	15.28	19.09	–9.13	13.59
Dioxathion	18.09	26.32	22.00	45.52	21.64
Diphenylamine	13.62	16.12	24.84	18.37	82.38
Disulfoton	19.41	19.70	21.10	1.48	8.72
Ethoprophos	14.01	13.25	17.59	–5.41	25.52
Azinphos-methyl	40.52	38.78	36.08	–4.3	–10.97
Etrimfos	20.25	24.70	23.58	21.98	16.43
Fenobucarb	13.32	15.96	12.58	19.79	–5.55
Folpet	27.76	29.18	28.55	5.12	2.85
Formothion	20.67	21.70	26.74	4.98	29.39
Iprodione	38.21	39.36	36.28	3.02	–5.04
Isoprocarb	11.67	12.01	15.42	2.95	32.15
Lenacil	35.25	33.54	32.64	–4.85	–7.42
Mecarbam	27.94	25.37	24.01	–9.2	–14.08
Methidathion	28.55	30.63	27.35	7.28	–4.21
Methiocarb	23.74	19.71	20.11	–16.98	–15.28
Metolachlor	24.77	24.36	22.91	–1.65	–7.52
Benfluralin	15.58	16.75	14.51	7.52	–6.90
o,p′-DDT	33.40	30.82	29.31	–7.73	–12.24
Oxyfluorfen	31.79	32.47	29.37	2.15	–7.61
Parathion	25.17	25.34	25.04	0.69	–.53
Parathion-methyl	22.00	23.87	24.26	8.48	10.27
Phorate	15.82	14.11	18.42	–10.82	16.45
Phosalone	40.65	36.89	38.99	–9.24	–4.07
p,p′-DDD	33.26	35.18	32.73	5.77	–1.59
p,p′-DDE	30.76	34.42	34.04	11.91	10.67
p,p′-DDT	35.46	33.51	32.49	–5.49	–8.39
Prometryn	22.95	22.05	25.72	–3.91	12.10
Prothiofos	30.41	30.72	27.94	1	–8.11
Quinalphos	27.84	30.06	29.87	7.96	7.28
Bifenthrin	39.16	39.85	34.12	1.77	–12.86
Quizalofob-*p*-ethyl	44.69	40.78	41.69	–8.75	–6.71
Terbufos	18.40	21.42	15.51	16.4	–15.68
Trifluralin	15.45	14.28	12.48	–7.58	–19.24
Bromophos-ethyl	28.85	34.90	25.48	20.98	–11.67
Prediction set					
Chlorofenson	29.80	36.34	32.56	21.95	9.26
Chlorobenzilate	32.82	32.15	31.23	–2.05	–4.86
Lambda-cyhalothrin	42.02	44.84	39.62	6.72	–5.71
Dichloran	16.62	15.57	13.55	–6.33	–18.50
Azinphos-ethyl	42.17	40.30	39.06	–4.43	–7.39
Edifenphos	35.03	30.38	31.23	–13.29	–10.84
EPN	38.60	31.10	37.48	–19.44	–2.89
Ethion	33.81	42.92	37.82	26.95	11.87
Fenitrothion	23.76	29.86	25.25	25.66	6.26
Fenthion	25.00	28.41	28.00	13.63	11.99
Fenvalerate	38.34	29.34	35.05	–23.46	–8.59
Bendiocarb	15.23	16.96	15.31	11.35	.56
Malathion	24.67	23.38	22.17	–5.23	–10.12
Metalaxyl	22.94	22.97	20.90	.12	-8.89
Metribuzin	21.57	18.97	22.92	–12.04	6.26
Oxadiazon	31.37	31.61	32.85	.75	4.72
Permethrin	43.01	40.88	37.87	–4.96	–11.95
Bifenox	39.71	36.70	35.64	–7.58	–10.24
Pirimicarb	20.91	14.00	22.91	–33.04	9.57
Propoxur	13.41	9.63	14.14	–28.21	5.44
Profenofos	30.61	35.30	34.58	15.32	12.97
Simazine	17.17	12.18	21.06	–29.09	22.66
Cis-tetrachlorovinphos	29.37	33.46	31.61	13.94	7.61
Triazophos	34.58	35.32	33.31	2.14	–3.68
Vamidothion	29.06	34.39	32.87	18.33	13.12

Experimental

Data set

The data set of the retention time of pesticides on DB-5ms fused-silica column was taken from the work by X.G. Chu (19). A collection of 70 pesticides was chosen as the data set and was randomly divided into two groups: a training set and a prediction set consisting of 45 and 25 molecules, respectively. The molecules in the data set including different pesticides are shown in Table I. The training set was used for the model generation and the prediction set was used for the evaluation of the generated model.

Descriptors

The molecular descriptors used to search for the best model of the retention time were calculated by the Dragon program

Table II. Specification of Multiple Linear Regression Models

Descriptors	Notation	Coefficient
Constant		57.368 (± 13.276)
First component WHIM index	E1v	48.248 (± 7.286)
Highest eigenvalue n.7 of burden matrix/weighted by atomic van der waals volume	BEHv7	7.566 (± 2.006)
Average connectivity index Chi-2	X2a	–210.247 (± 29.937)
3D-MoRSE signal 23 weighted by atomic Sanderson electronegativity	MoR23m	4.344 (± 1.917)
Principal moments of inertia B	PMIB	–736.915 (± 204.543)

Table III. The Values of the Descriptors that Were Used in this Work

Pesticide	E1v	BEHv7	X2a	MoR23m	PMIC
Training set					
Ametryn	0.4040	2.3320	0.3230	0.0520	0.009305
Bromopropylate	0.4800	2.6170	0.2960	–0.028	0.002905
Buprofezin	0.4420	2.5420	0.3080	–0.316	0.006056
Carbaryl	0.3860	2.2410	0.2930	–0.102	0.012358
Carbofuran	0.3080	2.4100	0.2970	–0.167	0.011960
Chinomethionate	0.3960	2.0480	0.2800	0.0460	0.008852
Atrazine	0.4090	1.9050	0.3310	–0.142	0.011929
Coumaphos	0.6200	2.5810	0.2890	–0.434	0.003270
Deltamethrin	0.6360	2.7870	0.2910	0.2260	0.002236
Demeton-s-methyl sulfone	0.4920	1.7370	0.3270	0.6460	0.006658
Dimethoate	0.3820	1.9210	0.3350	0.6440	0.010072
Dioxathion	0.3030	2.9690	0.3070	–0.171	0.003851
Diphenylamine	0.4740	1.6130	0.3080	–0.262	0.014148
Disulfoton	0.4280	1.9210	0.3410	0.6970	0.005683
Ethoprophos	0.4050	1.8830	0.3410	0.4700	0.011196
Azinphos-methyl	0.5050	2.5050	0.2840	0.1700	0.003986
Etrimfos	0.4030	2.4000	0.3010	–0.241	0.008058
Fenobucarb	0.3680	2.3710	0.3080	–0.266	0.015192
Folpet	0.5290	1.4510	0.2980	0.7790	0.007353
Formothion	0.4960	1.9130	0.3190	0.0850	0.010007
Iprodione	0.5890	2.4310	0.3000	0.2670	0.003937
Isoprocarb	0.3250	2.3940	0.3170	–0.209	0.015729
Lenacil	0.4760	2.3630	0.2740	–0.027	0.009429
Mecarbam	0.4070	2.4660	0.3240	0.1300	0.003714
Methidathion	0.4300	2.2820	0.2970	0.5340	0.006281
Methiocarb	0.3380	2.1650	0.3060	–0.007	0.008116
Metolachlor	0.3330	2.5530	0.2950	0.1080	0.009276
Benfluralin	0.4090	2.3100	0.3100	0.0720	0.005333
o,p′-DDT	0.4510	2.6020	0.3050	0.0860	0.005758
Oxyfluorfen	0.4460	2.6740	0.3050	–0.019	0.003308
Parathion	0.4380	2.1320	0.3140	0.0340	0.004637
Parathion-methyl	0.4150	2.1260	0.3070	–0.155	0.005958
Phorate	0.3800	1.9160	0.3410	0.4630	0.008695
Phosalone	0.5210	2.5360	0.2930	–0.106	0.003714
p,p′-DDD	0.5030	2.6110	0.3020	0.2780	0.005325
p,p′-DDE	0.5390	2.5500	0.3020	–0.163	0.005483
p,p′-DDT	0.5120	2.6110	0.3080	0.0810	0.005302
Prometryn	0.4530	2.4310	0.3320	0.1460	0.008685
Prothiofos	0.4560	2.4220	0.3130	0.4290	0.004118
Quinalphos	0.4570	2.4950	0.2960	–0.224	0.006828
Bifentrine	0.4910	2.7890	0.2900	–0.046	0.001542
Quizalofob-*p*-ethyl	0.5970	2.6990	0.2980	–0.407	0.001889
Terbufos	0.4680	2.4310	0.3540	0.3520	0.005462
Trifluralin	0.3780	2.3050	0.3110	–0.080	0.005426
Bromophos-ethyl	0.6530	1.9210	0.3110	–0.092	0.003690
Prediction set					
Chlorofenson	0.6160	2.4110	0.3060	–0.188	0.005205
Chlorobenzilate	0.4570	2.5910	0.2880	–0.263	0.007029
Lambda-cyhalothrin	0.6210	2.7890	0.2930	–0.162	0.001742
Dichloran	0.2930	0.9400	0.3110	–0.096	0.019515
Azinphos-ethyl	0.5460	2.6490	0.2920	0.1300	0.003563
Edifenphos	0.4820	2.6840	0.3070	–0.097	0.007582
EPN	0.4030	2.6830	0.2990	–0.002	0.004263
Ethion	0.6460	2.5960	0.3250	1.190	0.002842
Fenitrothion	0.5190	2.2660	0.3040	–0.283	0.006177
Fenthion	0.4470	2.4940	0.2980	–0.355	0.007058
Fenvalerate	0.4860	2.5110	0.3220	0.0520	0.004066
Bendiocarb	0.3430	2.4040	0.2970	–0.318	0.015365
Malathion	0.3630	2.5530	0.3210	0.1750	0.005548
Metalaxyl	0.2940	2.6690	0.2980	0.1130	0.008977
Metribuzin	0.3670	2.3960	0.3060	–0.103	0.012819
Oxadiazon	0.5370	2.4310	0.3040	–0.711	0.004153
Permethrin	0.5880	2.7440	0.2940	–0.422	0.002683
Bifenox	0.5560	2.6560	0.2980	–0.445	0.004074
Pirimicarb	0.4290	2.4310	0.3140	0.0340	0.010253
Profenofos	0.5860	2.3880	0.3130	0.0530	0.003844
Propoxur	0.3320	2.4060	0.3250	–0.291	0.016784
Simazine	0.4050	1.8460	0.3200	0.0270	0.015657
Cis-tetrachloro-vinphos	0.4860	2.4800	0.2950	–0.111	0.004901
Triazophos	0.5050	2.6490	0.2980	0.0140	0.005241
Vamidothion	0.5670	2.3520	0.3220	0.5910	0.004071

(20) and MOPAC program (AM1 semi empirical method) (21) on the basis of the minimum energy molecular geometries optimized by the Hyperchem package. Dragon is available software (by Milano Chemometrics and the QSAR Research Group) for the calculation of more than 1600 molecular descriptors. Subsequently, the method of stepwise multiple linear regression (MLR) was used to select the most important descriptors and to calculate the coefficients relating the descriptors to retention time. The descriptors that appear in the best MLR equation are shown in Table II. These descriptors are: first component WHIM index, highest eigenvalue n.7 of burden matrix / weighted by atomic van der Waals volume; average connectivity index Chi-2, 3D-MoRSE signal 23 weighted by atomic Sanderson electronegativity; and principal moments of inertia B. These descriptors were used as inputs for generated ANNs.

A detailed description of the theory behind of these descriptors has been adequately described elsewhere (22). E1v is one of the WHIM descriptors obtained as statistical indices of the atoms projected onto 3 principal components obtained from weighted covariance matrices of the atomic coordinates (23). Molecular descriptor obtained from the positive and negative eigenvalue of the adjacency matrix weighted the diagonal elements with atom weights are named BCUT descriptors (24). BEHv7 is the highest eigenvalue n.7 of burden matrix weighted by atomic van der Waals volume. Topological descriptors were calculated using two-dimensional representation of the molecules (25). MoR23m is one of the 3D-MoRSE descriptors that can be calculated by summing atomic properties viewed by different angular scattering functions. The 3D-MoRSE codes have great potential for representation of molecular structure. It is worth noting that they reflect the three-dimensional arrangement of the atoms of a molecule and do not reflect

the chemical bonds (26). Principal moments of inertia B (PMIB) was calculated using the MOPAC program (Version 6) (21).

ANN generation

The program for the feed-forward neural network that was trained by a back-propagation algorithm was written in FORTRAN 90. This network has five nodes in the input layer and one node in the output layer. Descriptors that appeared in the selected MLR model were used as inputs for the generated ANN, and its output was the retention time for the molecule of interest. The number of nodes in the hidden layer would be optimized. The initial weights were randomly selected from a uniform distribution that ranged between –0.3 and +0.3. The initial bias values were set to be one. These values were optimized during the network training. The value of each input was divided into its mean value to bring the values of the input variables into the dynamic range of the sigmoid transfer function in the ANN. Before training, the network was optimized for the number of nodes in the hidden layer, learning rates, and momentum. Then, the network was trained using the training set to optimize the values of weights and biases. Finally in order to evaluate the prediction power of the ANN, a trained network was employed to calculate the retention time for the prediction set.

Results and Discussion

The data set and corresponding observed and ANN predicted values of the retention time of all pesticides studied in this work are shown in Table I. Table II shows the best MLR models. It can be seen from Table II that five descriptors are used in the MLR model. These descriptors are: E1v, BEHv7, X2a, MoR23, and PMIB. Each of these variables encodes different aspects of the molecular structure. The calculated values of these descriptors are shown in Table III for all the molecules included in the data set.

The next step was the generation of the artificial neural network. Before training the network according to Figure 1, the parameters of the number of nodes in the hidden layer, weights and biases learning rates, and momentum values were optimized, which were 5, 0.7, 0.9, and 0.6 respectively. After optimization of the ANN parameters, the network was trained using a training set for the adjustment of weights and biase values. To control the over fitting of the network during the training procedure, the values of standard error of calibration (SEC) and standard error of prediction (SEP) were calculated and recorded to monitor the extent of the learning after each 100 iterations. Results showed that after 2800 iterations, the SEP values started to increase and over fitting began. To maintain the predictive power of the network at a desirable level, training was stopped at this point. For the evaluation of the

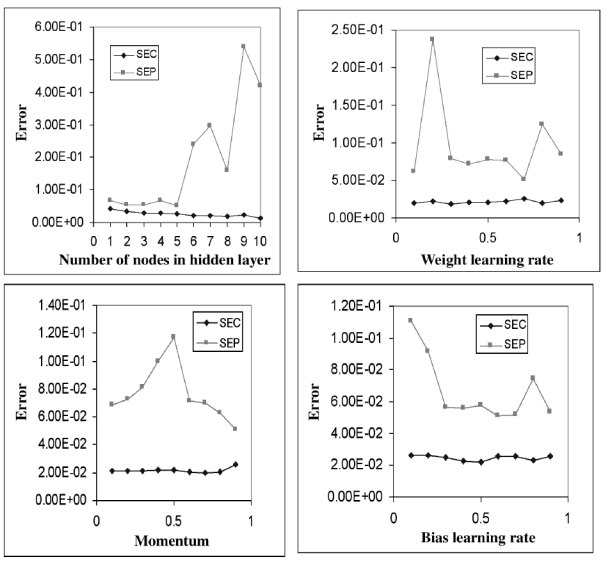

Figure 1. The parameters of the number of nodes in the hidden layer, weights and biases learning rates, and momentum against error.

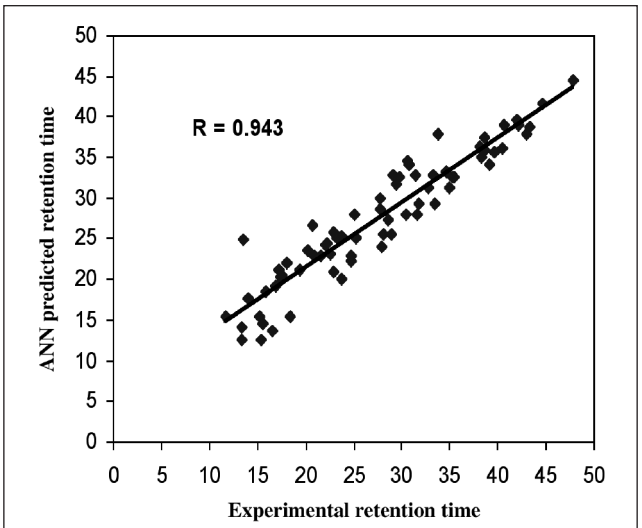

Figure 2. Plot of the ANN predicted retention time against the experimental values.

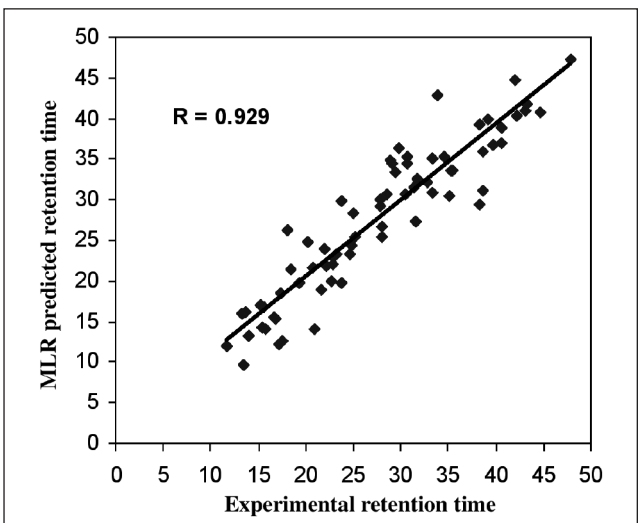

Figure 3. Plot of the MLR predicted retention time against the experimental values.

predictive power of the network, a trained ANN was used to predict the retention time of the pesticide included in the prediction set. Table I represents the experimental, MLR, and ANN predicted values of retention time for the training and prediction set compounds. The correlation coefficients of calibration and prediction for the MLR model were 0.958 and 0.882, and for the ANN model, they were 0.942 and 0.946, respectively, which shows the superiority of the ANN over the MLR model. The standard errors of calibration and prediction for the MLR model is 2.71 and 4.68 should be compared with the values of 3.41 and 2.90, respectively, for the ANN model. In the case of the ANN, the maximum and minimum relative errors for the predicted retention time are 82.38% and 0.53% for diphenylamine and parathion, respectively. However, with the exception of diphenylamine, the predicted values are in agreement with the results obtained by the experiment. Figures 2 and 3 show the plots of the ANN and MLR predicted versus the experimental values for the retention time of the training and prediction set. The residuals of the ANN and MLR calculated values of the retention time were propagated in both sides of zero line that indicates no systematic error exists in the development of the ANN.

Conclusion

The results of this study demonstrate that the QSPR method using the ANN techniques can generate suitable models for the prediction of retention time. The key strength of the neural networks is their ability to allow for flexible mapping of the selected features by manipulating their functional dependence implicitly, unlike regression analysis. Neural network handles both linear and nonlinear relationships without adding complexity to model. This capability offset the larger computing time required and complexity of the ANN method with respect to MLR.

References

1. C.M. Torres and Y. Pico, J. Mañes. Determination of pesticide residues in fruit and vegetables. *J. Chromatogr. A* **754:** 301 (1996).
2. A.I. Valenzula, D.S. Popa, M.J. Redondo, and J. Mañes. Comparison of various liquid chromatographic methods for the analysis of avermectin residues in citrus fruits. *J. Chromatogr. A* **918:** 59 (2001).
3. A.L. Simplicio and LV. Boas. Validation of a solid-phase microextraction method for the determination of organophosphorus pesticides in fruits. *J. Chromatogr. A* **833:** 35 (1999).
4. F.J. Garcia-March, G.M. Antón-fos, F. Pérez-Gimenez, M.T. Salabert-Salvador, R.A. Cercós-del-Pozo, and J.V. de Julián-Oritz. Prediction of chromatographic properties for a group of natural phenolic derivatives by molecular topology. *J. Chromatogr. A* **719:** 45 (1996).
5. T.F. Woloszyn and P.C. Jurs. Quantitative structure-retention relationship studies of sulfur vesicants. *Anal. Chem.* **64:** 3059 (1992).
6. M.H. Fatemi. Quantitative structure-property relationship studies of migration index in microemulsion electrokinetic chromatography. *J. Chromatogr. A* **1002:** 221 (2003).
7. R. Kaliszan. *Quantitative Structure–Chromatographic Retention Relationships*. Wiley-Interscience, New York, NY, 1987.
8. R. Kaliszan, A. Kaliszan, T.A.G. Noctor. Mechanism of retention of benzodiazepines in affinity, reversed-phase and adsorption high-performance liquid chromatography. *J. Chromatogr. A* **609:** 69 (1992).
9. S. Yang, J.G. Bumgarner, L.F.R. Kruk, and M.G. Khaledi. *Quantitative structure-activity relationships studies with micellar electrokinetic chromatography. J. Chromatogr. A* **721:** 323 (1996).
10. P. Lukkari, H. Vuorela, and M.L. Riekkola. Effect of buffer solution pH on the elution and separation of β-blockers by micellar electrokinetic capillary chromatography. *J. Chromatogr. A* **652:** 451 (1993).
11. M. Salo, H. Siren, P. Volin, S. Wiedmer, and H. Vuorela. Structure-retention relationships of steroid hormones in reversed-phase liquid chromatography. *J. Chromatogr. A* **728:** 83 (1996).
12. T. Vasiljevic, A. Onjia, D. Cokesa, and M. Lausevic. Optimization of artificial neural network for retention modeling in high-performance liquid chromatography. *Talanta* **64:** 785 (2004).
13. M. Jalali-Heravi and M.H. Fatemi. Simulation of mass spectra noncyclic alkanes and alkenes using artificial neural network.*Anal. Chim. Acta* **415:** 95 (2000).
14. S.L. Anker and P.C. Jurs. Quantitative structure-retention relationship studies of odor-active aliphatic compounds with oxygen-containing functional groups. *Anal. Chem.* **64:** 1157 (1992).
15. E. Marengo, M.C. Gennaro, and S. Angelino. Neural network and experimental design to investigate the effect of five factors in ion-interaction high-performance liquid chromatography. *J. Chromatogr. A* **799:** 47 (1998).
16. M. Jalali-Heravi and M.H. Fatemi. Artificial neural network modeling of Kovats retention indices for noncyclic and monocyclic terpens. *J.Chromatogr. A* **915:** 177 (2001).
17. W. Guo, Y. Lu, and X. Ming Zheng. The predicting study for chromatographic retention index of saturated alcohols by MLR and ANN. *Talanta* **51:** 479 (2000).
18. Y.L. Loukas. Artificial neural networks in liquid chromatography: efficient and improved quantitative structure-retention relationship. *J. Chromatogr. A* **904:** 119 (2000).
19. X.G. Chu, X.Z. Hu, and H.Y. Yao. Determination of 266 pesticide residues in apple juice by matrix solid-phase dispersion and gas chromatography-mass spectroscopy. *J. Chromatogr. A* **1063:** 201 (2005).
20. R. Todeschini, V. Consonni, A. Mauri, and M. Pavan. Milano Chemometrics and QSAR Reaserch Group,University of Milano-Bicocca,Version 3.0(2003).
21. J.J.P. Stewart. MOPAC, Semi Empirical Molecular Orbital Program, QCPE, Vol. 455 (1983), Research Laboratory, United States Air Force Academy, version 6 (1990).
22. R. Todeschini and V. Consonni. *Handbook of Molecular Descriptors*, Wiley–VCH, Weinheim (2000).
23. R. Todeschini, M. Lasagni, and E. Marengo. New molecular descriptors for 2D and 3D structures. *J. Chemom.* **8:** 263 (1994).
24. R.S. Pearman and K.M. Smith. Metric Validation and the Receptor-Relevant Subspace Concept. *J. Chem. Inf. Comput. Sci.* **39:** 28 (1999).
25. L.B. Kier and L.H. Hall. *Molecular Connectivity in Structure-Activity Analysis*. RSP-Wiley, Chichester, UK,1986
26. J.H. Schuur, P. Selzer, and J. Gasteiger. The Coding of the Three-Dimensional Structure of Molecules by Molecular Transforms and Its Application to Structure-Spectra Correlations and Studies of Biological Activity. *J. Chem. Inf. Comput. Sci.* **36:** 36 (1996).

Manuscript received February 25, 2006;
revision received February 14, 2007.