Javier Gotta
Sonia Keunchkarian
Cecilia Castells
Mario Reta

Laboratorio de Separaciones Analíticas, División Química Analítica, Facultad de Ciencias Exactas (UNLP), La Plata, Argentina

Received February 24, 2012
Revised June 18, 2012
Accepted June 21, 2012

Research Article

Predicting retention in reverse-phase liquid chromatography at different mobile phase compositions and temperatures by using the solvation parameter model

The prediction capability of the solvation parameter model in reverse-phase liquid chromatography at different methanol-water mobile phase compositions and temperatures was investigated. By using a carefully selected set of solutes, the *training set*, linear relationships were established through regression equations between the logarithm of the solute retention factor, log*k*, and different solute parameters. The coefficients obtained in the regressions were used to create a general retention model able to predict retention in an octadecylsilica stationary phase at any temperature and methanol-water composition. The validity of the model was evaluated by using a different set (the *test set*) of 30 solutes of very diverse chemical nature. Predictions of log*k* values were obtained at two different combinations of temperature and mobile phase composition by using two different procedures: (i) by calculating the coefficients through a mathematical linear relationship in which the mobile phase composition and temperature are involved; (ii) by using a general equation, obtained by considering the previous results, in which only the experimental values of temperature and mobile phase composition are required. Predicted log*k* values were critically compared with the experimental values. Excellent results were obtained considering the diversity of the test set.

Keywords: Linear solvation energy relationships (LSER) / Mobile phase effects / Retention mechanism / Reverse-phase liquid chromatography (RPLC) / Solvation parameter model / Temperature effects
DOI 10.1002/jssc.201200197

1 Introduction

Retention and its dependence with mobile phase composition and temperature in reverse-phase liquid chromatography (RPLC) were first studied by Melander et al. 30 years ago [1]. However, it is still an actual problem due to the difficulty of predicting retention with enough accuracy when typical chromatographic variables (mobile phase composition, stationary phase nature, pH, temperature, etc.) are changed. The final goal of a chromatographic separation is to get baseline resolution in a minimal analysis time; therefore, the computational prediction of retention factors to reduce the number of trial and error experiments is widely required.

One of the most successful models to understand and predict retention in liquid and gas chromatography is the "Solvation Parameter model" (SP model) developed by M.H. Abraham [2, 3]. This model can be placed within the frame of the "Linear Solvation Energy Relationships" (LSERs), since a multiparametric linear equation relates an appropriate form of the property under study (in this work will be the retention factor, *k*) and several *independent* solute parameters, each one considering a given type of solute–solvent interaction. The SP model (Eq. (1)) is now well established [4, 5] and it uses different solute descriptors, corresponding to a set of compounds (*training set*), to take into account the different intermolecular interactions involved in the retention process.

$$\log k = c + s\,S + a\,A + b\,B + v\,V + e\,E \quad (1)$$

Here, k is the chromatographic retention factor and the solute descriptors are as follows: S is the solute dipolarity/polarizability; A and B are the hydrogen-bond acidity and basicity, respectively; V is the McGowan characteristic volume [6] that accounts for the necessary energy to form the cavity within the solvent to accommodate the solute; and E is an excess molar refraction that accounts for polarizability interactions due to electron donor groups [2, 3, 7]. The intercept, c, and the regression coefficients s, a, b, v, and e, obtained from multivariable, simultaneous, least-squares regressions, contain chemical information since reflect the *difference* in

Correspondence: Dr. Mario Reta, Laboratorio de Separaciones Analíticas, División Química Analítica, Facultad de Ciencias Exactas (UNLP), 47 y 115 (1900) La Plata, Argentina
E-mail: mreta@quimica.unlp.edu.ar
Fax: +54–221-4254533

Abbreviations: RPLC, reverse-phase liquid chromatography; **RMSE**, root mean squared error

the complementary property to each solute parameter as follows [4, 8, 9]:

$$\log k = c + s'(s_s - s_m)S + a'(b_s - b_m)A + b'(a_s - a_m)B + v'(v_s - v_m)V + e'(e_s - e_m)E \quad (2)$$

where the subscripts "*s*" and "*m*" refers to the stationary and mobile phases, respectively; the coefficients s', a', b', v', and e' are fitting parameters [9, 10]. As an example, the complementary property to solute acidity, *A*, is the "solvent basicity" (in this case, the basicity of the stationary, b_s and mobile phases, b_m, respectively).

Wang and Carr [11, 12] have proposed two LSER models, one deduced from the Abraham model (Eq. (1)) termed "global LSER" and the other, termed "typical-conditions model" (TCM), based on principal component analysis over an initial experimental data set as starting point. Both models worked better than the local LSER. However, in the "global LSER" model, the test set consisted of 22 quite simple molecules and no temperature variations were considered.

The "hydrophobic subtraction model" (HSM) developed originally by Wilson et al. [12] is probably one of the best methods to predict RPLC retention at room temperature in C18 columns. However, the value of the SP model over the HSM is that both the solute descriptors and also, the SP coefficients have chemical meaning and can be used in a rational interpretation of the retention process.

It is well known that, in spite of the nonpolar nature of the C18 bonded phase, the *real* stationary phase also contains a significant amount of sorbed organic solvent and water [13, 14]. This means that the polarity–polarizability, hydrogen bond acidity, and hydrogen bond basicity (s_s, a_s and b_s in Eq. (2) or π_s^*, α_s and β_s in their original nomenclature) ought not to be zero as they are for a purely aliphatic material. Chromatographic and spectroscopic measurements show that the π_s^*, α_s, and β_s values of the stationary phase can be significantly >0 [15–19]; for instance, typical values for π_s^* range from 0.7 to 1.1 depending on mobile phase composition.

Temperature, which has not been considered in the previous cited models, is another important and convenient variable to improve resolution and analysis time in RPLC method development [20–26]. Poole et al. applied the SP model to study the influence of mobile phase and temperature on retention and selectivity in RPLC. They tested three mobile phases of similar strength in combination with a porous organic polymer stationary phase [27] and also, a polar-endcapped ODS column at different methanol-water mobile phase compositions and temperatures between 25 and 65°C [28]. However, in these works the authors did not check the prediction capability of the model by using a *test set*. In other work [29], Poole et al. studied the SP model at different mobile phase compositions and they used test solutes in an ODS phase, but the study was conducted at room temperature.

Within the context of the LSER approach, Carr et al. have examined the effect of the temperature and mobile phase composition on chromatographic RPLC retention of nonhydrogen bond donor solutes ($A = 0$) by using a procedure similar to the SP model [30], although they have not evaluated the prediction capability of the model by using a test set. They have used acetonitrile-water mixtures as mobile phases and temperatures ranging from 25 to 65°C.

Pappa-Louisi et al. [31] have obtained a mathematical model that combines mobile phase composition and temperature. In a first paper, they have compared several equations that account very well for the retention of six alkylbenzenes in acetonitrile-water mobile phase [31]. In a second paper, they have used the SP model to obtain a more general equation able to predict retention for any other solute, but the obtained equations have too many adjustable parameters and are quite complicated to be practical [32]. Those equations were assayed using seven compounds of pharmacological interest in acetonitrile-water and methanol-water mobile phases.

Rosés and collaborators explored two strategies to model retention of both neutral and ionized compounds in RPLC [33, 34]. These two models predicted retention with an accuracy close to that of the SP model. The authors studied several mobile phase compositions including several buffers at room temperature.

In this work, we have applied the SP model to obtain a general mathematical equation able to predict the retention of solutes at any methanol-water mobile phase composition in the range between 40 and 70% MeOH, and at any temperature from 30 to 70°C. We have carefully selected a series of 21 solutes for the training set to calibrate the system. The model was tested by comparing the experimental and calculated retention values for a set of 30 analytes of very different chemical nature, including compounds of pharmaceutical concern such as β-blockers and profens.

2 Experimental

2.1 Chemicals and materials

Methanol-buffered phosphate (pH 2.70; 25 mM) water mixtures were used as the mobile phases, which were filtered through 0.22-μm nylon membranes (Osmonics-Magna). Water was purified by means of a Milli-Q Purification System (Simplicity, Millipore, MA, USA). Analytes were purchased from Sigma-Aldrich (St. Louis, MO, USA).

Solutes were dissolved in a mixture of methanol-water (50:50), in a concentration of about 10 mg/mL; dilutions of approximately 1:100 of these solutions were prepared in each mobile phase.

2.2 Equipment and experimental measurements

An HP 1100 liquid chromatograph (Agilent Technologies, Palo Alto, CA, USA) equipped with a binary pump,

thermostat-controlled column compartment, degasser, and diode-array detector connected to the ChemStation software was used. The retention factors, k, were measured in a 75 × 4.6 mm i.d. (3.5 μm) Zorbax Eclipse SB-C18 column (Agilent Technologies) setting detection at 220 nm. Retention times are the average of triplicates, obtained in different days, and the relative standard deviations were below 1%. Holdup times were measured by coinjection of potassium bromide and setting detection at 210 nm. The extracolumn volume of the system was 60 μL, which is negligible as compared to the retention volume.

2.3 Solute-parameter calculations and multivariable least-squared regressions

Solute parameters for the training set and some for the test set were obtained from the literature [2, 3]. By using the Absolve module of the software ADME Boxes 5.0 Software (ACD/Labs/Pharma Algorithms Inc., Toronto, Canada), we have calculated the solute parameters for those solutes of the test set that were not found in the literature. The satisfactory performance of this software to predict reliable solute parameters have been previously shown [35]. Multivariable least-squared regressions were performed with Microsoft Office Excel 2007. The statistical significance of each term of the multiparametric equations were checked by the *P* values obtained with the SigmaPlot 4.01 software.

3 Results and discussion

3.1 Design of the SP model

The original SP model considers only nonionized analytes. The low mobile phase pH (2.7) was selected to avoid ionization of the residual silanol groups from the silica surface, and thus, to minimize electrostatic interactions with the solutes. The basic solutes *o*-toluidine, 3-chloroaniline, and 4-chloroaniline from the training set would be protonated at this low pH if the solvent had been pure water and the chromatographic run at room temperature. In the methanol/buffer mobile phase, the pK_a of amines decreases whereas pK_{a1} of phosphoric acid increases as the methanol content increases. Even more, when we take into account that the chromatographic measurements were conducted at 30, 50, and 70°C, the pKa of those amines decrease even more and the pK_{a1} of the acid is almost independent of temperature. The pK_a were taken from the literature or calculated from general equations [36–38]. From calculations (not shown) of pKa's and ionization enthalpies in methanol-water at the studied temperatures can be concluded that amines are neutral in almost all chromatographic conditions. Thus, electrostatic interactions, which are not explicitly included in the SP model, should be absent within the experimental conditions used in this study. The only ionized solutes are the aliphatic amines (β-blockers) included in the *test set*. Even though, for these amines the retention predictions are very good.

The solute parameters for analytes corresponding to the *training set* are reported in Table 1, and the respective logarithm of the retention factors, *k*, at each mobile phase composition and temperature are gathered in Table 2. The training set corresponds to the carefully selected compounds that were used in establishing the LSER coefficients of Eq. (1) for each mobile phase composition and temperature. Carr and Vitha [39] in a very clear and useful review have established and summarized a series of recommendations for the design, analysis, and interpretation of an LSER model to be statistically valid and to obtain information with chemical sense. First, the solutes that form the training set must span a wide range of solute parameters or, in other words, they must be chemically diverse. This requirement is fulfilled in this study: extreme values are 0.61 to 1.52, 0.51 to 1.5, 0 to 1.16, 0.09 to 0.59, and 0.77 to 1.48 for the *E*, *S*, *A*, *B*, and *V* parameters, respectively (Table 1). Second, the property to be studied should span at least one order of magnitude. This requisite is accomplished by selecting solutes with very different chemical properties. In this work, the log*k* values span almost two orders of magnitude within the same experimental condition (Table 2). Third, the descriptors must not exhibit significant covariance. Typically, correlation coefficients between two solute descriptors >0.5 or 0.6 are regarded as indicative of quite strong covariance, while values as high as 0.7 or 0.8 are unacceptable. Here, the covariance between the different solute descriptors was virtually nonexistent for the training set as is shown in Table 3 except for the *E* and *S* parameters for which some covariance was observed. This is expected since the *S* descriptor not only measures polarity but also polarizability interactions. However, that inconvenient is not a problem since the *e* coefficient resulted negligible (see Table 4) and statistically not significant according to the *P* values (>0.05). Fourth, because at least four parameters per descriptor are necessary and thus at least 20 solutes must be included in the training set, we used 21 solutes for this work.

The LSER coefficients from Eq. (1) are shown in Table 4. Solute parameters were considered independent of temperature within this temperature range. Very good regression coefficients and small standard deviations were obtained in all instances (Table 4). The obtained coefficients are typical values for RPLC stationary phases containing C18 or C8 groups with aqueous mobile phases [7, 39–41]: the two most influential intermolecular interactions affecting the retention process are the solute hydrogen-bond acceptor affinity (*negative b* term) and the cavity term that considers both dispersion interactions along with the necessary energy to break the hydrogen-bonded network to accommodate the solute within the cavity of the solvent (*positive v* term). The *e*-coefficient was statistically negligible in every regression, thus, the *E* parameter was not considered in the regression equations. The LSER coefficients of Table 4 and their variations with methanol composition and temperature are chemically interpreted as follows:

Table 1. Solute descriptors for the training set[a)].

Solute number	Solute	*E*	*S*	*A*	*B*	*V*
1	4-Chloroaniline	1.060	1.10	0.30	0.35	0.9390
2	Hydroquinone	1.000	1.00	1.16	0.60	0.8338
3	Ethylbenzene	0.613	0.51	0	0.15	0.9982
4	Naphthalene	1.340	0.92	0	0.20	1.0854
5	Bromobenzene	0.882	0.73	0	0.09	0.8914
6	Nitrobenzene	0.871	1.11	0	0.28	0.8906
7	Thymol	0.822	0.79	0.52	0.44	1.3387
8	Phenol	0.805	0.89	0.60	0.30	0.7751
9	2-Naphthol	1.520	1.08	0.61	0.40	1.1441
10	Benzonitrile	0.742	1.11	0	0.33	0.8711
11	Benzophenone	1.447	1.50	0	0.50	1.4808
12	Propiophenone	0.804	0.95	0	0.51	1.1548
13	Resorcinol	0.980	1.00	1.10	0.58	0.8338
14	Catechol	0.970	1.10	0.88	0.47	0.8338
15	Benzaldehyde	0.820	1.00	0	0.39	0.8730
16	Benzoic acid	0.730	0.90	0.59	0.40	0.9317
17	Benzyl benzoate	1.264	1.42	0	0.51	1.6804
18	*p*-Xylene	0.613	0.52	0	0.16	0.9982
19	4-Chloroacetanilide	0.980	1.50	0.64	0.51	1.2357
20	3-Chloroaniline	1.050	1.10	0.30	0.36	0.9390
21	*o*-Toluidine	0.970	0.90	0.23	0.59	0.9751

a) Poole et al. [42]; Abraham et al. [43]; Abraham et al. [44].

Table 2. log*k* values for the training set at different MeOH compositions and temperatures.

Solute[a)]	30°C				50°C				70°C			
	40%	50%	60%	70%	40%	50%	60%	70%	40%	50%	60%	70%
1	0.424	0.234	0.013	−0.189	0.390	0.145	−0.063	−0.279	0.308	0.079	−0.122	−0.325
2	−0.552	−0.761	−0.917	−0.998	−0.659	−0.805	−0.861	−0.987	−0.726	−0.905	−0.990	−1.081
3	1.780	1.378	0.983	0.611	1.565	1.195	0.850	0.460	1.364	1.010	0.681	0.382
4	1.788	1.362	0.941	0.566	1.562	1.161	0.801	0.447	1.313	0.955	0.627	0.338
5	1.534	1.164	0.787	0.442	1.330	0.982	0.670	0.340	1.127	0.806	0.516	0.245
6	0.788	0.495	0.209	−0.048	0.626	0.361	0.127	−0.114	0.482	0.231	0.014	−0.173
7	1.700	1.249	0.643	0.404	1.472	1.038	0.650	0.269	1.211	0.820	0.463	0.150
8	0.340	0.078	−0.165	−0.373	0.179	−0.044	−0.235	−0.430	0.047	−0.162	−0.346	−0.487
9	1.105	0.705	0.316	0.004	0.861	0.503	0.191	−0.102	0.639	0.324	0.031	−0.200
10	0.605	0.303	0.019	−0.221	0.467	0.189	−0.044	−0.276	0.325	0.070	−0.142	−0.308
11	1.713	1.226	0.770	0.386	1.501	1.025	0.641	0.279	1.237	0.833	0.477	0.165
12	1.091	0.728	0.389	0.095	0.932	0.592	0.301	0.018	0.758	0.448	0.178	−0.046
13	−0.284	−0.512	−0.696	−0.837	−0.404	−0.602	−0.742	−0.869	−0.539	−0.697	−0.856	−0.881
14	−0.038	−0.278	−0.466	−0.635	−0.162	0.369	−0.506	−0.686	−0.294	−0.462	−0.559	−0.661
15	0.586	0.302	0.033	−0.189	0.448	0.181	−0.038	−0.245	0.301	0.071	−0.207	−0.304
16	0.628	0.308	−0.009	−0.251	0.435	0.140	−0.096	−0.339	0.247	−0.009	−0.247	−0.424
17	—	1.708	1.170	0.697	1.981	1.463	1.003	0.571	1.699	1.220	0.799	0.426
18	1.817	1.426	1.021	0.655	1.599	1.226	0.891	0.531	1.385	1.033	0.714	0.406
19	0.864	0.521	0.188	−0.091	0.670	0.335	0.062	−0.202	0.446	0.159	−0.091	−0.306
20	0.583	0.333	0.064	−0.186	0.480	0.218	−0.008	−0.251	0.363	0.114	−0.112	−0.315
21	−0.071	−0.158	−0.234	−0.360	0.048	−0.097	−0.245	−0.362	0.059	−0.077	−0.077	−0.362

a) See Table 1 for solute identification.

Table 3. Covariance matrix for the solute parameters of the *training set*

	E	S	A	B	V
E	1				
S	0.62	1			
A	0.08	0.11	1		
B	0.31	0.57	0.54	1	
V	0.49	0.47	–0.31	0.25	1

- *The v coefficient*. The vV term can be dissected into at least two terms, a cavity and a dispersive term, the first one usually taken as the square of Hildebrand solubility parameter ($\delta_H^2 = -\Delta H_{vap}/V_M$, being ΔH_{vap} the molar enthalpy of vaporization and V_M the molar volume of the solvent) and the second one representing the susceptibility of the solvent to engage in London interactions [9]. In this study, the v coefficients are positive and high, indicating strong cohesivities in the hydroorganic mobile phases compared to cohesivity in the stationary phase and/or strong dispersive interactions in the stationary phase compared to mobile phase, a conclusion that is consistent with chemical intuition and with previous results [9, 39–41]. The v coefficient is proportional to the cohesivity of the mobile phase [9]. Thus, these coefficients decreases with the increase in the amount of organic solvent in the mobile phase due to the lower cohesivity of methanol as compared to water (δ_H = 29.7 MPa for methanol and 47.5 MPa for water) [41]. It also decreases with temperature [30], which can be attributed to a decrease in the solvent cohesivity as a consequence of a disruption of hydrogen bonding as temperature increases. It has also be considered that cohesivity of polar and nonpolar compounds diminishes with temperature, mainly because the molar volumes increase [45, 46].
- *The b coefficient*. It is negative and high, indicating that the stationary phase is much less acidic than the aqueous mobile phase ($a_s < a_m$ in Eq. (2)) [9, 39–41]. The mobile phase is formed by methanol and water, two hydrogen-bond donors much stronger than the aprotic bonded material of the stationary phase (ODS). The magnitude of the b coefficient also decreases with the amount of organic solvent in the mobile phase due to the lower acidity of methanol as compared to water (a_m decreases), and with temperature in agreement with previous observations with acetonitile-water mobile phases [30].
- *The a coefficient*. It is negative, which means that the stationary phase is less avid hydrogen-bond acceptor than the mobile phase ($b_s < b_m$ in Eq. (2)), in agreement with the lower basicity of the ODS phase studied in this work compared to that of the mobile phase. As it is typically observed in RPLC, this contribution is smaller than others and its dependence with changes in solvent composition is very small or nonexistent. This is attributed to a simultaneous and parallel increase of b_s and b_m terms as the amount of

Table 4. LSER coefficients of Eq. (1)[a]

% MeOH	30°C				50°C				70°C			
	40%	50%	60%	70%	40%	50%	60%	70%	40%	50%	60%	70%
c	− 0.54 ± 0.16	− 0.50 ± 0.09	− 0.48 ± 0.08	− 0.56 ± 0.06	− 0.46 ± 0.06	− 0.47 ± 0.06	− 0.49 ± 0.16	− 0.56 ± 0.06	− 0.42 ± 0.06	− 0.45 ± 0.06	− 0.47 ± 0.09	− 0.52 ± 0.07
s	− 0.48 ± 0.13	− 0.48 ± 0.09	− 0.40 ± 0.08	− 0.44 ± 0.06	− 0.53 ± 0.08	− 0.52 ± 0.06	− 0.48 ± 0.05	− 0.42 ± 0.05	− 0.55 ± 0.05	− 0.52 ± 0.06	− 0.52 ± 0.09	− 0.41 ± 0.07
a	− 0.14 ± 0.09	− 0.20 ± 0.06	− 0.28 ± 0.06	− 0.26 ± 0.04	− 0.27 ± 0.06	− 0.28 ± 0.04	− 0.27 ± 0.04	− 0.29 ± 0.04	− 0.34 ± 0.03	− 0.36 ± 0.04	− 0.38 ± 0.06	− 0.33 ± 0.05
b	− 3.13 ± 0.27	− 2.66 ± 0.18	− 2.12 ± 0.06	− 1.69 ± 0.12	− 2.57 ± 0.17	− 2.20 ± 0.01	− 1.86 ± 0.01	− 1.44 ± 0.11	− 2.17 ± 0.01	− 1.85 ± 0.11	− 1.39 ± 0.18	− 1.26 ± 0.013
v	3.06 ± 0.17	2.55 ± 0.09	1.98 ± 0.08	1.66 ± 0.06	2.71 ± 0.09	2.28 ± 0.07	1.87 ± 0.06	1.49 ± 0.06	2.41 ± 0.05	2.02 ± 0.06	1.64 ± 0.09	1.31 ± 0.07
e[b]	*− 0.11 ± 0.14*	*− 0.07 ± 0.09*	*0.005 ± 0.08*	*− 0.01 ± 0.06*	*− 0.03 ± 0.08*	*− 0.02 ± 0.06*	*− 0.01 ± 0.06*	*0.02 ± 0.06*	*0.005 ± 0.05*	*0.02 ± 0.06*	*0.07 ± 0.09*	*0.03 ± 0.07*
R^2	0.9804	0.9907	0.9890	0.9913	0.9925	0.9944	0.9940	0.9912	0.9946	0.9946	0.9822	0.9852
SD	0.12	0.07	0.07	0.05	0.07	0.05	0.05	0.05	0.05	0.05	0.07	0.06

a) Values obtained without the *E* solute parameter.
b) Values for the *e* coefficient of Eq. (1).

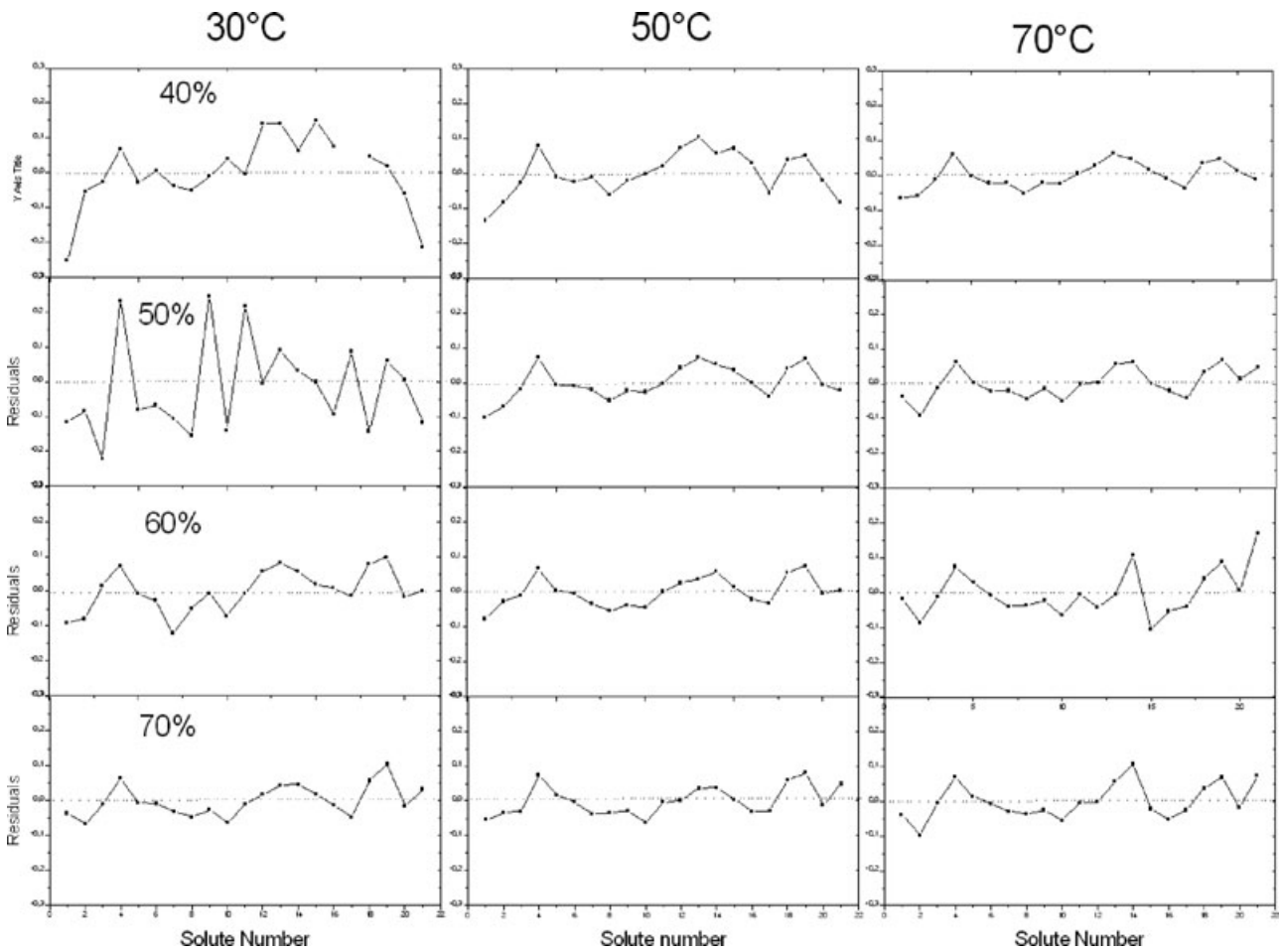

Figure 1. Residual plots for the multiple linear regressions (Eq. (1)) at the different mobile phase compositions and temperatures.

organic solvent is increased since the amount of sorbed organic solvent onto the stationary phase increases [39]. In this work, a slight increase of the *a* coefficient absolute value with temperature was observed. This is conceivable by considering that the b_s term decreases faster than the b_m term due to a desorption of solvent molecules from the stationary phase, and the concomitant decrease of its basicity, as temperature increases.

- *The s coefficient*. It is negative and small, indicative of the higher polarity of the mobile phase as compared with the stationary phase. This coefficient is also independent of changes in mobile phase composition and temperature, probably also due to a simultaneous change of the s_s and s_m terms in the same direction.

3.2 Evaluation of the model: residual analysis and prediction for the test set

The quality of the multiple linear regressions obtained in the previous section should not be evaluated only by the regression coefficients and the standard deviations. Two more accurate procedures to evaluate the LSER model are [39] as follows:

Procedure A: The *residual analysis*, which consists in plotting the differences between the experimental and calculated log*k* values (residuals) for each solute of the training set versus a number assigned to each solute in a systematic way. Residual plots are shown in Fig. 1, all in the same *y*-scale for better comparison. This type of plots are useful to detect some possible outliers in the regressions, which points could indicate either experimental errors or chemical interactions between the outlier compound and the biphasic system, not modeled by Eq. (1). If deviations of this type are present, they usually are not quite visible in plots of experimental versus calculated k values for the training set. The plots of Fig. 1 show that residuals are randomly distributed around zero and no clustering of solutes of the same chemical family occurs, indicating that the LSER coefficients of Table 4 are robust.

Procedure B: The prediction of the log*k* values of a separate set of solutes chemically different from the training set, the *test set*, by using the previously obtained LSER coefficients. In this study, the test set was made by selecting 30 solutes of very different polarity, hydrogen-bond

Table 5. Solute descriptors for the test set[a)]

Solute number	Solutes	*V*	*B*	*A*	*S*	*E*
1	4-Methylanisol	1.0569	0.30	0	0.77	0.699
2	Acebutolol	2.7556	2.10	0.90	2.42	1.600
3	Alprenolol	2.1587	1.44	0.15	1.09	1.250
4	Oxprenolol	2.2174	1.62	0.17	1.49	1.310
5	Propanolol[b)]	2.1480	1.42	0.17	1.43	1.880
6	Metoprolol	2.2604	1.76	0.17	1.33	1.170
7	Ibuprofen[b)]	1.7771	0.60	0.60	0.92	0.700
8	Suprofen	1.9026	0.82	0.57	1.89	1.510
9	Fenbufen	1.9779	1.05	0.62	1.80	1.780
10	Flurbiprofen	1.8389	0.60	0.57	1.51	1.500
11	Indoprofen	2.1100	1.17	0.57	2.30	1.920
12	Ketoprofen	1.9779	0.89	0.55	2.26	1.650
13	Vanillin	1.1313	0.69	0.29	1.33	1.040
14	Methylparaben	1.1313	0.45	0.69	1.37	0.900
15	Propylparaben	1.4131	0.45	0.69	1.35	0.860
16	4-Chloro-*m*-cresol	1.0384	0.22	0.67	1.02	0.920
17	3,4-Dichloroaniline	1.061	0.25	0.35	1.24	1.16
18	Benzene[b)]	0.7164	0.14	0	0.52	0.610
19	Toluene[b)]	0.8573	0.14	0	0.52	0.601
20	Propylbenzene[b)]	1.1391	0.15	0	0.5	0.604
21	Butylbenzene[b)]	1.2800	0.15	0	0.51	0.600
22	Chlorobenzene[b)]	0.8388	0.07	0	0.65	0.718
23	4-Nitrobenzene[b)]	1.1059	0.54	0.68	1.07	0.990
24	Aniline[b)]	0.8162	0.50	0.26	0.96	0.955
25	3-Nitroaniline[b)]	0.9904	0.35	0.40	1.71	1.200
26	α-Naphthylamine[b)]	1.1852	0.57	0.20	1.26	1.670
27	4-Nitrophenol[b)]	0.9493	0.26	0.82	1.72	1.070
28	2,4,6-Trichlorophenol[b)]	1.1423	0.20	0.48	0.94	0.960
29	3-Nitrophenol[b)]	0.9493	0.23	0.79	1.57	1.050
30	*p*-Phenylphenol[b)]	1.3829	0.40	0.59	1.41	1.560

a) Calculated with the ADME Boxes software.
b) Obtained from references indicated in Table 1.

acceptor and donor properties and hydrophilicity, from single solutes such as 4-chloroaniline, ethylbenzene, or phenol to more complexes, polar and bulky solutes such as profens and even protonated β-blockers. In Table 5, the test set with the corresponding solute parameters is depicted. Two procedures to get a function able to estimate log*k* at any mobile phase composition and temperature within the studied ranges were applied:

B.1: Several regressions between each LSER coefficient of Eq. (1) and different expressions of the mobile phase composition and temperature (φ, φ together with φ^2, $1/T$, $1/T$ and $1/T^2$, T, and φ/T) were assayed. The highest regression coefficients (r^2) and lowest standard deviations (SD) were obtained when φ, $1/T$ and φ/T were simultaneously used. Thus, each LSER coefficient of Table 4, *x*, were modeled as follows:

$$x = x_1 + x_2\phi + x_3/T + x_4\phi/T \quad (3)$$

Table 6. Dependence of the LSER coefficients and the intercept, *c*, with volumetric fraction (φ) and temperature (T) according to Eq. (3)

Fitting parameters	LSER coefficients				
	v	*b*	*a*	*s*	*c*
x_1	−7.2 ± 2	12.9 ± 3	−4.2 ± 1	−2.4 ± 1	1.8 ± 1
x_2	7.7 ± 4	−13.2 ± 6	4.5 ± 2	2.4 ± 2	−3.2 ± 2
x_3	3774.9 ± 750	−5600.6 ± 1195	1320.7 ± 345	556.3 ± 349	−710.5 ± 354
x_4	−3886.7 ± 1337	5573.0 ± 2130	−1535.9 ± 615	−709.0 ± 622	960.5 ± 631
R^2	0.9908	0.9790	0.8821	0.7733	0.7044
SD	0.059	0.095	0.028	0.027	0.028

Table 7. Experimental log*k* values for the test set and the corresponding calculated values by using the procedures indicated in Procedures **B.1** (Eq. (1) and (3)) and Procedure **B.2** (Eq. (6))

Solutes	65% MeOH, 40°C			Eq. (6)		45% MeOH, 60°C			Eq. (6)	
	$\log k_{exp}$	$\log k_{calc}$	Residuals[a)]	$\log k_{calc}$	Residuals[a)]	$\log k_{exp}$	$\log k_{cal}$	Residuals[a)]	$\log k_{cal}$	Residuals[a)]
4-Methylanisole	0.452	0.460	−0.007	0.483	−0.030	0.944	0.956	−0.012	0.899	0.045
Acebutolol	−0.749	−0.862	0.114	−0.837	0.088	−0.258	−0.197	−0.061	−0.497	0.240
Alprenolol	−0.186	0.097	−0.284	0.159	−0.346	0.383	0.799	−0.416	0.541	−0.159
Oxprenolol	−0.428	−0.318	−0.110	−0.265	−0.163	0.112	0.321	−0.208	0.047	0.065
Propanolol	−0.261	−0.038	−0.223	0.008	−0.269	0.306	0.633	−0.327	0.396	−0.089
Metoprolol	−0.525	−0.436	−0.089	−0.370	−0.155	0.040	0.195	−0.155	−0.113	0.153
Ibuprofen	0.863	0.957	−0.094	0.981	−0.118	1.653	1.717	−0.064	1.627	0.026
Suprofen	0.124	0.349	−0.226	0.343	−0.219	0.793	1.022	−0.229	0.934	−0.141
Fenbufen	0.407	0.076	0.332	0.082	0.325	0.818	0.720	0.098	0.585	0.233
Flurbiprofen	0.666	0.817	−0.152	0.817	−0.151	1.443	1.561	−0.118	1.497	−0.055
Ketoprofen	0.292	0.196	0.096	0.177	0.115	0.994	0.855	0.139	0.770	0.224
Indoprofen	0.140	−0.120	0.260	−0.128	0.268	—	—	—	—	—
Vanillin	−0.452	−0.466	0.014	−0.454	0.001	−0.124	−0.120	−0.004	−0.217	0.093
Methylparaben	−0.237	−0.137	−0.101	−0.142	−0.095	0.168	0.268	−0.099	0.231	−0.062
Propylparaben	0.197	0.377	−0.179	0.373	−0.176	0.776	0.940	−0.165	0.902	−0.127
4-Chloro-*m*-cresol	0.213	0.290	−0.077	0.290	−0.077	0.713	0.751	−0.037	0.741	−0.028
3,4-Dichloroaniline	0.132	0.263	−0.131	0.259	−0.127	0.646	0.720	−0.074	0.706	−0.059
Benzene	0.257	0.262	−0.005	0.287	−0.030	0.671	0.643	0.028	0.605	0.066
Toluene	0.513	0.514	−0.001	0.540	−0.027	0.992	0.974	0.019	0.936	0.056
Propylbenzene	0.977	1.008	−0.031	1.037	−0.059	1.627	1.624	0.002	1.584	0.043
Butylbenzene	1.228	1.256	−0.028	1.285	−0.057	1.961	1.950	0.011	1.910	0.051
Chlorobenzene	0.485	0.556	−0.071	0.574	−0.089	0.989	1.017	−0.029	0.999	−0.011
4-Nitrobenzene	−0.183	−0.218	0.035	−0.207	0.024	0.191	0.169	0.022	0.101	0.090
Aniline	−0.699	−0.501	−0.197	−0.482	−0.217	−0.538	−0.234	−0.304	−0.313	−0.225
3-Nitroaniline	−0.359	−0.271	−0.088	−0.292	−0.067	−0.007	0.069	−0.076	0.059	−0.066
α-Naphthylamine	−0.076	−0.088	0.012	−0.077	0.001	0.315	0.338	−0.023	0.258	0.057
4-Nitrophenol	−0.222	−0.291	0.068	−0.322	0.100	0.105	0.037	0.068	0.057	0.048
2,4,6-Trichlorophenol	0.590	0.599	–0.009	0.604	–0.014	1.144	1.140	0.004	1.125	0.019
3-Nitrophenol	−0.206	−0.160	−0.045	−0.186	−0.019	0.170	0.192	−0.022	0.210	−0.041
p-Phenylphenol	0.337	0.417	–0.080	0.411	−0.073	0.944	0.956	−0.012	0.899	0.045

a) Residuals = log*k* (experimental) – log*k* (calculated).

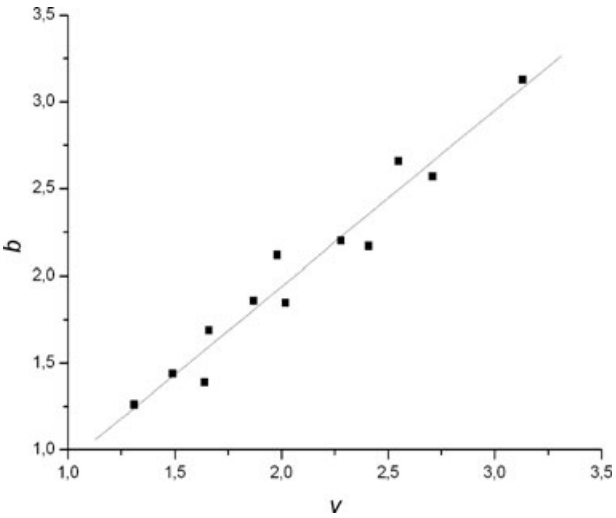

Figure 2. Relationship between *b* and *v* coefficients at the different temperatures and mobile phase compositions.

In Table 6, the results for the obtained regressions coefficients x_1, x_2, x_3, and x_4 along with the r^2 and SD values are shown. The corresponding *P* values were <0.05 except for the x_i coefficients when Eq. (3) applies to the *s* and *c* LSER coefficients, for which *P* values were between 0.15 and 0.28. The log*k* values for the test set were obtained by using two interpolated set of experimental conditions: 65% MeOH at 40°C and 45% MeOH at 60°C. In Table 7 the experimental and calculated log*k* values are shown. Predictions of log*k* are very good as indicated by their average absolute residuals of the log*k* values (differences between experimental and calculated), which are 0.102 and 0.113 for the first and second set of experimental conditions, respectively. The root mean squared error (RMSE) for the predicted values are 13 and 14% for the 65%-40°C and 45%MeOH-60°C experimental conditions, respectively. This indicates that the LSER coefficients are chemically significant and that the SP model generated here is, therefore, suitable for predicting retention factors for many solutes in this specific

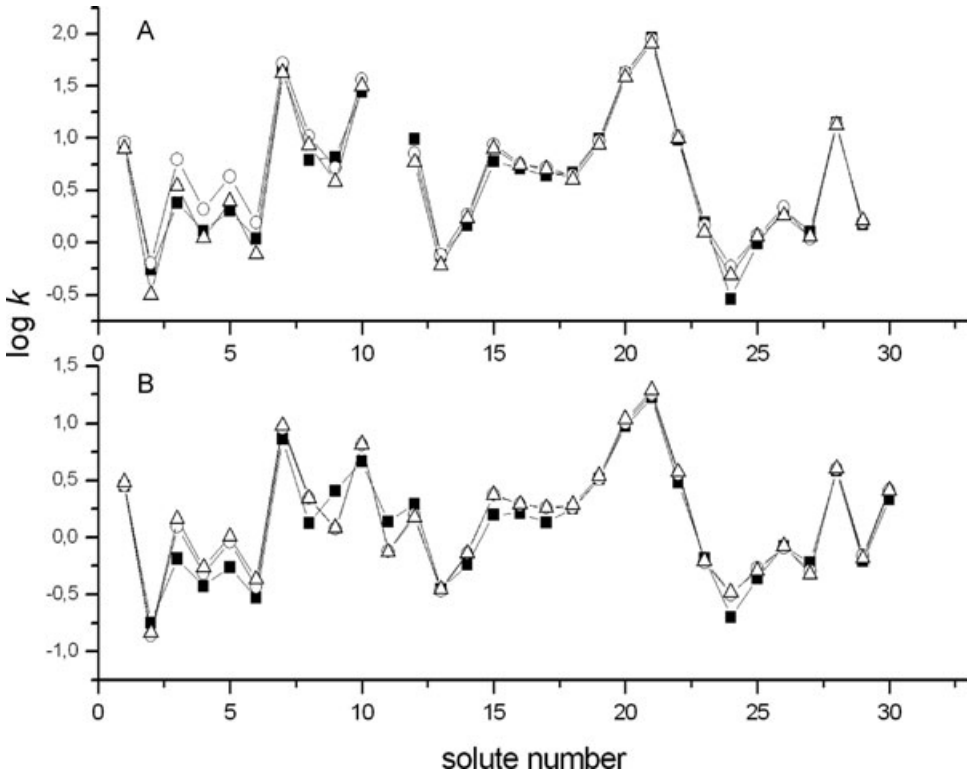

Figure 3. Experimental (■) and predicted log*k* values by calculating the LSER coefficients using the Eq. (3) and its fitting parameters of Table 6 (○), and by Eq. (6) (△) versus the solute number of Table 5. (**A**) 45% MeOH-60°C (**B**) 65% MeOH-40°C.

column at different temperatures and methanol mobile phase compositions.

B.2: In an attempt to reduce the number of experimental variables needed to predict retention, another approach is proposed. From Table 4, it can be observed that the *a* and *s* coefficients and also the intercept *c* (considered with their corresponding standard deviations and *P* values) remained virtually constant as mobile phase composition and temperature changed. This is in agreement with previous results obtained with methanol–water mobile phase in an ODS phase at room temperature [47, 48]. Thus, it is possible to average these coefficients and obtain a reduced equation as follows:

$$\log k = -0.49 - 0.48\mathrm{S} - 0.28\mathrm{A} + \mathrm{bB} + vV \quad (4)$$

Additionally, the linear dependence between *v* and *b* was previously observed for different mobile phase compositions, temperatures, and different stationary phases [9, 30, 49] but not observed e.g. at different mobile phase compositions and temperatures at the same time. Poole et al. proposed to linearly relate those LSER parameters at a given mobile phase composition or temperature [40]. Figure 2 shows the relationship between *b* and *v* coefficients at all mobile phase compositions and temperatures. Thus, we can linearly relate these coefficients at any mobile phase composition and temperature in the following way:

$$b = 0.115(\pm 0.07) - 1.0 \pm 0.1v$$
$$R^2 = 0.948; \mathrm{SD} = 0.133 \quad (5)$$

These values are surprisingly very close to those obtained in reference [9] for different stationary phases at a given mobile phase composition and temperature and, also, in reference [30] for a given stationary phase and different mobile phase compositions and temperatures. From these observations, it is possible to propose a single equation of chromatographic retention with only one single parameter, e.g. *v* or *b* applicable to any solute at any MeOH-water composition and temperature. The maximal accuracy will be in the used calibration zone between 40–70% MeOH and 30–70°C. By using Eq. (3) in combination with the fitting parameters gathered in Table 6 to calculate the *v* coefficient, and then Eq. (5) to obtain the corresponding *b* coefficient, it is possible to deduce the following general equation:

$$\log k = -0.49 - 0.48S - 0.28A + (-7.2 + 7.7\phi + 3774.9/T - 3886.7\phi/T)V + (7.4 \times 7.9\phi 3883.4/T + 3999.4\phi/T)B \quad (6)$$

Figure 3 compares the experimental log*k* values for the test set with the calculated ones by the two procedures mentioned in Procedures B.1 and B.2 at the two studied experimental conditions (65%-40°C and 45%-60°C) versus the solute number. Also, plots of experimental *k* values versus the calculated *k* ones are shown in Fig. 4. In these plots, a perfect prediction will be obtained if the intercept is zero and the slope is one. This is quiet well achieved as observed in the values shown within the figure. Also, the log*k* values and their residuals are gathered in Table 7. As can be

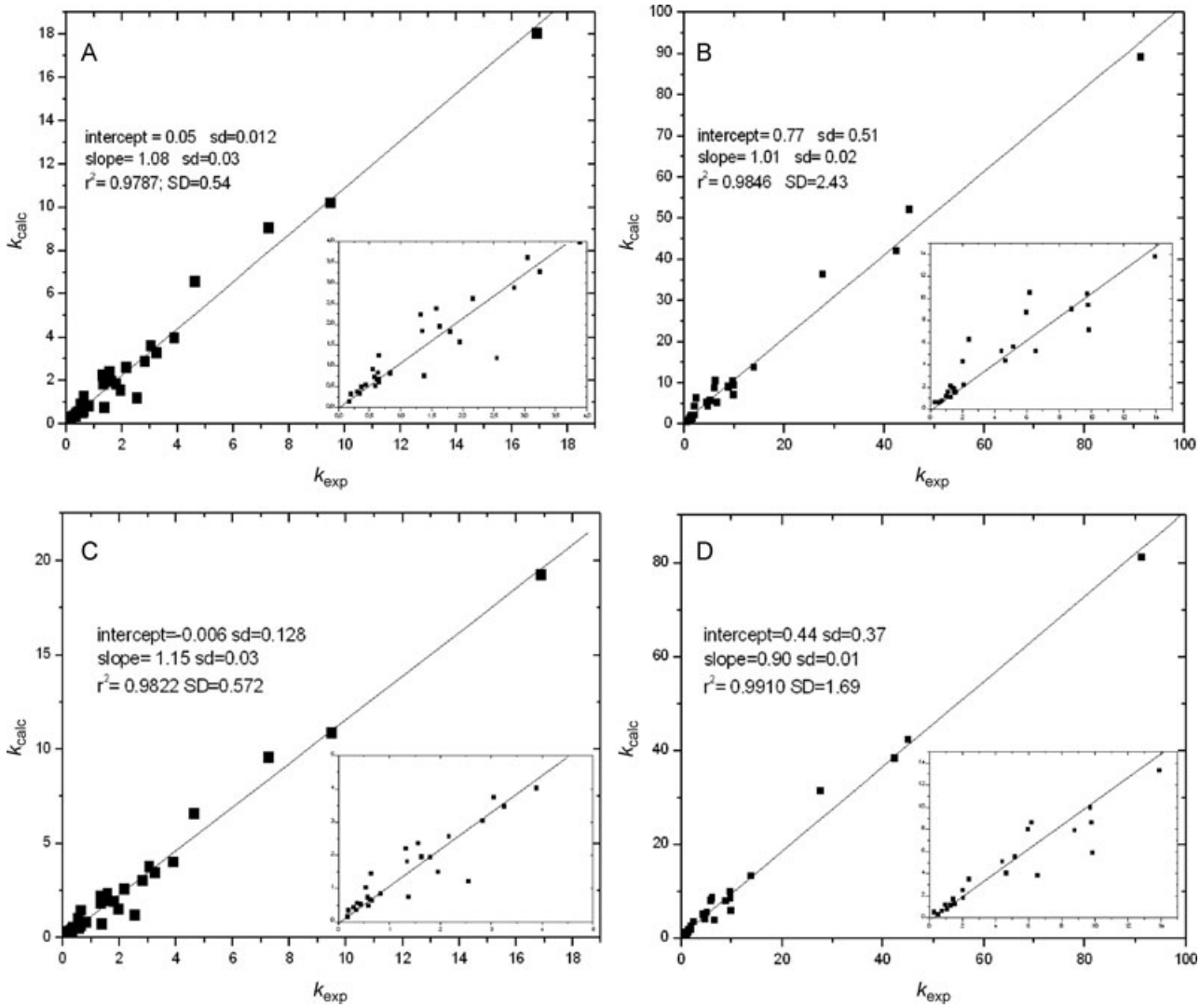

Figure 4. Experimental versus predicted k values by calculating the LSER coefficients using the Eq. (3) and its fitting parameters of Table 6 (plots A and B), and by Eq. (6) (plots C and D). (**A** and **C**): 45% MeOH-60°C; (**B** and **D**) 65% MeOH-40°C.

observed, the predicted retention by the two procedures is very good. Most of the residuals are lower than 0.02 log*k* units. The average absolute residuals of the log*k* values are 0,099 and 0,091 for the first and second set of experimental conditions, respectively. The RMSE for the predicted values are 14 and 12% for the 65%-40°C and 45%MeOH-60°C experimental conditions, respectively. These errors are a bit lower than the ones obtained with the previous model of Procedure B.1. Thus, method B.2 performs something better than the other procedure. Even though, some ionized compounds such as the five β-blockers have residuals within the experimental error. From all these results, the Eq. (6) can be considered as a practical and very accurate way to obtain retention times in the studied C18 phase at any methanol-water mobile phase composition and temperature near to the ranges studied in this work. These results also confirm that the experimental design made in Section 3.1 was accurate.

4 Concluding remarks

Retention of analytes of very different chemical nature under RPLC conditions was studied by using several methanol-water mobile phase compositions and temperatures through the use of the SP model. The LSER coefficients obtained by multiple linear regressions between the logarithm of the retention factor and different solute descriptors were critically analyzed. The model showed to be useful for predicting retention factors for 30 chemically different solutes in a C18 stationary phase at interpolated methanol mobile phase compositions and temperatures. The retention predictions were obtained for any solute by following two different procedures: (i) by calculating each LSER coefficient at a given mobile phase composition and temperature through the use of fitting parameters, or (ii) by using a general and practical equation in which only the mobile phase composition, temperature, and the solute parameters are necessary. Predicted

k values with both models were very close to the experimental data. However, the second model was slightly better than the first one.

Financial support from the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET, PIP2011–0777) and Agencia Nacional de Promoción Científica y Técnica (ANPCYT, PICT2007–00316) is gratefully acknowledged.

The authors have declared no conflict of interest.

5 References

[1] Melander, W. R., Stovekeen, J., Horváth, C., *J. Chromatogr. A* 1979, *185*, 111–127.

[2] Abraham, M. H., Grellier, P. L., Prior, D. V., Morris, J. J., Taylor, P. J, *J. Chem. Soc., Perkin Trans. 2* 1990, 521–529.

[3] Abraham, M. H., Grellier, P. L., Prior, D. V., Duce, P. P., Morris, J. J., Taylor, P. J., *J. Chem. Soc., Perkin Trans. 2* 1989, 699–711.

[4] Poole, C. F., Atapattu, S. N., Poole, S. K., Bell, A. K., *Anal. Chim Acta* 2009, *652*, 32–53.

[5] Kaliszan, R., van Straten, M. A, Markuszewski, M, Cramers, C. A., Claessens, H. A., *J. Chromatogr. A* 1999, *855*, 455–486.

[6] Kollie, T. O., Poole, C. F., *J. Chromatogr. A* 1991, *556*, 457–484.

[7] Poole, C. F., *The Essence of Chromatography*, Elsevier, Amsterdam 2003.

[8] Abraham, M. H., Zhao, Y. H., *Phys. Chem. Chem. Phys.* 2005, *7*, 2418–2422.

[9] Reta, M., Carr, P. W., Sadek, P. C., Rutan, S. C., *Anal. Chem.* 1999, *71*, 3484–3496.

[10] Li, J., Zhang, Y., Ouyang, H., Carr, P. W., *J. Am. Chem. Soc.* 1992, *114*, 9813–9828.

[11] Wang, A., Carr, P. W., *J. Chromatogr. A* 2002, *965*, 3–23.

[12] Wilson, N. S., Nelson, M. D., Dolan, J. W., Snyder, L. R., Carr, P. W., *J. Chromatogr. A* 2002, *961*, 171–193.

[13] Bocian, S., Fellinger, A., Buszewsky, B., *Chromatographia* 2008, *68*, 19–26.

[14] Bocian, S., Vajda, P., Felinger, A., Buszewsky, B., *Anal. Chem.* 2009, *81*, 6334–6346.

[15] Cheong, W. J., Carr, P. W., *J. Chromatogr. A* 1990, *499*, 373–393.

[16] Jones, J. L., Rutan, S. C., *Anal. Chem.* 1991, *63*, 1318–1322.

[17] Carr, P. W., Harris, J. M., *Anal. Chem.* 1987, *59*, 2546–2550.

[18] Lochmüller, C. H., Marshall, D. B., Wilder, D. R., *Anal. Chim. Acta* 1981, *130*, 31–43.

[19] Waksmundzka-Hajnos, M., Petruczynik, A., Cies, G., *J. Liq. Chromatogr. Relat. Technol.* 2007, *30*, 2473–2484.

[20] Gotta, J., Osorio Grisales, J., Reta, M. R., Castells, C. B., *J. Sep. Sci.* 2010, *33*, 2645–2653.

[21] Teutenberg, T., *High-Temperature Liquid Chromatography. A User's Guide for Method Development*, The Royal Society of Chemistry, Cambridge 2010.

[22] Tran, J. V., Molander, P., Greibrokk, T., Lundanes, E., *J. Sep. Sci.* 2001, *24*, 930–940.

[23] Zhu, P. L., Dolan, J. W., Snyder, L. R., *J. Chromatogr. A* 1996, *756*, 41–50.

[24] Zhu, P. L., Dolan, J. W., Snyder, L. R., Djordjevic, N. M., et al., *J. Chromatogr. A* 1996, *756*, 63–72.

[25] Zhu, P. L., Dolan, J. W., Snyder, L. R., Hill, D. W., et al., *J. Chromatogr. A* 1996, *756*, 51–62.

[26] Zhu, P. L., Snyder, L. R., Dolan, J. W., Djordjevic, N. M., et al., *J. Chromatogr. A* 1996, *756*, 21–39.

[27] Bolliet, D., Poole, C. F., *Analyst* 1998, *123*, 295–299.

[28] Kiridena, W., Poole, C. F., Koziol, W. W., *J. Chromatogr. A* 2004, *1060*, 177–185.

[29] Poole, C. F., Ahmed, H., Kiridena, W., DeKay, C., Koziol, W. W., *Chromatographia* 2005, *62*, 553–561.

[30] Carr, P. W., Doherty, R. M., Kamlet, M. J., Taft, R. W., Melander, W., Horvath, C., *Anal. Chem.* 1986, *58*, 2674–2680.

[31] Pappa-Louisi, A., Nikitas, P., Papachristos, K., Zisi, C., *J. Chromatogr. A* 2008, *1201*, 27–34.

[32] Jouyban, A., Soltanpour, S., Acree, W. E., Thomas Jr., D., Pappa-Louisi, A., *J. Sep. Sci.* 2009, *32*, 3898–3905.

[33] Espinosa, S., Bosch, E., Rosés, M., *J. Chromatogr. A* 2002, *945*, 83–96.

[34] Rosés, M., Subirats, X., Bosch, E., *J. Chromatogr. A* 2009, *1216*, 1756–1775.

[35] Padró, J. M. P., A., Agudelo Mesa, L. B., Reta, M., *Anal. Bioanal. Chem.* 2011, *399*, 2807–2820.

[36] Christensen, J. J., Hansen, L. D., Izzat, R. M., *Handbook of Proton Ionization Heats and Related Thermodynamic Quantities*, John Wiley & Sons, New York 1976.

[37] Rosés, M., Bosch, E., *J. Chromatogr. A* 2002, *982*, 1–30.

[38] Castells, C. B., Gagliardi, L. G., Ráfols, C., Rosés, M., Bosch, E., *J. Chromatogr. A* 2004, *1042*, 23–36.

[39] Vitha, M., Carr, P. W., *J. Chromatogr. A* 2006, *1126*, 143–194.

[40] Ali, Z., Poole, C. F., *J. Chromatogr. A* 2004, *1052*, 199–204.

[41] Wang, A., Tan, L. C., Carr, P. W., *J. Chromatogr. A* 1999, *848*, 21–37.

[42] Poole, C. F., Poole, S. K., Abraham, M. H., *J. Chromatogr. A* 1998, *798*, 207.

[43] Abraham, M. H., *Chem. Soc. Rev.* 1993, *22*, 73.

[44] Abraham, M. H., Chadha, H. S., Whiting, G. S., Mitchell, R. C., *J. Pharm. Sci.* 1994, *83*, 1085.

[45] Abraham, M. H., Rosés, M., Poole, C. F., Poole, F. K., *J. Phys. Org. Chem.* 1997, *44*, 151.

[46] Barton, A. F. M., *Handbook of Solubility Parameters and Other Cohesion Parameters*, CRC Press, Boca Raton, FL 1983.

[47] Kiridena, W., Atapattu, S. N., Poole, C. F., Koziol, W. W., *Chromatography* 2008, *68*, 11–17.

[48] Poole, C. F., Poole, S. K., *J. Chromatogr. A* 2002, *965*, 263–299.

[49] Torres-Lapasió, J. R., García-Alvarez-Coque, M. C., Rosés, M., Bosch, E., *J. Chromatogr. A* 2002, *955*, 19–34.

