ELSEVIER

Journal of Chromatography A, 954 (2002) 59–76

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Parabolic-Lorentzian modified Gaussian model for describing and deconvolving chromatographic peaks

R.D. Caballero, M.C. García-Alvarez-Coque, J.J. Baeza-Baeza*

Departamento de Química Analítica, Facultad de Química, Universitat de València, 46100 Burjassot, València, Spain

Received 22 November 2001; received in revised form 15 February 2002; accepted 20 February 2002

Abstract

A new mathematical model for characterising skewed chromatographic peaks, which improves the previously reported polynomially modified Gaussian (PMG) model, is proposed. The model is a Gaussian based equation whose variance is a combined parabolic-Lorentzian function. The parabola accounts for the non-Gaussian shaped peak, whereas the Lorentzian function cancels the variance growth out of the elution region, which gives rise to a problematic baseline increase in the PMG model. The proposed parabolic-Lorentzian modified Gaussian (PLMG) model makes a correct description of peaks showing a wide range of asymmetry with positive and/or negative skewness. The new model is shown to give better fittings than other models as the Li, log-normal or Pap–Pápai models, which have a different mathematical basis. The model parameters are also related to peak properties as the skewness and kurtosis. The PLMG model is applied to the deconvolution of peaks in binary mixtures of structurally related compounds that are highly overlapped (retention times in min): oxytetracycline (9.00)—tetracycline (10.20), sulfathiazole (3.67)—sulfachloropyridazine (3.93), and sulfisoxazole (5.14)—sulfapyridine (5.24). The use of non-linear least-squares calibration in combination with the PLMG model gave superior results than the classical multiple linear least-squares and partial least-squares regressions. The proposed method takes into account run to run changes in retention time that occur along the injection of standards and samples, and the possible interactions that exist between the coeluting compounds. This decreases significantly the quantitation errors.

Keywords: Skewed peaks; Modified Gaussian model; Peak modelling; Deconvolution

1. Introduction

Although chromatography is a powerful separation technique, full resolution of all components in a sample may not be feasible by optimising the experimental conditions, such as the column type or nature of the modifiers, and the mobile phase composition. Actually, a high probability exists of finding at least two overlapped peaks when complex samples are processed. In this case, a deconvolution technique must be applied to quantitate appropriately each analyte. However, the stability of the baseline, and the changes in retention time and peak shape among standards and samples, as well as the possible interactions between the solutes that elute at close times, limits the accuracy and precision of the deconvolution. These problems can be overcome by using a model, flexible enough to cope with the changes that affect the chromatographic peaks.

*Corresponding author.

E-mail address: juan.baeza@uv.es (J.J. Baeza-Baeza).

0021-9673/02/$ – see front matter
PII: S0021-9673(02)00194-2

The search of models that describe correctly the chromatographic peaks has been pursued intensively. In the 1960s, Fraser and Suzuki [1,2] used the Gauss and Cauchy functions to fit and deconvolve spectral absorption bands. Both functions were later modified to be applied in the description of chromatographic peaks. Today, several Gaussian modified functions are used routinely to model peaks with different asymmetry degrees [3,4].

In ideal conditions, a chromatographic peak is described by:

$$h(t) = H_0 \mathrm{e}^{-\frac{1}{2}\left(\frac{t-t_R}{\sigma}\right)^2} \quad (1)$$

where H_0 is the height at the maximum, t_R the retention time, and σ the standard deviation that measures the peak width. Peaks are, however, often skewed due to the complex interactions that are established between solute and stationary phase, and to extra-column processes. Several models based on the Gaussian function have been proposed to describe these deviations. Haarhoff and van der Linde [5], Fraser and Suzuki [2], Buys and Clerk [6], Chesler and Cram [7], Dondi et al. [8] developed some of the earliest models. The exponentially modified Gaussian model (EMG) has been used extensively [9–13]. Other more recent models are the generalised exponential [14], log-normal [15], exponential bi-Gaussian [16], coupled leading and trailing edge [17], Gaussian–Lorentzian [18], two-Gaussians [19], exponential Gaussian hybrid [20], and the Pap–Pápai function [21].

The polynomially modified Gaussian (PMG) model was proposed in our laboratory to improve the simulation and prediction of chromatograms [22], needed for a reliable optimisation of the resolution [23]. In this model, the deviations from ideality are interpreted as a change in the standard deviation as a function of time, according to a polynomial function:

$$h(t) = H_0 \mathrm{e}^{-\frac{1}{2}\left(\frac{t-t_R}{\sigma_0+\sigma_1(t-t_R)+\sigma_2(t-t_R)^2+\ldots}\right)^2} \quad (2)$$

This approach has demonstrated a great flexibility in the simulation of strongly tailed and fronted peaks. It has also been applied to the deconvolution of partially overlapped peaks in binary and ternary mixtures with good results [4,22], improving the performance of the EMG model, which is often taken as reference in modelling and resolution reports.

The peak models developed up-to-date usually fail in simulating the wide range of asymmetries and shapes that chromatographic peaks can adopt. Also, most models do not fit conveniently the baseline outside the elution region. Recently, an exponential Gaussian model has been reported [20], which describes satisfactorily highly skewed peaks, but its utility in the deconvolution of overlapped peaks has not been demonstrated. In a comparison study, Nikitas et al. [4] pointed out that the PMG model is the only one able to describe almost any peak. Nevertheless, this model has the drawback of the uncontrolled growth of the predicted signal outside the elution region, which departs from the experimental baseline. This problem is easily overcome, when possible, by making $h(t)=0$ outside the peak region. Nikitas et al. [4] proposed recently another solution, which was called the PMG2 model:

$$h(t) = H_0 \frac{\sigma_0}{\sigma_0 + \sigma_1(t-t_R) + \sigma_2(t-t_R)^2 + \ldots} \cdot \mathrm{e}^{-\frac{1}{2}\left(\frac{t-t_R}{\sigma_0+\sigma_1(t-t_R)+\sigma_2(t-t_R)^2+\ldots}\right)^2} \quad (3)$$

where $h(t \to \infty) = 0$, independently of the polynomial degree.

In this work, a new model that shows a great flexibility to describe chromatographic peaks, as the PMG model, and solves its drawback, is proposed. The model is again based on the Gaussian function where the variance is a parabolic-Lorentzian function. It describes correctly the elution region and baseline for peaks showing a wide range of asymmetries and kurtosis. Its good performance in the deconvolution of fused peaks of binary mixtures of structurally related compounds is demonstrated. Throughout the study, chromatographic conditions were chosen to obtain wide peaks with a strong distortion, in order to show the potentiality of the approach.

2. Theoretical section

2.1. Peak models

Several peak models based on Eq. (1), where

different functions substituted the variance, were examined (only these functions are shown below):

(i) Parabolic standard deviation modified Gaussian model (PSMG):

$$\sigma^2 = [\sigma_0 + m\,(t - t_R + d)^2]^2 \quad (4)$$

(ii) Fourth degree polynomial standard deviation modified Gaussian model (FSMG):

$$\sigma^2 = [\sigma_0 + m_1(t - t_R + d)^2 + m_2(t - t_R + d)^3 + m_3(t - t_R + d)^4]^2 \quad (5)$$

(iii) Parabolic variance modified Gaussian model (PVMG):

$$\sigma^2 = \sigma_0^2 + m(t - t_R + d)^2 \quad (6)$$

(iv) Fourth degree polynomial variance modified Gaussian model (FVMG):

$$\sigma^2 = \sigma_0^2 + m_1(t - t_R + d)^2 + m_2(t - t_R + d)^3 + m_3(t - t_R + d)^4 \quad (7)$$

(v) Parabolic-Lorentzian variance modified Gaussian model (PLMG):

$$\sigma^2 = \sigma_0^2 + m\frac{(t - t_R + d)^2}{1 + \dfrac{(t - t_R + r)^2}{w^2}} \quad (8)$$

PSMG and PVMG make a parabolic description of the standard deviation and variance, respectively, using five parameters (H_0, σ_0, t_R, m and d). FSMG and FVMG improve the description with fourth degree polynomials and contain seven parameters (H_0, σ_0, t_R, m_1, m_2, m_3 and d). PSMG and FSMG are two different forms of the PMG model, while PVMG and FVMG are modifications of the model in which the variance, instead of the standard deviation, is the polynomial function.

The PLMG model (Eq. (8)) combines a parabolic and a Lorentzian function to describe the variance profile (Fig. 1), and contains seven parameters (σ_0, m and d for the parabolic function, and r and w for the Lorentzian function, in addition to H_0 and t_R). The parabola has a minimum at $t = t_R - d$, and the Lorentzian function a maximum at $t = t_R + r$. For tailed peaks, the minimum of the parabola is located at

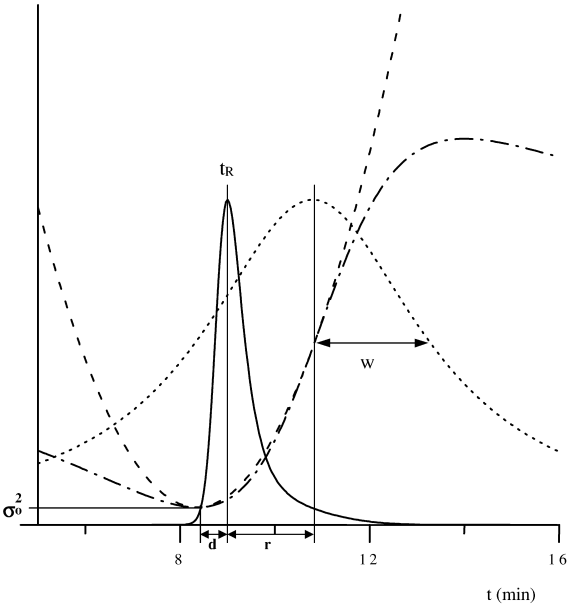

Fig. 1. Meaning of the parameters in the PLMG model (Eq. (8)).

times $t < t_R$ and, therefore, $d > 0$, whereas for fronted peaks, the minimum is found at times $t > t_R$ and $d < 0$. The slope of the parabola, and consequently the variance, increases with m. The Lorentzian function has been added to the model to decrease the variance growth out of the peak region.

Other models were checked for comparative purposes:

(vi) Li [24]:

$$h(t) = \frac{p}{\left(1 + e^{m(t_1 - t)}\right)^{\lambda_1} + \left(1 + e^{n(t_2 - t)}\right)^{\lambda_2} - 1} \quad (9)$$

(vii) Log-normal [15]:

$$h(t) = H_0 e^{\left\{\frac{\ln q}{\ln^2 a}\ln^2\left[\left(\frac{t - t_R}{W}\right)\left(\frac{a^2 - 1}{a}\right) + 1\right]\right\}} \quad (10)$$

(viii) Pap–Pápai [21]:

$$h(t) = H_0 e^{\left\{\left(\frac{4}{a^2} - 1\right)\left[\ln\left(1 + \frac{2a(t - t_R)}{D(4 - a^2)}\right) - \frac{2a(t - t_R)}{D(4 - a^2)}\right]\right\}} \quad (11)$$

Eq. (9) has seven parameters (m, n, p, t_1, t_2, λ_1 and λ_2), Eq. (10) five (H_0, t_R, a, q and W), and Eq. (11) four (H_0, t_R, a and D). In Eqs. (10) and (11), a is an asymmetry measurement, W the peak width, q

the ratio of the heights at a selected time and the maximum, and D the standard deviation.

2.2. Relationship between the PLMG model and peak shape

The meaning of the PLMG model can be understood if its parameters are conveniently related to peak shape measurements. During the elution, the PLMG model approximates to the PVMG model. Therefore, for simplicity, we will assume that the variance profile is parabolic in the peak region and will not consider the Lorentzian function, which cancels the variance out of the peak region. From Eq. (1):

$$\sigma^2 = -\frac{(t-t_R)^2}{2\ln\frac{h}{H_0}} \tag{12}$$

which for the PVMG model leads to:

$$p = -2\ln\frac{h}{H_0} = \frac{(t-t_R)^2}{\sigma_0^2 + m(t-t_R+d)^2} = \frac{(-A)^2}{\sigma_0^2 + m(d-A)^2} = \frac{(B)^2}{\sigma_0^2 + m(B+d)^2} \tag{13}$$

where A and B are the time distance between the maximum and the fronting and tailing edge of the chromatographic peak, respectively. Operating, the following is obtained:

$$(1-mp)A^2 + 2mpdA - p\sigma_m^2 = 0 \tag{14}$$

$$(1-mp)B^2 - 2mpdB - p\sigma_m^2 = 0 \tag{15}$$

where $\sigma_m^2 = \sigma_0^2 + md^2$ is the variance at the peak maximum ($t = t_R$). From Eqs. (14) and (15):

$$A = -\frac{mp}{1-mp}d + \sqrt{\left(\frac{mp}{1-mp}d\right)^2 + \frac{p\sigma_m^2}{1-mp}} \tag{16}$$

$$B = \frac{mp}{1-mp}d + \sqrt{\left(\frac{mp}{1-mp}d\right)^2 + \frac{p\sigma_m^2}{1-mp}} \tag{17}$$

The asymmetry degree can be defined as the ratio between the semiwidths difference $(B-A)$, at a height defined by p (Eq. (13)), and the width of an ideal Gaussian peak measured at the same height:

$$a = \frac{B-A}{2\sigma_m\sqrt{p}} = \frac{m\sqrt{p}}{1-mp}\cdot\frac{d}{\sigma_m} \tag{18}$$

For $p=1$ (which corresponds to a height ratio, $h/H=0.607$):

$$a = \frac{m}{1-m}\cdot\frac{d}{\sigma_m} \tag{19}$$

For a symmetrical peak, $a=0$, since the minimum of the parabola coincides with the peak maximum and $d=0$. For $a>0$, the peak has a positive skewness, and for $a<0$, the skewness is negative. The value of a indicates the percentage of asymmetry with respect to a pure Gaussian peak. Thus, $a=0.5$ or -0.5 means that 50% of the peak corresponds to the extra peak width due to the skewness. In other words, the deformation represents 50% of the Gaussian width.

On the other hand, the peak kurtosis can be defined as the change in peak width with respect to an ideal Gaussian peak:

$$c = \frac{B+A-2\sigma_m\sqrt{p}}{2\sigma_m\sqrt{p}} = \sqrt{\left(\frac{m}{1-mp}\frac{d}{\sigma_m}\right)^2 p + \frac{1}{1-mp}} - 1 \tag{20}$$

Again, for $p=1$:

$$c = \sqrt{\left(\frac{m}{1-m}\frac{d}{\sigma_m}\right)^2 + \frac{1}{1-m}} - 1 \approx \frac{1}{2}\left[\left(\frac{m}{1-m}\cdot\frac{d}{\sigma_m}\right)^2 + \frac{m}{1-m}\right] \approx \frac{m/2}{1-m} \tag{21}$$

where a Taylor series expansion has been done. When $c=0$, the peak has no kurtosis. For $c>0$ ($m>0$) the peak is leptokurtic (positive kurtosis, wider than a Gaussian) and for $c<0$ ($m<0$) the curve is platykurtic (negative kurtosis, flattened with respect to a Gaussian). For the studied peaks, $c>0$, and this value indicates the percentage of width change with respect to a pure Gaussian. Thus, if $c=0.5$, the width is 50% greater.

2.3. Deconvolution and quantitation methods

Three different methods were applied to quantitate the individual components in binary mixtures:

Method 1: The peaks were non-linearly fitted according to the PLMG model, taking the height of the peaks as the analytical parameter proportional to the concentration. The signal was treated as the addition of two individual peaks:

$$S_{(t)} = C_1 k_1 \mathrm{e}^{-\frac{1}{2}\frac{(t-t_{R1})}{\sigma_{0.1}^2 + m_1\frac{(t-t_{R1}+d_1)^2}{1+\frac{(t-t_{R1}-r_1)^2}{w_1^2}}}} + C_2 k_2 \mathrm{e}^{-\frac{1}{2}\frac{(t-t_{R2})}{\sigma_{0.2}^2 + m_2\frac{(t-t_{R2}+d_2)^2}{1+\frac{(t-t_{R2}-r_2)^2}{w_2^2}}}} + \beta_1 t + \beta_0 \quad (22)$$

A similar procedure has been used before for other peak models. As observed, the number of model parameters is large. For this reason, it is usually assumed that the profiles of the chromatographic peaks of the standard solutions do not change in the overlapped peaks. Also, according to Nikita et al. [4], this assumption leads to better results. However, if the experimental data showed a significant variability in peak shape, one or more peak shape parameters, mainly σ_0, could be included in the deconvolution model to account for theses changes.

In this work, the fitted parameters were limited to the retention times of the individual solutes and their heights. Two different approaches were assayed to quantitate appropriately the deconvolved solutes:

Method 1a: The signals of the compounds were considered independent. Therefore, the heights obtained from the deconvolution were interpolated in the calibration lines of each compound:

$$H_1 = a_1 C_1 + b_1$$
$$H_2 = a_2 C_2 + b_2 \quad (23)$$

H_i being the peak heights, C_i the concentration of the analytes, and a_i and b_i the calibration parameters.

Method 1b: An interaction between the two compounds in the mixture was assumed:

$$H_1 = a_{11} C_1 + a_{12} C_2 + b_1$$
$$H_2 = a_{21} C_1 + a_{22} C_2 + b_2 \quad (24)$$

The calibration parameters were obtained using standards of the individual compounds and mixtures different to that found in the analysed sample. For the standard mixtures, the values of H_1 and H_2 were obtained by applying the deconvolution procedure to their chromatograms. For each component, multiple linear least-squares regression was applied to obtain the calibration parameters a_{ij} and b_i.

Method 2: Multiple linear least-squares regression (MLR) was applied assuming that the shape and position of the chromatographic peaks do not change with the concentration. Accordingly, the signal of a mixture of two compounds was expressed as the sum of the individual peaks:

$$S(t) = C_1 \beta_1(t) + C_2 \beta_2(t) + \alpha_1(t) + \alpha_2(t) \quad (25)$$

where $\beta_1(t)$, $\beta_2(t)$, $\alpha_1(t)$ and $\alpha_2(t)$ are the calibration parameters for each compound at time t. The values of β and α were calculated by fitting to a straight-line the values of the experimental signal against the concentration, at each time, for the peaks obtained by injection of the individual standards. Once these parameters were known, Eq. (25) was fitted to the signal of the overlapped peaks in the mixture, to obtain the individual concentrations.

In this method, the main source of error is the change in retention times among injections. For this reason, to reduce the experimental errors, these times were corrected to use the same values for all calibration standards. These values were the time points closest to the means of the retention times in standards and samples.

Method 3: Two approaches were also applied using partial least-squares regression [25]. In Method 3a, the model was built using only the solutions of the individual standards. In Method 3b, the individual standards and mixtures of the two compounds were used. The retention times were corrected similarly as in Method 2. It should be noted that in Method 1, the retention times are fitting parameters during the deconvolution.

3. Experimental

3.1. Reagents

The probe compounds were the tetracyclines oxytetracycline chlorhydrate (OTC) and tetracycline chlorhydrate (TC), the sulfonamides sulfachloro-

pyridazine (SCP), sulfapyridine (SPD), sulfathiazole (STZ), and sulfisoxazole (SFZ) (Sigma, St. Louis, MO, USA), and the diuretic spironolactone (Searle, Madrid, Spain). Stock solutions containing 100 μg/ml of each compound were prepared by dissolving the solid reagents in 95% (v/v) ethanol (Prolabo, Fontenay, France). The working solutions of tetracyclines and sulfonamides were obtained by dilution with aqueous 1% acetic acid (Panreac, Barcelona, Spain) at pH 3.

The mobile phases were prepared with sodium dodecyl sulfate (SDS, 99% purity, Merck, Darmstadt, Germany) and 1-butanol, acetonitrile or 1-pentanol (Scharlab, Barcelona, Spain). The composition of the mobile phase was 0.05 *M* SDS–5% (v/v) 1-butanol for tetracyclines, 0.10 *M* SDS–6% (v/v) acetonitrile for sulfonamides, and 0.10 *M* SDS–0.5% (v/v) 1-pentanol for spironolactone. For tetracyclines and sulfonamides, the pH of the mobile phase was buffered at 3 with citric acid and sodium hydroxide (Panreac). Mobile phases and probe compound solutions were filtered through 0.45 μm nylon membranes (Micron Separations, Westboro, MA, USA), except sulfachloropyridazine for which cellulose acetate filters were used owing to adsorption problems with the nylon filters.

3.2. Apparatus

An Agilent HP 1100 (Palo Alto, CA, USA) liquid chromatograph equipped with an isocratic pump, an automatic injector and a UV–visible detector was used. The detection wavelength was set at 364 nm for tetracyclines, 275 nm for sulfonamides, and 254 nm for spironolactone. The signal was acquired through an HP 3396A integrator with the aid of the PEAK-96 program (Agilent, Avondale, PA, USA). The injection volume was 20 μl and the flow-rate, 1 ml/min. A Spherisorb ODS-2 column (125×4.6 mm I.D.) and precolumn (35×4.6 mm I.D.) with a 5 μm particle size (Scharlab) were employed.

4. Results and discussion

4.1. Reproducibility of peak parameters

The reliable deconvolution of chromatographic peaks needs the parameters describing peak position and shape be reproducible among injections, and do not depend excessively on solute concentration. Therefore, the reproducibility of peak parameters was first checked for the chromatographic system used in this work. For this purpose, triplicate injections of solutions containing 3 μg/ml of two probe compounds, OTC and TC, were made over a 13–14 h period, maintaining all experimental factors constant. These compounds gave peaks with low efficiencies and long tails in the selected elution conditions. Table 1 shows the measured retention times, retention factors (k), efficiencies (N), and asymmetry factors (B/A, where B and A were measured at 10% of peak height). The dead time was measured for each chromatogram.

The mean retention times, efficiencies and asymmetry factors were $t_R = 9.00$ and 10.20 min, $N = 274$ and 196, and $B/A = 2.42$ and 2.86, for OTC and TC, respectively. Changes in retention times and retention factors were ~1 and 2%, and ~1 and 4%, for OTC and TC, respectively. The larger variability in the retention factors was due to the additional uncertainty in the dead time measurement. The changes in peak shape parameters (N, A, B and B/A) were greater than 2%. In all cases, the reproducibility

Table 1
Peak parameters for OTC and TC obtained from non-consecutive injections

Time[a] (h)	t_R (min)	k	N[b]	A[c]	B[c]	B/A
Oxytetracycline						
1.0	8.88	7.46	288	0.52	1.25	2.41
7.6	8.98	7.55	277	0.53	1.29	2.43
12.8	9.04	7.56	271	0.56	1.31	2.36
13.1	9.13	7.43	260	0.54	1.35	2.48
Mean	9.00	7.50	274	0.54	1.30	2.42
ε_r (%)	1.2	0.9	4.4	3.0	3.2	2.1
Tetracycline						
3.1	10.15	8.09	197	0.59	1.70	2.90
9.7	10.16	8.68	182	0.62	1.78	2.86
13.5	10.12	8.64	192	0.60	1.72	2.86
13.8	10.35	8.86	213	0.63	1.77	2.83
Mean	10.20	8.57	196	0.61	1.74	2.86
ε_r (%)	1.1	3.9	6.6	3.0	2.2	0.9

[a] Time from the first injection.
[b] According to Foley and Dorsey [9].
[c] Measured at 10% of peak height.

among replicates (injections at close times) was excellent, with errors below 1%.

Similar results were achieved for both compounds in another study carried out with solutions of OTC and TC at six different concentrations, in the range 0.5 to 5.0 μg/ml. The change in retention times and retention factors was ~1 and 2–3%, respectively. The reproducibility was poorer for the shape parameters: ~1–5% for OTC and ~1–10% for TC. The following mean values were obtained: $t_R = 8.96$ and 10.10 min, $N = 284$ and 200, and $B/A = 2.42$ and 2.84, for OTC and TC, respectively. Similar reproducibilities resulted for the sulfonamides.

The use of retention times instead of retention factors is more convenient to describe the peak position. The reproducibility was certainly good. In contrast, peak shape may vary appreciably among injections. The peaks obtained for the standards during the calibration may differ remarkably from the peaks of the samples. This can affect seriously the success of the deconvolution. Therefore, the variability in the peak parameters was considered in the development of appropriate deconvolution methods.

4.2. Comparative study of peak models

New functions were examined to model the changes in variance during peak elution. For this purpose, the experimental variance profile in non-Gaussian chromatographic peaks was studied. Fig. 2 shows peaks for SFZ, SCP and TC, and the variance associated to the experimental signal calculated with Eq. (12) by keeping the retention time and height constant. The experimental signal was next fitted according to several modified Gaussian functions showing a polynomial standard deviation or variance (models i to iv), and a parabolic-Lorentzian variance (model v). These models were compared with the PMG2 model (Eq. (3)), and other different models proposed in the literature: the Li (Eq. (9)), log-normal (Eq. (10)), and Pap–Pápai (Eq. (11)) models. A fourth degree polynomial was selected as standard deviation in the PMG2 model, to include seven fitting parameters (H_0, t_R, and σ_0 to σ_4) as in the PLMG model. The Li model was also adapted to include only seven parameters. In the literature, the PMG, Li and log-normal models have been evaluated favourably against other models commonly used in chromatography, such as the EMG and Gram–Charlier models, and the series of Edgeworth–Cramer [4,24]. The models studied in this work are also simpler.

The regression coefficients, R, and the relative fitting errors, ε_{rf} (Eq. (26)), obtained in the description of the OTC and TC peaks according to the different models are given in Table 2:

$$\varepsilon_{rf}(\%) = \frac{\frac{\sum |S_i - \hat{S}_i|}{N}}{\frac{\sum S_i}{N}} \times 100 \quad (26)$$

where S_i and $\hat{S}_i$ are the experimental and model estimated signals, and N the total number of points.

The fittings were made for two time intervals around the peak maximum: 5σ and 38σ for OTC, and 6σ and 35σ for TC (σ was measured as the left semiwidth at 60% of peak height). The narrower interval (8.3–10.8 min for OTC, and 9.4–12.8 min

Table 2
Modelling of OTC and TC peaks[a]

Model	R	ε_{rf} (%)	R	ε_{rf} (%)
Oxytetracycline	8.3–10.8 min ($\pm 5\sigma$)		1.5–20 min ($\pm 38\sigma$)	
PSMG	0.99979	1.60	0.9966	17
FSMG	0.999994	0.21	0.9997	4.8
PVMG	0.99997	0.61	0.99978	5.4
FVMG	0.999995	0.23	0.99997	1.9
PLMG	0.999998	0.13	0.999981	1.6
PMG2	0.999995	0.24	0.99976	4.5
Li	0.99916	2.9	0.9986	9.8
Log-normal	0.9951	8.1	0.9962	15
Pap–Pápai	0.9919	11	0.9944	18
Tetracycline	9.4–12.8 min ($\pm 6\sigma$)		1.5–20 min ($\pm 35\sigma$)	
PSMG	0.99976	1.7	0.9963	16
FSMG	0.999966	0.64	0.99985	3.4
PVMG	0.999993	0.27	0.99981	4.1
FVMG	0.999995	0.23	0.999989	1.0
PLMG	0.999998	0.16	0.999994	0.75
PMG2	0.999979	0.48	0.99982	2.8
Li	0.9987	3.9	0.9980	11
Log-normal	0.9949	8.3	0.9960	15
Pap–Pápai	0.9902	12	0.9930	19

[a] Regression coefficients (R) and relative fitting errors (ε_{rf}, see Eq. (26)) are given.

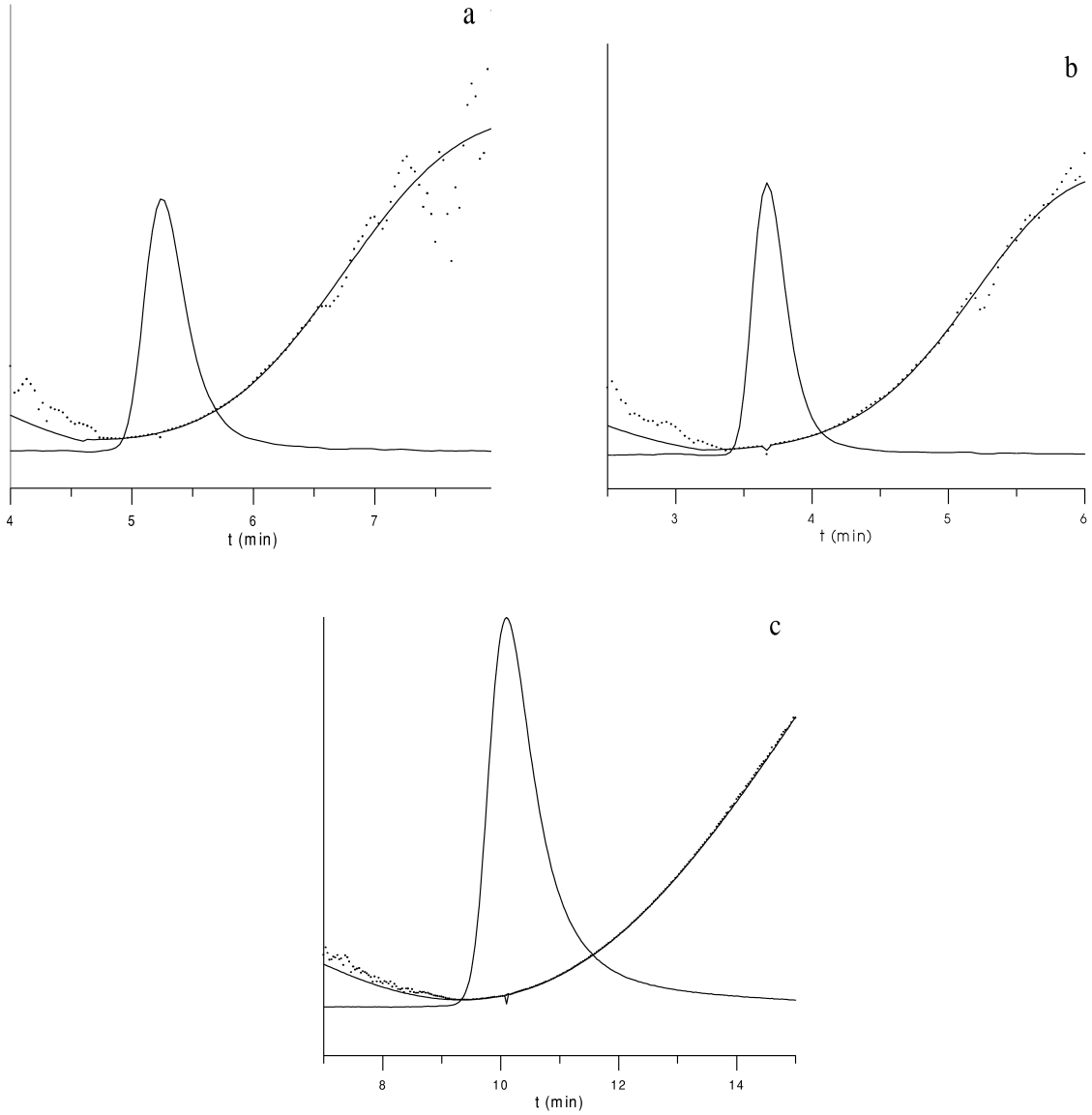

Fig. 2. Experimental chromatographic peaks and associated variance measured using Eq. (12) (points). The predicted variance according to Eq. (8) is also given (solid line). Compounds (B/A at 10% of peak height): (a) SFZ (1.65), (b) SCP (1.81), and (c) TC (2.86).

for TC) included the peak and a short baseline section. A longer baseline section was included in the wider interval (1.5–20 min for both compounds).

Figs. 3 and 4 show the behaviour of the models for OTC. In all cases, the experimental signal was fitted between 8 and 11 min, and the peak was modelled between 1.5 and 20 min. In this way, the model behaviour to describe the peak and baseline out of the fitting interval could be examined. The PLMG model gives the best peak description (Fig. 4b), and keeps the baseline within the experimental values. The PVMG model (Fig. 3b) has also a good behaviour, but the predicted baseline departs from the experimental signal. This model tends to $e^{-1/2m}$ at times far from the peak maximum. Therefore, if the peak is highly asymmetrical (large m), the

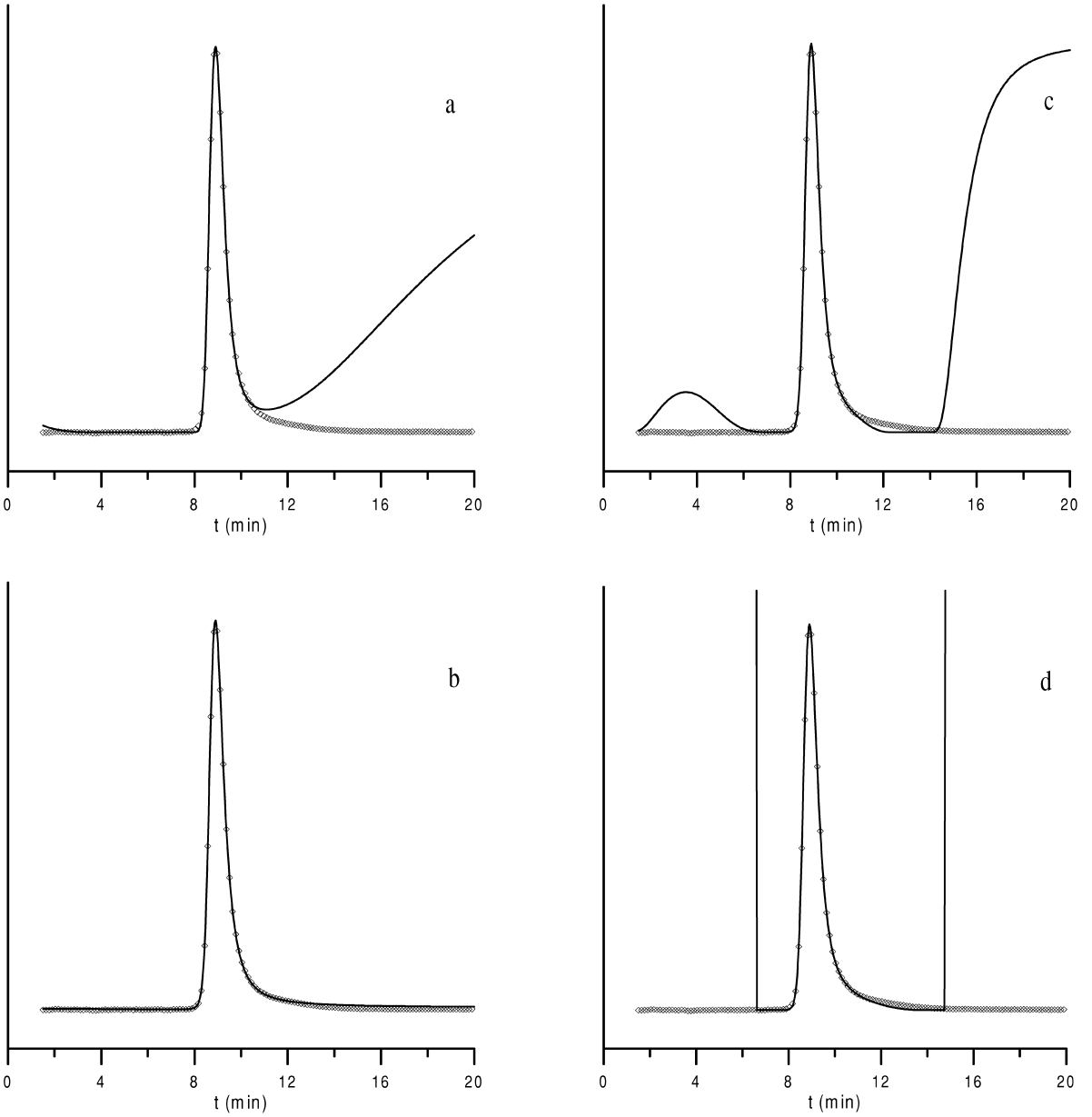

Fig. 3. OTC peak fitted to different models: (a) PSMG, (b) PVMG, (c) FSMG, and (d) FVMG. Points correspond to the experimental data (only 25% of the points are plotted) and the solid line is the modelled peak.

baseline is not recovered. The performance of the corresponding PMG model (Fig. 3a) is however improved. On the other hand, FSMG and FVMG fit the peak region well, but develop an undesirable growth of the baseline (Fig. 3c,d). Thus, for example, using the FVMG model, the baseline increases abruptly when the variance adopts negative values.

The Li model (Fig. 4c) was unable to describe the large skewness of the OTC peak. The logarithmic functions in Eqs. (10) and (11) do not describe the front of the peaks below a certain time (which depends on the width and asymmetry of the peak),

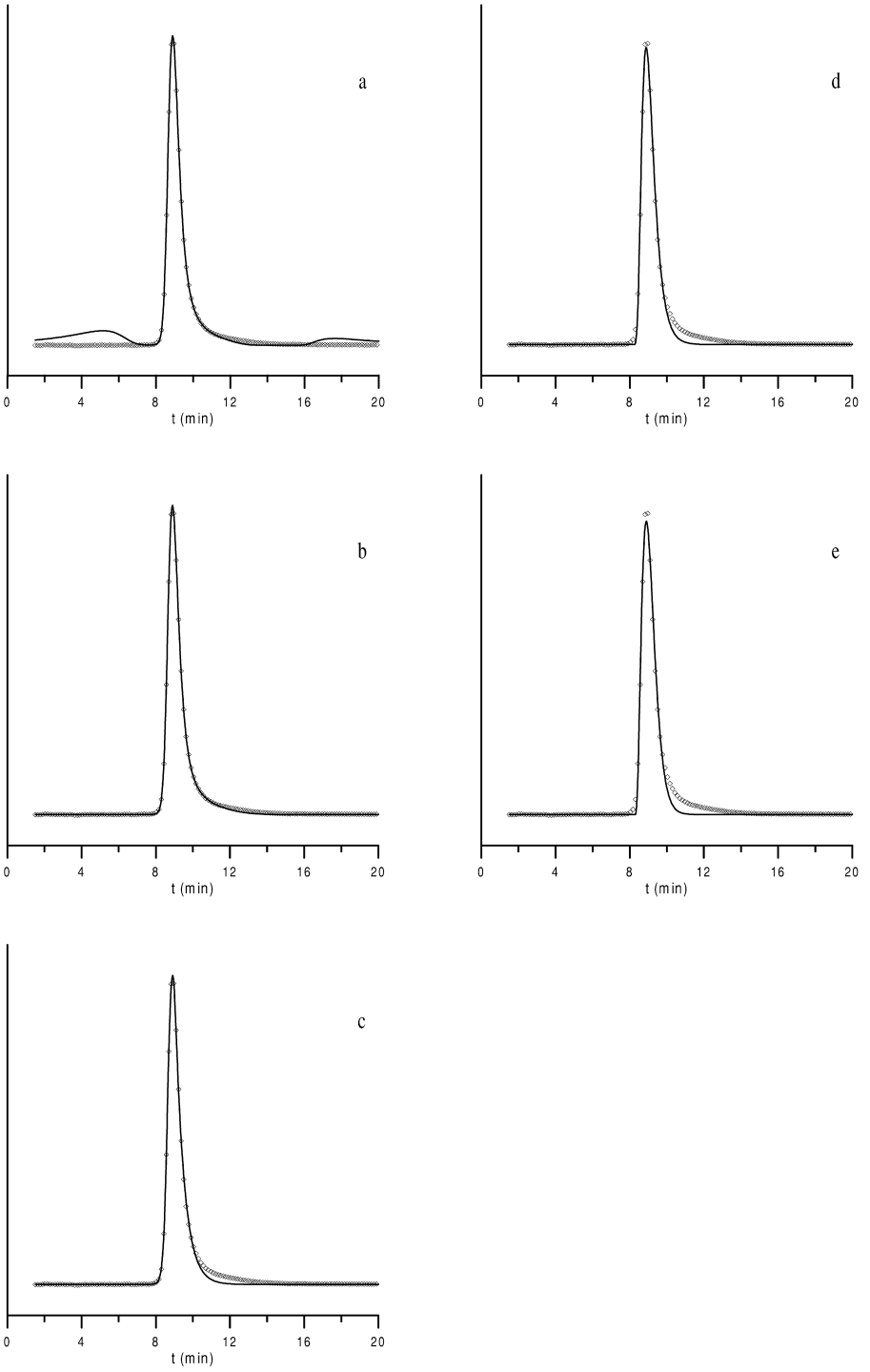

Fig. 4. OTC peak fitted to different models: (a) PMG2, (b) PLMG, (c) Li, (d) log-normal, and (e) Pap–Pápai. See Fig. 3 for details.

since the logarithm argument becomes negative. At these times the signal was set to zero (Fig. 4d,e).

The PLMG model was developed to control the anomalous baseline growth observed in the PMG models, keeping or enhancing their fitting ability. In the new function, the parabola describes the variance profile during peak elution, whereas the Lorentzian function cancels the uncontrolled growth of the baseline. In the limit, at times far from the maximum, the variance tends to a constant:

$$\sigma^2 = \sigma_0^2 + mw^2 \tag{27}$$

and the baseline tends to zero.

Both functions, parabola and Lorentzian, are depicted in Fig. 1 for the OTC peak. Its combination provides a great flexibility, allowing the fitting of peaks in a wide range of shapes, skewness both to the right and/or left, and diverse kurtosis. Fig. 5 illustrates experimental peaks with different asymmetry degrees, fitted with the PLMG model, and Fig. 6 shows peaks simulated with this model. The PLMG model can cope with high positive kurtosis degrees (Fig. 6a), and positive (Fig. 6b), or/and negative skewness (Fig. 6c). Finally, Fig. 2 illustrates the good fitting of the variance to the model.

4.3. Estimation of the PLMG model parameters

The peak model parameters can be related to the peak semiwidths, A and B, to achieve good initial estimates for the fitting procedure. For a pure Gaussian, the semiwidths measured for $p=1$ coincide with σ. Therefore, for the description of skewed peaks with the PLMG model (Eq. (8)), the smaller semiwidth at 60% of peak width (A or B) can be taken as the initial guess for σ_0, since it is less affected by the asymmetry. The initial values for the other model parameters can be obtained from Eqs. (16) and (17), assuming $p=1$. From these equations:

$$d = 2\frac{BA - \sigma_0^2}{B - A} \tag{28}$$

$$m = \frac{1}{1 + 2\dfrac{d}{B - A}} \tag{29}$$

The parameters of the Lorentzian function (r and w) should be taken to cancel the baseline increase. Thus, the Lorentzian maximum will be located within the peak. A good choice for r (difference between t_R and the Lorentzian maximum) is $d+(B-A)/2$, where A and B are measured at 10% of peak height ($p_{10}=4.6$). This leads to:

$$r = d + \frac{mp_{10}}{1 - mp_{10}}d = \frac{d}{1 - mp_{10}} \tag{30}$$

Finally, the Lorentzian width should comprise the whole peak. A good choice is:

$$w = 6\,(r + 3\sigma_0) \tag{31}$$

4.4. Deconvolution of overlapped peaks

4.4.1. Study of artificial samples

The PLMG model was first applied to the deconvolution of artificial overlapped peaks. The peak parameters were the following: $\sigma_{0,1}=0.20$, $\sigma_{0,2}=0.25$; $m_1=0.08$, $m_2=0.09$; $d_1=0.4$, $d_2=0.5$. The Lorentzian function parameters were evaluated using Eqs. (30) and (31): $r_1=0.63$, $r_2=0.85$; $w_1=7.4$, $w_2=9.6$. The simulated signal, S_i, was obtained by adding Gaussian errors to the theoretical points according to:

$$S_i = S_i^o + \varepsilon_i = S_i^o + n_i\sigma_S \tag{32}$$

where ε_i is the Gaussian error, σ_S the standard deviation of the measurements and n_i are numbers within the range -6 to $+6$ that approximately follow the normal distribution law. These numbers were generated by using the central limit theory [26]:

$$n_i = 6 - \sum z_j \tag{33}$$

where the sum is extended between $j=1$ and 12, and z_j are random numbers following a rectangular distribution in the range $0<z_j<1$. S_i^o is the signal without error obtained by addition of two modified Gaussian peaks including the baseline (Eq. (22)).

The peaks were simulated in the time interval 8–13 min (501 points with a step of 0.01 min). The baseline parameters were $\beta_0=\beta_1=0$, and $\sigma_S=0.01$, which gives an error of 2% for a height of 0.5 (peaks were simulated considering heights in the range $0.1<H<1$). Non-linear least-squares regression [27] was applied to deconvolve the simulated peaks according to Eq. (22). The deconvolution errors were

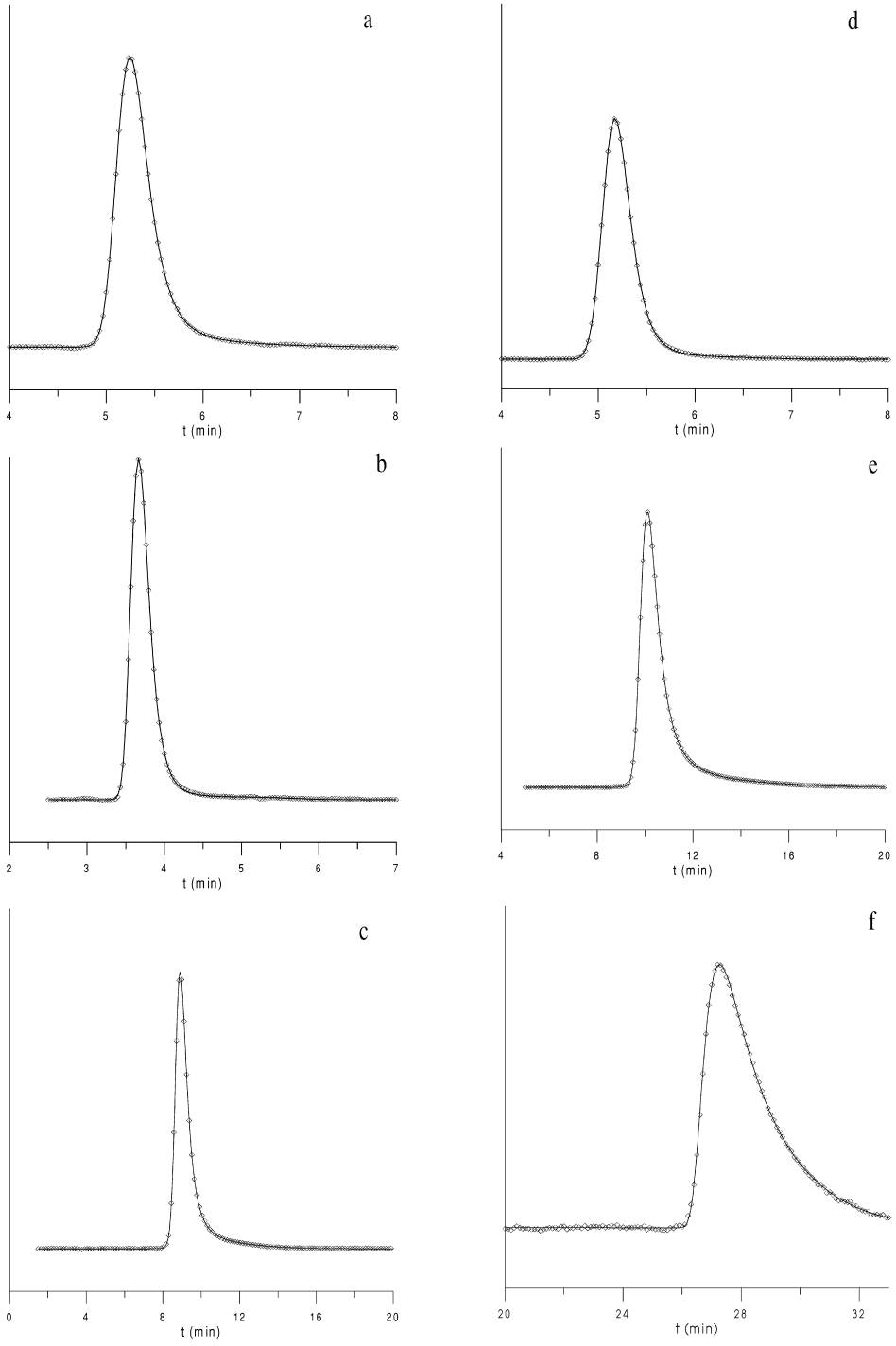

Fig. 5. Fitting of the PLMG model to the experimental peaks of several probe compounds (B/A at 10% of peak height): (a) SFZ (1.65), (b) SCP (1.81), (c) OTC (2.42), (d) SPD (2.75), (e) TC (2.86), and (f) spironolactone (5.60). See Fig. 3 for details.

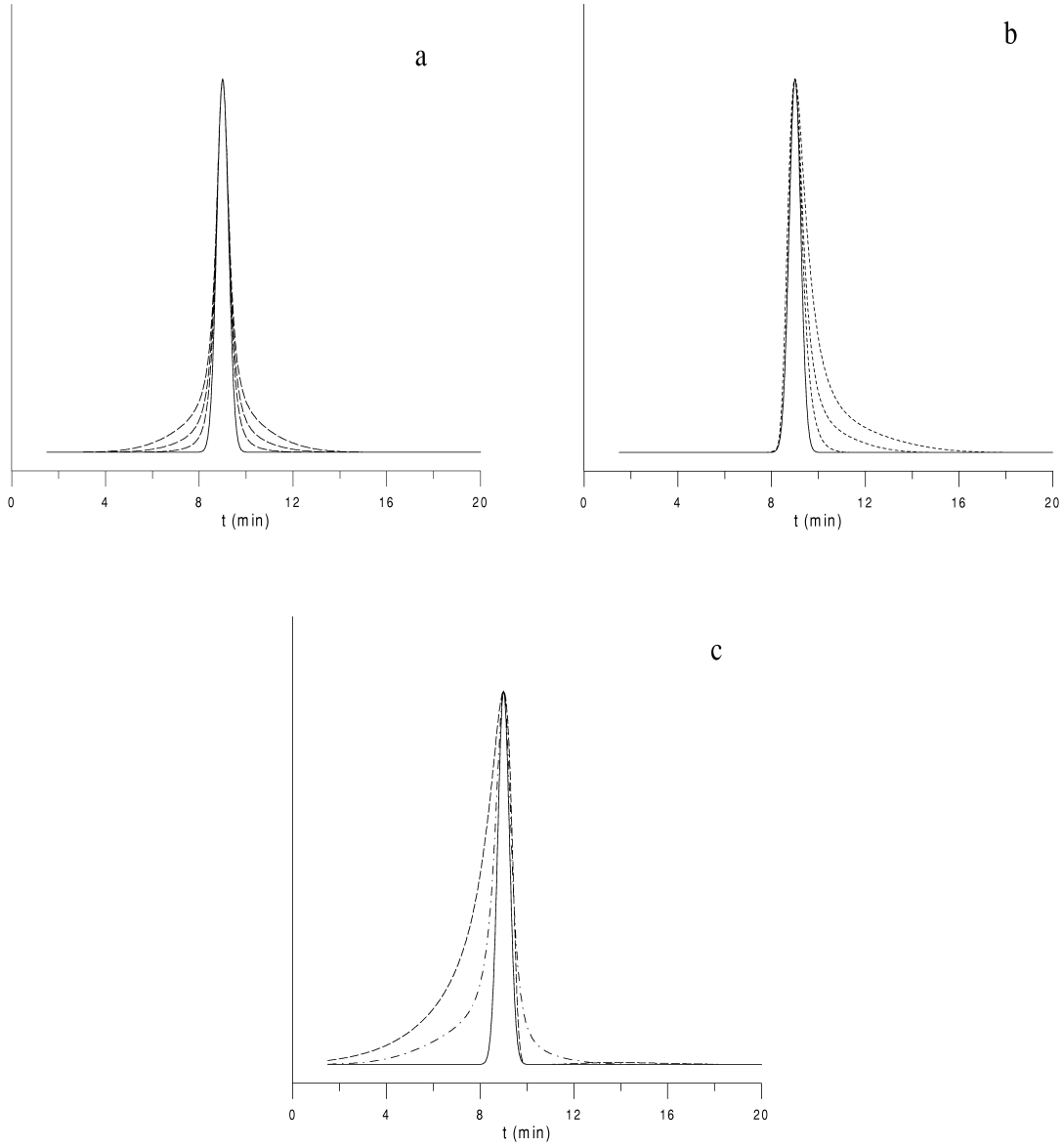

Fig. 6. Use of the PLMG model in the simulation of chromatographic peaks (dotted lines): (a) kurtosis ($c=0.33$, 0.72 and 1.41, according to Eq. (21)), (b) tailing asymmetry ($B/A=1.5$, 2.7 and 4.0), and (c) fronting asymmetry ($B/A=0.16$) or tailing-fronting asymmetry ($B/A=0.49$). The solid line corresponds to a Gaussian peak.

obtained as the mean of 18 simulations to reduce the effect of systematic trends due to the use of a discrete number of points.

The effect of peak overlapping (assuming similar heights) and of the ratio between peak heights (assuming retention times of 9.8 and 10.2 min) was studied. Only the peak heights (proportional to the concentration), and the retention times were fitted, keeping the remaining parameters constant. Table 3 shows the mean relative errors in the found heights. As expected, the quality of the results decreased at greater overlapping or peak height ratio, although the error for the bigger peak was almost constant. It should be indicated that the achieved errors are the

Table 3
Effect of peak separation and height ratio on the deconvolution of two compounds using the PLMG model

Retention time[a] (min)	ε_1 (%)[b]	ε_2 (%)[b]	Peak heights[c]	ε_1 (%)[b]	ε_2 (%)[b]
9.2–10.8	0.098	0.068	1.0–1.0	0.28	0.028
9.4–10.6	0.11	0.094	1.0–0.5	0.30	0.58
9.6–10.4	0.14	0.12	1.0–0.2	0.30	1.4
9.8–10.2	0.28	0.28	1.0–0.1	0.32	3.0
9.9–10.1	0.48	0.36	0.5–1.0	0.60	0.30
10.0–10.0	0.66	0.56	0.2–1.0	1.6	0.30
			0.1–1.0	3.2	0.30

[a] $H_1 = H_2 = 1$.
[b] Mean relative errors in heights ($n = 18$).
[c] $t_{R1} = 9.8$ min and $t_{R2} = 10.2$ min.

minimal expected for these determinations. Run to run peak shape variation in real experiments, produced in the sample preparation or elution, will decrease the accuracy.

The effect of the number of fitted parameters was also considered. In this study, peaks of the same height at 9.8 and 10.2 min were deconvolved. As observed in Table 4, the results were poorer when a greater number of parameters were fitted, due to overfitting.

The PLMG model was compared with the FVMG and FSMG models, using the same artificial example and assuming peaks of equal area. All model parameters were fitted. The mean relative errors were: ε_1 (peak at 9.8 min) = 0.7, 7.0 and 17%, and ε_2 (peak at 10.2 min) = 0.44, 4.8 and 9.4%, respectively. Therefore, the use of peak models giving poorer fittings (in this case, the FVMG and FSMG models) may yield significant systematic errors.

Finally, the effect of systematic errors in the peak model parameters was studied for the PLMG model (Table 5). Only the heights were optimised, keeping the remaining parameters constant. However, an uncertainty of 2% was assigned consecutively to each parameter, except to the retention times for which the error was 0.01 min. It is evident that the most critical parameters are the retention times, which can ruin the deconvolution. To avoid this, they should be always included in the fittings.

Table 4
Effect of the number of fitting parameters (PLMG model) on the deconvolution of two compounds[a]

Parameters	ε_1 (%)[b]	ε_2 (%)[b]
H_1, H_2	0.13	0.14
H_1, H_2, t_{R1}, t_{R2}	0.28	0.28
H_1, H_2, t_{R1}, t_{R2}, baseline (a, b)	0.36	0.32
H_1, H_2, t_{R1}, t_{R2}, baseline, σ_{01}, σ_{02}	0.40	0.28
H_1, H_2, t_{R1}, t_{R2}, baseline, σ_{01}, σ_{02}, m_1, m_2	0.44	0.38
H_1, H_2, t_{R1}, t_{R2}, baseline, σ_{01}, σ_{02}, m_1, m_2, r_1, r_2, w_1, w_2[c]	0.70	0.44

[a] $H_1 = H_2$, $t_{R1} = 9.8$ min, $t_{R2} = 10.2$ min.
[b] Mean relative errors in heights ($n = 18$).
[c] All model parameters were fitted.

Table 5
Importance of the errors in the parameters of the PLMG model on the deconvolution of two compounds[a]

Parameter[b]	ε_1 (%)[c]	ε_2 (%)[c]	Parameter[b]	ε_1 (%)[c]	ε_2 (%)[c]
No error	0.13	0.14	d_1	0.13	0.25
t_{R1}	14	19	d_2	0.14	0.21
t_{R2}	25	14	r_1	0.13	0.16
σ_1	0.25	0.23	r_2	0.13	0.15
σ_2	0.58	0.14	w_1	0.13	0.15
m_1	0.15	0.30	w_2	0.13	0.15
m_2	0.14	0.22	β_0	0.13	0.14
			β_1	0.14	0.15

[a] $H_1 = H_2$, $t_{R1} = 9.8$ min, $t_{R2} = 10.2$ min.
[b] The assigned error was 2%, except for t_{R1} and t_{R2} (0.01 min), β_0 (0.001) and β_1 (0.01).
[c] Mean relative errors in heights ($n = 18$).

4.4.2. Analysis of real samples

Finally, the deconvolution of real binary mixtures of OTC–TC, STZ–SCP and SPD–SFZ was examined. These compounds could not be resolved with the solvent systems used in our laboratory (SDS–butanol and SDS–acetonitrile), and showed different overlapping (Fig. 7). The number of experimental points was 361 for OTC–TC, which were evenly spaced in the range 6–18 min, 136 in the range 2.5–7 min for STZ–SCP, and 121 in the range 4–8 min for SPD–SFZ. Table 6 shows the chromatographic parameters for the six compounds, which

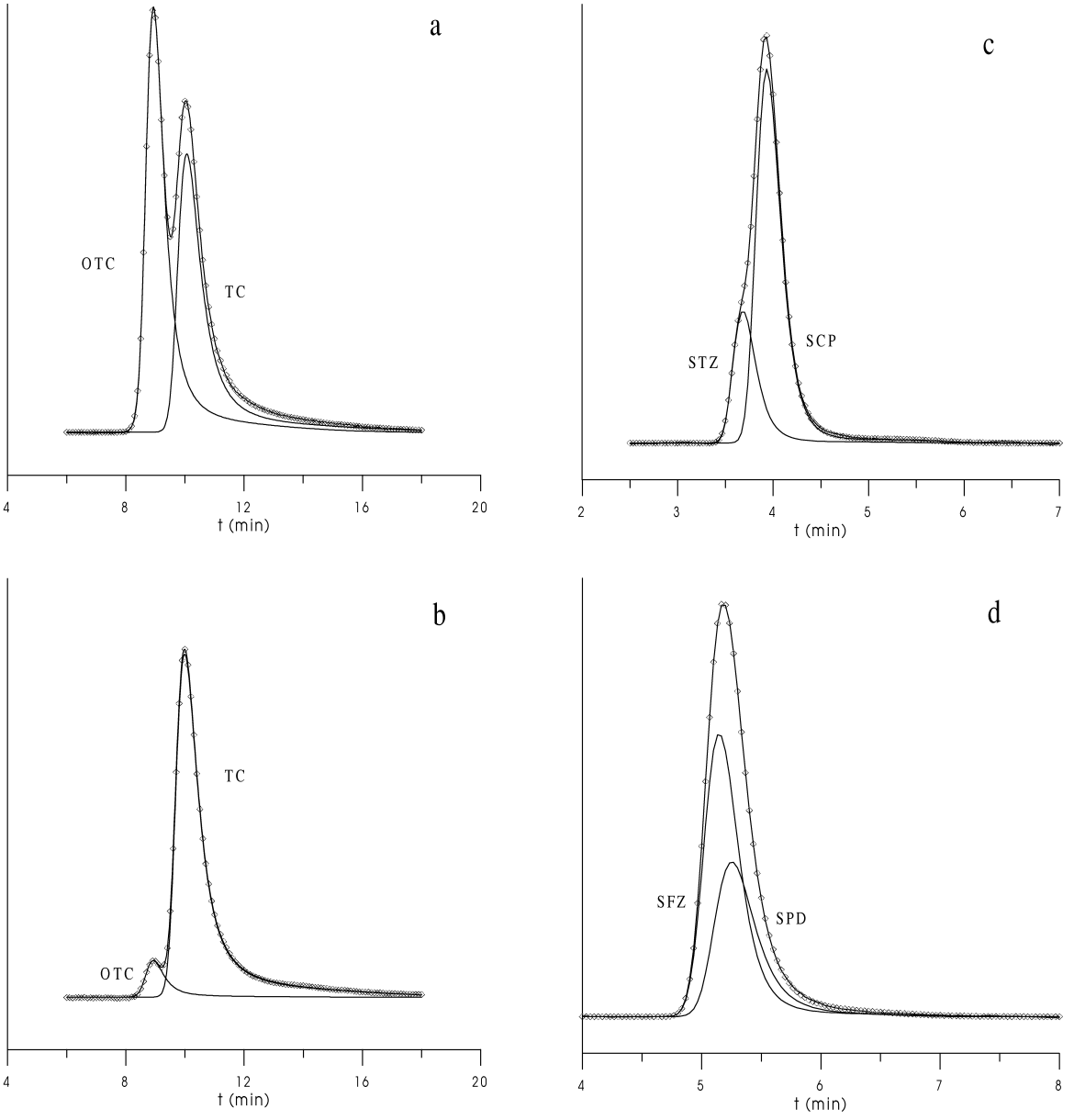

Fig. 7. Deconvolution of binary mixtures of: (a) OTC (3.0 μg/ml)–TC (2.0 μg/ml), (b) OTC (0.5 μg/ml)–TC (5.0 μg/ml), (c) STZ (2 μg/ml)–SCP (5 μg/ml), (d) SFZ (5 μg/ml)–SPD (3 μg/ml).

Table 6
Chromatographic parameters and resolution for the probe compounds

Compound	t_R	N^c	A^d	B^d	B/A	R_s^e
Oxytetracycline[a]	9.00	274	0.52	1.30	2.42	0.57
Tetracycline[a]	10.20	196	0.61	1.74	2.86	
Sulfathiazole[b]	3.67	584	0.14	0.39	2.79	0.46
Sulfachloropyridazine[b]	3.93	630	0.21	0.38	1.81	
Sulfisoxazole[b]	5.14	868	0.26	0.43	1.65	0.14
Sulfapyridine[b]	5.24	529	0.20	0.55	2.75	

[a] Mobile phase: 0.05 *M* SDS–5% 1–butanol at pH 3.0 (0.01 *M* citric acid buffer).
[b] Mobile phase: 0.10 *M* SDS–6% acetonitrile at pH 3.0 (0.01 *M* citric acid buffer).
[c] According to Foley and Dorsey [9].
[d] Measured at 10% of peak height.
[e] Chromatographic resolution measured for similar concentrations according to Eq. (34).

were obtained as the mean values of several experimental peaks. The same values within 1% error were calculated using Eqs. (16) and (17) when a PVMG model was assumed. The resolution was measured as follows:

$$R_s = \frac{2\Delta t_R}{W_1 + W_2} \quad (34)$$

where Δt_R is the difference between the retention times, and W_1 and W_2 the peak widths at 10% height. Table 7 gives the shape parameters obtained by fitting the peaks of the six compounds to the PLMG model.

The following solutions were injected:

(i) OTC–TC: Standard solutions at six concentrations for each compound, 0.5, 1.0, 2.0, 3.0, 4.0 and 5.0 μg/ml, and six solutions with mixtures of OTC–TC at different ratios, 0.5 : 5.0, 1.0 : 4.0, 2.0 : 3.0, 3.0 : 2.0, 4.0 : 1.0 and 5.0 : 0.5. Six replicate injections were done in all cases.

(ii) STZ–SCP: Standard solutions at five concentrations, 0.6, 1.1, 1.7, 2.2 and 2.8 μg/ml for STZ, and 1.0, 2.0, 3.0, 4.0 and 5.0 μg/ml for SCP, and three solutions with mixtures of STZ–SCP at different ratios, 1.7:5.0, 2.2:4.0 and 2.8:3.0 (triplicate injections).

(iii) SPD–SFZ: Standard solutions at five concentrations for each compound, 1.0, 2.0, 3.0, 4.0 and 5.0 μg/ml, and three solutions with mixtures of SPD–SFZ at different ratios, 5.0:3.0, 4.0:4.0 and 3.0:5.0 (triplicate injections).

The performance of method 1 to deconvolve real samples was compared with two other different strategies: a linear fitting of the experimental points at different times (method 2), and a partial least-squares regression (method 3). In all cases, the baseline was subtracted from the experimental signal previously to the application of the methods. In Method 1, the peak parameters were achieved by injection of individual standards for each compound. The mean parameters were adopted for the deconvolution, except for the height and retention time.

The relative errors obtained in the deconvolution of several probe mixtures are given in Table 8 for

Table 7
Parameters of the PLMG model (Eq. (8)) for the probe compounds

Compound	σ_0	m	d	r	w
Oxytetracycline	0.245	0.0890	0.555	1.93	4.23
Tetracycline	0.270	0.0932	0.770	2.35	7.66
Sulfathiazole	0.0790	0.0794	0.629	2.05	1.10
Sulfachlorpyridazine	0.0904	0.0729	0.610	2.22	1.31
Sulfisoxazole	0.115	0.0568	0.556	1.65	1.18
Sulfapyridine	0.125	0.0576	0.533	1.80	2.09

Table 8
Mean relative errors (%), referred to the component concentrations, obtained in the analysis of binary mixtures at different concentration ratios[a]

Mixture	Method[b]					Method[c]				
	1a	1b	2	3a	3b	1a	1b	2	3a	3b
OTC:TC										
0.5:5.0	9.3	4.2	27	21	18	1.7	0.5	3.1	3.1	1.4
1.0:4.0	7.6	3.5	9.2	8.7	5.6	0.9	0.5	1.2	1.3	0.9
2.0:3.0	1.2	1.1	2.8	4.1	3.7	3.3	2.4	4.2	4.5	3.2
3.0:2.0	1.3	0.8	1.9	1.6	1.4	1	1.1	2.6	2.4	1.8
4.0:1.0	1.7	1	2.6	2.6	1.7	4.4	2.6	7.4	6.3	5.1
5.0:0.5	2.5	2.3	3.1	3.2	2.6	12	6.3	15	14	7.8
STZ:SCP										
2.2:0.0	0.3	0.1	0.5	0.8	0.3	1.1	0.9	8.9	6.4	3.6
2.8:3.0	0.5	0.06	1.8	2.2	0.1	1.1	0.4	2.9	2.3	0.5
2.2:4.0	0.3	0.3	3.7	3.8	1.4	1.7	0.03	2.2	1.5	0.5
1.3:5.0	1.4	0.3	7.6	7.4	2.0	1.9	0.1	0.4	0.7	0.6
SPD:SFZ										
3.0:0.0	3.3	1.9	6.0	1.5	1.1	10	3.6	11	2.8	3.0
5.0:3.0	4.0	1.6	8.3	6.1	3.9	5.5	2.5	17	9.2	4.6
4.0:4.0	7.6	4.2	20	15	9.1	7.7	3.7	12	17	8.1
3.0:5.0	3.7	3.2	19	19	13	2.8	1.9	8.5	12	6.3

[a] Six replicates were made for OTC–TC, and three for STZ–SCP and SPD–SFZ. See text for the description of the methods.
[b] The errors correspond to the determination of OTC, STZ and SPD.
[c] The errors correspond to the determination of TC, SCP and SFZ.

different concentration ratios. The deconvolution of some peaks is shown in Fig. 7. MLR and PLS applied using individual standards gave similar results. The results of PLS were improved by the addition of mixtures to the calibration set. However, the use of a deconvolution method that includes an accurate description of the peaks lead to the best results. In this case, the global mean errors obtained in the deconvolution of the mixtures of OTC–TC, STZ–SCP and SPD–SFZ, using only the individual standards or the individual standards and mixtures were 3.5 and 1.8%, respectively, whereas the errors for PLS were 6.4 and 3.9%. Finally, the global mean error for MLR was 7.6%.

Consequently, the PLMG model applied to the modelling of chromatographic peaks permits the deconvolution of highly skewed overlapped peaks, for compounds at different concentration ratios. The results are better than those obtained using linear least-squares and partial least-squares regression. The proposed method takes into account the run to run changes in retention time that occur along the injection of standards and samples, and the possible interactions that exist between the coeluting compounds. This decreases significantly the deconvolution errors.

Acknowledgements

This work was supported by Project BQU2001-3047 (MCYT, Spain). R.D.C. is on leave from Universidad Nacional de Entre Ríos, Gualeguaychú (Argentina), and thanks the AECI of Spain for a research grant.

References

[1] R.D. Fraser, E. Suzuki, Anal. Chem. 38 (1966) 1770.
[2] R.D. Fraser, E. Suzuki, Anal. Chem. 41 (1969) 37.
[3] A.G. Stromberg, S.V. Romanenko, E.S. Romanenko, J. Anal. Chem. 55 (2000) 615.
[4] P. Nikitas, A. Pappa-Louisi, A. Papageorgiou, J. Chromatogr. A 912 (2001) 13.

[5] P.C. Haarhoff, H.J. van der Linde, Anal. Chem. 38 (1966) 573.
[6] T.S. Buys, K. de Clerk, Anal. Chem. 44 (1972) 1273.
[7] S. Chesler, S.P. Cram, Anal. Chem. 45 (1973) 1345.
[8] F. Dondi, A. Betti, G. Blo, C. Bighi, Anal. Chem. 53 (1981) 496.
[9] J.P. Foley, J.G. Dorsey, Anal. Chem. 55 (1983) 730.
[10] J.P. Foley, J.G. Dorsey, J. Chromatogr. Sci. 22 (1984) 40.
[11] D. Hanggi, P.W. Carr, Anal. Chem. 57 (1985) 2394.
[12] A. Berthod, Anal. Chem. 63 (1991) 1879.
[13] M.S. Jeansonne, J.P. Foley, J. Chromatogr. Sci. 29 (1991) 258.
[14] P.J. Naish, S. Hartwell, Chromatographia 26 (1988) 285.
[15] J. Olivé, J.O. Grimalt, J. Chromatogr. Sci. 29 (1991) 70.
[16] J.R. Torres-Lapasió, R.M. Villanueva-Camañas, J.M. Sanchis-Mallols, M.J. Medina-Hernández, M.C. García-Alvarez-Coque, J. Chromatogr. A 677 (1994) 239.
[17] J. Li, H.L. Pardue, Anal. Chem. 66 (1994) 3765.
[18] S. Le-Vent, Anal. Chim. Acta 312 (1995) 263.
[19] P. Papoff, A. Ceccarini, F. Lanza, N. Fanelli, J. Chromatogr. A 789 (1997) 51.
[20] K. Lan, J.W. Jorgenson, J. Chromatogr. A 915 (2001) 1.
[21] T.L. Pap, Zs. Pápai, J. Chromatogr. A 930 (2001) 53.
[22] J.R. Torres-Lapasió, J.J. Baeza-Baeza, M.C. García-Alvarez-Coque, Anal. Chem. 69 (1997) 3822.
[23] J.R. Torres-Lapasió, MICHROM Software, Marcel Dekker, New York, 2000.
[24] J. Li, Anal. Chem. 69 (1997) 4452.
[25] The Unscrambler 7.6, Camo, Oslo, 2000.
[26] T.H. Naylor, J.L. Bailintfy, D.S. Burdick, K. Chu, Computer Simulation Techniques, Wiley, New York, 1966.
[27] S.S. Rao, Engineering Optimization, Wiley, New York, 1996.