Asian Journal of Chemistry; Vol. 23, No. 6 (2011), 2571-2576

ASIAN JOURNAL OF CHEMISTRY

www.asianjournalofchemistry.co.in

Quantitative Structure-Retention Relationships Study of Phenols Using Neural Network and Classic Multivariate Analysis

MEHDI ALIZADEH[1], MAHMOOD CHAMSAZ[2] and SAEID ASADPOUR[2,*]

[1]Department of Chemistry, Gachsaran Azad University, Gachsaran, Iran
[2]Department of Chemistry, Ferdowsi University of Mashhad, Mashhad, P.O. Box 91735-654, Iran

*Corresponding author: E-mail: s.asadpour@gmail.com

(*Received*: 5 July 2010; *Accepted*: 10 February 2011) AJC-9613

A quantitative structure-retention relationship (QSRR) study, has been carried out on 50 diverse phenols in gas chromatography (GC) in a dual-capillary column system made of DB-5 (SE-54 bonded phase) and DB-17 (OV-17 bonded phase) fused-silica capillary columns by using molecular structural descriptors. Modeling of retention times of these compounds as a function of the theoretically derived descriptors was established by multiple linear regression (MLR), partial least squares (PLS) regression and artificial neural networks (ANN). Stepwise SPSS was used for the selection of the variables (descriptors) that resulted in the best-fitted models. For prediction retention times of compounds in DB-5 and DB-17 columns, three and four descriptors, respectively were used to develop a quantitative relationship between the retention times and structural properties. Appropriate models with low standard errors and high correlation coefficients were obtained. After variables selection, compounds randomly were divided into two training and test sets and MLR and PLS methods (with leave-one-out cross validation) and ANN used for building of the best models. The predictive quality of the QSRR models were tested for an external prediction set of 10 compounds randomly chosen from 50 compounds. The squared regression coefficients of prediction for the MLR, PLS and ANN models for DB-5 column were 0.9645, 0.9606 and 0.9808, respectively and also for DB-17 column were 0.9757, 0.9757 and 0.9875, respectively. Result obtained showed that non-linear model can simulate the relationship between structural descriptors and the retention times of the molecules in data sets accurately.

Key Words: Molecular descriptors, Retention time, Phenol, Quantitative structure-retention relationship, Artificial neural networks.

INTRODUCTION

As a potential alternative to expensive and time-consuming experimental trial and error approach traditionally adopted to optimize chromatographic separations, retention predictive models have received considerable attention in recent years[1]. An important property that has been extensively studied in QSAR is the chromatographic retention time. A quantitative structure-retention relationship (QSRR) study involves the prediction of chromatographic retention parameters using molecular structure. Quantitative structure-retention relationship studies are widely investigated in gas chromatography (GC) and high-performance liquid chromatography (HPLC)[2]. Chromatographic retention is a physical phenomenon that is primarily dependent on the interactions between the solute and the stationary phase. Molecular group contribution methods are widely employed to estimate gas chromatographic retention parameters[3].

Artificial neural networks (ANNs) are among the best available tools to generate non-linear models. Artificial neural networks are parallel computational devices consisting of groups of highly interconnected processing elements called neurons. Artificial neural networks (ANNs), was inspired by scientist's interpretation of the architecture and functioning of the human brain[4,5]. However, a methodology related to non-linear regression techniques was developed[6,7]. Reviews have been published concerning applications of ANN in different fields[8,9]. Recently, artificial neural networks (ANNs) have been used to a wide variety of chemical problems such as spectral analysis[10], prediction of dielectric constant[11] and mass spectral search[12]. Artificial neural networks have been applied to QSPR analysis due to its flexibility in modeling of non-linear problems, mainly in response to increase accuracy demands. They have been widely used to predict many physico-chemical properties[13-17]. Currently the phenols, a group of ubiquitous compounds in environmental samples mainly because of their heavy uses in the chemical industries, constitute an important class of ground water contaminants[18-20]. The phenolic micropollutants generally include chloro-, bromo-, nitro- and alkyl- phenols. Some of theses are either known or suspected endocrine disrupters. With the growing concern about the

water quality, trace-level analysis of these phenols has become important in recent years[21].

The main aim of the present work is development of a QSRR models by using ANN as non-linear method to predict the retention times of various phenols and comparison with MLR and PLS as linear methods.

In the present work, a QSRR study has been carried out on the GC retention times (t_R) for 50 diverse phenols by using structural molecular descriptors. The two linear methods, MLR and PLS and non-linear method, feed forward neural network with back-propagation training along with stepwise SPSS as variable selection software were used to model the retention times with the structural descriptors.

EXPERIMENTAL

The experimental data of the retention times data (t_R), for 50 chemical compounds including phenols were taken from Kim *et al.*[21], as shown in Table-1. The data set randomly was divided into two subsets in ANN, MLR and PLS: training and test sets including 40 and 10 compounds, respectively.

Multiple linear regression and partial least squares analysis: The multiple linear regression (MLR) is an extension of the classical regression method to more than one dimension[22]. Multiple linear regression calculates QSAR equation by performing standard multivariable regression calculations using multiple variables in a single equation. The stepwise multiple linear regression is a commonly used variant of MLR. In this case, a multiple-term linear equation is also produced, but not all independent variables are used. Each variable is added to the equation at a time and a new regression is performed. The new term is retained only if equation passes a test for significance. This regression method is especially useful when the number of variables is large and when the key descriptors are not known[23].

The PLS model will try to find a few PLS factors (also known as components or latent variables) that explain most of the variations in both predictors and responses. Factors that explain response variation provide good predictive models for new responses and factors that explain predictor variation which are represented by the observed values of the predictors. The partial least squares (PLS) regression method is well suited for problems with multicollinear predictor and response variables. Partial least squares is explained in detail in literature[24,25]. To obtain the PLS model with the best predictive performance, the number of PLS components that optimize the predictive ability of the model should be determined. This is typically done by cross-validation, a procedure in which the available data within the training set are split into several subgroups called validation sets. The prediction residual sum of squares (PRESS) for the test samples is determined as a function of the PLS components retained in the regression model which was formed with the training data. The procedure is usually repeated several times, with each subset in the training set being part of the test samples at least once[26].

Artificial neural networks (ANN): Principles, functioning and applications of artificial neural networks have been adequately described elsewhere[27,28]. The relevant principle of supervised learning in an ANN is that it takes numerical inputs (the training data) and transfers them into desired outputs. The input and output nodes may be connected to any other nodes within the network. The way in which each node transforms its input depends on the so-called 'connection weights' or ' connection strength' and bias of the node, which are modifiable. The output values of each node depend on both the weight strength and bias values. Training of the ANN can be performed by using the backpropagation algorithm. In order to train the network using the back propagation algorithm, the differences between the ANN output and its desired value are calculated after each training iteration and the values of weights and biases modified by using these error terms.

A three-layer feed-forward network formed by one input layer consisting of a number of neurons equal to the number of descriptors, one output neuron and a number of hidden units fully connected to both input and output neurons, were adopted in this study. The most used learning procedure is based on the back-propagation algorithm, in which the network reads inputs and corresponding outputs from a proper data set (training set) and iteratively adjusts weights and biases in order to minimize the error in prediction. To avoid overtraining and consequent deterioration of its generalization ability, the predictive performance of the network after each weight adjustment is checked on unseen data (validation set).

In this work, training gradient descent with momentum is applied and the performance function was the mean square error (MSE), the average squared error between the network outputs and the actual output.

The QSRR models for the estimation of the retention times of various compounds are established in the following six steps: molecular structure input and generation of the files containing the chemical structures stored in a computer-readable format; quantum mechanics geometry optimization with a semi-empirical method; structural descriptors computation; structural descriptors selection; structure-retention models generation with the multivariate methods and statistical analysis.

Computer hardware and software: All calculations were run on a Pentium IV personal computer with windows XP as operating system. The molecular 3D structures of data set were sketched using hyperchem (ver. 7.1), then each molecule was "cleaned up" and energy minimization was performed using the geometry. Optimization was done using semiempirical AM1 (Austin Model) Hamiltonian method. After optimization of structures, several descriptors are computed by hyperchem. Then 3D structures with lower energy conformers obtained by the aforementioned procedure were fed into dragon (ver. 5.2-2005) for calculation of the structural molecular descriptors (constitutional, topological, connectivity, geometrical, getaway and charge descriptors). Through these descriptors which having values exceeding 90 % zero or have equal values further than 90 % are not useful and should be removed. Then Descriptor selection was the accomplished by using Stepwise SPSS (SPSS Ver. 11.5, SPSS Inc.). PLS regression (PLS_Toolbox, version 2.1, Eigenvector Company) and other calculations were performed in the MATLAB (version 7.0, MathWorks, Inc.) environment.

RESULTS AND DISCUSSION

Descriptors selection: Generally the first step in variables selection is the calculation of the correlation between variables

TABLE-1
DATA SET AND CORRESPONDING OBSERVED AND (ANN, MLR, PLS) PREDICTED VALUES OF RETENTION TIMES (t_R)

No.	Name	DB-5 column				Name	DB-17 column			
		t_R (exp)	t_R (ANN)	t_R (PLS)	t_R (MLR)		t_R (Exp)	t_R (ANN)	t_R (PLS)	t_R (MLR)
		Training set					Training set			
1	Phenol	1388	1402.3	1400.7	1411.4	Phenol	1617	1632.8	1667.3	1667.3
2	*m*-Cresol	1484	1461.8	1485.1	1482.7	*o*-Cresol	1676	1712.2	1735.3	1735.3
3	*p*-Cresol	1492	1529.9	1517.8	1533.7	*m*-Cresol	1719	1724.8	1750.3	1750.3
4	2-Ethylphenol	1512	1552.7	1584.2	1579.5	*p*-Cresol	1731	1691.5	1711.2	1711.2
5	2-Chlorophenol	1529	1546.8	1553.4	1554.0	2,6-Dimethylphenol	1733	1785.7	1787.8	1787.8
6	3-Chlorophenol	1546	1540.1	1548.2	1545.9	2-Ethylphenol	1745	1762.2	1755.8	1755.8
7	2,5-Dimethylphenol	1546	1581.3	1606.6	1607.1	2-Chlorophenol	1804	1814.3	1824.2	1824.2
8	2,3-Dimethylphenol	1584	1550.3	1581.2	1567.4	3-Chlorophenol	1780	1811.2	1819.5	1819.5
9	2-Methoxyphenol	1586	1562.4	1570.3	1575.8	2,5-Dimethylphenol	1780	1763.0	1760.8	1760.8
10	4-Ethylphenol	1598	1583.8	1600.3	1604.5	2,3-Dimethylphenol	1836	1800.1	1805.2	1805.1
11	2,3,6-Trimethylphenol	1639	1644.5	1701.3	1685.4	2-Methoxyphenol	1903	1839.3	1825.5	1825.5
12	4-Isopropylphenol	1656	1643.9	1689.1	1682.1	2-*n*-Propylphenol	1821	1885.9	1859.5	1859.5
13	3,5-Dichlorophenol	1674	1701.2	1685.1	1666.8	4-Ethylphenol	1833	1779.7	1778.8	1778.8
14	2,4-Dichlorophenol	1674	1720.7	1718.9	1719.6	2,3,6-Trimethylphenol	1882	1852.1	1826.3	1826.3
15	2,3,5-Trimethylphenol	1676	1641.1	1697.3	1679.2	4-Isopropylphenol	1884	1868.6	1850.5	1850.5
16	4-*n*-Propylphenol	1686	1677.5	1704.5	1705.2	3,5-Dichlorophenol	1942	1952.1	1946.7	1946.7
17	2,3-Dichlorophenol	1710	1691.0	1693.5	1679.9	2,4-Dichlorophenol	1977	1961.0	1962.5	1962.5
18	3,4-Dichlorophenol	1722	1720.7	1718.9	1719.6	4-*n*-Propylphenol	1919	1923.2	1914.6	1914.6
19	2-Nitrophenol	1726	1714.1	1683.0	1691.8	2,3-Dichlorophenol	2002	1970.8	1977.9	1977.9
20	4-*tert*-Butylphenol	1727	1731.3	1779.9	1761.7	3,4-Dichlorophenol	1876	1958.9	1957.5	1957.5
21	3-Nitrophenol	1765	1776.3	1705.5	1726.9	4-*tert*-Butylphenol	1952	1967.2	1948.7	1948.7
22	4-Nitrophenol	1790	1742.3	1693.5	1708.2	3-Nitrophenol	2068	2047.6	2058.9	2058.9
23	4-*n*-Butylphenol	1794	1799.7	1811.3	1807.7	2,4,6-Trichlorophenol	2041	2081.1	2103.6	2103.6
24	2,4,5-Trichlorophenol	1827	1868.4	1874.8	1872.1	4-Nitrophenol	2110	2121.4	2124.4	2124.4
25	2,3,4-Trichlorophenol	1878	1833.0	1852.2	1836.9	4-*n*-Butylphenol	2026	2051.6	2057.5	2057.5
26	1-Naphtol	1895	1900.7	1824.3	1868.0	2,3,4-Trichlorophenol	2176	2092.8	2123.9	2123.9
27	2-Phenylphenol	1931	2038.2	2023.5	2053.2	1-Naphtol	2248	2236.0	2238.1	2238.1
28	Catechol	1961	1979.6	2006.5	2010.9	2-Phenylphenol	2306	2319.3	2384.3	2384.3
29	2,5-Dinitrophenol	1984	1988.4	1950.9	1966.4	Catechol	2298	2361.5	2307.5	2307.5
30	2,4-Dinitrophenol	2021	2009.9	1965.8	1989.5	2,4-Dinitrophenol	2419	2428.3	2492.8	2492.8
31	Hydroquinone	2099	2068.0	2034.0	2053.8	Hydroquinone	2411	2369.1	2305.1	2305.1
32	2-Methylresorcinol	2105	2121.4	2064.8	2055.3	Pentachlorophenol	2448	2448.2	2433.2	2433.2
33	Pentachlorophenol	2137	2137.7	2137.9	2106.4	4-Phenylphenol	2541	2532.9	2434.2	2434.2
34	4-Phenylphenol	2148	2037.7	2022.4	2051.5	Orcinol	2464	2442.4	2361.8	2361.8
35	Pyrogallol	2486	2547.4	2632.4	2626.1	Pyrogallol	2908	2912.6	2970.4	2970.4
36	Phloroglucinol	2692	2631.1	2624.0	2613.0	Phloroglucinol	2994	2990.5	3004.5	3004.5
37	3-Bromophenol	1643	1603.8	1608.5	1597.4	2-Bromophenol	1915	1907.2	1928.3	1928.3
38	4-Bromophenol	1652	1650.4	1641.2	1648.4	4-Bromophenol	1927	1930.4	1952.7	1952.7
39	2-Iodophenol	1720	1745.3	1727.6	1703.0	2-Iodophenol	2063	2019.5	2026.2	2026.2
40	3-Iodophenol	1760	1767.7	1722.5	1694.9	3-Iodophenol	2068	2093.8	2069.7	2069.7
		Test set					Test set			
41	*o*-Cresol	1446	1468.9	1490.3	1490.8	4-Chlorophenol	1804	1838.4	1855.0	1854.9
42	2,6-Dimethylphenol	1504	1550.3	1581.2	1567.4	4-Chloro-*m*-cresol	1917	1895.7	1900.7	1900.7
43	4-Chlorophenol	1555	1609.7	1580.9	1596.8	2,3,5-Trimethylphenol	1928	1933.7	1834.3	1834.3
44	2-*n*-Propylphenol	1591	1669.9	1701.4	1700.4	2-Nitrophenol	2089	2087.2	2100.0	2100.0
45	4-Chloro-*m*-cresol	1665	1642.4	1662.8	1663.3	2,4,5-Trichlorophenol	2087	2080.7	2100.9	2100.9
46	2,4,6-Trichlorophenol	1771	1831.3	1848.2	1830.6	2,5-Dinitrophenol	2375	2253.6	2453.8	2453.8
47	4,6-Dinitro-*o*-cresol	2084	2085.0	2096.9	2127.9	4,6-Dinitro-*o*-cresol	2463	2358.8	2482.1	2482.1
48	Orcinol	2166	2138.3	2056.4	2042.2	2-Methylresrcinol	2403	2286.3	2341.6	2341.6
49	2-Bromophenol	1618	1604.2	1613.7	1605.5	3-Bromophenol	1903	1857.9	1959.5	1959.5
50	4-Iodophenol	1775	1727.0	1755.1	1745.9	4-Iodophenol	2090	2104.4	2067.8	2067.8

TABLE-2
MOLECULAR DESCRIPTORS EMPLOYED FOR THE PROPOSED QSRR MODELS

	Number	Descriptor	Notation	Coefficient
DB-5 Column	1	Hydrophilic factor	hy	576.146
	2	Ghose-Crippen molar refractivity	amr	20.376
	3	Connectivity index chi-5	x5	152.561
		Constant		741.198
DB-17 Column	1	Hydrophilic factor	hy	579.471
	2	Detour index	w	1.733
	3	Qzz COMMA2 value / weighted by atomic masses	QZZm	2.007
	4	Dipole (debyes) sum Y	DSY	18.983
		Constant		1516.534

and with seeking property. In present case, to decrease the redundancy existed in the descriptors data matrix, the correlations of descriptors with each other and with the t_R of the molecules were examined and descriptors which showed high interrelation (*i.e.*, $r > 0.9$) with t_R and low interrelation (*i.e.*, $r < 0.9$) with each other were detected. For each class of the descriptor just one of them was selected for construction the final QSRR model and the rest were deleted. In second step, stepwise SPSS was used for variable selection. After these process for DB-5 and DB-17columns three and four descriptors were remained, respectively, that keeps most interpretive information for retention time. Table-2 shows descriptors and their coefficients that used in MLR method. A correlation analysis was carried out to evaluate correlations between selected descriptors with each other and with retention time (Table 3a-b).

TABLE-3a
CORRELATION MATRIX OF THE THREE DESCRIPTORS AND t_R USED IN THIS WORK FOR DB-5 COLUMN[a]

	hy	amr	x5	t_R
hy	1	-0.447	-0.302	0.759
amr		1	0.768	0.194
x5			1	0.296
t_R				1

[a]The definitions of the descriptors are given in Table-2.

TABLE-3b
CORRELATION MATRIX OF THE FOUR DESCRIPTORS AND t_R USED IN THIS WORK FOR DB-17 COLUMN[a]

	hy	w	QZZm	DSY	t_R
hy	1	-0.210	-0.217	0.118	0.729
w		1	0.398	-0.281	0.454
QZZm			1	-0.110	0.257
DSY				1	0.023
t_R					1

[a]The definitions of the descriptors are given in Table-2.

Artificial neural network optimization: A three-layer neural network was used and starting network weights and biases were randomly generated. Descriptors selected by stepwise method were used as inputs of network and the signal of the output node represent the retention time of phenols. Thus, networks have three and four neurons in input layer for DB-5 and DB-17 columns, respectively and one neuron in output layer. The networks performance was optimized for the number of neurons in the hidden layer (hnn), the learning rate (lr) of back-propagation, momentum and the epoch. As weights and biased are optimized by the back-propagation iterative procedure, training error typically decreases, but test error first decreases and subsequently begins to rise again, revealing a progressive worsening of generalization ability of the network. Thus training was stopped when the test error reaches a minimum value. Table-4 shows the architecture and specification of the optimized networks.

TABLE-4
ARCHITECTURE AND SPECIFICATION OF THE GENERATED ANNs

	DB-5 Column	DB-17 Column
Number of nodes in the input layer	3	4
Number of nodes in the hidden layer	7	4
Number of nodes in the output layer	1	1
Learning rate	0.2	0.3
Momentum	0.9	0.4
Epoch	6000	5000
Transfer function	Sigmoid	Sigmoid

Results of artificial neural network analysis and comparison with multiple linear regression and partial least squares: The non-linear QSRR model provided by the optimal neural networks is presented in Fig. 1(a-b) where computed or predicted retention time values are plotted against the corresponding experimental data. Fig. 2(a-b) shows a plot of residuals *versus* the observed retention time values. The substantial random pattern of this plot indicates that most of the data variance is explained by the proposed models.

The agreement between computed and observed values in ANN training and test sets are shown in Table-1. The statistical parameters calculated for the ANN, MLR and PLS models are presented in Table-5. Goodness of the ANN-based model is further demonstrated by the high value of the correlation coefficient R between calculated and observed t_R values for DB-5 and DB-17columns are (0.990, 0.985) and (0.994, 0.986) for training and test set, respectively. For comparison, a linear QSRR model relating retention times to the selected descriptors were obtained by means of MLR and PLS methods. With the purpose of MLR and PLS models built on the same subsets was used in ANN analysis.

Multiple linear regression (MLR) is one of the most commonly used modeling methods in QSRR. The colinearity problem of the MLR method has been overcome through the development of the partial least-squares (PLS) projections to latent structures method, which has been shown to be an efficient approach in monitoring many complex processes,

TABLE-5
STATISTICAL PARAMETERS OBTAINED USING THE ANN ,MLR AND PLS MODELS[a]

Ft	Fc	R^2t	R^2c	Rt	Rc	SEt	SEc	Model	
261.031	1942.695	0.970	0.981	0.985	0.990	40.670	37.107	ANN	DB-5 Column
142.400	927.516	0.947	0.961	0.973	0.980	49.979	52.669	PLS	
119.197	1032.028	0.937	0.964	0.968	0.982	55.222	50.131	MLR	
290.284	3022.320	0.973	0.988	0.986	0.994	32.40	35.571	ANN	DB-17 Column
152.903	1528.809	0.950	0.976	0.975	0.988	56.835	49.422	PLS	
152.980	1528.707	0.950	0.976	0.975	0.988	56.824	49.424	MLR	

[a]c refers to the calibration (training) set; t refers to test set; R is the correlation coefficient; R^2 is the correlation coefficient square and F is the statistical F value.

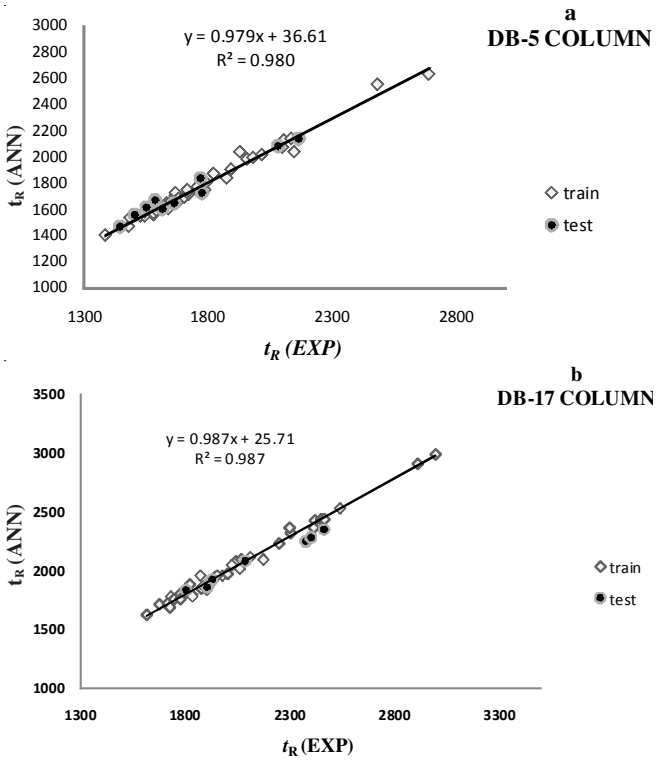

Fig. 1. Plots of predicted t_R estimated by ANN for DB-5 column (a) and ANN for DB-17 column (b) modeling *versus* experimental t_R compounds

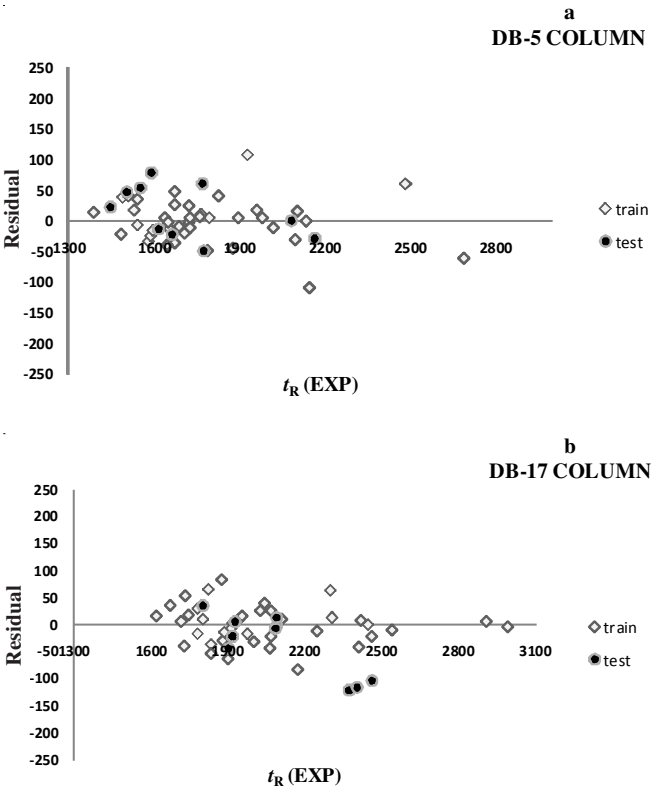

Fig. 2. Plots of residual *versus* experimental t_R in ANN for DB-5 column (a) and ANN for DB-17 column (b) models

reducing the high dimensional strongly cross-correlated data to a much smaller and interpretable set of principal components or latent variables. The number of significant factors for the PLS algorithm was determined using the cross-validation method. The optimum number of factors was concluded as the first local minimum in the PRESS *versus* number of factors. The best PLS models contained three and four selected descriptors in two and four latent variables space for DB-5 and DB-17 columns, respectively.

Comparison between statistical parameters in Table-5 reveals that non-linear ANN model produced better results with good predictive ability than linear models.

Conclusion

Quantitative structure-retention relationship analysis was performed on a series of phenols using ANN, MLR and PLS methods which correlate t_R values of these compound to their structural descriptors. According to the obtained results, it is concluded that the (hy, amr, x5) for DB-5 column and (hy, w, QZZm, DSY) for DB-17 column can be used successfully for modeling t_R property of the under study compounds. The statistical parameters of the built QSRR models were satisfactory showing the high quality of the chosen descriptors. High correlation coefficients and low prediction errors obtained confirm good predictive ability of ANN model. The proposed QSRR models with the simply calculated molecular descriptors can be used to estimate the chromatographic retention times for new compounds even in the absence of the standard candidates.

REFERENCES

1. R. Kaliszan, Quantitative Structure-Chromatographic Retention Relationships, John Wiley & Sons, New York (1987).
2. J. Ghasemi, S. Asadpour and A. Abdolmaleki, *Anal. Chim. Acta*, **588**, 200 (2007).
3. I. Teodora and I. Ovidiu, *Internet Elect. J. Mol. Design*, **1**, 94 (2002).
4. W.S. Mculloch and W. Pitts, *Bull. Math. Biophys*., **5**, 115 (1943).
5. D.E. Rumelhart, Parallel Distributed Processing. London: Mit Press (1982).
6. J. Zupan and J. Gasteiger, *Anal. Chim. Acta*, **248**, 1 (1991).
7. D.T. Manallack, D.D. Ellis and D.J. Livingstone, *J. Med. Chem*., **37**, 3758 (1994).
8. A. Guez and I. Nevo, *Clin. Chim. Acta*, **248**, 73 (1996).

9. V. Jakus, *Chem. Listy*, **87**, 262 (1993).
10. J.M. Vegas and P.J. Zufiria, Generalized Neural Network for Spectral Analysis: Dynamics and Liapunov Functions, Neural Networks, p. 17 (2004).
11. R.C. Schweitzer and J.B. Morris, *Anal. Chem. Acta*, **384**, 285 (1999).
12. C.S. Tong and K.C. Cheng, *Chemometr. Intell. Lab. Syst.*, **49**, 135 (1999).
13. F. Lui, Y. Liang and C. Cao, *Chemometr. Intell. Lab. Syst*., **81**, 120 (2006).
14. H. Golmohammadi and M.H. Fatemi, *Electrophoresis*, **26**, 3438 (2005).
15. E. Baher, M.H. Fatemi, E. Konoz and H. Golmohammadi, *Microchim. Acta*, **158**, 117 (2007).
16. M.H. Fatemi, *J. Chromatogr. A*, **1038**, 231 (2004).
17. M.H. Fatemi, *J. Chromatogr. A*, **955**, 273 (2002).
18. C. Schlett and B. Pfeifer, *Vom. Wasser*, **79**, 65 (1992).
19. J. Tremp, P. Mattrel, S. Fingler and W. Giger, *Water Air Soil Pollut.*, **68**, 113 (1993).
20. N. Masque, E. Pocurull, R.M. Marce and F. Borrull, *Chromatographia*, **47**, 176 (1998).
21. K.R. Kim and H. Kim, *J. Chromatogr. A*, **866**, 87 (2000).
22. R.H. Myers, Classical and Modern Regression with Application Pws-Kent Publishing Company: Boston (1990).
23. J. Ghasemi and Sh, Ahmadi, *Ann. Chim. (Rome)*, **97**, 69 (2007).
24. A. Lorber, L.E. Wangen and B.R. Kowalski, *J. Chemomet*., **1**, 19 (1987).
25. A. Hoskuldsson, *J. Chemom.*, **2**, 211 (1988).
26. H. Swierenga, A.P.D. Weijer, R.J.V. Wijk and L.M.C. Buydens, *Chemometr. Intell. Lab. Syst*., **49**, 1 (1999).
27. J. Zupan and J. Gasteiger, Neural Networks In Chemistry and Drug Design, Wiley-VCH Verlag, Weinheim (1999).
28. L. Fausett, Fundamentals of Neural Networks, Prentice Hall, New York (1994).