

Two- and three-parameter equations for representation of retention data in reversed-phase liquid chromatography

A. Pappa-Louisi, P. Nikitas*, P. Balkatzopoulou, C. Malliakas

Laboratory of Physical Chemistry, Department of Chemistry, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

Received 14 October 2003; received in revised form 9 January 2004; accepted 20 January 2004

Abstract

Two-parameter equations that describe the dependence of $\ln k$ upon φ , where k is the retention factor and φ the volume fraction of the organic modifier in the mobile phase, are examined in what concerns the underlying approximations and their performance to fit experimental data obtained from reversed-phase liquid chromatography. Using 293 experimental systems, it was found that the performance of these equations to describe $\ln k$ versus φ data is rather low, since the percentage of the systems that can be described satisfactorily ranges from 40 to 60% depending on the fitting equation. This percentage may be raised to 75%, if the discreteness effect is properly taken into account. A further improvement to 90% of the systems studied can be achieved only by the use of three-parameter equations, which may arise by refinements of the rough approximations of the two-parameter equations. Although the refinements do not lead always to better equations, we developed a new three-parameter expression of $\ln k$ that works more satisfactorily, since it combines simplicity, linearity of its adjustable parameters and the highest applicability.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Liquid chromatography; Retention models; Mathematical modelling

1. Introduction

The problem of the accurate representation of retention time, t_R , versus φ data, where φ is the volume fraction of the organic modifier in the mobile phase, is of high importance in liquid chromatography and in particular in designing proper optimisation techniques. In the conventional approach, the original t_R versus φ data is transformed into $\ln k$ versus φ data, where k is the retention factor, and these data are modelled using various empirical or strict theoretical equations. The tendency is to use as simple equations as possible in order to avoid numerical difficulties and reduce the number of experiments needed for an optimisation technique. In this trend, the two-parameter equation proposed by Johnson et al. [1,2] and based on the E_T scale for mobile phase polarity is extensively used for modelling retention data especially for practical optimisation and prediction techniques [1–14]. Here, we examine first whether two-parameter equations can actually be used for an accept-

able prediction of retention times in different φ values and second, if they can be modified to increase their predictive capabilities.

2. Theoretical part

2.1. Two-parameter equations

Retention in reversed-phase liquid chromatographic columns is ruled by the solute multiple interactions with both the stationary and the mobile phase constituents. Despite the complex nature of these interactions that may lead to different retention mechanisms, a common observation is that retention increases with the increase in mobile phase polarity. This observation has led Dorsey and co-workers to suggest the following linear relationship between $\ln k$ and the polarity of the mobile phase expressed through the E_T^N solvatochromic parameter [1,2]:

$$\ln k = m + nE_T^N \quad (1)$$

where m and n are adjustable parameters characteristic of the solute properties. Note that initially the polarity of the

* Corresponding author. Tel.: +30-2310-997773; fax: +30-2310-997709.

E-mail address: nikitas@chem.auth.gr (P. Nikitas).

mobile phase was expressed through the non-normalised $E_T(30)$ scale [1–6], which later was replaced by the normalised E_T^N parameter [7,8].

Another two-parameter equation comes from the adsorption model for the retention mechanism developed in [15] and it may be written as:

$$\ln k = a - \ln(1 + b\varphi) - \frac{c\varphi}{1 + b\varphi} \quad (2)$$

where a , b and c are adjustable parameters. Eq. (2) seems to be a three-parameter equation but, according to the theory [15,16], parameter b depends only upon the modifier. Consequently, it is, in fact, a two-parameter equation provided that a proper b value has been determined for a certain modifier.

Finally, it can be easily shown by numerical examples that the function $\ln(1 + b\varphi)$ can be effectively represented by the rational function:

$$f(\varphi) = \frac{q\varphi}{1 + b\varphi} \quad (3)$$

especially for $b < 2$. It is seen that the term $\ln(1 + b\varphi)$ of Eq. (2) may be absorbed by the last term of the right-hand side of this equation and this observation has led us to examine whether the following simple equation:

$$\ln k = a - \frac{c\varphi}{1 + b\varphi} \quad (4)$$

can be also used as a two-parameter equation by assuming a constant b value for each modifier. It is evident that when Eqs. (2) and (4) are applied to experimental data, parameters a , b and c are treated as adjustable parameters and therefore these parameters and especially parameter c are unlikely to take common values in the two equations. Eq. (4) is further simplified to:

$$\ln k = a' + \frac{c'}{1 + b\varphi} \quad (5)$$

because $c\varphi/(1 + b\varphi) = (c/b) - ((c/b)/(1 + b\varphi))$.

An interesting result arises from the comparison of Eqs. (1) and (5). If we take into account that the normalised E_T^N factor takes the value 1 at $\varphi = 0$, we readily obtain that $m + n = a' + c'$, which, by equating the right-hand sides of Eqs. (1) and (5), results in the following expression of E_T^N :

$$E_T^N = \frac{1 + p\varphi}{1 + b\varphi} \quad (6)$$

where p and b are constants that take certain values for each modifier. A similar expression for E_T^N has been proposed by Roses and Bosch [7,17] who showed that Eq. (6) represents very satisfactorily the experimental data.

Eq. (6) allows for an alternative expression of Eq. (1). If we change b to b_e in Eq. (6) for reasons explained further and divide $1 + p\varphi$ by $1 + b_e\varphi$, we obtain:

$$E_T^N = \frac{1 + p\varphi}{1 + b_e\varphi} = \frac{p}{b_e} + \frac{(b_e - p)/b_e}{1 + b_e\varphi} \quad (7)$$

which, by substitution into Eq. (1), yields:

$$\ln k = m' + \frac{n'}{1 + b_e\varphi} \quad (8)$$

where m' and n' are adjustable parameters, whereas b_e is determined from fitting E_T^N versus φ data to Eq. (6). Note that Eq. (8) is formally identical to Eq. (5) and differ only in the choice of the constants b and b_e . In Eq. (8), b_e necessarily takes the value determined from fitting E_T^N versus φ data to Eq. (6), whereas this limitation is not imposed on b of Eq. (5) which is treated as adjustable parameter. For this reason, we use different symbols for the constant appeared in the denominator of Eqs. (5) and (8). Note also that when Eq. (7) is adopted for the calculation of E_T^N values used in Eq. (1), then Eqs. (1) and (8) give identical results.

2.2. A critique of the two-parameter equations

The two-parameter Eqs. (1), (2), (4), (5) and (8) are extremely simplified expressions of $\ln k$ based on several rough approximations. For this reason, the applicability of such an equation to a certain class of solutes does not entail the validity of its model and no molecular information can be gained from its use.

A detail discussion on the approximations underlying Eq. (2) is presented in [16]. According to it, Eq. (2) is based on the adsorption model for the retention mechanism and assumes: (a) the independence of the molar volumes of the mobile phase constituents from the composition of this phase; (b) the validity of the Langmuir isotherm for the adsorption of the modifier on the hydrocarbon chains; and (c) the random approximation for the solute–solvent interactions at the adsorbed layer. It is seen that all these assumptions are rough approximations. In particular, the validity of the Langmuir isotherm assumes an ideal behaviour of the organic modifier–water mixture both at the adsorbed layer and in the mobile phase, an assumption that strongly contradicts with vapour–liquid equilibrium (VLE) tabulated data [18], which show significant deviations from the ideal behaviour in the mobile phase [16].

Eqs. (4) and (5) come from an empirical modification of Eq. (2), which, in fact, eliminates the entropy contribution to the retention mechanism (see further). Therefore, these equations are based on the same approximations we met in Eq. (2) and they additionally overlook entropy effects.

The limitations of Eq. (1) were first indicated by Cheong and Carr [6] who showed that the performance of this equation is good only over a narrow range of solvent composition. According to these authors, at least two solvent parameters are needed to account for the cavity formation and solute–solvent interactions in the mobile phase. Later, Barbosa and co-workers [10–13] studied several classes of solutes in acetonitrile–water mobile phases and observed two regions and therefore two straight lines at the plots of $\ln k$ versus E_T^N showing that the single Eq. (1) is incapable of describing the retention of these solutes. This behaviour

was attributed to the existence of three different structural regions in acetonitrile–water mixtures [10–13,19]. It is seen that Cheong and Carr attribute the limitations of Eq. (1) to the insufficient treatment of the various interactions in the mobile phase, whereas Barbosa et al. consider the different structural regions of acetonitrile–water mixtures as the main responsible factor of the limitations of Eq. (1). Our view on this issue is the following.

It is evident that the performance of Eq. (1) is associated with the approximations involved in its derivation. From the definition of $\ln k$, we have $\ln k = -\Delta G_r^\circ/RT$, where ΔG_r° is the standard free energy of retention. Noting that E_T^N is a measure of the mobile phase polarity, we readily conclude that the term nE_T^N of Eq. (1) is the energy contribution of the solute interactions with the constituents of the mobile phase. It is seen that Eq. (1) assumes that: (a) the only contribution to $\ln k$ comes from solute interactions in the mobile phase; and (b) these interactions can be represented by the product nE_T^N . Assumption (a) either disregards other contributions, like contributions from the differences in the molecular volumes of the organic modifier and the water in the mobile phase, entropy effects and contributions from the solute interactions with the constituents of the stationary phase, or it assumes that all these contributions are independent of φ and therefore they are included in the constant parameter m of this equation. In what concerns assumption (b), it obviously disregards the discreteness of the solute–solvent interactions. Therefore, the approximations on which Eq. (1) is based disregard at least the following effects:

- (i) differences in the molecular volumes of the organic modifier and the water in the mobile phase;
- (ii) discreteness effects in the various interactions;
- (iii) entropy effects;
- (iv) solute interactions with the constituents of the stationary phase.

2.2.1. Differences in the molecular volumes

In a recent paper [16], we have shown that, apart from other contributions, the differences in the molecular volumes of the organic modifier and the water in the mobile phase affect $\ln k$. Overlooking this contribution means that we accept the rough assumption that the molar volumes of the mobile phase constituents are independent of the composition of this phase, i.e. independent of φ . If we denote by $\ln k_c$ the part of $\ln k$ that is free from this contribution, then $\ln k$ and $\ln k_c$ are interrelated through the following equation [16]:

$$\ln k = \ln k_c + \ln \frac{1 - \alpha + \alpha x}{(1 - \alpha)(1 + \delta)} \quad (9)$$

where x is the organic modifier mole fraction, δ the percentage contraction of the mobile phase volume caused by the mixing of its constituents, and α given by:

$$\alpha = 1 - \frac{\rho_B/M_B}{\rho_w/M_w} \quad (10)$$

where ρ_B and ρ_w are the densities of the pure organic modifier and water, respectively, and M_B and M_w their molecular masses. The mole fraction x is related to the volume fraction φ through the following relationship [15]:

$$x = \frac{\varphi(1 - \alpha)}{1 + \delta - \varphi\alpha} \quad (11)$$

Note that δ has a very small contribution to the above equations [16]. Therefore, δ can be eliminated without any significant effect on the results. In this case, if Eq. (11) is substituted into Eq. (9) and take into account that $\ln k_c$ may be expressed by Eq. (1) or (8), we readily obtain that:

$$\begin{aligned} \ln k &= -\ln(1 - \varphi\alpha) + m + nE_T^N \\ &= -\ln(1 - \varphi\alpha) + m' + \frac{n'}{1 + b_e\varphi} \end{aligned} \quad (12)$$

which is also a two-parameter equation. The same correction may be made to Eqs. (2) and (5).

2.2.2. Discreteness effects

Each solute molecule, depending on the strength of the solute–water and solute–modifier interactions changes the average orientation and composition of the solvent (water + organic modifier) molecules that surround it. Thus, at the same mobile phase structurally different solute molecules “see”, in fact, a different solvent environment due to the discreteness of the solute–solvent interactions. This discreteness effect is totally ignored when the solute–solvent interactions are taken into account through the term nE_T^N in Eq. (1).

Qualitatively the discreteness effect is expected to have the following consequences. An organic solute molecule interacts attractively stronger with the molecules of the organic modifier than with the water molecules. Therefore, the stronger these interactions are, the higher the number of the modifier molecules that surrounds each solute molecule is expected to be. That is, the stronger the attractive solute–solvent interactions are, the higher the effective φ value is, resulting in a decrease in the effective E_T^N value. Schematically this consequence of the discreteness effect is shown in Fig. 1. Moreover, it is shown in Appendix A that if the effective E_T^N versus φ curve is still described by Eq. (6) (or Eq. (7)), parameter b (or b_e) should be higher than that of the original E_T^N versus φ curve.

Therefore, the discreteness effect is expected to influence the value of b in Eq. (5). Values of b , obtained when Eq. (5) is fitted to experimental data, close to the corresponding values of b_e determined from fitting Eq. (7) to experimental data is an indication that the discreteness effect is weak. This is expected for non-polar solutes with small molecules. However, as the molecular volume of a solute is increased, the solute–solvent interactions become stronger, because the number of the solute contacts with the surrounding molecules is increased. Therefore, the greater the molecular weight of a solute is, the poorer the performance

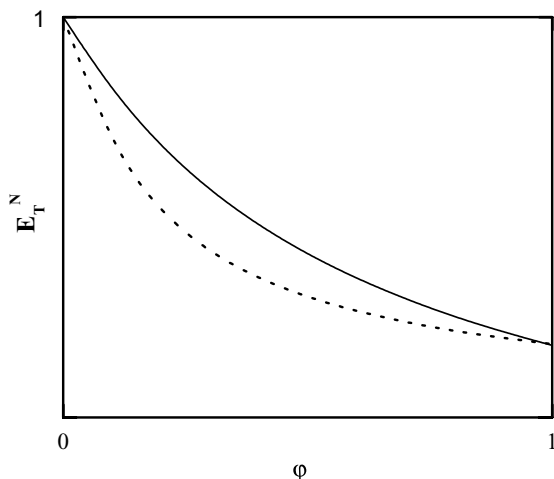


Fig. 1. Schematic plots of E_T^N vs. φ in the presence (\cdots) and absence ($—$) of discreteness effects.

of Eq. (1) is expected to be. In addition, if we like to account effectively for the discreteness effect, then we should use in Eq. (5), a constant b value per group of structurally similar solutes at a certain modifier rather than a constant b value at each modifier irrespective of the structure and the molecular weight of the solutes. In this case, we expect the greater the mean molecular weight of a group of structurally similar solutes is, the higher the b value used in Eq. (5) to be.

2.2.3. Entropy effects

According to the treatment in [15,16], entropy effects contribute to $\ln k$ only in case that the adsorption mechanism plays a dominant role in the solute retention. Then, the entropy contribution to $\ln k$ may be expressed as [15,16]:

$$\frac{\Delta S}{k} = \ln \left(\frac{1 - \theta}{1 - \varphi} \right) \quad (13)$$

where θ is the surface coverage of the hydrocarbon chains by the modifier molecules and k the Boltzmann's constant. The main approximation adopted in the derivation of this equation is that during the adsorption process each adsorbate (solute or modifier) molecule replaces from the adsorbed layer a cluster of water molecules, which always has dimensions equal to those of the adsorbate molecule. Note that there are a lot of experimental and theoretical evidences that this is a good approximation [20–23]. Thus, the only problem we have to overcome is to express θ in terms of φ . However, this relationship is, in fact, the adsorption isotherm concerning the adsorption of the modifier molecules on the chains of the stationary phase. Now if we like to keep as simple expression for $\ln k$ as possible, we are obliged to adopt the rough approximation of the Langmuir isotherm:

$$\frac{\theta}{1 - \theta} = \beta \frac{\varphi}{1 - \varphi} \quad (14)$$

where β is the adsorption equilibrium constant. Substitution of Eq. (14) into Eq. (13) yields:

$$\frac{\Delta S}{k} = -\ln[1 + (\beta - 1)\varphi] \quad (15)$$

It is seen that this contribution is taken into account in Eq. (2), which is derived from the adsorption model.

2.2.4. Contributions from the stationary phase

As a first approximation, the contribution of the solute interactions at the stationary phase to $\ln k$ may be accounted for by adding to the expression of $\ln k$, a term like qE_T^{Ns} , where E_T^{Ns} is the corresponding E_T^N factor of the stationary phase, i.e. a measure of the polarity of this phase. It is evident that the expression of E_T^{Ns} and therefore the contribution of the solute interactions at the stationary phase depend on the retention mechanism.

If the retention is due to partition, E_T^{Ns} is a constant, because the stationary phase is just the hydrocarbon chains [8]. In this case, the term qE_T^{Ns} is also a constant, which can be included in parameter m of Eq. (1). In contrast, if the retention is due to adsorption, a surface solution is formed on the chains and E_T^{Ns} is the measure of its polarity, i.e. a measure of the solvent (water + organic modifier) polarity at the surface solution of the stationary phase. Therefore, we may assume that E_T^{Ns} is given by an expression like Eq. (6) but with different b and p values, say b^s and p^s , and with the surface coverage θ of the modifier in place of φ . Thus, we are again forced to use the Langmuir isotherm, Eq. (14), to correlate θ with φ . Then, we readily find that E_T^{Ns} may be expressed as:

$$E_T^{Ns} = \frac{1 + p^*\varphi}{1 + b^*\varphi} \quad (16)$$

where $p^* = \beta(1 + p^s) - 1$ and $b^* = \beta(1 + b^s) - 1$.

It is evident that p^* and b^* depend exclusively upon the modifier but the problem is that they cannot be determined by independent experimental data. Therefore, we have to use additional approximations. The most radical approximation, coming from the low curvature of the E_T^N versus φ curves in the mobile phase, is to assume a linear dependence of E_T^{Ns} upon φ , i.e.:

$$E_T^{Ns} = q'\varphi \quad (17)$$

A more refined approach is to assume that for non-polar solutes with small molecules parameters b^s and p^s take values close to those of b_e and p of Eq. (7), because in this case the discreteness effect is small and the surface solution is likely to have a structure similar to that in the mobile phase, at least as a first approximation. In this case, we need only the value of the equilibrium constant β in order to calculate p^* and b^* values from $p^* = \beta(1 + p) - 1$ and $b^* = \beta(1 + b_e) - 1$ and therefore E_T^{Ns} values as a function of φ . This approach is followed and described in Section 4.

We should point out that the solute/stationary phase interactions can be treated explicitly by means of statistical

mechanics [24–28] or semi-statistical mechanical approaches [15,16]. However, in these cases the final expression of $\ln k$ is very complicated and contains a great number of adjustable parameters. Similar complicated expressions of $\ln k$ are expected if we consider specific effects, like the consequences of stationary phase anisotropy on the retention of shape-constrained solutes. The necessity to review and develop in this paper as simple equations as possible led us to disregard specific effects and adopt the above phenomenological treatment of the various contributions coming from stationary phase.

2.3. Three-parameter equations

If the above refinements of the two-parameter equations are taken into account, Eq. (1) is changed to the following expressions of $\ln k$ provided that retention is governed by the adsorption mechanism:

$$\ln k = m - \ln(1 - \alpha\varphi) - \ln[1 + (\beta - 1)\varphi] + nE_T^N + qE_T^{Ns} \quad (18)$$

$$\ln k = m' - \ln(1 - \alpha\varphi) - \ln[1 + (\beta - 1)\varphi] + \frac{n'}{1 + b_e\varphi} + \frac{q'}{1 + b^*\varphi} \quad (19)$$

or

$$\ln k = m' - \ln(1 - \alpha\varphi) - \ln[1 + (\beta - 1)\varphi] + \frac{n'}{1 + b_e\varphi} + q'\varphi \quad (20)$$

If the retention is due to partition, the above equations are reduced to Eq. (12). It is seen that the adsorption model yields expressions with at least four adjustable parameter (m' , n' , β and q'). This number is high enough, especially if we take into account that the approximations adopted for the derivation of these equations are in all cases rough approximations. This means that the above equations do not represent accurately the adsorption model for retention. Their fitting performance, which is expected to be high enough, is due to the great number of the adjustable parameters that counterbalance the various adsorption effects contributing to $\ln k$.

Therefore, for practical purposes there is no need to use the complete Eqs. (18)–(20). The term $\ln[1 + (\beta - 1)\varphi]$ may be absorbed by the term $n'/(1 + b_e\varphi)$, as shown earlier by Eq. (3) and as can be easily verified by simple numerical examples. In addition, the term $\ln(1 - \alpha\varphi)$ may be approximately deleted or in case of Eq. (19) it may be absorbed by the last term, $q'/(1 + b^*\varphi)$, for the same reason that the entropy term is absorbed by the rational function $n'/(1 + b_e\varphi)$. Thus, we obtain two three-parameter equations:

$$\ln k = m' + \frac{n'}{1 + b_t\varphi} + \frac{q'}{1 + b^*\varphi} \quad (21)$$

and

$$\ln k = m' + \frac{n'}{1 + b_t\varphi} + q'\varphi \quad (22)$$

It is evident that these equations are three-parameter equations only if b_t and b^* are known at each mobile phase. The process followed for the determination of these parameters is described in Section 4. The use of the symbol b_t instead of b_e arises from the observation that small alterations in the values of b_e improve the fitting performance of these equations.

Apart from Eqs. (21) and (22), we used as three-parameter equations Eqs. (2) and (5), by treating b as an adjustable parameter, and the conventional quadratic equation [16,29]:

$$\ln k = a + b\varphi + c\varphi^2 \quad (23)$$

At this point, it is worth noting the following. Some of the above equations are based on the partition model, Eqs. (1), (12) and (23), the rest on the adsorption model for retention. However, as stressed many times, these equations do not, in fact, express the properties of the partition or adsorption mechanism due to the rough approximations involved in their derivation. For this reason, they should be treated as simple mathematical equations applied to any system irrespective of the retention mechanism that governs the properties of this system. It is also evident that the results of the present paper and, in particular, the applicability or not of the equations we reviewed earlier, are not related to the clarification of the retention mechanism in reversed-phase chromatographic columns. The investigation on the origin of the retention mechanism follows different approaches, which are based on the elimination or the drastic reduction in the approximations or model assumptions adopted in the various tests (see, for example, [16,30,31]).

2.4. Fitting criteria

The performance of a certain equation to describe the experimental data may be estimated by the value of the standard deviation σ :

$$\sigma^2 = \sum_{i=1}^N \frac{(\ln k_{\text{exp},i} - \ln k_{\text{calc},i})^2}{N - p} \quad (24)$$

where $k_{\text{exp},i}$ is the i th experimental value of k , $k_{\text{calc},i}$ the corresponding value of k calculated from one of the above equations, N the total number of data points, and p the number of the adjustable parameters. Values of σ lower than 0.2 usually correspond to good fittings.

However, a good fitting of the $\ln k$ versus φ data does not entail an equally good fitting of the t_R versus φ data used in optimising or/and predictive techniques. In this case, the best statistical criterion of the fitting of t_R versus φ data by a certain equation is the standard deviations σ_t resulting from Eq. (24), if we replace $\ln k$ by t_R . Alternatively, in cases that t_R values are not available due to the lack of t_0 values,

where t_0 is the column dead time, a measure of the predictive capability of an equation may be the value of the following parameter:

$$s = \frac{1}{N} \sum_{i=1}^N \frac{|t_{\text{Rexp},i} - t_{\text{Rcalc},i}|}{t_{\text{Rexp},i}} = \frac{1}{N} \sum_{i=1}^N \frac{|k_{\text{exp},i} - k_{\text{calc},i}|}{1 + k_{\text{exp},i}} \quad (25)$$

If s is, for example, close to 0.1, then on an average the absolute error in the retention time is 10%, which is high enough. For example, for a retention time of 60 min the error may be ± 6 min or even higher. Thus, a value of s equal to or lower than 0.05 should be used for an acceptable fitting.

3. Experimental

In order to test the above equations, a wide number of solutes (293) in mobile phases modified with four different organic modifiers, methanol (MeOH), acetonitrile (ACN), isopropanol (iPrOH) and tetrahydrofuran (THF), were used. Table 1 shows all datasets examined in the present investigation. The majority of them were taken from literature [7–13,15,29–32].

In order to examine the effect of the nature of the organic modifier on the retention properties of a certain group of solutes, we studied eight catechol-related solutes, dopamine

(DA), serotonin (5HT), 3,4-dihydroxyphenylacetic acid (DOPAC), 5-hydroxyindole-3-acetic acid (HIAA), vanillyl-mandelic acid (VMA), 5-hydroxytryptophol (HTOH), 3,4-dihydroxyphenyl glycol (HPG) and homovanillic acid (HVA), using different hydroorganic mobile phases consisting of an aqueous phosphate buffer (pH 2.5) and all the above modifiers. The total ionic strength of the mobile phases was held constant at $I = 0.02$ M. All chemicals were used as received from commercial sources. Catechol-related compounds were available from Sigma or Aldrich. The liquid chromatography system consisted of a Shimadzu LC-10AD pump, a model 7125 syringe loading sample injector fitted with a 20 μ l loop (Rheodyne, Cotati, CA), a 250 mm \times 4 mm MZ-Analysentechnik column (5 μ m Inertsil ODS-3) thermostatted by a CTO-10AS Shimadzu column oven at 25 $^{\circ}$ C, and a Gilson EC detector (model 141) equipped with a glassy carbon electrode. The detection of the analytes was performed at 0.8 V versus the Ag/AgCl reference electrode. The eluent flow rate was varied from 0.5 to 1.5 ml/min depending on the mobile phase composition. The hold-up time, t_0 , was measured for every mobile phase composition by injection of water. It was found that t_0 changes in the experimental ranges of mobile phase compositions studied, except for the case of water-methanol solutions. The obtained experimental data in terms of $\ln k$ versus φ are shown in Table 2.

Table 1
Datasets examined in the present investigation

Set	Column	Modifier	Solute	Reference
1	Spherisorb C ₁₈	MeOH	15 Phenols	[7]
2	LiChrospher 100 RP18	MeOH	14 Benzene and 18 phenol derivatives	[8]
3	Hydrodecyl column	MeOH	17 Benzene derivatives	[32]
4	Heptadecafluorodecyl column	MeOH	17 Benzene derivatives	[32]
5	Kromasil C ₁₈	MeOH	Clari- and roxy-thromycin	[15]
6	Nucleosil 10-RP18	MeOH	Benzene, toluene, ethylbenzene	[29]
7	Inertsil ODS-3	MeOH	6 Non-polar benzene derivatives	[31]
8	LiChrosorb, C ₂	MeOH	As in set 7	[31]
9	Inertsil ODS-3	MeOH	8 Catechol- and indole-related compounds	Present work
10	LiChrospher 100 RP18	ACN	14 Benzene and 18 phenol derivatives	[8]
11	Ultrasphere ODS	ACN	9 Steroids	[9]
12	LiChrospher 100 RP18	ACN	10 Peptides	[10]
13	LiChrospher 100 RP18	ACN	6 Quinolones	[11]
14	LiChrospher 100 RP18	ACN	15 Diuretic compounds	[12]
15	LiChrospher 100 RP18	ACN	10 Peptide hormones	[13]
16	Inertsil ODS-3	ACN	6 Non-polar benzene derivatives	[31]
17	LiChrosorb C ₂	ACN	As in set 16	[31]
18	Kromasil C ₁₈	ACN	Clari- and roxy-thromycin	[15]
19	Inertsil ODS-3	ACN	8 Catechol- and indole-related compounds	Present work
20	Hypersil ODS	iPrOH	6 Non-polar benzene derivatives	[30]
21	Inertsil ODS-3	iPrOH	6 Non-polar benzene derivatives	[31]
22	LiChrosorb C ₂	iPrOH	As in set 20	[31]
23	Inertsil ODS-3	iPrOH	8 Catechol- and indole-related compounds	Present work
24	Nucleosil 10-RP18	THF	32 Aromatic compounds	[29]
25	Hypersil ODS	THF	6 Non-polar benzene derivatives	[30]
26	Inertsil ODS-3	THF	6 Non-polar benzene derivatives	[31]
27	LiChrosorb C ₂	THF	As in set 21	[31]
28	Inertsil ODS-3	THF	8 Catechol- and indole-related compounds	Present work

Table 2

Experimental retention values ($\ln k$) of catechol-related compounds in aqueous mobile phases modified with methanol, acetonitrile, isopropanol and tetrahydrofuran

φ	DA	HPG	5HT	VMA	DOPAC	HTOH	HIAA	HVA	t_0 (min)
Methanol–water									
0	1.128	1.330	2.676	2.296	3.813	4.432	4.819	5.358	1.844
0.02	0.749	1.018	2.238	1.960	3.405	3.968	4.334	4.861	1.844
0.05	0.343	0.667	1.747	1.593	2.984	3.456	3.806	4.324	1.844
0.10	−0.257	0.193	1.045	1.074	2.376	2.747	3.065	3.569	1.844
0.14	−0.681	−0.088	0.544	0.743	1.963	2.275	2.567	3.055	1.844
0.20	−1.063	−0.430	0.030	0.336	1.467	1.692	1.951	2.462	1.844
0.30	−1.631	−0.918	−0.817	−0.255	0.713	0.837	1.036	1.552	1.844
0.40	−1.925	−1.291	−1.603	−0.740	0.046	0.114	0.251	0.744	1.844
0.50	−2.117	−1.598	−1.836	−1.183	−0.594	−0.573	−0.486	−0.020	1.844
Acetonitrile–water									
0	1.136	1.338	2.683	2.303	3.820	4.438	4.826	5.364	1.832
0.02	0.476	0.814	1.914	1.748	3.095	3.630	3.985	4.409	1.813
0.06	−0.290	0.095	0.831	1.009	2.141	2.576	2.888	3.244	1.771
0.10	−0.757	−0.344	0.040	0.497	1.444	1.849	2.114	2.419	1.751
0.14	−1.445	−0.348	−0.327	0.152	0.919	1.325	1.542	1.797	1.692
0.20	−2.121	−0.699	−0.450	−0.222	0.326	0.724	0.883	1.074	1.634
0.30	−2.874	−1.106	−1.628	−0.314	−0.182	0.149	0.252	0.389	1.523
Isopropanol–water									
0	1.136	1.338	2.683	2.303	3.820	4.438	4.826	5.364	1.832
0.02	−0.313	0.385	1.044	1.116	2.413	2.860	3.151	3.523	1.762
0.04	−1.075	−0.156	0.192	0.525	1.718	2.057	2.305	2.661	1.761
0.06	−1.521	−0.441	−0.305	0.186	1.304	1.568	1.792	2.145	1.740
0.10	−2.171	−0.782	−0.787	−0.282	0.712	0.884	1.064	1.417	1.736
0.14	−3.097	−1.026	−1.254	−0.605	0.285	0.397	0.541	0.886	1.749
0.20	−3.451	−1.282	−2.007	−0.862	−0.132	−0.066	0.035	0.338	1.734
0.30	−5.827	−1.736	−3.429	−1.192	−0.680	−0.650	−0.566	−0.333	1.697
Tetrahydrofuran–water									
0	1.136	1.338	2.683	2.303	3.820	4.438	4.826	5.364	1.832
0.01	0.007	0.763	1.378	1.675	2.935	3.268	3.664	3.856	1.760
0.02	−0.419	0.539	0.892	1.443	2.634	2.849	3.249	3.381	1.749
0.04	−0.890	0.270	0.339	1.150	2.270	2.373	2.768	2.849	1.710
0.06	−1.226	0.096	−0.022	0.947	2.020	2.065	2.461	2.496	1.707
0.10	−1.584	−0.074	−0.261	0.700	1.720	1.717	2.104	2.100	1.657
0.14	−1.809	−0.141	−0.625	0.560	1.498	1.456	1.764	1.840	1.623
0.20	−2.138	−0.247	−1.145	0.298	1.223	1.105	1.492	1.387	1.577
0.30	−2.323	−0.299	−1.787	0.055	0.822	0.606	0.956	0.909	1.439

4. Data analysis

The analysis of data was carried out at Microsoft Excel spreadsheets using Solver for all fittings. The minimised quantity was the sum of squares of residuals, $SSR = \sigma^2(N - p)$, and independently the value of s from Eq. (25), because if s is used as a measure of the fitting performance of an equation, then the minimisation of s is expected to give better results. For fitting all data of a set of solutes to a certain equation, a suitable macro has been written. This macro: (a) calls Solver to find the fitting parameters for a certain solute by minimising SSR and the same procedure is followed by minimising s ; (b) stores the fitting parameters and the values of SSR, σ and s at a predefined region at the spreadsheet; and (c) changes the solute data and repeats steps (a) and (b) until all solutes have been treated.

Values of the E_T^N factor, necessary for testing Eq. (1), were taken from literature [6,8]. In particular, we used the E_T^N values suggested by Bosch et al. [8] for methanol–water solutions and the corrected E_T^{N*} values for the entire range of φ values for acetonitrile–water solutions used by the same authors. The above data were fitted to Eq. (7) and the obtained values of p and b_e are listed in Table 3. Note that

Table 3

Values of p and b_e of Eq. (7) obtained from fitting this equation to experimental E_T^N data

Mobile phase	p	b_e	σ
Methanol–water	0.40	0.83	0.0009
Acetonitrile–water	0.51	1.1	0.003
Isopropanol–water	0.51	1.6	0.009
Tetrahydrofuran–water	0.54	1.9	0.009

Table 4
Common values of parameter b calculated from Eq. (5)

Set	M_T	Common value of b per set of substances at each modifier			
		MeOH ^a (0.2) ^b	ACN ^a (2.0) ^b	iPrOH ^a (3.0) ^b	THF ^a (1.0) ^b
9, 19, 23, 28	180	2.0	7.0	15	25
7, 16, 21, 26	110	0	1.0	2.0	2.0
8, 17, 22, 27	110	0.2	1.0	4.5	1.5
5, 18	800	0.2	15		
2, 10	140	0.7	2.5		
1	150	0.2			
3	110	−0.4			
4	110	−0.2			
6	90	0			
11	300		8		
12	250		5		
13	330		11		
14	330		6		
15	1600		32		
20, 25	100			2.0	1.5
24	130				0.6

^a Modifier.

^b Common value of b for all substances at each modifier.

almost the same results have been suggested by Roses and Bosch [7] for the calculation of E_T^N values. E_T^N values for aqueous solutions of isopropanol and tetrahydrofuran were calculated from non-normalised $E_T(30)$ values taken from [6] using the normalisation equation suggested by Reichardt and Harbusch-Görnert [33]. The obtained E_T^N values were fitted to Eq. (7) for $\varphi \leq 0.8$ yielding the p and b_e values of Table 3. Therefore, for testing Eq. (1) the E_T^N values were obtained from Eq. (7) using the p and b_e values of Table 3. Note that in this case, Eq. (1) is identical to Eq. (8), which for applications needs only the value of b_e .

Eqs. (2) and (5) were treated as three- and two-parameter equations. In the latter case, we examined two sub-cases: parameter b takes a common value for all substances at each modifier or a common value per group of structurally similar substances at each modifier. These common values of b were determined as follows. Eqs. (2) and (5) were initially treated as three-parameter equations and b was determined by means of Microsoft Excel Solver for each solute and modifier. Then, for each modifier the mean value of b was calculated excluding any extreme values of b . The mean value of b thus obtained was further used to refit Eqs. (2) and (5) to the same experimental dataset, in order to estimate the fitting performance of this equation when b takes a constant value for all substances at a certain modifier. The same procedure was followed for the determination of a mean b value per group of structurally similar substances at each modifier. For simplicity, we treated approximately each set of solutes of Table 1 as a group of structurally similar substances. The common values of b calculated from Eq. (5) are given in Table 4. Eq. (2) gave similar results except for the case of set 19, where we found a very high b value ($b = 105$).

Finally, in order to determine the best values of b_t and b^* of Eq. (21), we worked as follows. We first selected

as model experimental systems the datasets 8, 17, 22 and 27, because: (a) the solutes of these sets have small and non-polar molecules; and (b) the study using a C₂ column ensures the validity of the adsorption model for retention; the short length of the carbon chains does not leave much space in the stationary phase for the solute molecules. Therefore, for these systems we may assume that b^s is close to b_e , because the discreteness effect should be small and the surface solution is likely to have a structure similar to that of the mobile phase. Under these assumptions, we fitted the experimental data of sets 8, 17, 22 and 27 to Eq. (19) using $b^* = \beta(1 + b_e) - 1$ and tried to determine a mean value for β . This has been succeeded to methanol and acetonitrile solutions, where we found $\beta = 3$ and 15, respectively, yielding $b^* = 4.5$ for methanol–water and $b^* = 30$ for acetonitrile–water mobile phases. Next using these values of b^* in Eq. (21), we examined if small alterations in the values of b_t around b_e yield better results. The values of b_t thus obtained together with the b^* values are given in Table 5. For mobile phases modified with isopropanol and tetrahydrofuran, the above method did not work, because Solver could not converge when Eq. (19) was used to fit the experimental data. For this reason, we attempted to determine parameters b_t and b^* of Eq. (21) by direct application of this equation to all experimental data in isopropanol–water

Table 5
Values of b_t and b^* suggested for use in Eqs. (21) and (22)

Mobile phase	b_t	b^*
Methanol–water	1	4.5
Acetonitrile–water	2	30
Isopropanol–water	2	40
Tetrahydrofuran–water	2	60

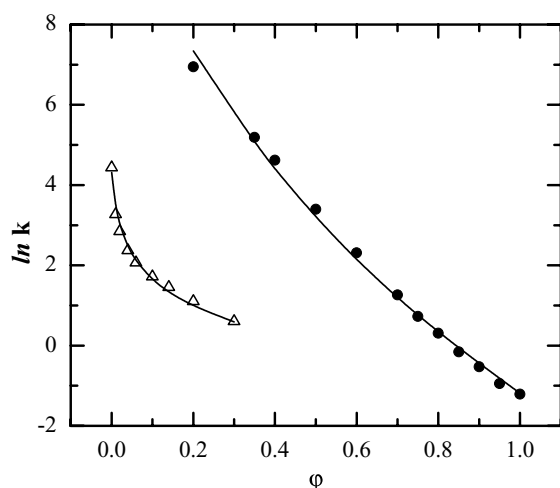


Fig. 2. Plots of $\ln k$ vs. φ for TE (●) retention in MeOH using C_2 column, and HTOH (Δ) retention in THF using C_{18} column. Points are experimental data and solid lines are constructed using the best fit of Eq. (1) ($\sigma = 0.1844$) and Eq. (2) ($\sigma = 0.144$), respectively.

and tetrahydrofuran–water mobile phases. The macro, we used for fittings, simplified significantly this attempt and the values of b_t and b^* obtained are listed in Table 5. Note that the same values of b_t were used in both Eqs. (21) and (22).

5. Results and discussion

Tables 6 and 7 show the number of systems exhibiting $\sigma < 0.2$ and $s < 0.05$. From these tables, several conclusions can be drawn. The first is related to the criteria $\sigma < 0.2$ and $s < 0.05$. It is seen that the performance of an equation depends on the above criteria. For example, for mobile phases modified with methanol, the performance of Eq. (1) is quite satisfactory if we adopt the criterion $\sigma < 0.2$, since it describes 79% of the systems, but it becomes very poor (79% becomes 39%) if we evaluate the fitting performance by the condition $s < 0.05$. In order to have a picture of the meaning of $\sigma < 0.2$, Fig. 2 depicts plots of $\ln k$ versus φ for: (a) *tert*-butylbenzene (TE) in mobile phases modified with methanol when a C_2 column is used; and (b) 5-hydroxytryptophol (HTOH) in THF mobile phases using a C_{18} column. In this figure, the points represent experimental data and the lines were calculated from the best fits of Eqs. (1) and (2), respectively. For these fits, the values of σ are 0.184 and 0.144, respectively. It is seen that the fitted equations describe quite satisfactorily the experimental data. Therefore, we may conclude that a fit characterised by $\sigma < 0.2$ is at least an acceptable fit. However, Tables 8 and 9 show that the predicted from Eqs. (1) and (2) retention times deviate significantly from the experimental values when $t_R > \sim 30$ min and consequently these equations fail to predict the retention of these substances at φ values that yield $t_R > \sim 30$ min.

In general, when we fit $\ln k$ versus φ data to an equation and then use the fitted equation to calculate t_R versus φ data, there is a problem with the accuracy of the predicted t_R values when t_R is high enough (at low φ values) for the following reason: The error $\delta(t_R)$ between predicted and calculated t_R values and the corresponding error in $\ln k$, $\delta(\ln k)$, are interrelated through the relationship:

$$d(\ln k) = \frac{dt_R}{t_0 k} \Rightarrow \delta(t_R) = (t_R - t_0)k \delta(\ln k) \quad (26)$$

Therefore, for the same error in $\ln k$ ($\delta(\ln k)$) imposed by the fitting procedure of the $\ln k$ versus φ data, the higher the value of t_R , the bigger the error in its predictive value, $\delta(t_R)$.

It is seen that the criterion $\sigma < 0.2$ as well as the $\ln k$ versus φ plots may be quite misleading in what concerns the performance of a fitting equation. Consequently, we may either reduce further the value of σ , for example $\sigma < 0.1$, or use the criterion $s < 0.05$. The latter has the advantage that it is directly related to the error in t_R , since it gives that the average absolute error in the retention time is less than 5%.

Note also that in general when we minimise SSR or σ to obtain the best fit of $\ln k$ versus φ data, the predicted t_R values are close to the experimental ones in the region of low t_R , whereas at high t_R values significant deviations between predicted and experimental t_R values may be detected. This behaviour is completely inverted if we minimise s to obtain the best fit of $\ln k$ versus φ data. Thus, it is a personal judgement not only the selection of the proper fitting equation but also the selection of the minimised quantity, SSR or s .

In what concerns the fitting capabilities of the equations we examined, it is seen that, in general, the performance of the two-parameter Eqs. (1), (8) and (12) is very poor, since, according to the s criterion, these equations describe satisfactorily only 44 and 40% of the systems studied, respectively. The performance of the two-parameter Eqs. (2) and (5) is somehow better: these equations give a good representation of 54 and 57% of the systems and this percentage is increased to 71 and 78%, respectively, if we use common b values for each set of solutes of Table 1. The latter percentages would be expected to be increased if we used a common b value per structurally similar group of solutes at each modifier. It is also evident that the above percentages are increased if we decrease the range of the φ values, whereas if we could increase the range of φ , the performance of all equations would deteriorate. From this point of view, our results agree with the observation made by Cheong and Carr [6] that Eq. (1) can give good results only over a narrow range of solvent composition.

From Tables 6 and 7 we observe that Eq. (12) is slightly worse than Eq. (1). At first sight this is an unexpected result, because Eq. (12) is a refinement of Eq. (1), as shown in the theoretical part. However, this inconsistency can be easily explained if we take into account that the performance of all simple equations studied in the present work is not associated with the model they represent. They are all based on such rough approximations that, in fact, they do not rep-

Table 6
Number of systems and their corresponding percentage exhibiting $\sigma < 0.2$ and $s < 0.05$ at each mobile phase

Equation	$\sigma < 0.2^a$	$s < 0.05^b$	$s < 0.05^c$	Percentages		
Methanol–water, number of solutes 106						
Two-parameter						
(1)≡(8)	84	41	33	79.2	38.7	31.1
(12)	76	37	27	71.7	34.9	25.5
(2) ^d	96	65	51	90.6	61.3	48.1
(2) ^e	104	79	77	98.1	74.5	72.6
(5) ^d	93	69	55	87.7	65.1	51.9
(5) ^e	104	81	77	98.1	76.4	72.6
Three-parameter						
(2) ^f	105	95	87	99.1	89.6	82.1
(5) ^f	105	95	87	99.1	89.6	82.1
(21)	104	94	88	98.1	88.7	83.0
(22)	105	87	83	99.1	82.1	78.3
(23)	105	94	90	99.1	88.7	84.9
Acetonitrile–water, number of solutes 104						
Two-parameter						
(1)≡(8)	63	48	44	60.6	46.2	42.3
(12)	71	58	51	68.3	55.8	49.0
(2) ^d	77	49	42	74.0	47.1	40.4
(2) ^e	89	67	63	85.6	64.4	60.6
(5) ^d	75	53	42	72.1	51.0	40.4
(5) ^e	97	84	78	93.3	80.8	75.0
Three-parameter						
(2) ^f	92	83	78	88.5	79.8	75.0
(5) ^f	99	97	87	95.2	93.3	83.7
(21)	101	100	89	97.1	96.2	85.6
(22)	100	89	86	96.2	85.6	82.7
(23)	94	81	82	90.4	77.9	78.8
Isopropanol–water, number of solutes 26						
Two-parameter						
(1)≡(8)	14	9	8	53.8	34.6	30.8
(12)	17	10	9	65.4	38.5	34.6
(2) ^d	17	10	9	65.4	38.5	34.6
(2) ^e	24	19	11	92.3	73.1	42.3
(5) ^d	18	11	9	69.2	42.3	34.6
(5) ^e	24	18	11	92.3	69.2	42.3
Three-parameter						
(2) ^f	24	25	17	92.3	96.2	65.4
(5) ^f	24	24	15	92.3	92.3	57.7
(21)	24	22	19	92.3	84.6	73.1
(22)	18	18	16	69.2	69.2	61.5
(23)	18	10	10	69.2	38.5	38.5
Tetrahydrofuran–water, number of solutes 57						
Two-parameter						
(1)≡(8)	28	31	24	49.1	54.4	42.1
(12)	14	13	3	24.6	22.8	5.3
(2) ^d	39	34	29	68.4	59.6	50.9
(2) ^e	48	44	42	84.2	77.2	73.7
(5) ^d	37	35	29	64.9	61.4	50.9
(5) ^e	48	46	42	84.2	80.7	73.7
Three-parameter						
(2) ^f	50	51	48	87.7	89.5	84.2
(5) ^f	47	51	48	82.5	89.5	84.2
(21)	53	55	51	93.0	96.5	89.5
(22)	43	48	45	75.4	84.2	78.9
(23)	46	46	44	80.7	80.7	77.2

^a Minimisation of SSR and then calculation of σ .

^b Minimisation of s .

^c Minimisation of SSR and then calculation of s .

^d b takes a common value from Table 3 for all substances in each modifier.

^e b takes a common value per group of substances at each modifier from Table 4.

^f b is treated as adjustable parameter.

Table 7
Number of systems and their corresponding percentage exhibiting $\sigma < 0.2$ and $s < 0.05$ for all systems (293) studied

Equation	$\sigma < 0.2^a$	$s < 0.05^b$	$s < 0.05^c$	Percentages		
Two-parameter						
(1)≡(8)	189	129	109	64.5	44.0	37.2
(12)	178	118	90	60.8	40.3	30.7
(2) ^d	229	158	131	78.2	53.9	44.7
(2) ^e	265	209	193	90.4	71.3	65.9
(5) ^d	223	168	135	76.1	57.3	46.1
(5) ^e	273	229	208	93.2	78.2	71.0
Three-parameter						
(2) ^f	271	254	230	92.5	86.7	78.5
(5) ^f	275	267	237	93.9	91.1	80.9
(21)	282	271	247	96.2	92.5	84.3
(22)	266	242	230	90.8	82.6	78.5
(23)	263	231	226	89.8	78.8	77.1

^a Minimisation of SSR and then calculation of σ .

^b Minimisation of s .

^c Minimisation of SSR and then calculation of s .

^d b takes a common value from Table 3 for all substances in each modifier.

^e b takes a common value per group of substances at each modifier from Table 4.

^f b is treated as adjustable parameter.

Table 8
Experimental and predicted from Eq. (1) retention times (t_R in min) that correspond to Fig. 2

	φ									
	0.40	0.50	0.60	0.70	0.75	0.80	0.85	0.90	0.95	1.00
t_{exp}	259.60	78.06	28.06	11.49	7.76	5.98	4.69	4.03	3.51	3.28
t_{pred}	207.89	64.19	23.92	10.88	7.93	6.10	4.93	4.16	3.65	3.30
δt^a (min)	51.71	13.87	4.14	0.61	0.17	0.12	0.23	0.13	0.14	0.02
% δt^b	19.92	17.76	14.76	5.35	2.20	1.98	4.95	3.33	3.89	0.61

^a Absolute difference between predicted and experimental retention times.

^b Percentage error in the predicted retention time.

resent the retention mechanism, whatever this mechanism is. Therefore, a refinement of an approximation involved in these equations or the inclusion of a certain effect does not necessarily entail the improvement of the fitting performance of the modified equation. In fact, this improvement is a blind process, which is verified or not from the application of the modified equation to as many systems as possible. In this respect, the better performance of Eq. (5) when we use a common b values for each structurally similar group of solutes at each modifier may show an effective way to account for discreteness effects. The fact that b as a rule increases

with the increase in the mean molecular weight of the solutes at the various sets (see Table 4 and Fig. 3) is a strong evidence that this is due to the discreteness effect. However, another significant factor responsible for the performance of Eq. (5) should be the fact that when b is determined for each set of solutes, Eq. (5) is no more a purely two-parameter equation; we first treat it as a three-parameter equation and next determine the mean value of b for each set.

The performance of Eq. (5) is increased considerably if we use it as a three-parameter equation. In general, all the three-parameter equations give good results, since their

Table 9
Experimental and predicted from Eq. (2) retention times (t_R in min) that correspond to Fig. 2

	φ									
	0	0.01	0.02	0.04	0.06	0.1	0.14	0.2	0.3	
t_{exp}	156.9	47.98	31.95	20.05	15.17	10.88	8.59	6.34	4.08	
t_{pred}	138.68	57.73	37.08	21.68	15.73	10.20	7.72	5.79	4.03	
δt^a (min)	18.22	9.75	5.13	1.63	0.56	0.68	0.87	0.55	0.04	
% δt^b	11.61	20.32	16.06	8.14	3.69	6.25	10.10	8.67	1.10	

^a Absolute difference between predicted and experimental retention times.

^b Percentage error in the predicted retention time.

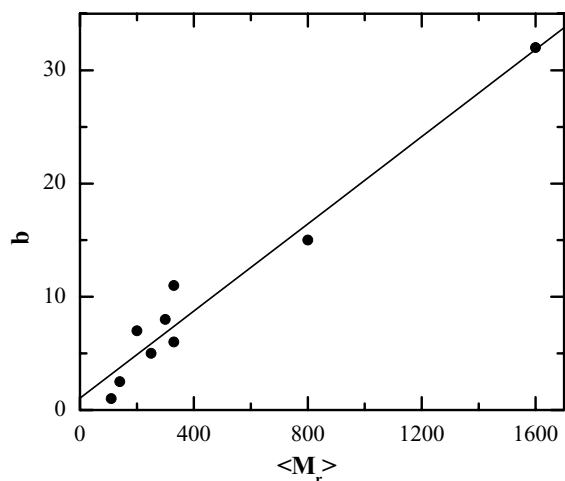


Fig. 3. Linear relationship between parameter b calculated from fitting Eq. (5) to each set of Table 1 when ACN is used as modifier and the mean molecular weight of the solutes of these sets.

applicability ranges from 80 to 90%, according to the s criterion. From the three-parameter equations, the best performance is exhibited by Eq. (21). This equation combines two advantages: the linearity of the adjustable parameters of Eq. (23), provided that b_t and b^* are known at a certain modifier, and the high fitting capabilities of the rational functions, like Eq. (5). It is also free from the convergence problems of the non-linear Eqs. (2) and (5). For example, such problem appeared when we fitted Eq. (2) to sets 11 and 12. In addition, the predicted $\ln k$ versus φ curves by Eq. (21) are quite smooth and free from physically meaningless portions, which are usually met in polynomial fittings. The superiority of Eq. (21) in relation to the conventional Eq. (23) is shown in Fig. 4.

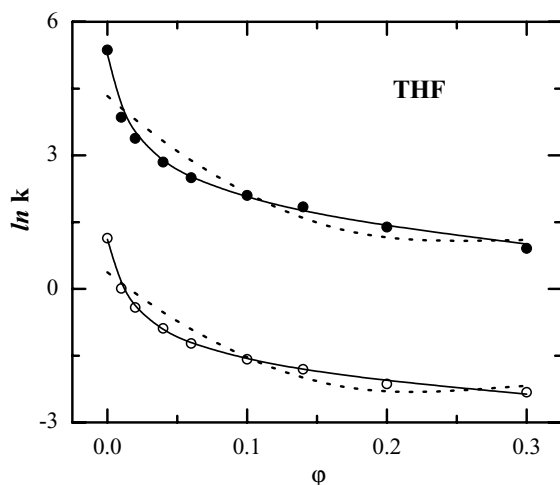


Fig. 4. Plots of $\ln k$ vs. φ for VMA (●) and DA (○) retention in THF. Points are experimental data, solid lines correspond to the best fit of Eq. (21), and dotted lines to the best fit of Eq. (23).

6. Conclusions

The performance of two-parameter equations to fit $\ln k$ upon φ experimental data is quite poor. Thus, the conventional Eq. (1) (identical to Eq. (8)) describes only 44% of the systems studied if the performance criterion $s < 0.05$ is adopted. The results are improved if we use Eq. (5) and b takes either a common value for each modifier or even better a common value per group of structurally similar solutes at each modifier. In the latter case, which is associated with the discreteness effect in the solute–solvent interactions, the applicability is extended to ca. 78% of the systems studied. The low performance of the two-parameter equations is attributed to the rough approximations involved in their derivation. However, if we like to keep the expression of these equations as simple as possible for practical purposes, then any modification of the two-parameter equations involves additional rough approximations. Despite this limitation, we found that a new three-parameter expression of $\ln k$, Eq. (21) works more satisfactorily, since it combines simplicity, linearity of the adjustable parameters m' , n' and q' and the highest applicability describing satisfactorily more than 90% of the systems. Note that in this equation the general b_t and b^* parameters depend exclusively upon the modifier. Their values for the four common reversed-phase organic modifiers, MeOH, ACN, iPrOH and THF, are listed in Table 5, whereas for other solvents they should be determined following the procedure suggested in this paper.

Finally, we should point out that the above results are indicative. Thus, the performance of all equations is considerably improved if we use narrow ranges of φ values, whereas it deteriorates if we consider that the accuracy of 5% in retention times is insufficient for component identification and make more stringent this criterion. For example, if we adopt the criterion $s < 0.025$, then the applicability of the three-parameter equations falls below 60% of the systems studied, whereas the corresponding percentage of Eq. (1) decreases from 44% ($s < 0.05$) to 19%. Therefore, we should be very careful when we choose a certain equation in an optimisation technique.

Appendix A

From Fig. 1, we obtain that:

$$(E_T^N)^{\text{ef}} = \frac{1 + p^{\text{ef}}\varphi}{1 + b_e^{\text{ef}}\varphi} \leq \frac{1 + p\varphi}{1 + b_e\varphi} = E_T^N \quad (\text{A.1})$$

where $(E_T^N)^{\text{ef}}$ is the effective E_T^N factor due to the discreteness effect. Note that the inequality is valid throughout the range of φ values apart from $\varphi = 0$ and $\varphi \approx 1$. At $\varphi \approx 1$, we have $(1 + p^{\text{ef}})/(1 + b_e^{\text{ef}}) = (1 + p)/(1 + b_e)$, which yields:

$$p^{\text{ef}} = \frac{(1 + p)(1 + b_e^{\text{ef}})}{1 + b_e} - 1 \quad (\text{A.2})$$

In addition, inequality (A.1) gives: $p^{\text{ef}} + b_e + b_e p^{\text{ef}} \varphi < p + b_e^{\text{ef}} + p b_e^{\text{ef}} \varphi$, which in the region of very small φ values ($\varphi \approx 0$) results in:

$$p^{\text{ef}} + b_e < p + b_e^{\text{ef}} \quad (\text{A.3})$$

Now substitution of p^{ef} from Eq. (A.2) into Eq. (A.3) and rearrangement yields:

$$\frac{(p - b_e)(b_e^{\text{ef}} - b_e)}{1 + b_e} < 0 \quad (\text{A.4})$$

which shows that b_e^{ef} is higher than b_e ($b_e^{\text{ef}} > b_e$), because p is always lower than b_e since the inequality $p - b_e < 0$ is the necessary and sufficient condition for the E_T^N versus φ curve to be concave up. For simplicity, in the main text, b_e^{ef} is denoted by b .

References

- [1] B.P. Johnson, M.G. Khaledi, J.G. Dorsey, *Anal. Chem.* 58 (1986) 85.
- [2] B.P. Johnson, M.G. Khaledi, J.G. Dorsey, *J. Chromatogr.* 384 (1987) 221.
- [3] J.J. Michels, J.G. Dorsey, *J. Chromatogr.* 435 (1988) 85.
- [4] M.M. Hsieh, J.G. Dorsey, *J. Chromatogr.* 631 (1993) 61.
- [5] P. T. Ying, J.G. Dorsey, *Talanta* 38 (1991) 237.
- [6] W.J. Cheong, P.W. Carr, *Anal. Chem.* 61 (1989) 1524.
- [7] M. Roses, E. Bosch, *Anal. Chim. Acta* 274 (1993) 147.
- [8] E. Bosch, P. Bou, M. Roses, *Anal. Chim. Acta* 299 (1994) 219.
- [9] D. Barron, J.A. Pascual, J. Segura, J. Barbosa, *Chromatographia* 41 (1995) 573.
- [10] J. Barbosa, V. Sanz-Nebot, I. Toro, *J. Chromatogr. A* 725 (1996) 249.
- [11] J. Barbosa, R. Berges, V. Sanz-Nebot, *J. Chromatogr. A* 719 (1996) 27.
- [12] J. Barbosa, R. Berges, V. Sanz-Nebot, *Chromatographia* 51 (2000) 417.
- [13] V. Sanz-Nebot, I. Toro, J. Barbosa, *J. Chromatogr. A* 870 (2000) 335.
- [14] J.R. Torres-Lapasio, M.C. Garcia-Alvarez-Coque, M. Roses, E. Bosch, *J. Chromatogr. A* 955 (2002) 19.
- [15] P. Nikitas, A. Pappa-Louisi, P. Agrafiotou, *J. Chromatogr. A* 946 (2002) 33.
- [16] P. Nikitas, A. Pappa-Louisi, P. Agrafiotou, *J. Chromatogr. A* 946 (2002) 9.
- [17] E. Bosch, M. Roses, *J. Chem. Soc., Faraday Trans.* 88 (1992) 3541.
- [18] J. Gmehling, U. Onken, W. Arlt, *Vapor-Liquid Equilibrium Data Collection, Aqueous-Organic Systems, Chemistry Data Series, vol. I*, DECHEMA, Frankfurt/Main, 1981.
- [19] Y. Marcus, Y. Migron, *J. Phys. Chem.* 95 (1991) 400.
- [20] P. Nikitas, *J. Electroanal. Chem.* 263 (1988) 147.
- [21] P. Nikitas, *Electrochim. Acta* 38 (1993) 2663.
- [22] P. Nikitas, A. Pappa-Louisi, *J. Phys. Chem.* 94 (1990) 361.
- [23] R. Guidelli, M.L. Foresti, *J. Electroanal. Chem.* 197 (1986) 103.
- [24] L. Peichang, L. Xiaoming, *J. Chromatogr.* 292 (1984) 169.
- [25] M. Jaroniec, D.E. Martire, *J. Chromatogr.* 351 (1986) 1.
- [26] R. Boehm, D. Martire, *J. Phys. Chem.* 84 (1980) 3620.
- [27] D. Martire, R. Boehm, *J. Phys. Chem.* 87 (1983) 1045.
- [28] M. Borowko, *J. Colloid Interface Sci.* 102 (1984) 519.
- [29] P.J. Schoenmakers, H.A.H. Billiet, L. de Galan, *J. Chromatogr.* 218 (1981) 261.
- [30] W.J. Cheong, P.W. Carr, *J. Chromatogr.* 499 (1990) 373.
- [31] P. Nikitas, A. Pappa-Louisi, P. Agrafiotou, *J. Chromatogr. A*, submitted for publication.
- [32] P.S. Sadek, P. Carr, R.M. Doherty, M.J. Kamlet, R.W. Taft, M.H. Abraham, *Anal. Chem.* 57 (1985) 2971.
- [33] C. Reichardt, E. Harbusch-Görnert, *Liebigs Ann. Chem.* (1983) 721.