ELSEVIER

Journal of Chromatography A, 727 (1996) 27–38

JOURNAL OF CHROMATOGRAPHY A

Partial least-squares and principal component regression of multi-analyte high-performance liquid chromatography with diode-array detection[1]

A. Garrido Frenich, M. Martínez Galera, J.L. Martínez Vidal*, M.D. Gil García

Department of Hydrogeology and Analytical Chemistry, Faculty of Experimental Sciences, University of Almería, 04071 Almería, Spain

First received 17 May 1995; revised manuscript received 3 October 1995; accepted 6 October 1995

Abstract

Three multivariate calibration methods, partial least-squares (PLS-1 and PLS-2) and principal component regression (PCR), were applied to the simultaneous determination of cypermethrin, fenvalerate and *cis*- and *trans*-permethrin by HPLC. Several preprocessing algorithms for pretreatment of data in chemometric approaches are discussed. Mean centering and the selection of the chromatogram region to realize the calibration were found to be advantageous.

Keywords: Principal component analysis; Chemometrics; Computer-assisted chromatography; Partial least-squares regression; Multivariate calibration; Mobile-phase composition; Calibration methods; Pesticides; Pyrethroids; Cypermethrin; Fenvalerate; Permethrin

1. Introduction

The separation of overlapped peaks is essential in every area of separation science, including chromatography and electrophoresis. The degree of peak separation will strongly influence the reliability of an analytical method [1]. There are several publications concerned with the evaluation of the accuracy of analytical methods in chromatographic situations where the peaks are overlapped [2–5].

Workers in chromatography are confronted with the occurrence of partially resolved peaks due to co-elution of solutes in the sample or to similarity between their retention times. Further, in complex sample matrices such as biological, agricultural or natural products, it may not be possible to avoid overlapping peaks. This problem becomes more important when the analyte is at a concentration level near the detection limit. Traditionally, this type of problem has been solved by modifying the experimental conditions by trial and error until the aforesaid errors are minimized. Thus, different mobile or stationary

[1] Paper presented at the Annual Meeting of the Spanish Chromatography Group, 7-as Jornadas de Análisis Instrumental, Madrid, April 3–6, 1995.

* Corresponding author.

0021-9673/96/$15.00 © 1996 Elsevier Science B.V. All rights reserved
SSDI 0021-9673(95)01150-1

phases (columns) or even working techniques (e.g., isocratic, gradient elution) are tested, which are time-consuming and involve the consumption of expensive solvents.

The advent of multi-dimensional detection systems and the affordability of personal computers provided with software that allows storage and subsequent processing of chromatographic data have fostered developments of new experimental procedures for the characterization of unresolved peaks. There are different alternatives for computer-assisted resolution of chromatographic peaks, such as the following. (a) Techniques for fitting the chromatographic peak to known functions. These are perhaps the most accurate chromatographic curve deconvolution methods that do not require calibration standards. Like every other technique, however, curve fitting relies on some basic assumptions: the number of peaks present in the overlapped group must be known (peak purity) and good estimates of the true peak-shape parameters are required. Curve fitting is performed by mathematical regression techniques; modelled peaks, shapes or curves are fitted to the overlapped peaks using non-linear least-squares. Solutions involving comparison of logarithmic spectra [6] or the use of chemometric methods have been proposed in liquid chromatography (LC) with diode-array detection (DAD) [7–11]. (b) Integration by tracing a line perpendicular to the baseline from the valley between two peaks, or one line joining the valley and the end of the second peak (skimming) by computing the area of each one separately from the two zones thus established. (c) Use of derivative techniques. They are characterized by reduced band widths, which implies a potential improvement in resolution. The fundamental properties, limitations and potential applications of time-domain derivatives in LC have been evaluated [12]. This procedure has been applied to the resolution of several mixtures [12–16].

Other procedures have been also described. Thus, Campins Falcó et al. [17] developed a procedure named the "H-point standard additions method" in which the analytical signal absorbance values (peak height) registered at the retention time of the analyte were used. Cladera et al. [18] utilized multiple linear regression (MLR) in the resolution and quantification of binary and ternary mixtures of phenol compounds, which presented overlapped signals in HPLC. Hayashi et al. [19] proposed a one-dimensional Kalman filter to resolve partially overlapped chromatographic peaks using a one-dimensional empirical model based on prior measurements of peak shape and location. Later, Hayashi and Rutan [20] examined the accuracy and precision of the adaptive Kalman filter using computer simulations of chromatographic situations, where a known peak overlaps with an unknown (interferent) peak. A Kalman filter based on repetitive filtering of diode-array spectra obtained across a chromatogram has also been developed [21]. Principal component regression (PCR) and partial least-squares regression (PLS) were applied to multivartiate analysis of overlapped peaks in gas chromatography [22].

In the present work, PLS and PCR multivariate calibration methods were applied to resolve highly overlapped chromatographic peaks. Both are examples of indirect calibration methods, i.e. they do not require individual spectra/chromatograms of each analyte and interferent to be known in advance, but all expected phenomena must be spanned in the calibration set. They offer full spectrum advantages. Each method needs a calibration step where the relationship between the spectra/chromatograms and the component concentrations is deduced from a set of reference samples, followed by a prediction step in which the results of the calibration are used to determine the component concentrations from the sample spectrum/chromatogram.

PLS and PCR are based on the regression of chemical concentrations on latent variables or factors. PLS differs from PCR in that it uses the concentration data from the training set and the spectral data in modelling, whereas PCR only uses the spectral data. Hence PLS can reduce the influence of dominant but irrelevant factors, and in some cases yields models of lower dimensionality, in order to achieve better correlations with concentrations during prediction. PLS also

has the advantage of being able to model a number of analytes simultaneously, the so-called PLS-2 approach. These chemometric techniques have been discussed in more detail elsewhere [23–28].

Multivariate calibration methods, and specifically PLS and PCR, have been found useful for quantitative analysis in various spectroscopic techniques. For example, they have been applied to UV–Visible absorption spectra [29–38], fluorescence spectra [39–42] and infrared spectra [43–50]. Also, they have been successfully applied to flow injection data [51–57] and polarographic data [58]. The great advantage of multicomponent analysis using multivariate calibration is the speed of the method of determination of the components of interest in a mixture, because of a separation step can be avoided.

Pyrethroid insecticides are very effective compounds against agricultural pests with short lifetimes and relatively low mammalian toxicity [59,60]. HPLC has been used for the analysis of pyrethroids [61] and also the separation of pyrethroid enantiomers [62,63]. This paper describes the development of a combined HPLC–DAD system and direct data treatment using PLS and PCR for simultaneous multi-analyte determination of the components of a mixture of cypermethrin, fenvalerate and *cis*- and *trans*-permethrin. The implications of a number of preprocessing techniques and calibration criteria on the training data are presented in detail

2. Experimental

2.1. Instrumentation

A Waters (Milford, MA, USA) Model 990 liquid chromatographic system, equipped with a Model 600E constant-flow pump, a Rheodyne six-port injection valve with a 20-μl sample loop and a Model 990 photodiode-array detector, was used. The detector was interfaced with an Olivetti PCS-386 personal computer using Waters Model 991 software and a Waters Model 990 plotter. The absorbance (A), wavelength (λ), and time (t) were digitized using the Waters Model 991 software, which allows representation and storage of absorption spectra obtained at preset times. An IBM 486-DX microcomputer, provided with a Grams/386 software package and PLSplus V2.1G [64], was used for treatment of data. A conversion program written in Array-Basic with the object of transferring the files obtained with the Waters Model 991 software to an ASCII XY format, which allows the manipulation of these files with the Grams/386 software, was used.

The chromatographic separations were performed using a Hypersil C_{18} column (15×0.46 cm I.D.; 5 μm particle size).

2.2. Chemicals

Analytical standards of cypermethrin and fenvalerate (99%) were obtained from Riedel-de Haën (Seelze, Germany) and permethrin (24.6% *cis* and 73.4% *trans*) was supplied by Dr. Ehrenstorfer (Augsburg, Germany). Standard solutions of these compounds were prepared by dissolving the appropriate amounts in acetonitrile (ACN). HPLC-grade glacial acetic acid and analytical-reagent grade ACN and methanol (MeOH) obtained from Riedel-de Haën were also used. Milli-Q water, obtained from a Millipore (Bedford, MA, USA) Milli-Q filtration/purification system, was used.

2.3. Procedure

A calibration matrix for cypermethrin, fenvalerate and *cis–trans*-permethrin using a fifteen-sample set in the range 0–10 μg ml^{-1} was performed. Volumes of 20 μl were injected onto the chromatographic system and the chromatographic separations were performed on a C_{18} column with a mobile phase ACN–water (85:15, v/v) at a flow-rate of 1.5 ml min^{-1}. The solvents were filtered daily through a 0.45-μm cellulose acetate (water) or polytetrafluoroethylene (ACN and MeOH) membrane filter, and degassed with helium during and before use. A mean-centering pretreatment of data was applied. The optimized calibration matrices, in the chromatographic region between 150 and 210 s, calculated by appli-

cation of PLS and PCR methods, were used to determine cypermethrin, fenvalerate and *cis–trans*-permethrin in the prediction set.

3. Results and discussion

3.1. Optimization of the mobile phase

The analytical separation of cypermethrin, fenvalerate and *cis–trans*-permethrin using a Hypersil C_{18} column (15×0.46 cm I.D.) was investigated. Aqueous MeOH or aqueous ACN were initially tested as mobile phases for performing the separation. The three pesticides are easily eluted by both aqueous MeOH and aqueous ACN mobile phases, the retention times being slightly shorter when ACN is used as organic modifier. Table 1 summarizes the retention times (t_R), capacity factors (k) and resolution values (R) for mobile phases with different percentages of ACN. The R values are <1.

Fig. 1 shows the chromatograms corresponding to cypermethrin using ACN–water (85:15 and 75:25, v/v) as the mobile phase. Cypermethrin shows two peaks when the percentage of water in the mobile phase is >20%, because of an isomerization process. Analogous behaviour has been explained by other workers [65].

Fig. 2 shows the chromatograms corresponding to permethrin using ACN–water (95:5 and 85:15, v/v) as the mobile phase. It can be seen that with a higher percentage of organic modifier in the mobile phase, permethrin shows only one peak. On the other hand, the composition of the mobile phase did not affect to the peak corresponding to fenvalerate. The behaviour of the three pesticides did not change significantly when the mobile phase was MeOH–water.

No differences in the behaviour of these pesticides were observed on changing the pH of the mobile phase with 0.1 *M* acetate buffer (pH 5). In the light of these results, we chose ACN–water (85:15, v/v) as the mobile phase in order to achieve the partial separation of *cis*- and *trans*-permethrin.

The mobile phase flow-rate did not significantly affect the peaks width or the analytical signal (peak area or peak height) in the range tested (1–2 ml min^{-1}). A flow-rate of 1.5 ml min^{-1} was selected.

A typical 3D spectrochromatogram obtained with the optimized chromatographic system is shown in Fig. 3, representing 210 absorption spectra measured at 1-s intervals over the period 0–210 s. Because of the highly overlapping peaks, conventional measures of the different analytical signals (area or height of chromatographic peaks) cannot be realized. With the aim of resolving the ternary mixture, several different chemometric approaches were evaluated.

3.2. Calibration

A training set of fifteen samples (C1–C15) and an independent prediction set were taken; the

Table 1
Influence of ACN in the mobile phase on the retention times (t_R), capacity factors (k) and resolution (R) values[a] for cypermethrin, fenvalerate and *cis–trans*-permethrin (flow-rate 1.5 ml min^{-1})

ACN (%)	Cypermethrin		Fenvalerate		*trans*-Permethrin		*cis*-Permethrin		$R_{ci,fe}$	$R_{fe,t\text{-}pe}$	$R_{t\text{-}pe,c\text{-}pe}$
	t_R (min)	k	t_R (min)	k	t_R (min)	k	t_R (min)	k			
75	4.88	6.18	5.42	6.97	5.69	7.37	6.41	6.20	0.90	0.50	0.88
80	3.46	4.16	3.76	4.61	4.03	5.01	4.48	3.48	0.67	0.60	0.85
85	2.60	3.00	2.76	3.25	2.99	3.60	3.30	1.87	0.50	0.67	0.80
90	1.99	1.34	2.09	2.22	2.26	2.48	2.35	1.12	0.33	0.62	0.26
95	1.65	1.32	1.69	1.38	1.85	1.61	–	–	0.22	0.91	–

[a] $R_{ci,fe}$ = resolution cypermethrin–fenvalerate; $R_{fe,t\text{-}pe}$ = resolution fenvelerate–*trans*-permethrin; $R_{t\text{-}pe,c\text{-}pe}$ = resolution *trans*-permethrin–*cis*-permethrin.

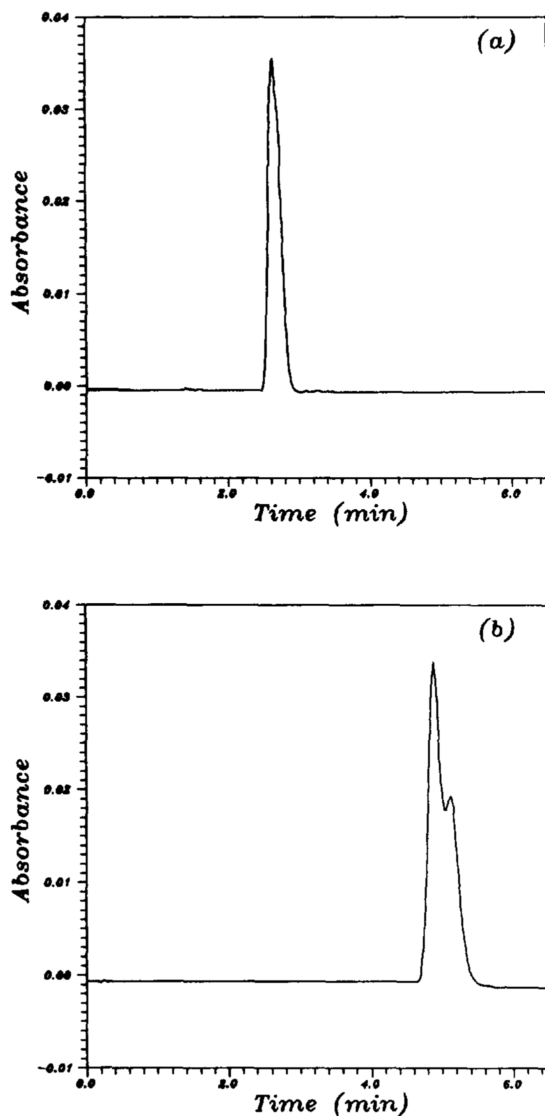

Fig. 1. Chromatogram of a sample containing 6 μg ml^{-1} of cypermethrin. Flow-rate, 1.5 ml min^{-1}. Mobile phase: (a) ACN–water (85:15, v/v); (b) ACN–water (75:25, v/v).

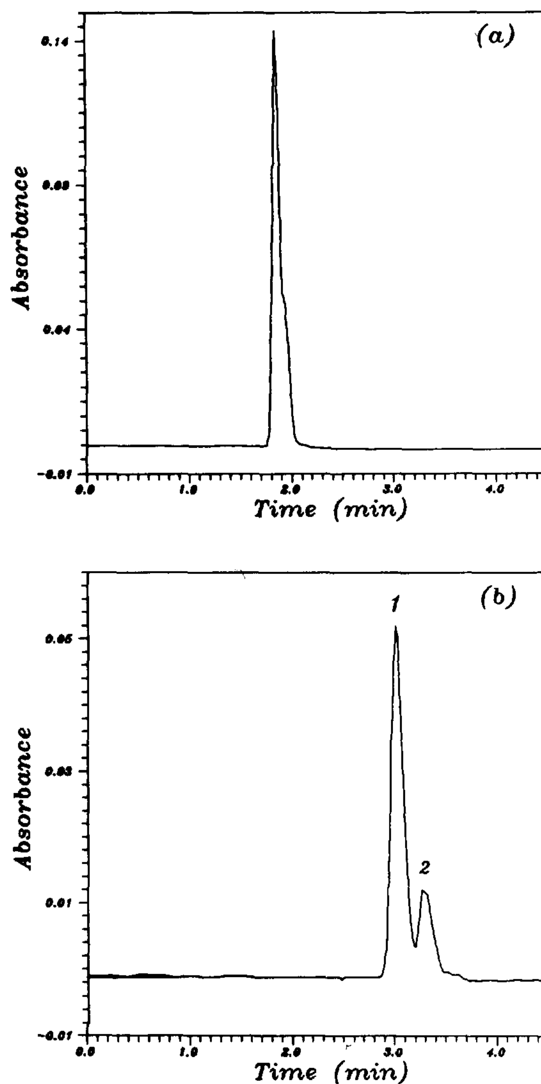

Fig. 2. Chromatogram of a sample containing 6 μg ml^{-1} of total permethrin. Flow-rate, 1.5 ml min^{-1}. Mobile phase: (a) ACN–water (95:5, v/v); (b) ACN–water (85:15, v/v). Peaks: 1 = *trans*-permethrin; 2 = *cis*-permethrin.

concentrations are given in Tables 2 and 3, respectively.

The optimum dimensionality of the PCR and PLS methods was selected as that with the fewest number of factors such that the PRESS (prediction error sum of squares) is not significantly greater than the PRESS from the model that yields a minimum PRESS. The F statistic was used to make the significance determination. Empirically was determined that an F-ratio probability of 0.75 is a good choice [26].

The PRESS is defined as

$$\mathrm{PRESS} = \sum_{i=1}^{N} (\hat{x}_i - x_i)^2$$

where N is the number of samples, x_i is the true concentration of sample i and $\hat{x}_i$ is the predicted

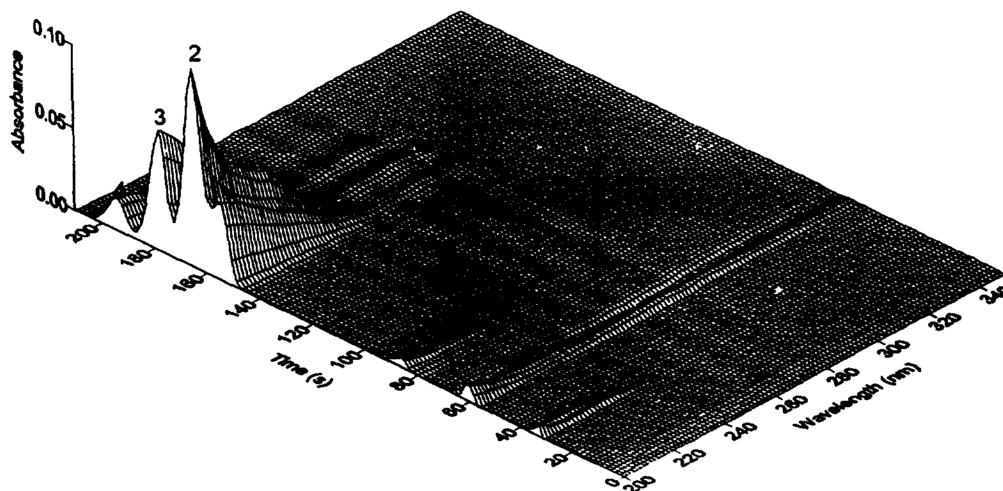

Fig. 3. Three-dimensional spectrochromatogram of a sample of (1) 6 μg ml^{-1} of cypermethrin, (2) 6 μg ml^{-1} of fenvalerate, (3) 4.5 μg ml^{-1} of *trans*-permethrin and (4) 1.5 μg ml^{-1} of *cis*-permethrin.

concentration of sample *i*. The PRESS was calculated in all cases using a cross-validation method, leaving out one sample at a time, in order to model the system without overfitting the concentration data [28,66]; thus the concentration of the sample left out was predicted using the $N-1$ model for all N samples. The prediction ability of the methods for each analyte is expressed in terms of the root mean square difference (RMSD):

Table 2
Concentration data for the calibration set

Standard	Cypermethrin (μg ml^{-1})	Fenvalerate (μg ml^{-1})	*trans*-Permethrin (μg ml^{-1})	*cis*-Permethrin (μg ml^{-1})
C1	0.00	5.00	3.75	1.25
C2	5.00	0.00	3.75	1.25
C3	4.00	8.00	4.50	1.50
C4	6.00	4.00	6.00	2.00
C5	6.00	8.00	3.00	1.00
C6	8.00	4.00	4.50	1.50
C7	8.00	6.00	3.00	1.00
C8	6.00	6.00	4.50	1.50
C9	2.10	10.00	3.00	1.00
C10	4.00	1.70	7.50	2.51
C11	5.00	5.00	0.00	0.00
C12	4.00	6.00	6.00	2.00
C13	5.80	3.00	1.80	0.60
C14	4.00	5.00	1.50	0.50
C15	10.00	4.00	1.50	0.50
C16	3.20	2.60	1.50	0.50
C17	3.50	3.10	2.40	0.80
C18	9.50	5.00	3.30	1.10
C19	8.00	5.50	4.00	1.33
C20	5.00	2.00	1.20	0.40

Table 3
Concentration data for the prediction set

Test no.	Cypermethrin (μg ml^{-1})	Fenvalerate (μg ml^{-1})	*trans*-Permethrin (μg ml^{-1})	*cis*-Permethrin (μg ml^{-1})
T1	8.0	4.0	3.0	1.0
T2	3.0	4.0	3.5	1.2
T3	2.0	10.0	3.0	1.0
T4	4.0	2.0	7.5	2.5
T5	2.0	3.0	4.5	1.5
T6	4.0	4.5	1.5	0.5
T7	4.0	6.0	6.0	2.0
T8	4.0	8.0	4.5	1.5
T9	6.0	4.0	6.0	2.0
T10	6.0	8.0	3.0	1.0
T11	8.0	4.0	4.5	1.5
T12	8.0	6.0	3.0	1.0

$$\text{RMSD} = \left[\frac{1}{N} \sum_{i=1}^{N} (\hat{x}_i - x_i)^2 \right]^{0.5}$$

In the process of PLS-1 modelling, the covariance between the spectral scores and a single analyte is maximized. This often leads to the loadings of the first PLS-1 factor approximating the pure component spectrum of the analyte under examination. The PLS-2, however, maximizes the covariance between the spectral scores and a linear combination of a number of variables. In our case three variables are considered, although four components are present, because the concentration of the *cis*-permethrin component in the mixture varies as a constant function of that of *trans*-permethrin; this is known as collinearity. When this situation occurs, the spectral decomposition cannot determine any difference between the two components across the training set and hence models them as a single factor. A simple visual aid to identifying this potential problem is to plot one component concentration versus another. If the samples fall on a straight line, the concentrations are linearly correlated, although the ideal case would be a symmetrical shape.

Although PCR and PLS are linear methods, in a real spectroscopic or chromatographic application there may be sources of non-linearities, e.g., chemical interactions or non-linear responses in the detector at certain wavelengths. If non-linearities are present, they may be modelled by the inclusion of extra latent variables (factors) in the regression model [28,67] and this could explain the need of the four factors to describe a three-component system. Nevertheless, some non-linearities may be corrected by external methods (transformation of the data, limiting the span of the regression model) while there are non-linearities that are not compensated. To solve this problem, different algorithms of non-linear expansions of PLS regression have been described [68–70] in addition to a method based on local modelling in PCR [71–73].

3.3. Preprocessing

Different methods were used for the pretreatment of data to remove effects of variations in instrumental conditions, signal noise, etc. Mean centering involves the subtraction of the variable mean from the individual variable values. This should almost always be used as it provides better mathematical accuracy in calculating the PLS or PCR model. Scaling performs an operation which divides each data point by the data point standard deviation calculated from all the training chromatograms. This tends to amplify regions of small variations relative to regions with large variations and this is generally applicable to systems in which small peaks of interest

are present among much larger peaks. Baseline correction, as the name suggests, performs a linear baseline correction over each defined region, being useful for samples where the baselines are unstable from sample to sample. By smoothing, random noise in the chromatograms is removed and the signal-to-noise ratio (S/N) is increased. Derivative signals can enhance resolution, because the differentiation resolves complex peaks due to signal overlap and eliminates baseline shift, but generally lead to a decrease in S/N with each derivatization. Smoothing and differentiation were done by the convolution algorithm of Savitzky and Golay [74].

The effect of these preprocessing techniques on the RMSD of the calibration matrix for PCR, PLS-2 and PLS-1 is shown in Table 4. Mean centering had a small but beneficial effect on this data set, because it reduces the PLS-1 model dimensionality and RMSD value for cypermethrin. Also, it can be seen that scaling had a detrimental effect on the PLS-1 model dimensionality and on the RMSD values. On the other hand, the baseline correction and the smoothing did not have a beneficial effect on this data set. The smoothing of the data was selected to decrease the amount of noise while maintaining minimal broadening of the peaks. As our data set presents narrow peaks, when the number of smoothing points is increased the signal amplitude decreases slightly and the peaks are broader with a significant loss of resolution. Finally, both the first and second derivatizations had a detrimental overall effect on the RMSD of this data set.

The PCR and PLS-2 models were also built using the preprocessed data and, as expected, resulted in dimensionality and RMSD values very similar to those for PLS-1, in agreement with other workers [32,36,37].

Table 4
Effect of various preprocessing techniques on the relative prediction errors of PLS-1, PLS-2 and PCR models

Model	Pre-processing technique	RMSD[a]			
		Cypermethrin	Fenvalerate	*trans*-Permethrin	*cis*-Permethrin
PLS-1	None	0.10 (5)	0.21 (3)	0.12 (3)	0.04 (3)
	Mean-centering (MC)	0.09 (4)	0.20 (3)	0.12 (3)	0.04 (3)
	MC + baseline correction	0.09 (5)	0.20 (3)	0.12 (3)	0.04 (3)
	MC + scaling	0.78 (5)	1.10 (5)	0.38 (5)	0.13 (5)
	MC + smoothing	0.07 (5)	0.38 (3)	0.14 (3)	0.05 (3)
	MC + ^{1}D	0.12 (4)	0.36 (3)	0.15 (3)	0.06 (3)
	MC + ^{2}D	0.13 (4)	0.36 (3)	0.15 (3)	0.72 (3)
PLS-2	None	0.20 (3)	0.21 (3)	0.12 (3)	0.04 (3)
	Mean-centering (MC)	0.18 (3)	0.21 (3)	0.12 (3)	0.04 (3)
	MC + baseline correction	0.22 (3)	0.20 (3)	0.12 (3)	0.04 (3)
	MC + scaling	0.95 (5)	1.48 (5)	0.45 (5)	0.15 (5)
	MC + smoothing	0.18 (3)	0.21 (3)	0.12 (3)	0.04 (3)
	MC + ^{1}D	0.12 (4)	0.37 (4)	0.14 (4)	0.05 (4)
	MC + ^{2}D	0.13 (4)	0.36 (4)	0.26 (4)	0.74 (4)
PCR	None	0.19 (3)	0.22 (3)	0.13 (3)	0.04 (3)
	Mean-centering (MC)	0.19 (3)	0.21 (3)	0.12 (3)	0.04 (3)
	MC + baseline correction	0.23 (3)	0.20 (3)	0.12 (3)	0.04 (3)
	MC + scaling	0.87 (5)	1.98 (5)	0.54 (5)	0.18 (5)
	MC + smoothing	0.19 (3)	0.21 (3)	0.12 (3)	0.04 (3)
	MC + ^{1}D	0.12 (4)	0.37 (4)	0.14 (4)	0.05 (4)
	MC + ^{2}D	0.13 (4)	0.36 (4)	0.16 (4)	0.74 (4)

[a] The number of factors is given in parentheses.

Table 5
Effect of time interval on the RMSD values by the PLS-2 method

Time interval (s)	No. of factors	RMSD			
		Cypermethrin	Fenvalerate	*trans*-Permethrin	*cis*-Permethrin
0–210	3	0.18	0.21	0.12	0.04
150–210	3	0.18	0.21	0.12	0.04

3.4. Selection of the region for the analysis and averaging

The effect of the size of the chromatographic data matrix on the prediction ability was studied in two ways. First, two regions were selected for analysis, the first between 0 and 210 s, which implies working with 210 experimental points (as the chromatograms are digitalized every 1 s), and the second between 150 and 210 s, which implies working with 60 experimental points. The second region was taken into account because it is the zone with the maximum analytical information from the mixture components of interest. These regions selected from the original data set were used to build PLS-1, PLS-2 and PCR models after mean-centering. In Table 5 are given the RMSD values obtained by PLS-2. It can be seen that the prediction error is not significantly changed. However, the shortest region is selected because by reducing the size of the regions used the amount of memory and time necessary to perform all the calibration calculations are reduced.

In the second case, the data set was reduced by averaging the chromatographic variables before mean-centering. The results obtained by PLS-2 are shown in Table 6. It can be observed that the RMSD values do not change when the data set is averaged to 30 points. When the number of points is reduced to ten, the RMSD values are strongly increased. Similar results were obtained using the PLS-1 and PCR methods.

3.5. Calibration design and prediction

The effect of the size of the calibration set on the RMSD and on the predictions of the independent test set was determined by reducing and increasing the number of samples in the experimental design. In this way, three calibration matrices were built, the first with twenty samples (from C1 to C20), the second with fifteen samples (from C1 to C15) and the third with ten samples (from C1 to C10), whose concentrations are presented in Table 2. The results obtained by the PLS-2 method, given in Table 7, show a small increase in the RMSD as the number of samples in calibration is reduced. Nevertheless, the predictions obtained were poor

Table 6
Effect of time averaging on the RMSD values by the PLS-2 method

No. of seconds averaged[a]	No. of factors	RMSD			
		Cypermethrin	Fenvalerate	*trans*-Permethrin	*cis*-Permethrin
0 (180)	3	0.18	0.21	0.12	0.04
2 (90)	3	0.18	0.21	0.12	0.04
3 (60)	3	0.17	0.21	0.12	0.04
6 (30)	3	0.15	0.19	0.15	0.05
18 (10)	4	0.59	0.82	0.25	0.09

[a] Values in parentheses are the number of data points used.

Table 7
Effect of sample number in the calibration matrix by the PLS-2 method

No. of samples	No. of factors	RMSD			
		Cypermethrin	Fenvalerate	*trans*-Permethrin	*cis*-Permethrin
20	6	0.08	0.15	0.11	0.02
15	3	0.18	0.21	0.12	0.04
10	4	0.19	0.25	0.12	0.04

Table 8
Recoveries of cypermethrin (Cyp), fenvalerate (Fen) and *cis–trans*-permethrin (Per) in the prediction set depending on samples number utilized for the calibration matrix built

Test No.	10 samples (matrix)				15 samples (matrix)				20 samples (matrix)			
	Cyp	Fen	*t*-Per	*c*-Per	Cyp	Fen	*t*-Per	*c*-Per	Cyp	Fen	*t*-Per	*c*-Per
T1	117.1	75.0	53.7	54.0	100.5	104.2	99.3	99.0	100.5	103.5	102.0	99.0
T2	142.7	74.8	59.1	59.5	108.0	108.0	97.1	98.3	107.7	107.3	100.3	99.1
T3	101.0	99.3	98.7	98.0	102.5	100.0	98.3	98.0	103.0	99.8	97.0	98.0
T4	100.3	91.0	99.3	99.6	100.2	95.7	99.8	99.4	100.3	88.0	100.4	100.0
T5	110.0	107.0	106.9	107.3	109.9	108.0	106.2	106.7	110.0	107.7	106.7	106.7
T6	97.3	102.0	109.3	108.0	98.3	106.0	106.0	106.0	98.3	105.8	103.3	106.0
T7	100.0	94.7	97.2	97.5	99.7	97.0	98.7	99.0	98.8	98.2	97.8	98.5
T8	100.0	100.4	99.1	99.3	99.5	101.9	100.9	100.7	989.3	102.1	100.0	100.7
T9	100.5	96.5	100.0	100.0	100.3	96.0	107.0	101.0	105.0	95.3	101.2	101.0
T10	100.7	100.8	101.0	101.0	100.7	100.4	99.7	100.0	101.5	102.3	95.7	97.0
T11	98.8	98.5	100.4	100.7	98.9	97.8	100.0	100.0	99.3	97.0	100.7	107.0
T12	101.0	101.7	97.7	98.8	101.5	102.5	96.7	96.0	101.5	102.2	95.7	97.0

when the size of the calibration set was reduced to ten samples (Table 8). In addition, the use of the fifteen-sample calibration set reduces the time necessary to perform the experimental work and all the calibration calculations.

4. Conclusions

The effect of some preprocessing techniques and the RMSD values of the calibration matrix are similar for the PLS-1, PLS-2 and PCR calibration methods. Mean-centering and the selection of a region of the chromatogram to realize the calibration was found to be advantageous, whereas smoothing, baseline correction and averaging did not have a beneficial effect and differentiation had a detrimental effect. On reducing the number of calibration standards used in modelling, the error of prediction increased and the prediction on some samples was significantly worse.

References

[1] D.L. Massart, S.N. Vandeginste, S.N. Deming, Y. Michotte and L. Kaufman, Chemometrics, Elsevier, Amsterdam, 1988.
[2] A. Westerberg, Anal. Chem., 41 (1969) 1770.
[3] L.R. Snyder, J. Chromatogr. Sci., 10 (1972) 200.
[4] J.P. Foley, J. Chromatogr., 384 (1987) 301.
[5] N. Dyson, Chromatographic Integration Methods, Royal Society of Chemistry, Cambridge, 1990, p. 50.
[6] A.F. Fell, H.P. Scott, R. Gill and A.C. Moffat, J. Chromatogr., 273 (1987) 3.
[7] A.G. Wright, J.C. Barridge and A.F. Fell, Chromatographia, 24 (1987) 533.
[8] E. Sánchez, L.S. Ramos and B.R. Kowalski, J. Chromatogr., 385 (1987) 151.
[9] B.M. Vandeginste, J. Chemometr., 1 (1987) 57.
[10] J.G.D. Marr, B.J. Clark and A.F. Fell, Anal. Proc., 25 (1988) 150.

[11] G.G.R. Seaton, J.G.D. Marr, B.J. Clark and A.F. Fell, Anal. Proc., 23 (1988) 424.
[12] A.A. Fasanmade and A.F. Fell, Anal. Chem., 61 (1989) 720.
[13] A.F. Fell, H.P. Scott, R. Gill and A.C. Moffat, J. Chromatogr., 282 (1983) 123.
[14] A. Muñoz de la Peña, F. Salinas, T. Galeano and A. Guiberteau, Anal. Chim. Acta, 234 (1990) 263.
[15] J.A. Jimena García, J. Giménez Plaza and J.M. Cano Pavón, J. Liq. Chromatogr., 17 (1994) 277.
[16] P. Parrilla, M. Martínez Galera, J.L. Martínez Vidal and A. Garrido Frenich, Analyst, 119 (1994) 2231.
[17] P. Campins Falcó, F. Bosch Reig, R. Herraez-Hernández and A. Sevillano-Cabeza, Anal. Chim. Acta, 268 (1992) 73.
[18] A. Cladera, E. Gómez, J.M. Estela and V. Cerdá, J. Chromatogr. Sci., 30 (1992) 453.
[19] Y. Hayashi, T. Shibazaki and M. Uchiyama, Anal. Chim. Acta, 201 (1987) 185.
[20] Y. Hayashi and S.C. Rutan, Anal. Chim. Acta, 271 (1993) 91.
[21] T. Barker and S.D. Brown, J. Chromatogr., 469 (1987) 77.
[22] J.F. Faigle, R.J. Poppi, I.S. Scarminio and E. Bruns, J. Chromatogr., 539 (1991) 123.
[23] H. Martens and T. Naes, Trends Anal. Chem., 3 (1984) 204.
[24] H. Martens and T. Naes, Trends Anal. Chem., 3 (1984) 266.
[25] K.R. Beebe and B.R. Kowalski, Anal. Chem., 59 (1987) 1007A.
[26] D.M. Haaland and E.V. Thomas, Anal. Chem., 60 (1988) 1193.
[27] E. Sánchez and B.R. Kowalski, J. Chemometr., 2 (1988) 247.
[28] H. Martens and T. Naes, Multivariate Calibration, Wiley, Chichester, 1989.
[29] R. Jones, M.J. Orchad and K. Hall, J. Pharm. Biomed. Anal., 3 (1985) 335.
[30] R. Tauler, A. Izquierdo Ridora and E. Casassas, Anal. Quim., 87 (1991) 571.
[31] W.P. Carey and L.E. Wangen, Chemometr. Intell. Lab. Syst., 10 (1991) 245.
[32] P. MacLaurin, P.J. Worsfold, M. Crane and P. Norman, Anal. Proc., 29 (1992) 65.
[33] A. Espinosa-Mansilla, A. Muñoz de la Peña, F. Salinas and A. Zamoro, Anal. Chim. Acta, 258 (1992) 47.
[34] A. Espinosa-Mansilla, A. Muñoz de la Peña, F. Salinas and M. Martínez Galera, Anal. Chim. Acta, 276 (1993) 141.
[35] A. Espinosa-Mansilla, A. Muñoz de la Peña, F. Salinas and M. Martínez Galera, Fresenius' J. Anal. Chem., 347 (1993) 371.
[36] I. Durán-Merás, A. Muñoz de la Peña, A. Espinosa-Mansilla and F. Salinas, Analyst, 118 (1993) 807.
[37] M. Martínez Galera, J.L. Martínez Vidal, A. Garrido Frenich and P. Parrilla, Analyst, 119 (1994) 1189.
[38] A. Garrido Frenich, M. Martínez Galera, J.L. Martínez Vidal, P. Parrilla and M.D. Gil García, Fresenius' J. Anal. Chem., submitted for publication.
[39] W. Lindberg, J.A. Persson and S. Wold, Anal. Chem., 55 (1983) 643.
[40] R. Jones, T.J. Coomber, J.P. McCormick, A.F. Fell and B.J. Clark, Anal. Proc., 25 (1988) 381.
[41] B. Pedersen and H. Martens, in L. Munck (Editor), Fluorescence Analysis in Foods, Longman, Harlow, 1989, Ch. 13.
[42] A. Munñoz de la Peña, I. Durán-Merás, M.D. Moreno, F. Salinas and M. Martínez Galera, Fresenius' J. Anal. Chem., 348 (1994) 1765.
[43] P.M. Fredericks, J.B. Lee, P.R. Osborn and D.A. Swinkels, J. Appl. Spectrosc., 39 (1985) 303.
[44] P.M. Fredericks, J.B. Lee, P.R. Osborn and D.A. Swinkels, J. Appl. Spectrosc., 39 (1985) 311.
[45] D.M. Haaland, Anal. Chem., 60 (1988) 1208.
[46] E.V. Thomas and D.M. Haaland, Anal. Chem., 62 (1990) 1091.
[47] B.C. Erickson, R.J. Pell and B.R. Kowalski, Talanta, 38 (1991) 1459.
[48] W. Wei, Z. Mo and S. Yao, Anal. Chim. Acta, 251 (1991) 143.
[49] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch and C. de la Pezuela, Talanta, 40 (1993) 1671.
[50] O. Faix and J.H. Böttcher, Holzforschung, 47 (1993) 45.
[51] I. Lukari and W. Lindberg, Anal. Chim. Acta, 211 (1988) 1.
[52] M.J.P. Gerritsen, K. Kateman, M.A.J. van Opstal, W.P. van Bennekom and B.G.M. Vandeginste, Anal. Chim. Acta, 241 (1990) 23.
[53] W. Lindberg, G.D. Clark, C.P. Hanna, D.A. Whitman, G.D. Christian and J. Ruzicka, Anal. Chem., 62 (1990) 849.
[54] D.A. Whitman, M.B. Seasholtz, G.D. Christian, J. Ruzicka and B.R. Kowalski, Anal. Chem., 63 (1991) 775.
[55] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, M. Redón and J. Riba, Anal. Chim. Acta, 259 (1992) 219.
[56] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, J. Riba and E. Rovira, Talanta, 40 (1993) 261.
[57] P. MacLaurin, P.J. Worsfold, P. Norman and M. Crane, Analyst, 118 (1993) 617.
[58] A.J. Reviejo, F.J. Buyo, J.M. Pingarrón and J.L. Peral, Electroanalysis, 5 (1993) 303.
[59] J.E. Casida and L.O. Ruzo, Pestic. Sci., 11 (1980) 257.
[60] J.P. Demoute, Pestic. Sci., 27 (1989) 375.
[61] R.B. Leidy, T.J. Sheets and L.A. Nelson, Beitr. Tabakforsch. Int., 13 (1986) 191.
[62] G.R. Cayley and B.W. Simpson, J. Chromatogr., 356 (1986) 123.
[63] N. Oi, H. Kitahara and R. Kira, J. Chromatogr., 515 (1990) 441.
[64] GRAMS-386 Software Package, Version 2.0, and Add-on Application PLS plus Version 2.1G, Galactic Industries, Salem, NH.
[65] J.P. Kutter and T.J. Class, Chromatographia, 33 (1992) 103.

[66] M. Stone, J. R. Statist. Soc., 36 (1974) 111.
[67] T. Naes and T. Isaksson, NIR News, 4 (1993) 14.
[68] I.E. Frank, Chemometr. Intell. Lab. Syst., 8 (1990) 109.
[69] A. Hoskuldsson, J. Chemometr., 6 (1992) 307.
[70] S. Wold, Chemometr. Intell. Lab. Syst., 14 (1992) 71.
[71] T. Naes, B.R. Kowalski and T. Isaksson, Anal. Chem., 62 (1990) 664.
[72] T. Naes and T. Isaksson, Appl. Spectrosc., 46 (1992) 34.
[73] A.H. Aastveit and P. Maurum, Appl. Spectrosc., 47 (1993) 463.
[74] A. Savitzky and M.J.E. Golay, Anal. Chem., 36 (1964) 1627.