ELSEVIER

Emerging approaches to estimate retention factors in high performance liquid chromatography

José María Bermúdez-Saldaña, Laura Escuder-Gilabert, Rosa María Villanueva-Camañas,
María José Medina-Hernández, Salvador Sagrado *

Departamento de Química Analítica, Universitat de València, C/Vicente Andrés Estellés s/n, E-46100 Burjassot, Valencia, Spain

Abstract

The retention factor is one of the most universally used parameters in chromatography. The errors associated with the conventional ways to determine the retention factor of compounds in liquid chromatography are studied and compared with those corresponding to new approaches. The later avoid the use of extra-column time and hold-up time values, which have proven to be tedious and ambiguous. Simulations and real data, used to examine the accuracy of four different approaches (two classic and two new), suggest that the new approaches could be considered more satisfactory than the classic ones.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Retention factor estimations; Hold-up time measurements; Extra-column time correction; 'Relative' retention factors

1. Introduction

Most of the aspects of the hold-up volume concept in column chromatography have been deeply investigated: utility and importance (i.e. determination of system-dependent chromatographic parameters as for instance the retention factor), practical problems of its use, different methods to estimate it, etc. [1–3] and references therein. Accurate determination of retention factors in liquid chromatography is important for theoretical studies (i.e. as a thermodynamic value for comparing types of columns), quantitative structure-retention relationships, QSRRs, modeling biopartitioning of compounds (i.e. quantitative retention-activity relationships, QRARs), system suitability issues (i.e. quality assurance), etc.

Unfortunately, the almost universal retention factor estimation has two main drawbacks: (i) It does not consider the extra-column time, which is the retention time contribution due to the injector, detector and connections, introducing a systematic error in k estimation [1]. (ii) It involves the use of hold-up volume (or time) values, whose measures has proven to be controversial, ambiguous and difficult, particularly in reversed-phase liquid chromatography [2], and it is associated to poor reliability in long-term studies [3].

In order to reduce the errors associated to this approach, several recommendations have been suggested. IUPAC [1] has recommended the use of k values corrected for extra-column time. This is particularly important in the cases that extra-column time is relatively large, compared with the experimental gross hold-up time values, as in the cases of using some connecting tube to preheating the mobile phase prior the column [4]. Otherwise, 'gross retention factor' are obtained [3]. Additionally, workers are urged to critically examine the gross hold-up time values they measure, to insure they are at least physically possible [2].

On the other hand, an alternative approach has been proposed. It avoids the dependence between k estimation and the experimental measurement of the hold-up time and extra-column time, which are substituted by the experimental measurement of the gross retention time (a more reliable parameter) of a reference compound [3]. The reference com-

* Corresponding author. Tel.: +34 96 354 4878; fax: +34 96 354 4953.
 E-mail address: sagrado@uv.es (S. Sagrado).

pound is chromatographed during the same working session that the test compounds. The reliability of the approach, compared with the classical ones, has been tested by means of precision studies, analysis of factors affecting retention factor estimation and uncertainty calculations. These assays were performed in micellar liquid chromatography under intermediate precision conditions (inter-day assays, different equipments, column lengths and mobile phase flow rates).

This article proposes a modification of the approach based on the use of one reference compound. Additionally, a new approach based on the use of two reference compounds is introduced. Simulation studies are performed in order to establish and compare the accuracy of classical and new approaches. Finally, the approaches are tested with an available real data set from the literature. The approaches are ranked according to their accuracy and practical utility for routine work. Some recommendations for using the new approaches are outlined.

2. Theory

Four different approaches (two classic and two new) have been studied. The nomenclature related to chromatographic parameters is consistent with IUPAC recommendations [1]. For clarity, we have used the term k to represent the 'true' (correct), retention factor or the 'accepted reference' retention factor of a test compound, while k with a superscript is used to represent a k-estimate by any approach. See Appendix for definition of all of the symbols used in the equations and tables considered in this paper.

2.1. Classic approaches to estimate k

In estimating the retention factor of a test compound two equations can be considered:

$$k^{\mathrm{g}} = \frac{t_{\mathrm{R}}^{\mathrm{g}} - t_{\mathrm{M}}^{\mathrm{g}}}{t_{\mathrm{M}}^{\mathrm{g}}} \tag{1}$$

$$k^{\mathrm{t}} = \frac{t_{\mathrm{R}}^{\mathrm{g}} - t_{\mathrm{M}}^{\mathrm{g}}}{t_{\mathrm{M}}^{\mathrm{g}} - t_{\mathrm{ext}}} \tag{2}$$

where k^{g} (gross retention factor estimate) and k^{t} (true retention factor estimate) are estimates of k (true retention factor) and depend on the gross retention time of the test compound ($t_{\mathrm{R}}^{\mathrm{g}}$), the gross hold-up time ($t_{\mathrm{M}}^{\mathrm{g}}$) and, in the case of Eq. (2), the extra-column time (t_{ext}). Eq. (1) (approach 1) represents the almost universal, but biased, approach to determine the retention factor of test compounds, while Eq. (2) (approach 2) represents an unbiased estimation of the true retention factor. In fact, this equation re-written for k, $k = (t_{\mathrm{R}}^{\mathrm{g}} - t_{\mathrm{M}}^{\mathrm{g}})/(t_{\mathrm{M}}^{\mathrm{g}} - t_{\mathrm{ext}})$, correspond to the correct definition of the retention factor [1]. Unfortunately, the errors associated mainly with the experimental measure of $t_{\mathrm{M}}^{\mathrm{g}}$ and

t_{ext} affect the accuracy of k^{g} and k^{t} estimations [3]. According to Eqs. (1) and (2), errors in $t_{\mathrm{R}}^{\mathrm{g}}$ should be less important in k estimation.

2.2. New approach using a reference compound

The use of a reference (a selected compound, r1) can be an alternative to Eqs. (1) and (2). For instance, if the reference r1 is chromatographed in the same conditions that the test compound and combining Eq. (1) corresponding to both reference and test compound [3], we can derive

$$k^{\mathrm{r1}} = \left(\frac{t_{\mathrm{R}}^{\mathrm{g}}}{t_{\mathrm{R1}}^{\mathrm{g}}} \right)(1 + k_1) - 1 \tag{3}$$

where k^{r1} is an estimate of k for the test compound, which besides its $t_{\mathrm{R}}^{\mathrm{g}}$, uses the gross retention time of the reference r1 ($t_{\mathrm{R1}}^{\mathrm{g}}$) and its retention factor (k_1).

Unfortunately, the true retention factor for r1 is not known and must be estimated experimentally, for instance in a given experiment carefully designed for this reference. k_1 could be estimated using approach 1 and used as a constant for further experiments, as Escuder-Gilabert et al. suggested [3]. In this case, k^{r1} would be a 'gross retention factor' estimate. On the other hand, k_1 can be also estimated by means of Eq. (2) (new proposal of this paper; approach 3). In this case, the approach 3 estimations should be closer to those from approach 2, then improving their accuracy. However, such k_1-estimates are obtained in a single laboratory and then they are subject to laboratory errors. Ideally, k_1 should be an 'accepted reference' value (for which the nomenclature k_1 in Eq. (3) would be consistent). In the absence of an 'accepted reference' value, we suggest that k_1 is estimated 'in-house' by means of approach 2 (but for simplicity we will use the term k_1 in Eq. (3) for that estimation). Therefore, the impact of errors in k_1 over the accuracy of k^{r1} has to be considered.

2.3. New approach based on two reference compounds

It is also possible to use two references (two selected compounds r1 and r2). In this case, we can combine Eq. (2) (not Eq. (1) as in approach 3) corresponding to the test compound and the two references (new proposal of this paper; approach 4) to derive an unbiased estimation of the true retention factor

$$k^{\mathrm{r2}} = \frac{k_2(t_{\mathrm{R}}^{\mathrm{g}} - t_{\mathrm{R1}}^{\mathrm{g}}) + k_1(t_{\mathrm{R2}}^{\mathrm{g}} - t_{\mathrm{R}}^{\mathrm{g}})}{t_{\mathrm{R2}}^{\mathrm{g}} - t_{\mathrm{R1}}^{\mathrm{g}}} \tag{4}$$

where k^{r2} is an estimate of k for the test compound, which besides its $t_{\mathrm{R}}^{\mathrm{g}}$, uses the gross retention times of the references r1 and r2 ($t_{\mathrm{R1}}^{\mathrm{g}}$ and $t_{\mathrm{R2}}^{\mathrm{g}}$) and their retention factors (k_1 and k_2). As before, until 'accepted reference' values are available, we suggest that k_1 and k_2 are estimated 'in-house' using Eq. (2) in a given statistically consistent experiment and used

as constants for further experiments. As before, the impact of errors in k_1 and k_2 over the accuracy of k^{r2} has to be considered.

Using approach 4 (and approach 3) the chromatographer does not need the experimental measure of t_M^g and t_{ext} in routine work, once k_1 and k_2 have been estimate for that chromatographic conditions. Moreover, the use of approach 4 has additional benefits. Once t_{R1}^g and t_{R2}^g have been measured in a new experimental chromatographic condition in which k is not expected to vary, it is possible to estimate t_M^g and t_{ext} for that condition by combining the Eq. (2) re-written for k_1 and k_2

$$t_M^g = \frac{t_{R1}^g k_2 - t_{R2}^g k_1}{k_2 - k_1} \tag{5}$$

$$t_{ext} = \frac{t_{R1}^g (k_2 + 1) - t_{R2}^g (k_1 + 1)}{k_2 - k_1} \tag{6}$$

In some instances, these estimations could serve as a way to check the reliability of the k estimation using k_1 and k_2 values. For instance, a negative t_{ext} or t_M^g values will suggest an error in the application of Eq. (4). All these considerations make approach 4 very attractive.

3. Experimental

3.1. Simulation study

In order to compare the classic and new approaches, we simulate an experimental situation in which two references, r1 and r2, and a set of 69 test compounds are injected in a given chromatographic condition (i.e. C18 column, 5-μm particles, 150 mm × 4.6 mm column dimensions and flow-rate = 1 mL min^{-1}). Some chromatographic parameters were fixed: t_{ext}, t_M, k_1 and k_2 for the references and t_R for the 69 test compounds. From t_R, k values were computed as $k = (t_R - t_M)/t_M$ (equivalent to Eq. (2) re-written for k). These data are shown in Table 1 (simulation case 1) and were assumed to be 'true' values. From them, other parameters were calculated (assuming to be experimental values): $t_M^g (= t_M + t_{ext})$, $t_R^g (= t_R + t_{ext})$, $t_{R1} (= t_M(1 + k_1))$, $t_{R2} (= t_M(1 + k_2))$, $t_{R1}^g (= t_{R1} + t_{ext})$, $t_{R2}^g (= t_{R2} + t_{ext})$ as well as k_1^g and k_2^g applying Eq. (1) (see Table 1).

Using these simulated experimental values, k-estimates for the set of 69 test compounds were performed applying Eqs. (1)–(4). In order to compare the four approaches in terms of accuracy, we provoke some errors on the parameters that appear in these equations. Since the 'true' k values of the test compound has been fixed, the error in percentage, $\%E = 100$ (k estimated $- k)/k$, for any approach was computed as a measure of its accuracy.

For simplicity, errors associated to the gross retention times were assumed negligible respect to those of the other parameters in Eqs. (1)–(4) (a situation expectable in practi-

Table 1
Conditions used in the accuracy study

Parameters	Simulation[a]		Real data[b]
	Case 1[c]	Case 2[d]	
t_{ext} (min)	**0.055**	**0.03**	**0.0733**
t_M (min)	**0.8**	**0.2**	0.8417
k_1	**10**[e]	**10**	12.3027[f]
k_2	**1**	**20**	0.5636
t_M^g (min)	0.85	0.23	**0.915**
t_{R1} (min)	8.8	2.2	11.1969
t_{R2} (min)	1.6	4.2	1.31611
t_{R1}^g (min)	8.85	2.23	11.2702
t_{R2}^g (min)	1.65	4.23	1.38941
k_1^g	9.41	8.696	11.317
k_2^g	0.941	17.391	0.518

The initial fixed values used to calculate the other data are shown in bold case.

[a] True retention times (t_R) of test compounds were set from **1.5** to **35.5** min each 0.5 min.

[b] The original data vector for the selected set of compounds, from which 'accepted reference' k values [4] were calculated, was log k = [**−0.249**; **−0.07**; **0.055**; **0.33**; **0.521**; **0.675**; **0.888**; **1.09**; **1.322**; **1.552**]. In Fig. 6a, the eighth and first compounds were selected as r1 and r2 references, respectively.

[c] Conditions used in Figs. 1–3.

[d] Conditions used in Fig. 5.

[e] In Fig. 4 k_1 was set to **2**. New values were $t_{R1} = 2.4$; $t_{R1}^g = 2.45$ and $k_1^g = 1.88235$.

[f] In Fig. 6b and c, the second ($k_1 = 0.851138$) and first compounds were selected as r1 and r2 references, respectively. New values were $t_{R1} = 1.5581$; $t_{R1}^g = 1.38941$ and $k_1^g = 0.782954$.

cal work), and they were not considered in this study. We focused the simulation study in the more critic parameters, characteristic of each equation: (i) t_M^g, which affects Eqs. (1) and (2); (ii) t_{ext}, which affects Eq. (2) (the cases of correlated and inversely correlated t_{ext}- and t_M^g-errors were studied); (iii) k_1, which affects Eqs. (3) and (4) and (iv) k_2, which affects Eq. (4) (the cases of correlated and inversely correlated k_1- and k_2-errors were studied). The provoked errors on these parameters should allow distinguishing between the quality of the approaches in terms of $\%E$.

Nine simulations (n) were performed provoking errors in the t_M^g, t_{ext}, k_1 and k_2 parameters in the range ±10% as follows: $n = 1, -10\%$; $n = 2, -7.5\%$, $n = 3, -5\%$; $n = 4, -2.5\%$, $n = 5, 0\%$ (no error); $n = 6, +2.5\%$; $n = 7, +5\%$; $n = 8, +7.5\%$ and $n = 9, +10\%$. In addition, to simulate the case of inversely correlated errors between t_M^g and t_{ext}, in Eq. (2), and k_1 and k_2, in Eq. (4), the errors in t_{ext} and k_2 were also simulated in the opposite way (from +10% to −10%). Finally, estimations of t_M^g and t_{ext} associated to approach 4 (Eqs. (5) and (6)) were performed.

In order to check the effect of a change in the chromatographic conditions, a second case was simulated (new column-length = 50 mm and flow-rate = 1.5 mL min^{-1}; assuming to correspond to a new chromatographic situation respect to the simulation case 1). The new data are shown in Table 1 (simulation case 2). All calculations were performed using routines developed in MATLAB 5.3 (Mat-

lab Ver. 5.3.0.10183 (R11), ©The Mathwoks Inc., Natick, MA).

3.2. Real data

Wilson et al. [4] reported data for a large set of compounds chromatographed in carefully controlled chromatographic conditions and using procedures to 'minimize experimental error'. For instance, the calculated log k data have been corrected for the extra-column time. At this point, the k values reported by these authors will be considered in this work as 'accepted reference values'. The reported chromatographic data (column 1 in Table 3 of that paper [4]; C_{18} column 5-μm particles, 150 mm × 4.6 mm column dimensions; GL Inertsil ODS-3; acetonitrile-water 50% mobile phase, temperature 35 °C) of ten selected solutes (compounds 16, 18, 24, 22, 13, 1, 2, 3, 5, and 6) covering a wide range of log k values, were used to calculate the corresponding 'accepted reference' k values. Two compounds of that set were used as references (r1 and r2).

The authors also reported the hold-up time in those chromatographic conditions (here used as t_M^g value), and declared an extra-column volume up to 110 μL (from which t_{ext} was calculated taking into account the flow-rate used by the authors: 1.5 mL min^{-1}). These values are shown in Table 1 (Real data). Unfortunately, the authors do not report the experimental t_R or t_R^g values, so they were computed from the available k data. From the k vector, we calculated the t_R ($=t_M(1+k)$) values (from Eq. (2) re-written for k), and then, the corresponding t_R^g ($=t_R+t_{ext}$) values, which were used as experimental data to perform the k-estimations by means of Eqs. (1), (3) and (4).

4. Results and discussion

4.1. Accuracy of conventional approaches

Fig. 1 shows the percentage of error in k estimation, %E values, corresponding to the four approaches as function of the gross retention time of the test compounds, t_R^g (Table 1 simulation case 1 conditions). Fig. 1a shows the %E values of approach 1 (based on k^g by applying Eq. (1)) when the simulated errors in t_M^g are within the ±10% range. Even with no error in t_M^g (simulation $n=5$), there is an error in k^g estimation ($k^g < k$), which is t_R^g-independent. The magnitude of this error can be calculated as %$E = -100 t_{ext}/t_M^g = -5.88$%. This equation can be derived from the definition of %E ($=100(k^g - k)/k$) and the Eqs. (1) and (2) (the last re-written for k instead of k^t).

The %E values obtained in the case of errors in t_M^g (simulations $n=1$–9, except $n=5$) depend on t_R^g. %E values go from -24.85% to 17.27% for $t_R^g = 1.55$ min and from -14.65% to 4.83% for $t_R^g = 35.55$ min. As can be observed, there is a trend to obtain negative %E values using approach 1. This provokes the singular situation that when a relatively low

negative error in t_M^g exists (i.e. simulations $n=3$ or $n=4$) the minimum error is achieved (%E values within the ±5% range), since the error in t_M^g partially compensates the omission of t_{ext}.

Fig. 1b shows the %E values of approach 2 (based on k^t by applying Eq. (2)) when the simulated errors in t_{ext} and t_M^g are equal (correlated) within the ±10% range. As can be expected, accurate estimations (%$E=0$) was found in the case of no errors in t_{ext} and t_M^g (simulation $n=5$). The %E values obtained in the case of errors in the parameters depend on t_R^g. %E values go from -20.13% to 24.60% for $t_R^g = 1.55$ min and from -9.31% to 11.38% for $t_R^g = 35.55$ min.

4.2. Accuracy study of new approaches

Fig. 1c shows the %E values of approach 3 (based on k^{r1} by applying Eq. (3)) when the simulated errors in the reference retention factor, k_1, are within the ±10% range. With no error in k_1 (simulation $n=5$) there is an error in k estimation, positive for low-retained test compounds (%$E=5.89$% for $t_R^g = 1.55$ min), and slightly negative for high-retained test compounds (%$E=-0.43$% for $t_R^g = 35.55$ min). The %E values obtained in the case of errors in k_1 depend on t_R^g. %E values go from -14.12% to 25.91% for $t_R^g = 1.55$ min and from -9.70% to 8.83% for $t_R^g = 35.55$ min.

Some particular situations can be found for approach 3. For test compounds retained as the reference r1 ($t_R^g \approx t_{R1}^g$), the %E values are of the same sign and similar magnitude that the provoked k_1-error. On the other hand, if k_1 is estimated by means of Eq. (1) (as k_1^g) and used in Eq. (3), the estimated retention factor for the test compounds is just k^g, being this approach equivalent to Eq. (1) from the accuracy point of view. Therefore, is preferable the use of Eq. (2) to estimate k_1, as we recommend and performed in this work.

Fig. 1d shows the %E values of approach 4 (based on k^{r2} by applying Eq. (4)) when the simulated errors in the two references retention factors, k_1 and k_2, are equal (correlated) within the ±10% range. Accurate estimations (%$E=0$) were found in the case of no errors in k_1 and k_2 (simulation $n=5$). This confirms the unbiased nature of Eq. (4) (as occurs with Eq. (2)). The %E values obtained in the case of errors in the parameters were t_R^g-independent (in contrast to approach 2, Fig. 1b, and the other approaches). These %E values are equal (sign and magnitude) than the provoked k_1- and k_2-errors.

In addition, results of Eq. (4) can be combined with results of Eqs. (5) and (6) in order to test their reliability. Fig. 2 shows the t_{ext} and t_M^g estimates based on approach 4. As can be observed, the estimation of t_M^g is accurate in all simulations. In contrast, t_{ext} estimation is accurate only for the case of no error in k_1 and k_2 (simulation $n=5$). For errors in k_1 and k_2 under -7.5% (simulations $n=1$ and $n=2$) a negative t_{ext}-estimate was obtained, suggesting inconsistency in k^{r2}-estimates.

Comparing the %E versus t_R^g plots in Fig. 1, the approaches could be ranked (highest accuracy order) as: Eq. (4) > (Eq.
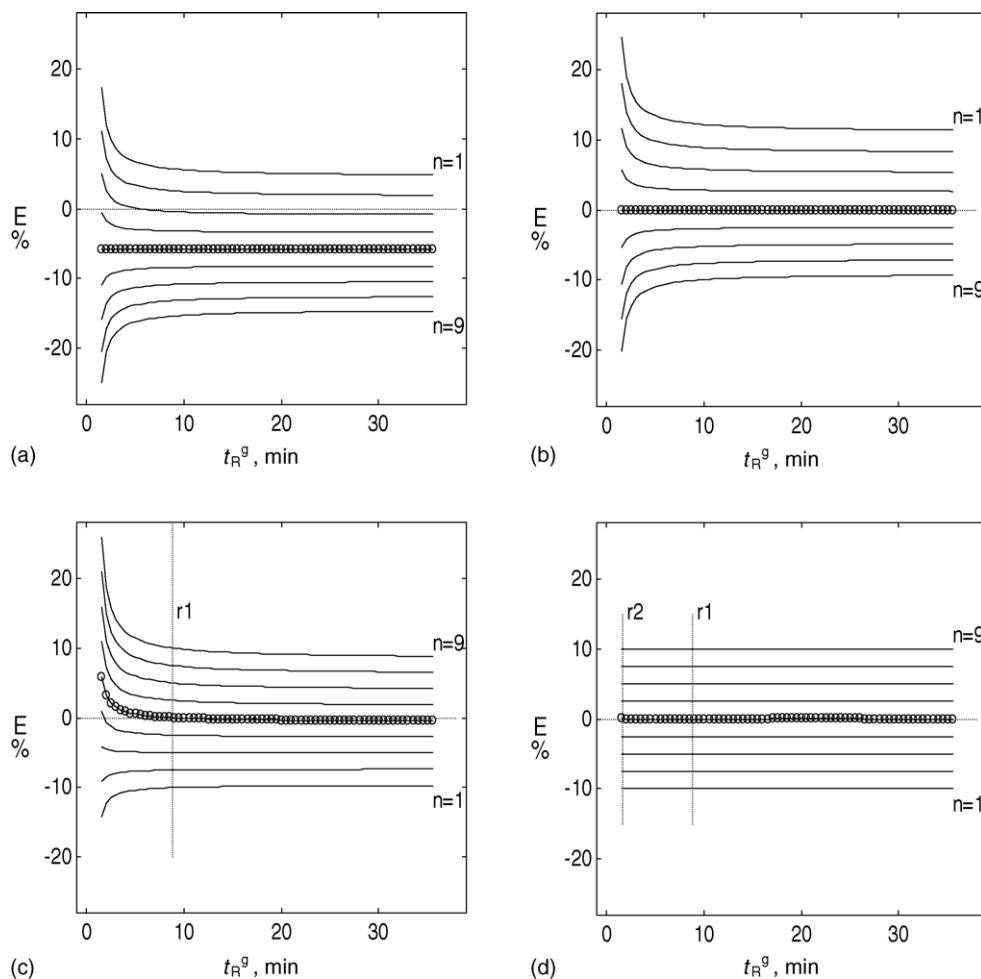
Fig. 1. %E (percentage of error in k estimation) vs. t_R^g (gross retention time of the test compounds, in min) for nine simulated errors in the parameters characteristic to each approach (from simulation $n = 1$, error $= -10\%$ to $n = 9$, error $= 10\%$). In the case of no error in the parameters (simulation $n = 5$), the symbol (o) was superimposed. Conditions used correspond to the simulation case 1 in Table 1. (a) %E values of approach 1 (Eq. (1)) with errors provoked on t_M^g. (b) %E values of approach 2 (Eq. (2)) with correlated errors provoked on t_M^g and t_{ext}. (c) %E values of approach 3 (Eq. (3)) with errors provoked on k_1. The gross retention time corresponding to the r1 reference has been indicated (vertical dashed line). (d) %E values of approach 4 (Eq. (4)) with correlated errors provoked on k_1 and k_2. The gross retention times corresponding to the r1 and r2 references has been indicated (vertical dashed line). n is used to indicate the simulation number.

(3) $\sim$ Eq. (2)) $>$ Eq. (1). This rank assumes equality of errors in the parameters characteristic of each approach. However, from the practical point of view we can also propose the following rank: (Eq. (3) $\sim$ Eq. (4)) $>$ (Eq. (2) $\sim$ Eq. (1)), attending to the number of parameters to be determined and their corresponding ambiguity. In order to complete the study, it will be valuable to consider other aspects affecting some approaches, such as the effect of correlation between errors in the case of approaches 2 or 4, the selection of reference compounds in the case of approaches 3 and 4 and studying the impact of a change in the chromatographic conditions to all the approaches.

4.3. Effect of correlation of errors

Previously, we have considered the situation of correlated t_{ext}- and t_M^g-errors (approach 2) or k_1- and k_2-errors (approach

4). This situation should be considered normal; at least that systematic errors associated to the experimental determination of these parameters have the same sign. However, considering the intrinsic difficulties to reliably determine t_{ext} and t_M^g and the contribution of imprecision of any measurement process to the accuracy [5], the possibility of uncorrelated errors was considered. The extreme case, inversely correlated t_{ext}- and t_M^g-errors or k_1- and k_2-errors, was considered in this work.

In the case of Eq. (2), the impact over %E of inversely correlated errors in t_M^g and t_{ext} did not offer notable changes (not shown) respect to the case of correlated errors (Fig. 1b). In contrast, in the case of Eq. (4), the impact over %E of inversely correlated errors in k_1 and k_2, offered a different view, as shown in Fig. 3a, respect to the correlated errors case (Fig. 1d). Now, the %E values are t_R^g-dependent (except for the simulation $n = 5$ and for $t_R^g = 3.2457$ min, an inflex-
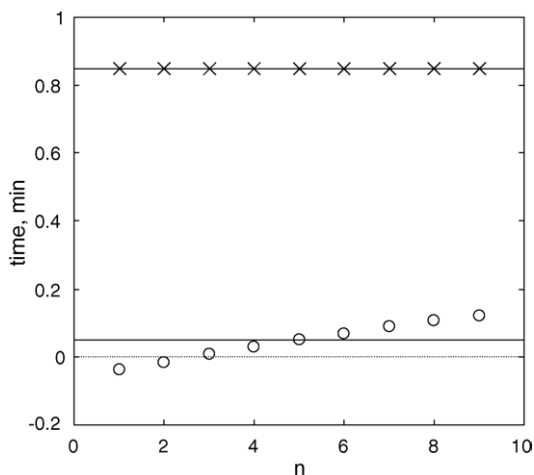
Fig. 2. Comparison between estimated (x) gross hold-up time, t_M^g, and (o) extra-column time, t_{ext}, by means of Eqs. (5) and (6) in Fig. 1d conditions and their corresponding true values (solid line). n is used to indicate the simulation number as in Fig. 1.

ion point, where $\%E = 0$). The $\%E$ values range was $\pm 13.17\%$ for $t_R^g = 1.55$ min and ± 11.71 for $t_R^g = 35.55$ min. As shown Fig. 3b, both the t_{ext} and t_M^g estimations from Eqs. (5) and (6) were different for each simulation (in contrast to Fig. 2). The firsts simulations (from $n = 1$ to $n = 4$) gave negative t_{ext} values when Eq. (6) is applied, thus suggesting inconsistency in the use of Eq. (4). Then, the particular (a priori less probable) situation of inversely correlated k_1- and k_2-errors situate approach 4 somewhat closer to approaches 2 or 3 in terms of accuracy.

4.4. Selection of reference compounds

Approaches 3 and 4, implies the selection of one and two reference compounds, respectively. The impact of the reten-

tion degree of these references over $\%E$ was examined by means of simulation. A low-retained compound was selected as r1 reference ($k_1 = 2$). As shown Fig. 4a, this situation had benefits for the approach 3 performance, compared with the previous one ($k_1 = 10$; Fig. 1c). In fact, a decrease of the $\%E$ values along the t_R^g range is observed. This new situation had no consequences over the approach 4 results, which coincide with those in Fig. 1d, except in the case of inversely correlated k_1- and k_2-errors, as shown Fig. 4b. Comparing the actual situation ($k_1 = 2$, $k_2 = 1$; Fig. 4b) with the previous one ($k_1 = 10$, $k_2 = 1$; Fig. 3a) an increase of $\%E$ values, particularly important for high-retained test compounds is observed.

4.5. Effect of the experimental chromatographic conditions

In some instances, it could be necessary a change in the chromatographic conditions in order to avoid experimental problems. For instance, if a set of high-retained test compounds (i.e. high hydrophobic molecules in reversed phase chromatography) have to be chromatographed, it could be convenient to use a shorter column and/or to increase the flow-rate to avoid large analysis time. Theoretically, a change in these parameters should not represent a change in k. A new chromatographic situation was simulated with a shorter column and higher flow rate (Table 1, simulation case 2) than the previous simulated (Table 1, simulation case 1), thus providing lower t_{ext} and t_M^g values. The impact of these changes on the four approaches was examined.

All the $\%E$ values corresponding to approach 1 were systematically displaced to negative values (not shown) respect to those obtained previously (Fig. 1a). For instance, the constant $\%E$ value in the case of no error in t_M^g (simulation $n = 5$), was now $\%E = -100 t_{ext}/t_M^g = -13.04\%$. This result is notably worse than the value $\%E = -5.88\%$ previ-
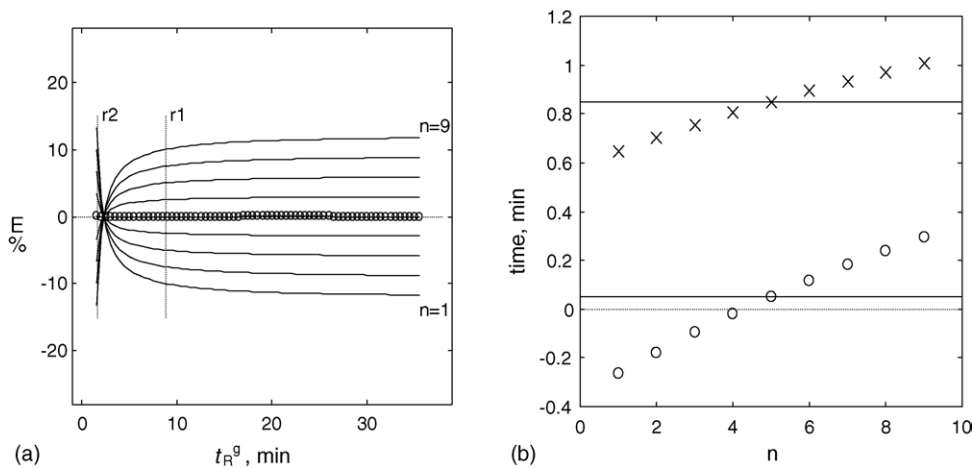

Fig. 3. Effect of inversely correlated errors in k_1 and k_2 (errors in k_2 are opposite to errors in k_1) over the estimations of approach 4: (a) $\%E$ vs. time (min). (b) t_M^g and t_{ext}. See Figs. 1 and 2 for further details.
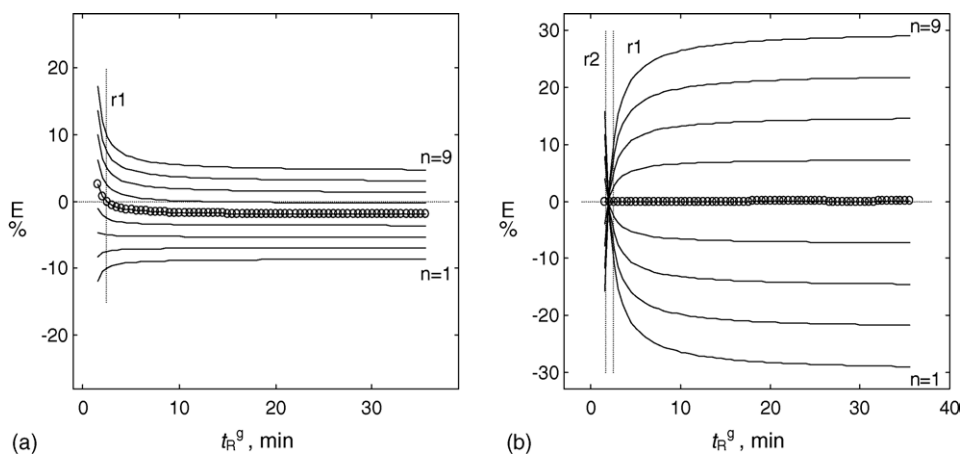
Fig. 4. Effect of the use of a low-retained r1 reference over %E values (vs. t_R^g) associated to: (a) approach 3 and (b) approach 4 with inversely correlated errors in k_1 and k_2 (k_2 is the same as in Figs. 1d and 3). See Fig. 1 for further details.

ously obtained. This fact is consequence of the relatively higher value of t_{ext} respect to t_M^g in the new simulated conditions. In contrast, there were no especial remarks for the results obtained using approach 2 (not shown) respect to those obtained previously (Fig. 1b).

In the case of approach 3, the %E values (not shown) were closer to those displayed in Fig. 4a than in Fig. 1c. This can be attributed to the already described effect of using a low-retained r1 reference (see Section 4.4). In the new chromatographic conditions, although $k_1 = 10$ (as in Fig. 1c) the t_{R1}^g is quite low ($t_{R1}^g = 2.23$ min) due to the short-column and high-flow rate simulated, which resembles the situation in Fig. 4a. This confirms that the value of t_{R1}^g is determinant on the %E values of approach 3 (low t_{R1}^g values for r1 reference are recommendable).

The new chromatographic conditions had no consequences over approach 4 results, which coincide with those in Fig. 1d, except in the case of inversely correlated k_1- and

k_2-errors, as shown Fig. 5a. The %E values corresponding to Eq. (4) in this particular case can be compared with those obtained in the initial conditions (i.e. Figs. 3a and 4b). The results shown in Fig. 5a were similar to those in Fig. 4b, particularly for high-retained test compounds. The value ($|4.23 - 2.23| = 2.00$ min) corresponding to Fig. 5a is closer to that in Fig. 4b ($|1.65 - 2.45| = 1.20$ min) than in Fig. 3a ($|1.65 - 8.85| = 7.20$ min). This indicates that $|t_{R2}^g - t_{R1}^g|$ is determinant on the %E values of approach 4 in the case of inversely correlated k_1- and k_2-errors.

Fig. 5b shows the effect of changes in the column length and flow rate on t_M^g and t_{ext} estimations by means of Eqs. (5) and (6) in the case of inversely correlated k_1- and k_2-errors. Except in simulation $n = 5$, errors in both estimates were found. For the case of positive k_1-errors combined with negative k_2-errors (simulations $n = 6$ to $n = 10$), negative t_{ext} and t_M^g values were obtained, thus pointing doubts over the estimations made from Eq. (4).
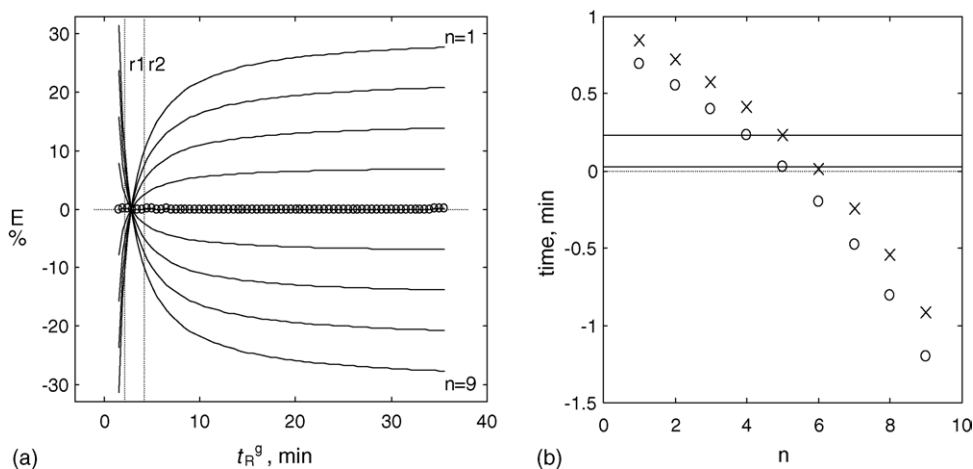

Fig. 5. Effect of the modification of the chromatographic conditions (column length and mobile phase flow-rate; simulation case 2 in Table 1) over the estimations of approach 4 with inversely correlated errors in k_1 and k_2. (a) %E vs. t_R^g. (b) t_M^g and t_{ext}. See Figs. 1 and 2 for further details.
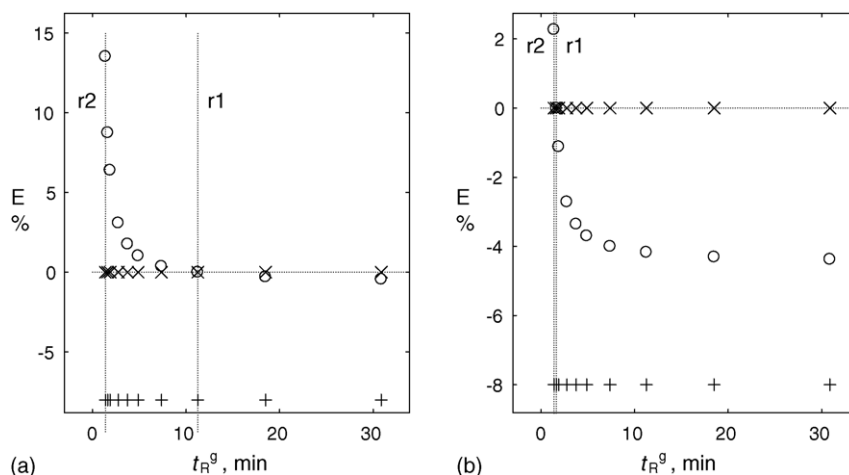
Fig. 6. %E values corresponding to real data calculations based on (+) approach 1, (o) approach 3 and (x) approach 4. Conditions used correspond to the real data in Table 1. (a) High-retained r1 reference and low-retained r2 reference (conditions favorable to approach 4 and unfavorable to approach 3). (b) Low-retained r1 reference and low-retained r2 reference (conditions favorable to approach 3 and unfavorable to approach 4).

4.6. Comparison of approaches using real data

Fig. 6 shows the results of applying approaches 1, 3 and 4 to estimate the retention factor of a set of ten compounds whose experimental log k values (assumed as 'accepted reference' data) were available [4]. Eq. (2), re-written for k, was used to calculate the corresponding t_R^g values using the available k, t_M^g and t_{ext} values [4] (Table 1, real data), so approach 2 was not included in the study. As can be observed, the %E values of approach 1 are almost constant (%$E \approx -8\%$). This result is consistent with those obtained in the simulation studies (i.e. Fig. 1a, simulation $n = 5$) and suggests insignificant error in t_M^g, otherwise, a dependence between %E and t_R^g would be observed. The %E values associated to approach 3 are close to zero for high-retained compounds, showing a positive trend (up to %$E = +15\%$ for the least-retained compound) as t_R^g decreases. This result is also consistent with those obtained in the simulation studies (i.e. Fig. 1c, sim-

Table 2

Comparison of approaches to estimate the retention factor (k) of compounds from their gross retention time (t_R^g) in a given chromatographic conditions (stationary and mobile phases) but changing the column length and flow rate

Approach Eq., estimate	Data[a]	Main advantages	Main limitations or difficulties
1: Classic Eq 1, k^g (biased)	t_M^g	t_{ext} is not required	It is biased except for $t_{ext} = 0$ Large errors are associated with low t_M^g values Reliable t_M^g is difficult to estimate (and ambiguous) [2] It needs to control t_M^g changes It is not recommendable for long-term studies and intermediate precision conditions [3]
2: True Eq. (2), k^t	t_M^g t_{ext}	It estimates the true k It is available (but not practical) for long-term studies	Reliable t_{ext} and t_M^g are difficult to estimate (and ambiguous) [2] It needs to control t_{ext} and t_M^g changes (tedious task)
3: Relative to one reference, r1 Eq. (3), k^{r1} (biased)	t_R^g	t_M^g or t_{ext} are not required[b]	It does not estimate the true k (errors can be minimized selecting a r1 reference with low t_{R1}^g value)
	k_1	It should be useful for long-term studies	
4: Relative to two references, r1 and r2 Eq. (4), k^{r2}	t_{R1}^g	It estimates the true k	It shows low robustness for large uncorrelated errors in k_1 and k_2 (errors can be minimized selecting r1 a r2 references with large $\|t_{R2}^g - t_{R1}^g\|$ values)
	k_1 t_{R2}^g	t_M^g or t_{ext} are not required[b] It should be useful for long-term studies	
	k_2	Estimation of t_M^g or t_{ext} is feasible using Eqs. (5) and (6) (with practical or control aims)	

[a] Data required besides t_R^g.

[b] Once k_1 (and k_2 in the case of approach 4) have been established in a given experimental chromatographic condition and used as constant in any conditions in which a change in the retention factor is not expectable. For establishing this values approach 2 is recommended in a careful statistically consistent experiment.

ulation $n = 5$) and suggests an accurate k_1 value, otherwise, %E values far from zero would be observed for high t_R^g values. Finally, Fig. 6a also shows the %E values associated to approach 4, which are close to zero for all the compounds. This result is also consistent with those obtained in the simulation studies (i.e. Fig. 1d, simulation $n = 5$) and suggests that both k_1 and k_2 are accurate, otherwise, %E values far from zero would be observed.

These results correspond to a relatively high-retained r1 reference and a low-retained r2 reference (see Table 1, real data). Such conditions are more favorable for approach 4 than for approach 3. Fig. 6b corresponds to a low-retained r1 reference (see Table 1, real data), which is a preferable situation for approach 3, according to the simulation study. In fact, the %E values for the least-retained compound notably decreases (%$E \approx 2\%$) while the %E values for high-retained compounds tend to moderate negative values (%$E \approx -4\%$). This effect is consistent with those obtained in the simulation studies (i.e. Fig. 4a, simulation $n = 5$). In the conditions of Fig. 6b the absolute $t_{R2}^g - t_{R1}^g$ difference is lower than in Fig. 6a, which should limit the approach 4 performance. However, as shown Fig. 6b the %E values corresponding to Eq. (4) are still close to zero, suggesting an excellent behavior of this approach.

In a real situation, the reliability of using Eq. (4) with the selected k_1 and k_2 values could be checked a priori by predicting t_{ext} and t_M^g values from the experimental t_{R1}^g and t_{R2}^g measurements based on Eqs. (5) and (6). For instance, in the case of Fig. 6b (a worse case situation for approach 4) the estimations obtained applying Eqs. (5) and (6) were $t_{ext} = 0.0733$ min and $t_M^g = 0.9150$ min. These values, which coincide with the values indicated by the authors[4] (see Table 1, real data), would suggest the a priori accuracy of the k_1 and k_2 values used in Eq. (4), and therefore, the reliability of k estimations of the test compounds from their corresponding k^{r2} values.

5. Conclusions

Table 2 summarizes the main features of four (two classic and two new) approaches for estimating the retention factor of a compound, consistent with the results obtained in this study. Simulations are based on provoking errors in the parameters characteristic of each approach (second column in Table 2), assuming that these errors are systematic. However, the magnitude of the simulated errors ($\pm 10\%$) might also account for random errors; therefore, the conclusion of this paper should have reasonable validity in view of being extrapolated to the routine work.

From a mathematical point of view, the approaches 2 and 4 are preferable, due to the biased nature of Eq. (3), and particularly, Eq. (1). However, the key point to rank the approaches, in practice, is the magnitude of the errors associated to these parameters in routine. In this sense, it could be reasonable to rank the approaches as: Eq. (4) > Eq.

(3) > Eq. (2) > Eq. (1). This can be justified attending to the less ambiguity and higher reliability of t_{R1}^g and t_{R2}^g measurements used in approaches 3 and 4, compared with t_M^g and t_{ext} used in approaches 1 and 2 (particularly, in long-term studies). Additionally, an extra-advantage of approach 4 over the other approaches is that it permits its partial auto-control offered by combining Eqs. (4)–(6).

Simulation studies show that in the case of correlated errors in k_1 and k_2 (the normal expected situation), approach 4 is the unique one that provides errors in the k estimations (%E) independents of the retention of the test compound (t_R^g), which is very convenient for the accuracy of the lowest-retained compounds, for which the rest of approaches tend to fail. The consistency of approach 4 has been confirmed here studying real data.

However, the reliability of approach 4 (or 3) strongly depends on the quality associated to the retention factors of the reference compounds (k_1 and k_2). Therefore, the proposed ranking of approaches would be more suitable if accepted reference k_1 and k_2 values were available. On the other hand, even if the k_1 and k_2 values are obtained in the laboratory by means Eq. (2), the internal consistency between k estimations of compounds in a long-term sense should be guaranteed, due to the effect of standardization respect to a reference. This should minimize the impact of changes in the columns/equipment (new column, column length, flow rate, etc.), but also prevent errors due to lack of control (drifts, fluctuations, etc.). In fact, the continuous monitoring of the reference compounds represents itself an internal quality control scheme of the chromatographic system.

Appendix A. Symbols used in equations and tables

k	true or 'accepted reference' retention factor
k_1	true or 'accepted reference' retention factor of the reference compound r1
k_2	true or 'accepted reference' retention factor of the reference compound r2
k^g	gross retention factor estimate by Eq. (1)
k^t	retention factor estimate by Eq. (2)
k^{r1}	retention factor estimate by Eq. (3), relative to the use of the reference compound r1 with available k value (k_1)

k^{r2} retention factor estimate by Eq. (4), relative to the use of two reference compounds r1 and r2 with available k values (k_1 and k_2).

t_{ext} extra-column time

t_M hold-up time ($t_M = t_M^g - t_{ext}$)

t_M^g gross hold-up time

t_R retention time ($t_R = t_R^g - t_{ext}$)

t_R^g gross retention time

t_{R1}^g gross retention time of the reference compound r1

t_{R2}^g gross retention time of the reference compound r2

References

[1] J.A. García Domínguez, J.C. Díez-Masa, Pure Appl. Chem. 73 (2001) 969.

[2] C.A. Rimmer, C.R. Simmons, J.G. Dorsey, J. Chromatogr. A 965 (2002) 219.

[3] L. Escuder-Gilabert, J.M. Bermúdez-Saldaña, R.M. Villanueva-Camañas, M.J. Medina-Hernández, S. Sagrado, J. Chromatogr. A 1033 (2004) 247.

[4] N.S. Wilson, M.D. Nelson, J.W. Dolan, L.R. Snyder, R.G. Wolcott, P.W. Carr, J. Chromatogr. A 961 (2002) 171.

[5] L.A. Currie, Anal. Chim. Acta 391 (1999) 105.