ELSEVIER

Analytica Chimica Acta 463 (2002) 75–88

ANALYTICA CHIMICA ACTA

www.elsevier.com/locate/aca

Comparative study of net analyte signal-based methods and partial least squares for the simultaneous determination of amoxycillin and clavulanic acid by stopped-flow kinetic analysis

A. Muñoz de la Peña [a,*], A. Espinosa-Mansilla [a], M.I. Acedo Valenzuela [a], H.C. Goicoechea [b], A.C. Olivieri [c]

[a] *Departamento de Química Analítica, Facultad de Ciencias, Universidad de Extremadura, 06071 Badajoz, Spain*
[b] *Cátedra de Química Analítica I, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Ciudad Universitaria, 3000 Santa Fe, Argentina*
[c] *Departamento de Química Analítica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531,2000 Rosario, Argentina*

Received 1 August 2001; received in revised form 24 April 2002; accepted 24 April 2002

Abstract

A comparative study about advantages and limitations of net analyte signal (NAS)-based methods (NBMs) and partial least squares (PLS) calibration in kinetic analysis has been performed. The different multivariate calibration methods were applied to the determination of binary mixtures of amoxycillin and clavulanic acid, by stopped-flow kinetic analysis. The reactions of oxidation of these compounds with cerium(IV), in sulphuric acid medium, were monitored by following the changes on the fluorescence of the oxidation products, in stopped-flow mode. The differences on the kinetic profiles obtained at $\lambda_{ex} = 256$ nm and $\lambda_{em} = 351$ nm, were used to determine mixtures of both compounds by multivariate calibration of the kinetic data, using PLS-1, a modification of hybrid linear analysis (HLA) and net analyte pre-processing combined with classical least squares (NAP/CLS) methods. The NBMs allowed the selection of optimal time data regions by calculating the minimum error indicator function (EIF), improving the results and making NBMs very convenient for the analysis. In addition, the use of the net analyte signal concept allows the calculation of the analytical figures of merit, limit of detection (LOD), sensitivity and selectivity, for each component.

Keywords: Stopped-flow kinetic analysis; Simultaneous determination; Amoxycillin; Clavulanic acid; Net analyte signal-based methods

1. Introduction

Multicomponent analysis is gaining popularity for the simultaneous determination of mixtures of compounds in several fields. The main techniques proposed are based in the use of spectroscopic data. With modern instruments, numerous spectral data can be recorded and easily digitized, and different mathematical approaches have been proposed to deal with these overdetermined systems [1].

The area of multicomponent determinations is one of the most promising in the context of kinetic analysis methods [2,3]. However, only in a few cases, these multivariate calibration methods have been applied to the analysis of kinetic data.

* Corresponding author. Tel.: +34-924289375; fax: +34-924289300.
E-mail address: arsenio@unex.es (A. Muñoz de la Peña).

0003-2670/02/$ – see front matter
PII: S0003-2670(02)00373-2

Mixtures of reacting analytes can be resolved from differences in their rate of reaction, and consequently kinetic profiles, with a common reagent. In the past, the method of proportional equations has been commonly applied for binary mixtures resolution in first- or pseudo-first order kinetic reactions [4]. However, only a small fraction of the kinetic data collected was used, which led to a poor precision.

Procedures based on least squares fitting [5], Kalman filtering [6,7] and nonlinear least squares fitting [8,9] have been also proposed. Multicomponent analysis in time-dependent chemical systems applying factor analysis based methods, as principal component regression (PCR) or partial least squares (PLS), correct the effects of component interactions in a mixture, but have been scarcely used to date in connection with kinetic systems [10–20].

Alternatively, second order methods which use simultaneously two measurements orders, time and spectra, have been reported. Among them, multiway data analysis, as PARAFAC [21] or nPLS [22–25] methods are expected to give improved results in certain cases, as single wavelength experimental curves (kinetic curves) may be less informative than multiwavelength-kinetic data, if differences exist in the spectra of the analytes of interest.

The analytical application of these multivariate calibration methods require human decisions, the most difficult being the number of factors involved. Different criteria, including empirical functions [26–28] and methods based on theoretical study of experimental errors have been proposed [29–31]. Xu and Schechter developed a new algorithm for factor analysis, in order to eliminate the necessity of choosing an optimum factor number during calibration [32]. The method is based on the useful concept of net analyte signal (NAS), previously developed by Lorber [33], which lead to a new family of multivariate calibration methods, of interest at the present moment to explore the possibilities and limitations of these approaches in general and, more specifically, in kinetic analysis methods. One of these methods is hybrid linear analysis (HLA) [34], which can be applied provided a very accurately measured pure spectrum of the analyte is available.

Two additional NAS-based methods (NBMs) have been recently introduced, which do not require the pure spectrum to be known. One of then was introduced by Goicoechea and Olivieri [35], and it is known as HLA/GO. Although in its method [32], denoted as HLA/XS, Xu and Schechter used all the factors for prediction, in order to build a method free from optimum factor estimation, Olivieri and Goicoechea suggested selecting the optimum number by the cross-validation procedure of Haaland and Thomas [36]. In this way, the method can be used even when the number of samples exceeds the recommend limit of one third the number of sensors.

A final possibility is net analyte pre-processing combined with classical least squares (NAP/CLS) [37]. One of the advantages of NBMs is that the possibility of simultaneous determination can be studied separately for each component and optimization can be carried out accordingly.

In addition, the analyst is not only interested in the final output of the calibration, rather, figures of merit characterizing the whole calibration process are required. Lorber [38] presented a generalization that allows estimation of figures of merit for multivariate data. The estimation of those figures of merit was restricted to the classical calibration model (when the pure spectra, or concentrations of all components in the calibration set, are known). Later on [39], NAS calculation was extended to inverse calibration models, that only require the knowledge of the concentrations of the analyte of interest in the calibration set. In this context, one of the most interesting characteristics of the NAS concept is the possibility of calculation of the analytical figures of merit and method performance, error propagation, signal to noise, limit of detection (LOD), precision, accuracy, sensitivity and selectivity for each component.

On the other hand, PLS and HLA were introduced as full-spectrum methods. They provide considerable improvement in analytical precision and accuracy as compared to other methods that are restricted to a small number of data points (e.g. inverse least squares) [36,40,41]. However, more and more evidence, from simulations and from experiments, shows that these full-spectrum algorithms could also benefit from wavelength selection [42,43]. It has been demonstrated that the analytical performance is governed to a large extend by the uncertainty in the NAS of a component. The derived function, minimum error indicator, has been proposed as a criterion for data selection [44]. Although NAS concepts have been

initially defined for spectra (multiwavelength determinations), they are easily adapted to any other type of multivariate signal, as the time-dependent data used in this study.

In this paper, a comparative study of NBMs and PLS multivariate calibration has been performed, applied to the stopped-flow kinetic determination of amoxycillin and clavulanic acid. The reaction of oxidation of these compounds with cerium(IV) in sulfuric acid medium has been previously studied fluorimetrically [45], and the simultaneous determination of both compounds has been performed by PLS multivariate calibration [46].

2. Experimental

2.1. Apparatus and software

An SLM Aminco Bowman Series 2 luminescence spectrometer connected to a PC microcomputer with the AB2 software which runs on the OS2 operating system, was used for fluorescence measurement. The instrument incorporates the MilliFlow stopped-flow reactor, allowing the study of changes in luminescence reactions when two reactants are vigorously forced through the mixing chamber and suddenly stopped into the observation cell. The MilliFlow consists of two fill syringes, two drive syringes, an observation cell (path length of 2 mm), a stop syringe, a stop block and an exhaust and fill valve levers. Hamilton gasting syringes of 2.5 ml (drive syringes) were used to contain the two reactant solutions. The syringes are made from controlled, inner-diameter borosilicate glass with precision machined Teflon plunger tips (these pistons are simultaneously driven by air-operated plungers). Thermostatic equipment permits a constant temperature between 10 and 45 °C in the MilliFlow stopped-flow reactor to be maintained.

All kinetic curves were measured in random order with respect to analyte concentrations, and those corresponding to the calibration set were recorded in different days with respect to the validation set. They were saved in ASCII format, and transferred to a PC Pentium 450 microcomputer for subsequent manipulation by either PLS or NAS-based programs. PLS, HLA/GO and NAP/CLS were applied with the program MULTIVAR, written in Visual Basic 5.0 [47] and available at fttp://fbioyf.unr.edu.ar/cientifico/multivar.exe. The Sigmaplot 5.0 software was used for regression analysis and treatment of data. Sensor selection was carried out by a moving-window strategy: the predicted error sum of square (PRESS) was minimized for calibrations [48] and calculation of the minimum error indicator function (EIF) in unknown samples [35]. Both procedures were implemented with the program MULTIVAR.

2.2. Reagents

Amoxycillin and clavulanic acid stock solutions were prepared by weighting the required amount of the respective compounds (amoxyxillin from Sigma and lithium clavulanate kindly donated by Smith Kline Beecham laboratories) and dissolved in purified water (from a Milli-Q water system of Millipore). Clavulanic solutions were daily prepared. Cerium sulphate (Merck) and sulphuric acid (Suprapur, Merck) were also used.

2.3. Commercial samples

2.3.1. Pharmaceuticals

Augmentine (per packet: 250 mg of amoxycillin trihydrated, 62.5 mg of clavulanic acid, 2.5 g of sucrose and excipients) from Smith Kline Beecham S.A.; Pangamox (per packet: 875 mg of amoxycillin trihydrated, 125 mg of clavulanic acid, 2.5 g of sucrose and excipients) from Alonga, Clamoxyl (per injection: 500 mg of amoxycillin), and Euplecanic (per pill: 500 mg of amoxycillin trihydrated, 125 mg of clavulanic acid and excipients) from Uriach-Biophorm, were acquired in a local pharmacy.

2.4. Methodology

One-drive syringe was filled with a solution containing amoxycillin and clavulanic acid in concentrations between 0.7 and 4.0 $\mu g\,ml^{-1}$ in 0.1 M sulphuric acid. The other syringe was filled with a solution containing 8×10^{-5} M cerium(IV) in 0.1 M sulphuric acid. The stopped-flow reactor was prepared for the acquisition of the kinetic curve. The instrument was set up as follow: $\lambda_{ex} = 256$ nm and $\lambda_{em} = 351$ nm (bandpass 4 nm), detector voltage 750 V. The kinetic curve was scanned up to 100 s, with a resolution of 30 ms. Three

Table 1
Composition of the calibration set for applying PLS-1, HLA/GO and NAP/CLS methods

Number of calibration sample	Amoxycillin ($mg\,l^{-1}$)	Clavulanic acid ($mg\,l^{-1}$)
1	0.00	0.00
2	0.00	0.76
3	2.28	0.00
4	4.00	2.40
5	0.70	2.60
6	2.60	0.70
7	2.28	0.76
8	2.40	0.40
9	2.40	2.40
10	1.00	1.00
11	1.00	4.00
12	4.00	0.00
13	4.00	0.00
14	2.00	3.00
15	3.00	1.00

replicates were recorded for each concentration. The optimized calibration models, calculated by application of PLS-1, HLA/GO and NAP/CLS, were applied to analyze the spectra of the samples, and calculate the concentrations of amoxycillin and clavulanic acid in their mixtures or in the pharmaceutical formulations assayed.

2.5. Calibration and validation sets

To resolve the mixture of amoxycillin and clavulanic acid, a calibration set was constructed by preparing 15 randomized calibration samples. Table 1 shows the composition of the binary mixtures used in the calibration set (the concentrations values are refereed to the concentrations in the syringes).

Similarly, a validation set with 15 randomized validation samples was prepared, with the concentrations of amoxycillin and clavulanic acid reported in Table 3.

3. Theory

3.1. Notation

The following matrices and vectors will be used throughout the present work: an $I \times J$ data matrix $\boldsymbol{R}$ composed of the calibration responses of I samples at J sensors, a $J \times 1$ vector $\boldsymbol{s}_k$ containing the pure spectrum of analyte k at unit concentration and an $I \times 1$ vector $\boldsymbol{c}_k$ of calibration concentrations of analyte k.

3.2. Net analyte signal

The NAS for analyte k ($\boldsymbol{r}_k^*$) is defined as the part of its spectrum which is orthogonal to the space spanned by the spectra of all other analytes [38,39]. In general, for inverse calibration methods, it is given by the following equation:

$$\boldsymbol{r}_k^* = [\boldsymbol{I} - \boldsymbol{R}_{-k}(\boldsymbol{R}_{-k})^+]\boldsymbol{r} = \boldsymbol{P}_{\mathrm{NAS},k}\boldsymbol{r} \quad (1)$$

where $\boldsymbol{r}$ is the spectrum of a given sample (when $\boldsymbol{r}$ is the spectrum $\boldsymbol{s}_k$ of pure k at unit concentration, Eq. (1) becomes $\boldsymbol{s}_k^* = \boldsymbol{P}_{\mathrm{NAS},k}\boldsymbol{s}_k$), $\boldsymbol{I}$ a $J \times J$ unitary matrix, $\boldsymbol{R}_{-k}$ an $J \times A$ column space spanned by the spectra of all other analytes except k ($\boldsymbol{R}_{-k}^+$ is the pseudoinverse of $\boldsymbol{R}_{-k}$ and A is the number of spectral factors used to build the model), and $\boldsymbol{P}_{\mathrm{NAS},k}$ is a $J \times J$ projection matrix which projects a given vector onto the NAS space. In factor based methods, such as PCR, PLS and NBMs, this latter statement is true only to the extent that the selected factors are representative of the spectra of all other analytes.

3.3. NAS-based methodologies

Although the PLS algorithm is well known, the recently introduced NBMs [32,34,35] deserve some comments. There are several NBMs that may be considered:

(1) The HLA technique [34] involves constructing a blank data matrix $\boldsymbol{R}_{-k} = [\boldsymbol{R} - \boldsymbol{c}_k(\boldsymbol{s}_k)^{\mathrm{T}}]$, and using the first significant A eigenvectors of $((\boldsymbol{R}_{-k})^{\mathrm{T}}\boldsymbol{R}_{-k})$ in order to define the projection matrix $\boldsymbol{P}_{\mathrm{NAS},k}$. The concentration of component k in an unknown sample is obtained from its spectrum $\boldsymbol{r}$ as

$$\boldsymbol{c}_k = \frac{(\boldsymbol{s}_k)^{\mathrm{T}}\boldsymbol{P}\boldsymbol{r}}{(\boldsymbol{s}_k)^{\mathrm{T}}\boldsymbol{P}\boldsymbol{s}_k} = \frac{(\boldsymbol{s}_k)^{\mathrm{T}}\boldsymbol{P}\boldsymbol{P}\boldsymbol{r}}{(\boldsymbol{s}_k)^{\mathrm{T}}\boldsymbol{P}\boldsymbol{P}\boldsymbol{s}_k} = \frac{(\boldsymbol{s}_k^*)^{\mathrm{T}}\boldsymbol{r}_k^*}{||\boldsymbol{s}_k^*||^2} \quad (2)$$

which is the basis of the prediction step in NBMs.

The optimum number of factors A can be obtained by the well known cross-validation procedure. It should be noticed that HLA can be applied provided a very accurately measured spectrum of pure k is available, and thus, it cannot be used

when interactions among sample components occur, or when $\boldsymbol{s}_k$ is unknown.

(2) The method developed by Xu and Schechter [32] is another recently introduced NBM, which does not require the pure spectrum $\boldsymbol{s}_k$ to be known. The following procedure is carried out in order to obtain the matrix $\boldsymbol{R}_{-k}$. Each calibration spectrum is divided by its concentration of analyte k, c_{ik} (excluding those samples in which k is absent to avoid a division by zero), and then the average of the resulting spectra is calculated as

$$\bar{\boldsymbol{s}}_{\text{cal}} = \frac{1}{I'}\sum_{i=1}^{I}\frac{\boldsymbol{r}_{i,\text{cal}}}{c_{ik}} \tag{3}$$

where I' is the number of calibration samples for which $c_{ik} \neq 0$. The contribution of the average spectrum $\bar{\boldsymbol{s}}_{\text{cal}}$ is subtracted from the data matrix by the following operation:

$$\boldsymbol{R}_{-k} = \boldsymbol{R} - \boldsymbol{c}_k(\bar{\boldsymbol{s}}_{\text{cal}})^{\text{T}} \tag{4}$$

The least squares approximation to $\boldsymbol{s}_k$ is subsequently used for the calculation of $\boldsymbol{s}_k^*$:

$$\boldsymbol{s}_k^* = \boldsymbol{P}_{\text{NAS},k}\boldsymbol{s}_{k,\text{LS}} = \boldsymbol{P}_{\text{NAS},k}\left[\frac{\boldsymbol{R}^{\text{T}}\boldsymbol{c}_k}{(\boldsymbol{c}_k)^{\text{T}}\boldsymbol{c}_k}\right] \tag{5}$$

where $\boldsymbol{c}_k$ should be mean centered.

Notice that Eq. (5) will approximate $\boldsymbol{s}_k^*$ even if the least squares fit $\boldsymbol{s}_{k,\text{LS}}$ contains contributions from the spectra of other analytes, provided the matrix $\boldsymbol{P}_{\text{NAS},k}$ is able to cancel out these latter contributions. Although Xu and Schechter suggest using all the I factors of $\boldsymbol{R}_{-k}$ for prediction, in order to build a method free from optimum factor estimation, one can select the optimum number A by cross-validation [35], then the method is HLA/XS. It works in a similar manner to HLA and can be used even when the number of samples exceeds the recommended limit of one third the number of sensors [32].

(3) The third variant of NBMs was described by Goicoechea and co-workers (HLA/GO) and showed excellent results when analyzing complex pharmaceutical mixtures [35,48]. In this variant, the mean (uncentered) calibration spectrum is first obtained:

$$\bar{\boldsymbol{r}}_{\text{cal}} = \frac{1}{I}\sum_{i=1}^{I}\boldsymbol{r}_{i,\text{cal}} \tag{6}$$

where $\boldsymbol{r}_{i,\text{cal}}$ is the spectrum for the ith calibration sample. Then the contribution of analyte k is subtracted from the data matrix $\boldsymbol{R}$ in the following way:

$$\boldsymbol{R}_{-k} = \boldsymbol{R} - \frac{\boldsymbol{c}_k(\bar{\boldsymbol{r}}_{\text{cal}})^{\text{T}}}{\bar{c}_{k,\text{cal}}} \tag{7}$$

where $\bar{c}_{k,\text{cal}}$ is the mean (uncentered) calibration concentration of analyte k. The calculation of $\boldsymbol{s}_k^*$ is then carried out with the following equation:

$$\boldsymbol{s}_k^* = \boldsymbol{P}_{\text{NAS},k}\left[\frac{(\bar{\boldsymbol{r}}_{\text{cal}})^{\text{T}}}{\bar{c}_{k,\text{cal}}}\right] \tag{8}$$

(4) NAP/CLS [37,47]. This variant involves using the least squares approximation $\boldsymbol{s}_{k,\text{LS}}$ both to obtain the matrix $\boldsymbol{R}_{-k}$ through Eq. (9) and $\boldsymbol{s}_k^*$ through Eq. (5):

$$\boldsymbol{R}_{-k} = \boldsymbol{R} = \boldsymbol{c}_k\boldsymbol{s}_{k,\text{LS}} \tag{9}$$

This method can be shown to consist of the following steps: (a) pre-processing the raw data matrix $\boldsymbol{R}$ by projecting it onto the space orthogonal to that spanned by all analytes except k, leading to the net analyte data matrix $\boldsymbol{R}_k^*$ (i.e. $\boldsymbol{R}_k^* = \boldsymbol{P}_{\text{NAS},k}\boldsymbol{R}$), and (b) correlating this latter matrix to the analyte concentrations $\boldsymbol{c}_k$ through a CLS procedure [37,47].

3.4. Figures of merit

Selectivity, sensitivity and LOD can be calculated and used for method comparison or to study the quality of a given analytical technique. The LOD is not strictly necessary in the present case, but only for the analysis of impurities [49]. The selectivity is usually defined in multivariate calibration by resorting to NAS calculations [50,51]:

$$\text{SEL} = \frac{||\boldsymbol{s}_k^*||}{||\boldsymbol{s}_k||} \tag{10}$$

Details on the calculations of the NAS in PLS and other multivariate methods are given in [35]. On the

other hand, the sensitivity is given by

$$\mathrm{SEN} = \frac{1}{||\boldsymbol{b}_k||} \tag{11}$$

where $\boldsymbol{b}_k$ is the vector of final regression coefficients appropriate for component k, which can be obtained by any multivariate method.

The LOD in multivariate calibration has been calculated according to the following expression [50]:

$$\mathrm{LOD} = 3||\epsilon||||\boldsymbol{b}_k|| \tag{12}$$

where $||\epsilon||$ is a measure of the instrumental noise and may be estimated, in turn, by registering several spectra for blank samples, calculating the norm of the NAS for each sample and the corresponding S.D. The latter can be taken as an approximation to $||\epsilon||$.

Another parameter, that may be useful for method comparison, is the analytical sensitivity γ [35,52]. It may be defined, in analogy to univariate calibration, as the quotient:

$$\gamma = \frac{\mathrm{SEN}}{||\epsilon||} \tag{13}$$

and allows one to compare analytical methods regardless of the specific technique, equipment, and scale employed and establishes the minimum concentration difference (γ^{-1}) which is statistically discernible by the method across the dynamic range where it is applicable.

4. Results and discussion

4.1. Kinetic curves

Amoxycillin or clavulanic acid (non fluorescent itself) react with cerium(IV) in sulphuric acid medium to give a fluorescent product, when heating the solutions. The fluorescent products have been reported to be the complexes formed between the oxidation product of amoxycillin or clavulanic acid, and the cerium(III). In previous papers, the physico-chemical conditions for the development of the reaction have been investigated [45,46]. The oxidation products in both cases present excitation and emission maxima at 256 and 351 nm, respectively which are the same as the excitation and emission maxima of cerium(III) (Fig. 1). An optimum temperature of 43 °C has been chosen for the development of the reaction. A 0.1 M sulphuric acid concentration was chosen to prepare the

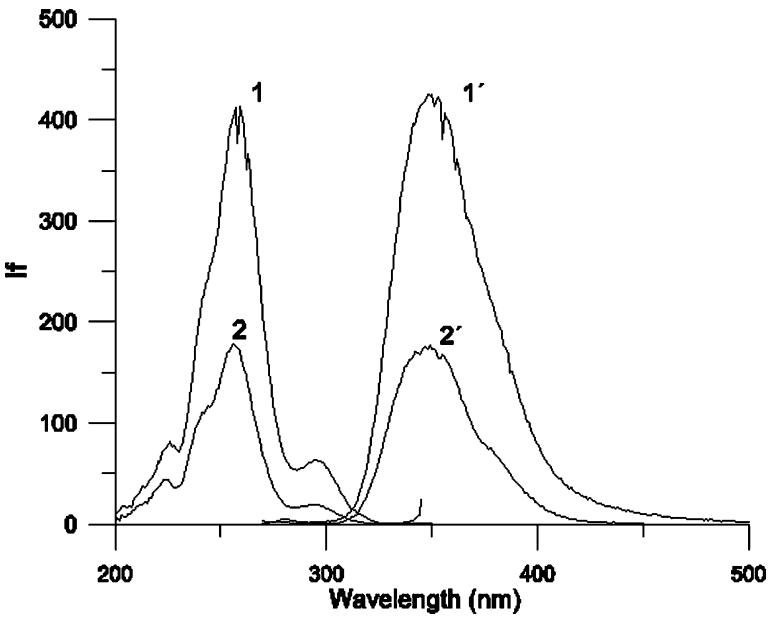


Fig. 1. Emission ($\lambda_{em} = 351$ nm) and excitation ($\lambda_{ex} = 256$ nm) spectra for the following systems: curves 1 and 1′, amoxycillin–cerium(IV); curves 2 and 2′, clavulanic acid-cerium(IV). [H_2SO_4] = 0.1 mol l^{-1}; [clavulanic acid] = 10 μg ml^{-1}; [amoxycillin] = 10 μg ml^{-1}; [cerium(IV)] = 3.2×10^{-5} mol l^{-1}; $T_a = 60$ °C.

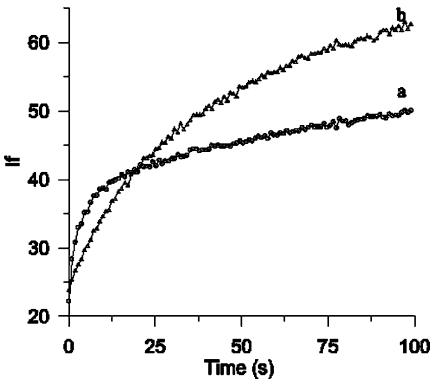


Fig. 2. Kinetic curves obtained for: (a) amoxycillin (11.70 μg ml^{-1}); (b) clavulanic acid (7.69 μg ml^{-1}).

solutions of the analytes and cerium(IV), from which the syringes are filled, been 8×10^{-5} M the optimum cerium concentration.

Fig. 2 shows the kinetic curves of the products of reaction of amoxycillin and clavulanic acid with cerium(IV) in the selected conditions. The reaction of oxidation of amoxycillin is completed in 10 s. For the clavulanic acid, the reaction is slower and in 100 s, the fluorescence is still increasing. It proves the high overlapping of the analytical signals.

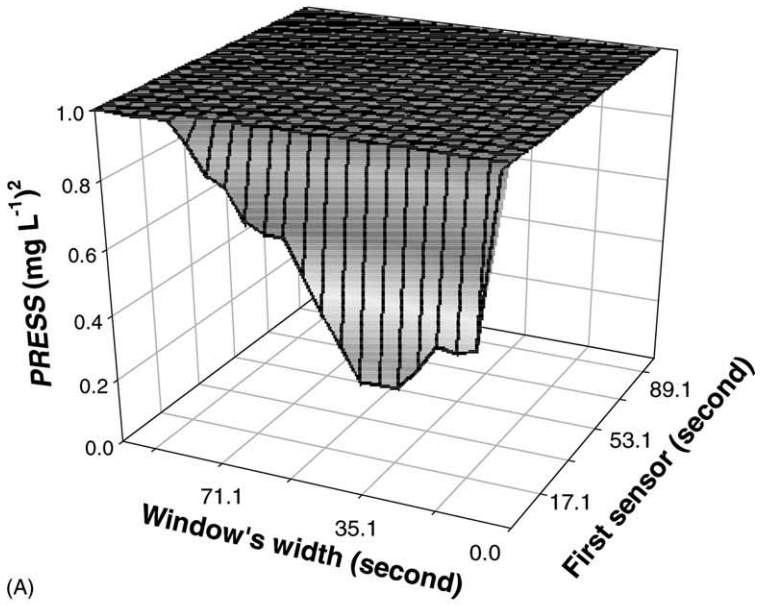
4.2. Optimization of the PLS-1, HLA/GO and NAP/CLS models

A popular multivariate method for resolving the present task, which has become a routine in pharmaceutical analyses, is PLS. An alternative approach involves the use of any of the available (NBMs, see the following sections).

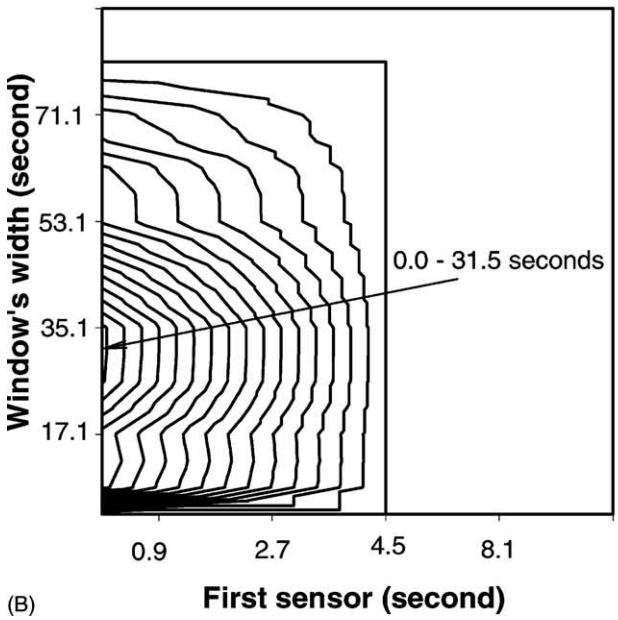
Kinetic curves for the samples corresponding to the calibration set (shown in Table 1) were recorded in the range 0–100 s (112 points), and subjected to PLS-1, HLA/GO and NAP/CLS analyses. The optimum number of factors to be used within the PLS-1 and NBMs algorithms is an important parameter to achieve better performance in prediction. This allows one to model the system with the optimum amount of information, avoiding overfitting. The cross-validation procedure was applied, consisting of systematically removing one of the training samples in turn, and using only the remaining ones for construction of the latent factors and regression [40]. The optimum numbers of factors are shown in Table 2. This latter table also gives the values of the optimal kinetic data regions used for all the multivariate methods in the corresponding calibrations, as well as other important statistical parameters and figures of merit such as the square of the

Table 2
Optimum number of factors, calibration statistical parameters and figures of merit when applying PLS-1, HLA/GO and NAP/CLS methods

Component	Statistical parameters and figures of merit	Multivariate method		
		PLS-1	HLA/GO	NAP/CLS
Amoxycillin	Sensor range (s)	0–31.5	0–31.5	0–31.5
	Factors	3	2	2
	RMSD (mg l^{-1})	0.212	0.209	0.211
	REP (%)	9.80	9.64	9.75
	R^2	0.9708	0.9717	0.9711
	Selectivity	0.102	0.106	0.099
	Sensitivity	4.22	4.36	4.08
	Analytical sensitivity (γ^{-1}, mg l^{-1})	0.036	0.034	0.037
	LOD (mg l^{-1})	0.23	0.23	0.25
Clavulanic acid	Sensor range (s)	1–78.3	1–78.3	1–78.3
	Factors	3	2	2
	RMSD (mg l^{-1})	0.205	0.328	0.204
	REP (%)	12.14	20.28	12.07
	R^2	0.9786	0.9431	0.9788
	Selectivity	0.154	0.135	0.143
	Sensitivity	10.7	9.35	9.89
	Analytical sensitivity (γ^{-1}, mg l^{-1})	0.014	0.016	0.015
	LOD (mg l^{-1})	0.09	0.11	0.10



(A)



(B)

Fig. 3. (A) Three-dimensional plot of the calculated values of PLS-1 calibration PRESS corresponding to amoxycillin, as a function of first sensor and window width. The values of PRESS were calculated using the optimum number of calibration clavulanic acid–cerium(IV); factors in each spectral region. (B) Two-dimensional contour plot of the top figure, showing the minimum PRESS at sensor 1 with a window of 36 sensors (corresponding to the spectral range 0–31.5 s).

correlation coefficient (R^2), the root mean square difference (RMSD), the relative error of prediction (REP (%)), the sensitivity (SEN), the selectivity (SEL), the analytical sensitivity (γ^{-1}) and the LOD.

The region of sensors to be selected is a critical step for increasing the predictive ability of multivariate analysis, and should ideally eliminate both uninformative and/or highly correlated data. In the present report, we have applied a moving-window strategy to the calibration set itself, in order to find the most informative range in the time profile by localization of the minimum PRESS [43]. Fig. 3A shows an example for amoxycillin using PLS-1 calibration, in the form of a three-dimensional plot of the calculated values of PRESS for the calibration set as a function of first sensor and window width. The values of PRESS were calculated using the optimum number of calibration factors in each time profile region. Fig. 3B is the corresponding two-dimensional contour plot of Fig. 3A. A visual inspection of both figures indicates that the smaller PRESS is located in the time region starting at sensor 1 (0 s) with a window of 36 sensors (i.e. in the range 0–31.5 s).

The obtained values for the present calibrations show that HLA/GO and NAP/CLS yields slightly better results for both analytes, in comparison with PLS-1, but the observed differences are not statistically significant.

Table 3
Validation results when applying PLS-1, HLA/GO and NAP/CLS on the validation set

Sample	Added ($mg\,l^{-1}$)	PLS-1		HLA/GO		NAP/CLS	
		Found ($mg\,l^{-1}$)	Recovery (%)	Found ($mg\,l^{-1}$)	Recovery (%)	Found ($mg\,l^{-1}$)	Recovery (%)
Amoxycillin							
1	0.80	0.70	87.3	0.65	81.3	0.69	86.7
2	2.40	2.13	88.7	2.20	91.7	2.10	87.5
3	2.40	2.75	114.7	2.70	112.4	2.79	116.2
4	2.00	1.92	96.1	1.98	99.0	1.88	93.9
5	4.00	4.25	106.3	4.21	105.2	4.31	107.7
6	4.00	3.83	95.9	3.85	96.4	3.85	96.4
7	1.00	1.08	107.8	0.98	97.9	1.09	109.2
8	1.00	1.05	105.2	1.11	111.1	1.00	100.0
9	4.00	4.00	100.0	4.01	100.2	4.04	101.1
10	4.00	3.98	99.4	3.97	99.3	4.01	100.3
11	4.00	3.99	99.7	4.11	102.6	3.99	99.7
12	2.00	1.81	90.6	1.85	92.5	1.79	89.3
13	2.00	1.82	91.1	1.88	93.9	1.79	89.7
14	3.00	3.15	104.9	3.10	103.3	3.18	105.0
15	3.00	3.02	100.7	3.01	100.4	3.04	101.4
Clavulanic acid							
1	0.76	0.63	83.5	0.89	117.1	0.64	84.9
2	4.00	3.89	97.2	3.77	94.2	3.92	98.0
3	0.00	0.02	–	−0.02	–	−0.03	–
4	4.00	2.98	74.6	2.71	67.9	3.05	76.3
5	0.70	0.71	101.7	0.47	67.9	0.62	89.1
6	2.80	2.46	87.8	2.15	76.9	2.43	86.8
7	1.00	1.03	103.4	1.39	139.0	1.02	101.8
8	4.00	3.98	99.4	4.01	100.1	4.04	100.9
9	1.00	0.98	98.0	0.68	68.4	0.92	91.8
10	1.00	0.98	98.2	0.65	65.5	0.92	92.4
11	0.00	0.58	–	−0.14	–	0.57	–
12	2.00	2.16	108.1	2.03	101.6	2.21	110.2
13	2.00	2.08	104.1	1.91	95.7	2.13	106.4
14	1.00	0.89	89.0	0.76	76.3	0.84	84.0
15	1.00	0.90	90.1	0.73	72.6	0.86	86.3

4.3. Predictions for the validation set

The optimal PLS-1, HLA/GO and NAP/CLS calibrations found in the manner described previously for each analyte were then applied to the prediction of the concentrations of the components in the 15 synthetic samples corresponding to the validation set, with the results collected in Table 3. As can be seen, the recoveries are reasonably good, although the worst results are obtained when applying HLA/GO for clavulanic acid. A convenient way to establish if bias is absent for the determination of both components when using PLS-1 or NBMs is to draw the elliptic joint confidence region (EJCR) for the slope and intercept

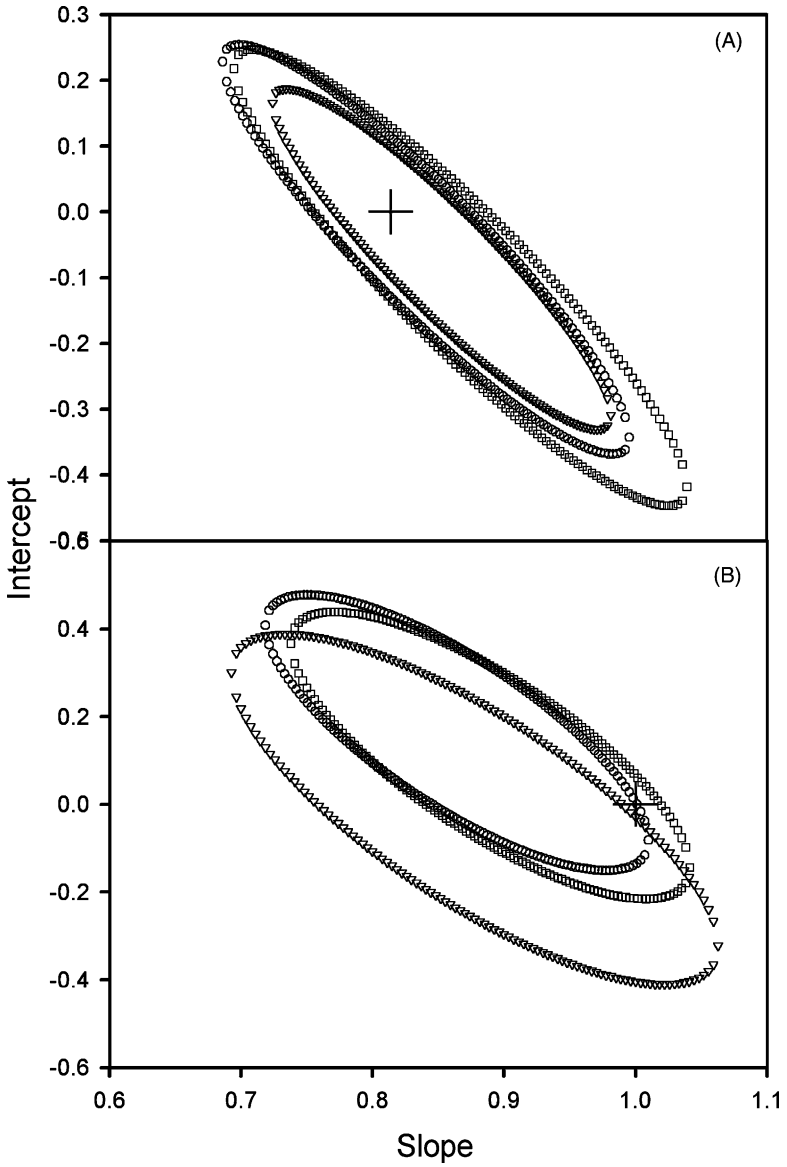


Fig. 4. EJCR for the slope and intercept for the validation data set (including 15 values), as calculated from determination of: (A) amoxycillin and (B) clavulanic acid. Both by using PLS-1 (circle), HLA/GO (triangle) and NAP/CLS (square) analysis. The cross mark indicates the theoretical point (1, 0).

when plotting c_{pred} versus c_{act}. Conventional individual confidence intervals for these parameters can lead to erroneous conclusions when carried out independently of each other, because this ignores their strong mutual correlation [53]. Fig. 4A and B show these regions for the determination of amoxycillin and clavulanic acid, respectively. As can be seen, all the ellipses contain the theoretically expected value of (1, 0) for the first component, although the corresponding one to HLA/GO presents a reduced size, indicating a lower dispersion. Regarding the second component, only the ellipse corresponding to NAP/CLS contains the value (1, 0). This statistical test allows us to detect an improvement in the predictive ability of HLA/GO for amoxycillin and of NAP/CLS for clavulanic acid, as compared to PLS-1.

With the purpose of studying the performance of the proposed multivariate methods under controlled conditions, simulated data were produced in the following way. Kinetic profiles at unit concentration for each analyte were calculated by smoothing the corresponding time-dependent experimental data. A 16 sample calibration set was then produced with a four-level full factorial design (the concentrations were in the range 0–4 mg l^{-1} for both analytes), and a 50 sample validation set was produced with random concentrations within the range employed for calibration. Random Gaussian noise and random base line drifts were added to all profiles in order to mimic the experimentally obtained data. All these calculations were done using suitable Matlab 5.3 routines.

Table 4 shows the results obtained by applying the different multivariate methods to the simulated data. As can be seen, if the added Gaussian noise is very low, then comparably good results are obtained for both analytes with all methods. In particular, the number of calibration factors is consistent with the dimensionality of the model in terms of number of chemical constituents. As expected, when the noise added to the kinetic profiles is comparable to that experimentally observed, the results match those found in the real cases herein studied. The number of required spectral components is also increased, indicating that profile overlapping and/or noise are responsible for the experimental results.

4.4. Predictions for commercial samples

Table 5 shows the results obtained for four commercial samples that were analyzed with the optimized

Table 4
Comparative study of the influence of noise in kinetic profiles when applying PLS-1, HLA/GO and NAP/CLS methods to simulated calibration and validation data[a]

		Method					
		PLS-1		HLA/GO		NAP/CLS	
		1	2	1	2	1	2
Amoxycillin							
Calibration	Factors	2	3	1	2	1	2
	RMSECV	0.052	0.449	0.052	0.430	0.052	0.422
	REP (%)	1.25	11.1	1.24	10.7	1.24	10.4
Validation	RMSEP	0.056	0.510	0.056	0.451	0.056	0.477
	REP (%)	1.21	11.3	1.21	9.90	1.21	10.5
Clavulanic acid							
Calibration	Factors	2	3	1	2	1	2
	RMSECV	0.018	0.47	0.018	0.49	0.018	0.46
	REP (%)	0.40	11.1	0.39	11.6	0.40	10.9
Validation	RMSEP	0.021	0.57	0.021	0.61	0.021	0.55
	REP (%)	0.43	13.1	0.43	14.2	0.43	12.6

[a] Noise level 1 corresponds to adding very low Gaussian random noise to all kinetic profiles (typically on the order of 1% of the maximum signal). Noise level 2 corresponds to adding sufficient Gaussian random noise in order to mimic the experimental results (ca. 10% of the maximum signal).

Table 5
Results obtained when analyzing commercial samples by applying PLS-1, HLA/GO and NAP/CLS

Commercial sample	Reported	PLS-1		HLA/GO		NAP/CLS	
		Found	Recovery (%)	Found	Recovery (%)	Found	Recovery (%)
Amoxycillin ($mg\,l^{-1}$)							
Augmentine	2.50	2.72	108.8	2.65	106.0	2.77	110.7
	3.75	3.82	101.9	3.75	100.0	3.87	103.3
Pangamox	3.71	2.94	79.3	2.98	80.5	2.98	80.5
	4.45	3.56	80.0	3.57	80.2	3.57	80.2
Eupeclanic	3.82	3.14	82.2	3.20	83.7	3.14	82.2
Clamoxyl	3.00	2.72	90.2	2.76	91.9	2.70	90.1
Clavulanic acid ($mg\,l^{-1}$)							
Augmentine	0.65	0.38	58.5	0.33	51.7	0.46	71.5
	0.97	0.63	64.7	0.63	65.2	0.63	65.2
Pangamox	0.55	0.54	99.0	0.44	83.2	0.45	83.3
	0.65	0.56	85.4	0.59	90.5	0.59	90.5
Eupeclanic	0.99	0.74	74.9	0.79	80.1	0.79	80.1
Clamoxyl	0.00	–	–	–	–	–	–

models. These are acceptable for the determination of amoxycillin with both PLS-1 and NBMs, although in the case of clavulanic acid the recoveries are very poor. The use of NBMs offers an alternative approach, which involves the use of a moving-window strategy in order to alleviate the effect of interferences by selecting optimal kinetic data regions for specific samples [54].

In the present work, this kinetic data region selection was performed by calculating an EIF as a function of a moving window, using information from the NAS regression plot (NASRP) for each particular sample [54,55]. The latter is a plot of the elements of the sample vector $\boldsymbol{r}_k^*$ versus those of $\boldsymbol{s}_k^*$. The expression for EIF used in the present context is

$$\mathrm{EIF} = \frac{\mathrm{Var}(||\boldsymbol{r}_k^*|| - ||\boldsymbol{r}_{k,\mathrm{true}}^*||)^{1/2}}{||\boldsymbol{r}_{k,\mathrm{true}}^*||} \tag{14}$$

where $||\boldsymbol{r}_k^*||$ is the norm of the net analyte profile for the test sample and $||\boldsymbol{r}_{k,\mathrm{true}}^*||$ is the latter without errors. As defined in Eq. (14), EIF is an approximation to the relative error in the determination of $||\boldsymbol{r}_k^*||$. Although $||\boldsymbol{r}_{k,\mathrm{true}}^*||$ is not known, it has been shown that EIF can

Table 6
Results obtained when analyzing clavulanic acid in commercial samples by applying NAP/CLS, selecting the sensor regions with the minimum error indicator—moving-window criteria

Commercial sample	Reported ($mg\,l^{-1}$)	Region (s)	EIF	Found ($mg\,l^{-1}$)	Recovery (%)
Augmentine	0.65	0–88	0.043	0.46	75.5
		9–88	0.035	0.77	118.0
	0.97	0–88	0.091	0.63	65.2
		9–62.1	0.076	0.76	78.3
Pangamox	0.55	0–88	0.069	0.45	83.3
		4.5–88	0.020	0.58	106.4
Eupeclanic	0.99	0–88	0.16	0.79	80.1
		9–88	0.026	0.93	93.9

be reasonably approximated by [54]:

$$\mathrm{EIF} = \frac{[s_{\mathrm{fit}}^2(1 + (N^2 s_{\mathrm{fit}}^2/4||\boldsymbol{r}_k^*||^2))]^{1/2}}{||\boldsymbol{r}_k^*||} \quad (15)$$

where s_{fit} is the S.D. of the best fitted straight line to the NASRP (in a given region) and N is the number of points in the latter plot. In this work, the vectors $\boldsymbol{r}_k^*$ and $\boldsymbol{s}_k^*$ were calculated from NAP/CLS.

Table 6 shows the results obtained when sensor selection was applied. As can be seen, a substantial improvement is obtained when the samples are analyzed in the most convenient kinetic regions, where the effect of interferences is minimized. This makes the NBMs more convenient that PLS-1 for analyzing both components in the studied pharmaceutical samples. However, the application of PLS-1 in the optimum profile regions selected by the EIF method gives results which favorably compare to those shown in Table 6 for NAP/CLS.

5. Conclusions

The contents of amoxycillin and clavulanic acid were simultaneously determined, using kinetic data from the oxidation reactions of these compounds with cerium(IV) in sulphuric acid medium. The fast oxidation reactions were fluorimetrically monitored in stopped-flow mode. PLS-1, HLA/GO and NAP/CLS multivariate calibration methods were applied. A validation set of synthetic mixtures as well as several commercial pharmaceuticals were studied. The performance of the investigated multivariate methods was found to be of comparable quality for the experimental system under study. However, the methods based on NAS calculation provide a convenient way of selecting variables, in this case the optimal kinetic profile, for calibration and prediction, on the basis of the search of the minimum EIF, based on the minimization of the PRESS, calculated as a function of a moving kinetic data window. This feature selection procedure allows to build the calibration models using data which are relevant to the target parameters to be predicted, improving the results and making NBMs very convenient for the analysis. Its constitutes a clear advantage over traditional variable selection methods, in which an easily understandable parameter is minimized, i.e. the linearity of the plot of NAS as a function of net analyte sensitivity.

Acknowledgements

Financial support from CICYT of Spain (Project PB98-0999), Consejería de Educación, Ciencia y Tecnología de la Junta de Extremadura (Project IPR00A036), CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas), the University of Rosario, the Agencia de Promoción Científica y Tecnológica (Project no. 06-00000-01765) and University of Litoral (Project CAI+D 2002, no. 219) is gratefully acknowledged.

References

[1] H. Martens, T. Naes, Multivariate Calibration, Wiley, Chichester, 1989.
[2] B.M. Quencer, S.R. Crouch, CRC Crit. Rev. Anal. Chem. 24 (1993) 243.
[3] M. Silva, Analyst 118 (1993) 681.
[4] H.A. Mottola, Kinetic Aspects of Analytical Chemistry, Wiley, New York, 1988.
[5] G.M. Ridder, D.W. Margerum, Anal. Chem. 49 (1977) 2098.
[6] P.D. Wentzell, M.I. Kazayannis, S.R. Crouch, Anal. Chim. Acta 224 (1989) 263.
[7] R. Xiong, A. Velasco, M. Silva, D. Pérez Bendito, Anal. Chim. Acta 251 (1991) 313.
[8] S. Abe, T. Saito, M.S. Suda, Anal. Chim. Acta 181 (1986) 203.
[9] A. Cladera, E. Gómez, J.M. Estela, J. Cerdá, V. Cerdá, Anal. Chim. Acta 272 (1993) 339.
[10] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, J. Riba, E. Rovira, Talanta 40 (1993) 261.
[11] A. Cladera, E. Gómez, J.M. Estela, V. Cerdá, Anal. Chem. 65 (1993) 707.
[12] J. Havel, F. Jiménez, R.D. Bautista, J.J. Arias León, Analyst 118 (1993) 1355.
[13] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, J. Riba, Anal. Chem. 66 (1994) 2905.
[14] A. Espinosa-Mansilla, A. Muñoz de la Peña, F. Salinas, A. Zamoro, Quim. Anal. 13 (1994) 106.
[15] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, M. Redón, Anal. Chem. 67 (1995) 4477.
[16] J.M. García, A.I. Jiménez, J.J. Arias, K.D. Khalaf, A. Morales-Rubio, M. de la Guardia, Analyst 120 (1995) 313.
[17] I. Durán-Merás, A. Espinosa Mansilla, F. Salinas López, Analyst 120 (1995) 2567.
[18] Y.L. Xie, J.J. Baeza-Baeza, G. Ramis-Ramos, Anal. Chim. Acta 321 (1996) 75.

[19] M. de la Guardia, K.D. Khalaf, B.A. Hasan, A. Morales-Rubio, J.J. Arias, J.M. García-Fraga, A.I. Jiménez, F. Jiménez, 121 (1996) 1321.
[20] Z. Bouhsain, S. Garrigues, M. de la Guardia, Analyst 121 (1996) 1935.
[21] J.C.G. Esteves da Silva, C.J.S. Oliveira, Talanta 49 (1999) 889.
[22] R. Bro, J. Chemom. 10 (1996) 47.
[23] M. Azubel, F.M. Fernández, M.B. Tudino, O.E. Trocoli, Anal. Chim. Acta 398 (1999) 93.
[24] A.K. Pettersson, B. Karlberg, Anal. Chim. Acta 354 (1997) 241.
[25] J. Coello, S. Maspoch, N. Villegas, Talanta 53 (2000) 627.
[26] H.M. Cartwright, J. Chemom. 1 (1987) 111.
[27] R.I. Shrager, R.W. Hendler, Anal. Chem. 54 (1982) 1147.
[28] T. Rossi, I.M. Warner, Anal. Chem. 54 (1986) 810.
[29] A.G. González, D. González-Arjona, Anal. Chim. Acta 314 (1995) 251.
[30] D. González-Arjona, J.A. Mejías, A.G. González, Anal. Chim. Acta 295 (1994) 19.
[31] E.R. Malinowski, J. Chemom. 4 (1990) 102.
[32] L. Xu, I. Schechter, Anal. Chem. 69 (1997) 3722.
[33] A. Lorber, Anal. Chem. 58 (1986) 1167.
[34] A.J. Berger, T.W. Koo, I. Itzkan, M.S. Feld, Anal. Chem. 70 (1998) 623.
[35] H.C. Goicoechea, A.C. Olivieri, Anal. Chem. 71 (1999) 4361.
[36] D.M. Haaland, E.V. Thomas, Anal. Chem. 60 (1988) 1193.
[37] H.C. Goicoechea, A.C. Olivieri, Chemom. Intell. Lab. Syst. 56 (2001) 73.
[38] A. Lorber, Anal. Chem. 58 (1986) 1167.
[39] A. Lorber, K. Faber, B.R. Kowalski, Anal. Chem. 69 (1997) 1620.
[40] D.M. Haaland, E.V. Thomas, Anal. Chem. 60 (1988) 1202.
[41] E.V. Thomas, D.M. Haaland, Anal. Chem. 62 (1990) 1091.
[42] M. Otto, T. George, Anal. Chim. Acta 200 (1987) 379.
[43] J.H. Kalivas, N. Roberts, J.M. Sutter, Anal. Chem. 64 (1992) 1200.
[44] L. Xu, I. Schechter, Anal. Chem. 68 (1996) 2392.
[45] P. Izquierdo, A. Gómez Hens, D. Pérez Bendito, Analyst 118 (1993) 707.
[46] A. Muñoz de la Peña, M.I. Acedo Valenzuela, A. Espinosa-Mansilla, R. Sánchez Máqueda, Talanta 56 (2002) 635.
[47] H. C Goicoechea, A.C. Olivieri, Trends Anal. Chem. 19 (2000) 599.
[48] M.S. Collado, V.E. Mantovani, H.C. Goicoechea, A.C. Olivieri, Talanta 52 (2000) 909.
[49] J.M. Green, Anal. Chem. 68 (1996) 305A.
[50] R. Boqué, F.X. Rius, Chemom. Intell. Lab. Syst. 32 (1996) 11.
[51] K.S. Booksh, B.R. Kowalski, Anal. Chem. 66 (1994) 782A.
[52] L. Cuadros Rodriguez, A.M. García Campaña, C. Gimenez Linares, M. Román Ceba, Anal. Lett. 26 (1993) 1243.
[53] A.G. González, M.A. Herrador, A.G. Asuero, Talanta 48 (1999) 729.
[54] H.C. Goicoechea, A.C. Olivieri, Analyst 124 (1999) 725.
[55] H.C. Goicoechea, A.C. Olivieri, Anal. Chim. Acta 453 (2002) 289.