

PREDICTION GAS CHROMATOGRAPHIC RETENTION TIMES ON PONA COLUMN FOR SOME PETROLEUM COMPONENTS BASED ON LINEAR SOLVATION RELATIONSHIPS

N.E.Moustafa

Egyptian Petroleum Research Institute - Nasr City- Cairo – Egypt

Received December 5, 2007, accepted March 15, 2008

Abstract

Linear solvation energy relationship (LSER) was tested as predictive model for gas chromatographic retention times of hydrocarbons in petroleum condensate sample. The retention times of these compounds were correlated with its LSER descriptors by multi-linear regression analysis (MLRA). The best one and five parameter multi-linear regression models showed good predictive ability. The five parameters model showed more acceptable predictive ability than one parameter model particularly with the aromatics. The predictive ability of the obtained models was tested towards some unknown substituted benzenes. The previous published boiling point model was used as confirmed model for more accurate retention time's prediction.

Keywords: LSER, gas chromatography, petroleum condensate oil.

1. Introduction

Petroleum fractions are complex hydrocarbon mixtures containing hundreds of different aliphatic and aromatic components in widely differing concentrations. Although Many GC techniques are developed as potential methods for separation such huge number of isomers containing petroleum sample but these techniques suffer from some limitations or disadvantageous. The difficulties in many isomers identification arises from the unavailability of these isomers as standards.

The solvation parameter model in a form suitable for characterizing the retention properties of gas chromatographic stationary phases is given by Eqn. (1) [1-3].

$$\text{Log SP} = c + r R^2 + S \pi_2^H + \alpha \sum \alpha_2^H + b \sum \beta_2^H + l \text{Log } L^{16} \quad (1)$$

SP is some free energy related solute property such as a gas-liquid partition coefficient (K_L), retention factor (k), specific retention volume, or relative adjusted retention time. The remainder of the equation is made up of product terms called system constants (r, s, a, b, l) and solute descriptors ($R^2, \pi_2^H, \alpha_3^H, \beta_2^H, \log L^{16}$). Besides $\log L^{16}$ the other solute descriptors are R^2 the excess molar refraction (in $\text{cm}^3 \cdot 10^{-1}$), π_2^H the ability of the solute to stabilize a neighboring dipole by virtue of its capacity for orientation and induction interactions, and α_2^H and β_2^H the solute's effective hydrogen-bond acidity and basicity, respectively. The system constants describe the complementary stationary phase properties to the solute descriptors. The r constant establishes the capacity of the stationary phase for interactions with solute n- or p-electron pairs; the s constant the stationary phase's capacity for dipole-type interactions; the a constant characterizes the stationary phase's hydrogen-bond basicity (the ability to interact as a base); the b constant the stationary phase's hydrogen-bond acidity; and the l constant incorporates contributions from stationary phase cavity formation and solute-stationary phase dispersion interactions. The system constants are determined by multiple linear regression analysis of experimental $\log K_L$ or $\log t$ values for a group of varied solutes with known solute descriptors. The number of solutes should be sufficient to establish the statistical and chemical validity

of the model. The system constants are more than mere regression constants and contain the entire chemical information for the system. Validation of the system constants, therefore, requires that they meet both statistical and chemical tests of acceptability. Nowadays, not only LSER descriptors are reported for a huge number of compounds but also many methods have been reported for determination solute descriptors^[4-6].

Although linear solvation energy relationships have similarly been defined for gas and liquid chromatography data, LSER has not gained general usage in gas chromatography, but in liquid chromatography, where LSER is used to predict retention data, to predict physical properties of solutes and classify chromatographic columns^[7]. The aim of the present work is the testing the accuracy of LSER relationship as gas chromatographic retention time's predictive model for some unknowns, based on its reported descriptors, in the complex petroleum sample.

2. Experimental

The examined condensate oil was a sample from the Karanis-1x field of the Khalda Co., Egypt. Claus-500 Perkin-Elmer gas chromatograph with FID was used for analysis. The original condensate oil was analyzed on a PONA 100 m x 0.25 mm x 0.25 μ m capillary column for determination of its total composition from C₂-C₃₆⁺. using multistage program: 100 °C for 5 min ramp at 5 °C/min at 150 °C then ramp at 12 °C /min till 250 °C and remain isothermal for 5 min then ramp at 40 °C/min till 300 °C and isothermal for 110 min. Helium carrier gas 1.20 ml/min; Sample size 1 μ l and injector and detector temperatures 350 °C. The sample injected at least three times at split ratio 1:30 to give an acceptable retention time deviation of 0.047 min per peak. The standard compounds were obtained commercially from various sources. The peak assignments in the chromatogram of a condensate sample were established based on retention times that matched previously analyzed standards, which were purchased from Aldrich Co.

3. Results and discussions

The complementary stationary phase system constants are identified as the contribution from cavity formation and dispersion interactions, *l*, the contribution from interactions with solute n- or p-electrons, *e*, the contribution from dipole– dipole and dipole–induced dipole interactions, *s*, the contribution from hydrogen-bond basicity *a*, and *b* the contribution from hydrogen-bond acidity. The system constants are determined by multiple linear regression analysis (MLR) of experimental log_t values for a group of varied solutes selected to establish the statistical and chemical validity of the model. Table 1 gives the values of LSER descriptors for the investigated solutes.

3.1 Correlation of the retention times

Eq. 1 gives the stationary phase constants for the investigated 44 solutes using solutes descriptors given in Table 1 by multi-linear regression analysis. This equation regression coefficients are $r = 0.9456$, $SSR = 0.583$ and $s.d. = 2.1087$. The ratio SSE can be used as an indicator of outliers: Its values are between 0 ("bad") and 1 ("good"): the values below 0.70 might indicate the presence of influential outliers. Accordingly, this equation straight line relationship has a considerable number of outliers and can not used for retention time's prediction for the unknown component in the chromatogram. It is observed that the light components are the reason for this linear non-fitness. Accordingly, Eq. 2 represents LSER equation for a subset of 36 components (Fig.1). It is observed the improvements in the values of the regression coefficients. Eq. 4 is the most fitted obtained equation from the point of view of retention time prediction ($SSR = 0.964$). The equation is characterized by a lower number of outliers compared with the others (Fig.2). Accordingly, equations 2 and 4 are applied for retention time prediction. Table 2 gives the values of the retention times calculated by Eqs. 2-4 in comparison with the experimental ones. It can be observed that Eq. 4 is the best fit model from the point of view of retention time's prediction for saturates. Whereas, Eqs. 2 and 3 are the best models for aromatics retention times' prediction because in this case the lower cumulative value of the other descriptors cannot be neglected particularly with the heavy aromatics. The deviations, Δ , between the experimental and predicted retention times by Eq. 4 are given in Table2.

3.2 Physical interpretation of retention correlation equations

The validity of Eqs. 2-4 can be represented in term of the contributions of the different system constants in the retention of the given solute according to their meanings. It is observed from Eq.1

that, the poly(dimethylsiloxane) stationary phase has significant dipolarity / polarizability repulsion ($s = -0.138$) and is hydrogen-bond basic and acidic where $a = 2.400$ and $b = 3.950$, respectively. Electron-pair repulsions are important ($r = -5.150$). b and r account for specific solute adsorption at the liquid–solid interface^[4], which depends on liquid coating percent and column inner surface pretreatments. The stationary phase has a cohesive degree corresponding to l equal 5.68.

The poly(dialkylsiloxane) polymers have a helical structure with the alkyl groups turned out^[4] and furthermore methyl groups have a lower degree of steric shielding on the siloxane bond. For this relatively non-polar structure, it is expected that l constant is the main effective constant in the given solute retention. This was confirmed by F - and t statistics values for the descriptors of Eqs. 1 and 2 (Table 3). High value of the F -statistic of the L descriptor implies that the higher contribution degree of this descriptor in the given correlation. The comparatively lower F values for the other descriptors indicate the possibility of the neglecting of their values in the predicted retention time values of the saturated compounds. The F -values for R_2 and π_2^H descriptors rationalize the more accuracy of Eqs. 2 and 3 as predictive model for retention times of heavy aromatics. This indicating that, the lower cumulative value of these descriptors cannot neglected with heavy substituted aromatics for more accurate retention time value.

$$-Y = -4.490 - 0.138 \cdot R_2 - 5.150 \cdot \pi_2^H + 2.400 \cdot \sum \alpha_2^H + 3.950 \cdot \sum \beta_2^H + 5.680 \cdot \log_{10} L^{16} \quad (1)$$

$$R = 0.9456 \quad n = 44 \quad SSE = 0.583 \quad s.d. = 2.1087 \quad F = 64.16$$

$$-Y = -0.401 + 4.070 \cdot R_2 - 4.180 \cdot \pi_2^H - 1.500 \cdot \sum \alpha_2^H + 3.320 \cdot \sum \beta_2^H + 3.800 \cdot \log_{10} L^{16} \quad (2)$$

$$R = 0.9745 \quad n = 36 \quad SSE = 0.742 \quad s.d. = 0.4893 \quad F = 107.39$$

$$-Y = -0.213 + 2.620 \cdot R_2 - 1.420 \cdot \pi_2^H + 3.760 \cdot \log_{10} L^{16} \quad (3)$$

$$R = 0.9745 \quad n = 36 \quad SSE = 0.829 \quad s.d. = 0.4446 \quad F = 188.28$$

$$-Y = -0.393 + 3.910 \cdot \log_{10} L^{16} \quad (4)$$

$$R = 0.9687 \quad n = 36 \quad SSE = 0.964 \quad s.d. = 0.4098 \quad F = 486.52$$

The high F -value for a term implies that this term is useful for the given model.

The Student's t -values characterize the relative importance of descriptors in a particular.

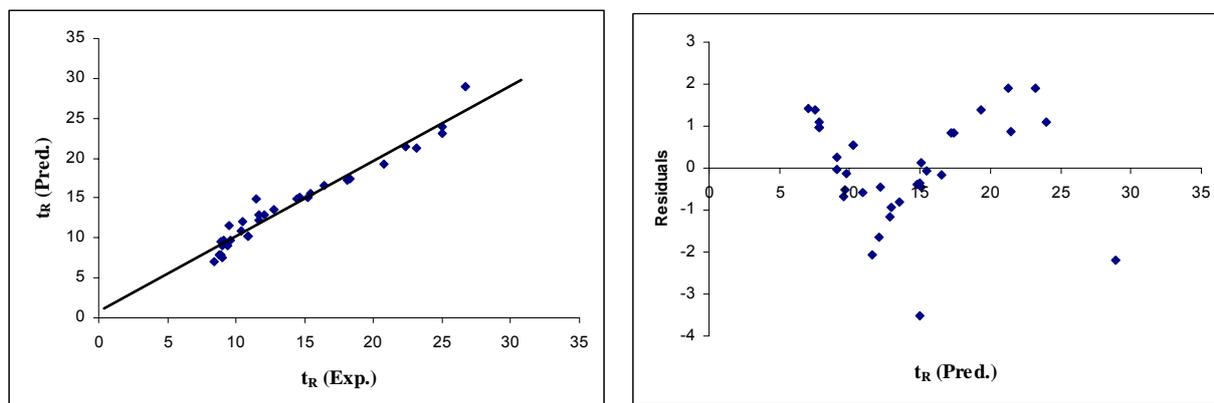


Fig. 1: (a) Predicted t_R values by Eq. 2 vs. experimental t_R values. (b) Plot of residuals vs. experimental retention times for 44 petroleum components.

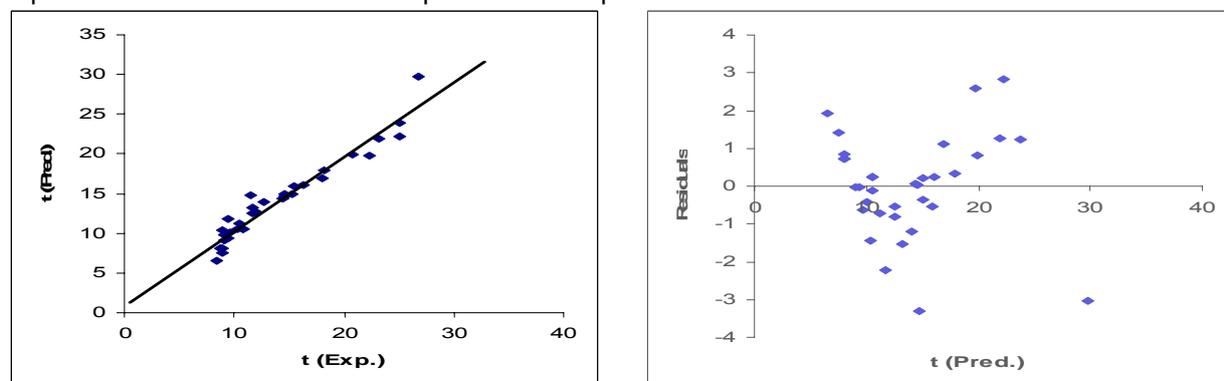


Fig. 2: (a) Predicted t_R values by Eq. 4 vs. experimental t_R values. (b) Plot of residuals vs. experimental retention times for 38 petroleum components.

Table1 The values of LSER descriptors for the investigated solutes.

Component	t_R (Exp.)	R_2	π_2^H	$\Sigma\alpha_2^H$	$\Sigma\beta_2^H$	$\log_{10} L^{16}$
Methane	8.150	0.000	0.000	0.000	0.000	-0.323
Ethane	8.210	0.000	0.000	0.000	0.000	0.492
Propane	8.310	0.000	0.000	0.000	0.000	1.050
i-Butane	8.430	0.000	0.000	0.000	0.000	1.615
n-Butane	8.520	0.000	0.000	0.000	0.000	1.615
2,2-Dimethylpropane	8.520	0.000	0.000	0.000	0.000	1.820
2-Methylbutane	8.790	0.000	0.000	0.000	0.000	2.162
i-Pentane	8.790	0.000	0.000	0.000	0.000	2.162
n-Pentane	8.910	0.000	0.000	0.000	0.000	2.162
Cyclohexane	10.480	0.310	0.000	0.000	0.000	2.964
2-Methylpentane	9.360	0.000	0.000	0.000	0.000	2.503
n-Hexane	9.620	0.000	0.000	0.000	0.000	2.668
n-Heptane	10.840	0.000	0.000	0.000	0.000	2.809
2,4-Dimethylpentane	10.840	0.000	0.000	0.000	0.000	2.809
2,5-Dimethylhexane	11.710	0.000	0.000	0.000	0.000	3.308
2,3,4-Trimethylpentane	11.680	0.000	0.000	0.000	0.000	3.481
Methylcyclohexane	11.460	0.260	0.100	0.000	0.000	3.877
n-Octane	12.780	0.000	0.000	0.000	0.000	3.677
n-nonane	15.420	0.000	0.000	0.000	0.000	4.182
n-Decane	18.250	0.000	0.000	0.000	0.000	4.686
n-Undecane	20.720	0.000	0.000	0.000	0.000	5.191
n-Dodecane	23.140	0.000	0.000	0.000	0.000	5.696
n-Tridecane	25.060	0.000	0.000	0.000	0.000	6.200
n-Tetradecane	26.720	0.000	0.000	0.000	0.000	7.714
n-octacosane	54.280	0.000	0.000	0.000	0.000	13.780
Dotricontane	84.050	0.000	0.000	0.000	0.000	15.790
Hexatricontane	126.880	0.000	0.000	0.000	0.000	17.740
Benzene	10.380	0.610	0.520	0.000	0.140	2.786
Toluene	12.050	0.601	0.520	0.000	0.140	3.325
Ethybenzene	14.440	0.613	0.510	0.000	0.150	3.778
m-Xylene	14.650	0.600	0.580	0.000	0.150	3.940
o-Xylene	15.210	0.600	0.580	0.000	0.150	3.940
p-Xylene	14.650	0.600	0.520	0.000	0.160	3.839
n-Propylbenzene	16.380	0.604	0.500	0.000	0.150	4.230
1,2,4-Trimethylbenzene	18.060	0.600	0.560	0.000	0.190	4.441
Naphthelene	22.360	1.340	0.920	0.000	0.200	5.161
1-methylnaphthalene	25.060	1.340	0.900	0.000	0.200	5.789
Ethanol	8.430	0.250	0.420	0.370	0.480	1.485
1-Propanol	8.970	0.240	0.420	0.370	0.480	2.031
1-Butanol	9.140	0.220	0.420	0.370	0.480	2.601
1-Pentanol	9.510	0.220	0.420	0.370	0.480	3.106
2-Propanol	8.430	0.212	0.360	0.330	0.560	1.764
Methylketone	8.930	0.166	0.680	0.000	0.510	2.755
2-Methylpropanol	9.020	0.217	0.390	0.370	0.480	2.413

Table 2 The comparison between the calculated and predictive t_R for the investigated components.

Component	t_R (Exp.)	t_R (pred.) Eq. 2	t_R (pred.) Eq. 3	t_R (pred.) Eq. 4	Δ^*
Methane	8.15				
Ethane	8.21				
Propane	8.31				
i-Butane	8.43				
n-Butane	8.52				
2,2-Dimethylpropane	8.52				
2-Methylbutane	8.79	7.81	7.92	8.06	0.730
i-Pentane	8.79	7.81	7.92	8.06	0.730
n-Pentane	8.91	7.81	7.92	8.06	0.850
Cyclohexane	10.48	12.12	11.74	11.20	-0.716
2-Methylpentane	9.36	9.11	9.20	9.39	-0.034
n-Hexane	9.62	9.74	9.82	10.04	-0.419
n-heptane	10.84	10.27	10.35	10.59	0.250
2,4-Dimethylpentane	10.84	10.27	10.35	10.59	0.250
2,5-Dimethylhexane	11.71	12.17	12.23	12.54	-0.831
2,3,4-Trimethylpentane	11.68	12.83	12.88	13.22	-1.538
Methylcyclohexane	11.46	14.97	14.90	14.77	-3.306
n-Octane	12.78	13.57	13.61	13.98	-1.204
n-nonane	15.42	15.49	15.51	15.96	-0.539
n-Decane	18.25	17.41	17.41	17.93	0.321
n-Undecane	20.72	19.32	19.31	19.90	0.816
n-Dodecane	23.14	21.24	21.20	21.88	1.262
n-Tridecane	25.06	23.16	23.10	23.85	1.211
n-Tetradecane	26.72	28.91	28.79	29.77	-3.049
Benzene	10.38	10.96	11.12	10.50	-0.120
Toluene	12.05	12.97	13.13	12.61	-0.558
Ethybenzene	14.44	14.82	14.87	14.38	0.061
m-Xylene	14.65	15.09	15.35	15.01	-0.362
o-Xylene	15.21	15.09	15.35	15.01	0.198
p-Xylene	14.65	14.99	15.06	14.62	0.033
n-Propylbenzene	16.38	16.54	16.56	16.15	0.234
1,2,4-Trimethylbenzene	18.06	17.21	17.26	16.97	1.089
Naphthalene	22.36	21.48	21.40	19.79	2.574
1-Methylnaphthalene	25.06	23.95	23.79	22.24	2.818
Ethanol	8.43	7.58	7.46	7.55	1.422
1-Propanol	8.97	9.66	9.55	9.78	-0.637
1-Butanol	9.14	11.58	11.45	11.75	-2.242
1-Pentanol	9.51	7.02	6.46	6.50	1.926
2-Propanol	8.43	9.59	9.62	10.38	-1.449
Methylketone	8.93	9.06	8.87	9.04	-0.022
2-Methypropanol	9.02	10.96	11.12	10.50	-0.120

* The deviation for Eq. 4

Table 3 F and t statistics for the descriptors of Eqs. 1 and 2.

term	F-Eq. 1	t Stat	F-Eq.2	t Stat
const		-2.21		-0.51
R_2	8.38E-05	-0.01	2.26	1.50
π_2^H	0.0356	-0.19	0.726	-0.85
$\sum \alpha_2^H$	0.00651	0.08	0.0761	-0.28
$\sum \beta_2^H$	0.0123	0.11	0.263	0.51
$\log_{10} L^{16}$	305.97	17.49	386.25	19.65

4. Application

The predictive abilities of Eqs. 2 and 4 are tested for retention times of some unknown substituted aromatics. Table 4 gives the predicted retention times for some unknowns' substituted benzenes according to the Eqs 2 and 4. The previous published boiling point model was used as confirmed model [8]. The bracketing compounds particularly n-alkanes, concentration, stationary phase non-polarity, boiling point, model predictive error and previous identified chromatograms all were used as assistance means in these theoretical identifications. Fig. 3 shows the chromatogram of some from these theoretically identified substituted benzenes.

Table 4
The predictive t_R for some unknown substituted benzenes.

Component	Eq.-b.p.*	Eq. 2	Eq. 4
n-Butylbenzene	19.50	18.38	18.10
n-Pentylbenzene	21.83	20.28	20.06
n-Hexylbenzene	23.99	22.19	21.97
Bi/Diphenyl	27.08	24.56	23.12
2,5-Dimethylhexane	11.91	12.54	12.54
2,3,4-Trimethylpentane	12.27	13.22	13.22
Styrene	15.56	14.61	14.68
1,2,3-Trimethylbenzene	18.67	17.29	16.97
1,2,4-Trimethylbenzene	18.15	17.21	16.97
1,3,5-Trimethylbenzene	17.60	17.38	16.97
n-Propylbenzene	17.02		
Isopropylbenzene	16.28		
4-Ethyltoluene	17.32		
1,3,5-Triethylbenzene	22.93		
lindane	18.98	16.66	
Indene	19.44		
Naphthalene	23.14		19.79
1-Methylnaphthalene	25.92	21.48	22.24
2-Methylnaphthalene	25.54	23.95	
Acenaphthylene	29.59		
Acenaphthene	29.59	27.00	24.90
Fluorene	31.15	27.94	26.67
Phenathrene	35.81	32.56	29.20
Anthracene	35.81	33.01	29.45
1-Methylphenanthrene		36.98	
Fluoranthene	39.45	38.40	34.12
Pyrene	42.47		34.13
Chrysene	47.04		
Benzo[a]pyrene	51.95		
Benzo[e]pyrene	51.74		
Benzo[ghi]preylene	52.47		
Indeno[1,2,3-cd]pyrene	56.21		
Dibenzo[a,h]anthracene	54.96		

* -Y = - 27.940 + 0.104*B.P , R = 0.998, n = 19, SSE = 0.991, s.d. = 0.153, F = 5046.40 [8]

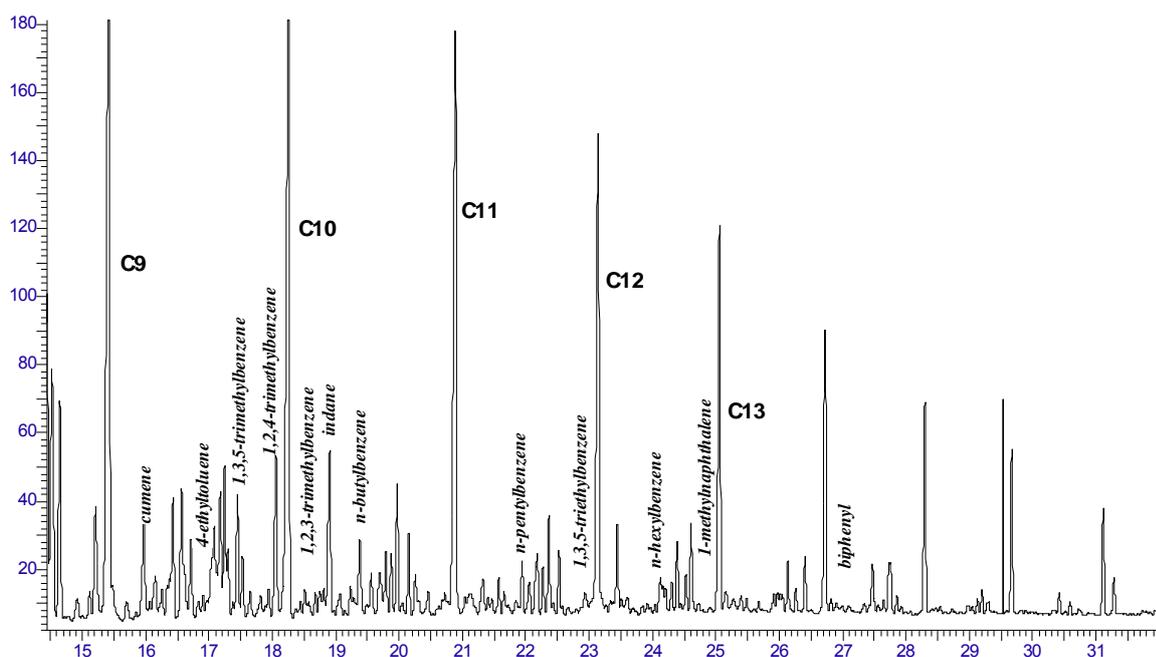


Fig. 3: The chromatogram of some theoretically identified substituted benzenes.

5. Conclusions

To obtain the PONA stationary phase LSER equation, the experimental retention times of separated 44 petroleum condensate components were correlated by MLRA with its reported descriptors. The obtained LSER model showed a lower accuracy for retention time prediction of petroleum condensate components. Three models for subset of 38 components are derived based on one to five LSER descriptors. The simple LSER model contains only one descriptor (I) is the best fit model for retention time prediction of saturates. The five based model descriptors (R_2 , π_2^H , $\sum \alpha_2^H$, $\sum \beta_2^H$ and $\log_{10} L^{16}$) should be used as predictive model with aromatics. This model gives theoretically identified retention times of acceptable accuracy with some unknown substituted benzenes.

References

- [1] Abraham, M. H., Poole, C. F. and Poole, S. K. J. Chromatogr. A, 842, 79 (1999).
- [2] Poole, S. K. and Poole, C. F. Anal. Commun., 33, 353 (1996).
- [3] Poole, C. F., Kollie, T. O. and Poole, S. K., Chromatographia, 34, 281 (1992).
- [4] Li, Q, Poole, C. F., Kiridena, W. and. Koziol, W. W., Analyst, 125, 2180–2188 (2000).
- [5] Niederer, C., Goss, K. and Schwarzenbach, R. P., Supporting Informations, Environmental Science and Technology, 2006.
- [6] Mark Vitha, M., Carr, P. W, J. of Chromatography A, 1126, 143–194 (2006).
- [7] Abraham, M.H. Ibrahim, A. and Zissimos, A.M., J. Chromatogr. A 1037, 29 (2004).
- [8] Moustafa, N. E., Chromatographia, Accepted (2007).