ELSEVIER

Chemometrics and Intelligent Laboratory Systems 27 (1995) 211–220

Chemometrics and intelligent laboratory systems

Kinetic spectrophotometric resolution of binary mixtures using three-way partial least squares

Yu-Long Xie [1], Juan José Baeza-Baeza, Guillermo Ramis-Ramos *

Department of Analytical Chemistry, University of Valencia, E-46100 Burjassot (Valencia), Spain

Received 6 June 1994; accepted 4 October 1994

Abstract

Three-way partial least squares (PLS) was applied to kinetic–spectrophotometric data. The coupling reaction of diazotized sulfanilamide with *o*-, *m*- and *p*-amino benzoic acid (ABA), and with orciprenaline (ORC), to give azodyes was monitored. Three binary mixtures of substrates, i.e., *o*-ABA/ORC, *m*-ABA/*p*-ABA and *o*-ABA/*m*-ABA, with different values of the rate constant ratio and spectra which overlapped seriously were studied. The spectra of the mixtures were scanned with a 2 nm resolution every 30 s during ca. 15 min. The data sets contained from 30×36 to 30×48 time–wavelength data. Nine mixtures of each binary combination of substrates were used for calibration, thus the three-way calibration data sets contained from $9 \times 30 \times 36$ to $9 \times 30 \times 48$ concentration–time–wavelength data. The two-way PLS modelling was constructed on the basis of single wavelength kinetic curves, and the three-way PLS modelling was applied to series of three-way data arrays consisting of a number of selected wavelengths each (up to the whole spectra). The results based on three-way data arrays were better than that of ordinary PLS, particularly with mixtures having both a low rate constant ratio and small spectral differences.

1. Introduction

Together with the introduction of advanced instrumentation capable of generating multi-dimensional data arrays, three- and multi-way methods of data analysis have provoked the interest of chemists and many applications have been described [1–15]. In quantitative analysis, a calibration set with data taken at increasing concentrations of the analytes is necessary. Therefore, concentration provides one dimension of the signal–concentration data array. To construct a three-way array, the other two dimensions can be provided by two-dimensional instrumentation including excitation–emission fluorimetric scans, or separation techniques coupled to UV–visible, infrared, or mass spectrometric detection. Another way of generating a two-dimensional data array is to follow a chemical reaction with an instrument providing uni-dimensional scans, e.g., a diode-array UV–visible spectrophotometer. The kinetic–spectrophotometric information obtained, together with the multivariate calibration at several concentrations, gives rise to a three-way data array which can be useful to resolve mixtures of compounds with very similar properties.

* Corresponding author.

[1] On leave from the Department of Chemistry, Xiangtan University, Xiangtan, People's Republic of China.

0169-7439/95/$09.50 © 1995 Elsevier Science B.V. All rights reserved
SSDI 0169-7439(94)00069-7

Multicomponent determinations is an area of great interest in the context of kinetic analysis [16]. Since many years ago, mixtures of analytes have been resolved on the basis of their different reaction rates by the method of proportional equations [17,18]. However, only a small fraction of the data collected is used, which leads to poor precision. More recently, simultaneous nonlinear regression of the rate constants and the initial concentrations of the analytes has been used [19–21]. Also, the Kalman filter is widely used for the kinetic determination of analytes in mixtures, and many applications have been described [22–27]. The linear form of the Kalman filter requires invariant rate constants to be assumed from run to run which is a serious drawback, since the rate constants are functions of many experimental conditions. By employing the nonlinear form of the filter, known as the extended Kalman filter, invariant rate constants are not required and the rate constants and other calibration factors can be incorporated into the filter model [27–29]. An algorithm which uses the Kalman filter for the simultaneous determination of analytes using first- and second-order kinetic data has been developed [30].

To adequately apply these methods, however, the model should be known completely, and the parameters of the model, e.g., rate constants, molar absorptivities and so on, should be accurately determined during calibration. These are 'hard-modelling' methods in which an ignored factor introduces large errors. This happens frequently in kinetic analysis when factors such as side reactions, catalytic effects or a temperature drift are not included in the kinetic model. In multicomponent analysis the possible interactions between the analytes which can reinforce or reduce the overall signal can also occur. To consider one or several of these factors can extraordinarily complicate both the model and the calibration procedure, thus hindering application of hard-modelling methods to real samples. Thus, owing to the complexity of the real world, the scope and applicability of hard-modelling methods is severely limited.

On the contrary, no limits derived from the complexity of the system are dictated by the so-called soft-modelling methods. Thus, soft-modelling methods can be universally applied to systems of all types without almost any previous knowledge about the system, e.g., the nature and even the number of significant parameters (such as how many unknown concentrations) involved in the system can be ignored. In soft-modelling methods, such as principal component regression (PCR) and two- and multi-way partial least squares (PLS) regression, no model is previously assumed, and an empirical model is derived from the data themselves. Usually, the model is constructed from a series of mixtures known as the calibration set. Very often in PCR and PLS a linear approach is assumed. In linear soft-modelling, the relationships between the evaluated parameters and the measured signal should be linear, or at least it should be possible to obtain linear relationships by applying some kind of transformation to the signal, to the evaluated parameters or to both the signal and the parameters. Otherwise, an error which is partitioned among the evaluated parameters is produced. These calibration techniques have been scarcely applied to kinetic data [31,32], especially to three-way kinetic–spectral data arrays.

The two-way PLS calibration method has been widely used to multicomponent analysis of equilibrium systems, and also the PLS technique has been extended to the treatment of three-way data array by Wold et al. [2], which has been extensively used in chromatography [8–10]. In this work, binary mixtures of compounds were resolved by two- and three-way PLS using kinetic data taken at a single wavelength, and two-way kinetic–spectral data provided by a photodiode array detector, respectively. The potential of these techniques for the kinetic determination of mixtures of drugs was evaluated. The coupling reaction of diazotized sulfanilamide with the *o*-, *p*- and *m*-amino benzoic acids, and with orciprenaline to yield intensely coloured azo dyes, was monitored. Binary mixtures of the substrates with different ratios of the rate constants, and partially overlapped spectra of the products, were studied. The advantage of the three-way PLS relative to ordinary two-way PLS is discussed.

2. Theoretical

Lowercase bold characters are used for column vectors, uppercase bold characters for two-way matrices, and underlined uppercase bold characters for three-way matrices. The transpose of a matrix or a

vector is represented by the superscript T. Unless otherwise stated, lowercase and uppercase plain characters are used as running indices and to indicate the number of dimensions of the vectors or matrices, respectively. Lowercase and uppercase plain characters are also used for scalars.

In linear PLS no model is required, but a linear relationship between the response, i.e., the absorbance, and the evaluated parameters, i.e., the initial concentrations of the analytes, should exist. Let us assume a mixture containing L active substrates, C_l, which couples with a single reagent, R, following a first- or pseudo-first order reaction:

$$C_l + R \rightarrow P_l$$

In the systems used in the experimental part, the change of absorbance is due only to the change in the concentration of the products, P_l. If it is assumed that the absorbances are additive and follow the Lambert–Beer law, we can write:

$$A_{i,j} = \sum S_{i,l} c_l \left[1 - \exp(-k_l t_j)\right] \quad (1)$$

where $A_{i,j}$ is the change of absorbance of the mixture at the time t_j ($j = 1, \ldots, J$) and the wavelength λ_i ($i = 1, \ldots, I$), $S_{i,l}$ is the sensitivity (the molar absorptivity multiplied by the optical path length) of P_l at λ_i, c_l is the initial concentration of C_l ($l = 1, \ldots, L$), and k_l is the corresponding rate constant. If the model of Eq. (1) is followed, a linear relationship between $A_{i,j}$ and c_l exists. However, this linear relationship is preserved in many other instances, as far as both the Lambert–Beer and additivity laws and the first- or pseudo-first order behaviors are followed. Other conditions are not required. Thus, for instance if the analytes also absorb, Eq. (1) should be replaced by:

$$A_{i,j} = \sum c_l \left[S_{i,l} + (S'_{i,l} - S_{i,l})\exp(-k_l t_j)\right] \quad (2)$$

where $S'_{i,l}$ is the sensitivity with respect to the analytes. Furthermore, if the products are not stable and are hydrolysed following a first order law, Eq. (1) is replaced by:

$$A_{i,i} = \sum S_{i,l} c_l \frac{k_{1,l}}{k_{2,l} - k_{1,l}} \times \left[\exp(-k_{1,l} t_j) - \exp(-k_{2,l} t_j)\right] \quad (3)$$

where $k_{1,l}$ and $k_{2,l}$ represent the first-order rate constants for the formation and hydrolysis of the P_l products, respectively. In both Eqs. (2) and (3) a linear relationship between $A_{i,j}$ and c_l is preserved. Therefore, any of the Eqs. (1), (2) or (3) can be written as:

$$\mathbf{A} = \sum \mathbf{s}_l c_l \mathbf{k}_l = \mathbf{SCK} \quad (4)$$

where $\mathbf{A}$ is a $I \times J$ matrix which contains all the wavelength–time information, $\mathbf{s}_l$ is the sensitivity vector of P_l, and $\mathbf{k}_l$ is a vector containing the kinetic information for the lth component and whose generic element is $\{1 - \exp(-k_l t_j)\}$ or any other function of t_j. Thus, $\mathbf{S}$ ($= \{\mathbf{s}_l\}$) is an $I \times L$ matrix containing the spectral information of the analytes, $\mathbf{K}$ ($= \{\mathbf{k}_l\}$) is an $L \times J$ matrix which represents the kinetic information, and $\mathbf{C}$ is the $L \times L$ diagonal matrix of the initial concentrations of the analytes.

Eqs. (1)–(4) are not used as models for the calibration and evaluation processes, but only as a way to show that a linear relationship between the response and the concentrations to be estimated exists. An alteration of the model, such as the situation represented by Eqs. (2) or (3) with respect to Eq. (1), would bring much trouble to model-based methods, but causes no effects when soft-modelling techniques are used, because no explicit model is necessary to be adopted in the latter. To construct the prediction model, a calibration set constituted by N mixtures of known composition and different concentrations of the substrates should be measured. From this $N \times I \times J$ three-way data array, the following three-way response model is obtained:

$$A_{nij} = \sum S_{il} C_{nl} K_{jl} \quad (5)$$

or:

$$\underline{\mathbf{A}} = \sum \mathbf{s}_l \otimes \mathbf{c}_l \otimes \mathbf{k}_l \quad (6)$$

where $\underline{\mathbf{A}}$ is a $N \times I \times J$ three-way response data array, and A_{nij} is its generic element; $\mathbf{c}_l$ is the vector of the concentration of the lth analyte in the N mixtures, and the symbol $\otimes$ represents the tensor product (outer product) [2].

The PLS based on two blocks of matrices is well known and widely used in analytical chemistry. The matrices (e.g., $\mathbf{X}$ and $\mathbf{Y}$) are decomposed into score matrices and loading matrices, with the constraint that the score vectors of the same component (factor, or

latent variable) are connected by the so-called inner relation. The PLS model for matrix form data is:

$$\mathbf{X} = \Sigma \mathbf{t}_i \mathbf{p}_i^{\mathrm{T}} + \mathbf{E} = \mathbf{T}\mathbf{P}^{\mathrm{T}} + \mathbf{E}$$

$$\mathbf{Y} = \Sigma \mathbf{u}_i \mathbf{q}_i^{\mathrm{T}} + \mathbf{F} = \mathbf{U}\mathbf{Q}^{\mathrm{T}} + \mathbf{F} \quad (7)$$

$$\mathbf{U} = \mathbf{T}\mathbf{B} + \mathbf{H}$$

and the prediction equation for **Y** is:

$$\mathbf{Y} = \mathbf{T}\mathbf{B}\mathbf{Q}^{\mathrm{T}} + \mathbf{F} \quad (8)$$

where the typical elements of the **T** and **U** matrices are t_{ns} and u_{ns}, respectively.

The generalization of PLS to the three-way data arrays $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$ was described by Wold et al. [2]. Assuming that $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$ consist of the predictors and the values to be predicted, respectively, three-way PLS leads to the following model:

$$X_{nij} = \Sigma t_{ns} p_{sij} + e_{nij};\ n = 1, \ldots, N; i = 1, \ldots, I;$$
$$j = 1, \ldots, J$$
$$Y_{nij} = \Sigma u_{ns} q_{sij} + f_{nij};\ n = 1, \ldots, N; i = 1, \ldots, M;$$
$$j = 1, \ldots, P \quad (9)$$

which can be written in matrix–tensor form:

$$\underline{\mathbf{X}} = \mathbf{T} \otimes \underline{\mathbf{P}}^{\mathrm{T}} + \underline{\mathbf{E}}$$

$$\underline{\mathbf{Y}} = \mathbf{U} \otimes \underline{\mathbf{Q}}^{\mathrm{T}} + \underline{\mathbf{F}} \quad (10)$$

and

$$\mathbf{U} = \mathbf{T}\mathbf{B} + \mathbf{H} \quad (11)$$

where $\underline{\mathbf{P}}$ is a three-way matrix with typical element p_{sij}, and $\underline{\mathbf{Q}}$, $\underline{\mathbf{E}}$ and $\underline{\mathbf{F}}$ are analogous to $\underline{\mathbf{P}}$. The orders of $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$ can be different, and in this study, $\underline{\mathbf{Y}}$(**C**) had a lower order than $\underline{\mathbf{X}}$(**A**). The PLS decomposition of two blocks of matrices and the three-way data array is illustrated in Fig. 1.

To estimate the parameters in the three-way model of Eq. (10), Wold suggested to unfold the three-way array in the direction which leaves the first mode intact, which in our case is the concentration mode [1,2]. Thus, the model parameters t_{ns}, p_{sij}, u_{ns} and q_{sij} are estimated on the basis of the unfolded data matrices **A** and **C**.

Following the analogy between this procedure and the procedure which is normally used to process kinetic data monitored on a single wavelength, the slice of the three-way response $\underline{\mathbf{A}}$ referred to one wavelength was picked out and used to construct the ordinary (two-way) PLS calibration model.

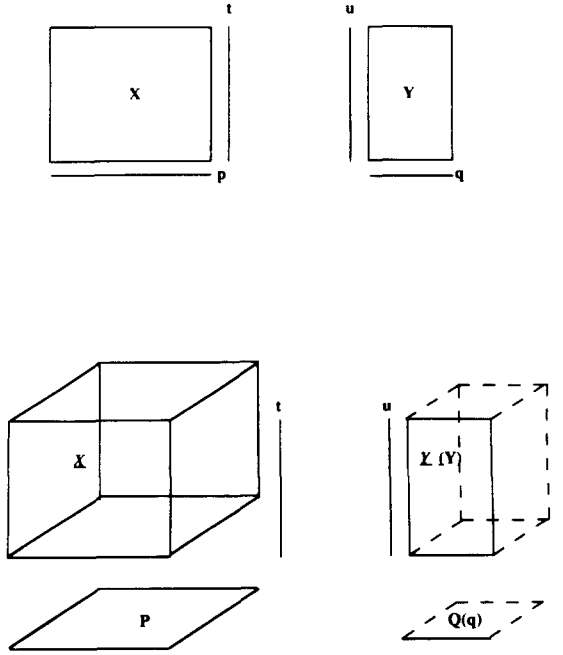

Fig. 1. A schematic view of the PLS modelling for two- and three-way data arrays: PLS model with one latent variable for two blocks of matrices **X** and **Y** (upper part) and for the three-way data array $\underline{\mathbf{X}}$ and data array $\underline{\mathbf{Y}}$ (or **Y**) (lower part). The dashed lines represent the possible absence of the third dimension.

The NIPALS algorithm [33] was used to calculate the latent variables in both three-way and ordinary PLS, and cross validation (by the leave-one-out method) was adopted to determine the number of latent variables to be retained.

3. Experimental

3.1. Apparatus

An HP 8452A diode array spectrophotometer (Hewlett-Packard, Palo Alto, CA, USA) provided with a 1-cm quartz cell was used. The pH values were adjusted with a Crison MicroPH 2001 pH meter provided with a combined glass electrode. A 500 μl piston pipette was used to start the reactions. A 486 IBM compatible microcomputer was used to control the spectrophotometer and to acquire and treat the data.

Data acquisition began 10 s after starting the reactions. The data files produced by the HP 89531A operation software (Hewlett-Packard) were processed by the author's own programmes which were written in MATLAB (MathWorks, Sherborn, MA, USA).

3.2. Reagents, solutions and procedures

Reagent grade *o*-, *m*- and *p*-amino benzoic acid (ABA) (Merck, Darmstadt, Germany), orciprenaline (ORC) (kindly donated by Boehringer-Ingelheim, Barcelona, Spain), sulfanilamide (Sigma, St. Louis, MO, USA), sulfamic acid, sodium nitrite and sodium dodecyl sulphate (SDS) (Fluka, Buchs, Switzerland), and citric acid (Panreac, Barcelona, Spain) were used. Distilled demineralized water (Barnstead, Sybron, Taunton, MA, USA) was used throughout. The molecular structure of the reagents is shown in Fig. 2. The coupling reaction is

$$\underset{\text{Substrate}}{\text{H–Ar–R}} + \underset{\text{Electrophilic agent}}{\text{Ar'–N}\equiv\text{N}^+}$$

$$\rightarrow \underset{\text{Azodye}}{\text{Ar'–N=N–Ar–R}} + \text{H}^+$$

To buffer the pH, mixtures containing citric acid and sodium hydroxide were prepared. The final buffer concentration was 0.25 mol l^{-1} in all cases. To prepare the stock solutions of the substrates, 17.5 mg *o*-ABA, 17.5 mg *m*-ABA, 18.5 mg *p*-ABA and 56.3 mg ORC were dissolved in 1 ml ethanol, and diluted

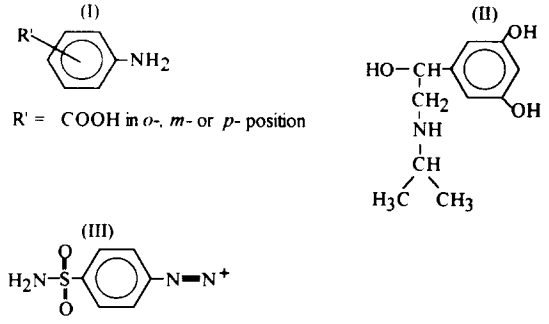

Fig. 2. Molecular structure of the reagents: (I) amino benzoic acids; (II) orciprenaline; (III) diazonium ion of sulfanilamide.

Table 1
First-order rate constants for *o*-, *m*-, and *p*-ABA, and ORC at different pH values in a 2% SDS medium

Substrates [a]	λ_{max} (nm)	pH	$k \pm s_k$ ($s^{-1} \times 10^{-3}$) [b]
o-ABA	370	4.10	6.31 ± 0.12
ORC	406	4.10	1.20 ± 0.03
m-ABA	360	3.90	8.16 ± 0.40
p-ABA	368	3.90	1.67 ± 0.11
o-ABA	370	3.80	4.48 ± 0.08
m-ABA	360	3.80	7.14 ± 0.07

[a] ABA, amino benzoic acid; ORC, orciprenaline.
[b] Obtained from three solutions with different substrate concentrations.

with water to 50 ml. The 4×10^{-2} mol l^{-1} sulfanilamide stock solution was prepared in 0.3 mol l^{-1} HCl. The 0.2 mol l^{-1} $NaNO_2$, 0.5 mol l^{-1} sulfamic acid and 20% SDS solutions were made with water.

To prepare the 1×10^{-2} mol l^{-1} diazonium ion solution, 12.5 ml sulfanilamide was introduced in a 50 ml volumetric flask, 15 ml $NaNO_2$ solution was added, the mixture was allowed to react for 10 min, 15 ml sulfamic acid was added to destroy the excess nitrite, and after 15 min the solution was made up to the mark with water. The diazonium ion solution was renewed daily.

Binary mixtures of the substrates were prepared by introducing aliquots of the corresponding stock solutions, 20 ml buffer solution and 2.5 ml SDS into a 25 ml volumetric flask, and the volume was completed with water. A 2.25 ml volume was transferred into a dry 1-cm quartz cell, the reaction was started by injecting 0.25 ml diazonium ion solution, and a keyboard button was simultaneously pressed to start the data acquisition period ($t = 0$). Mixing was facilitated by bubbling the solution in the cell four times with the piston pipette. The spectra were scanned every 30 s, from $t = 10$ s to $t = 910$ s. One point every 2 nm was obtained (the data acquisition resolution of the HP8452A spectrophotometer). The blank was prepared in the same way in the absence of the substrate. The blank absorbance matrix was always subtracted from the sample absorbance matrices before processing the data. Owing to the perturbations and

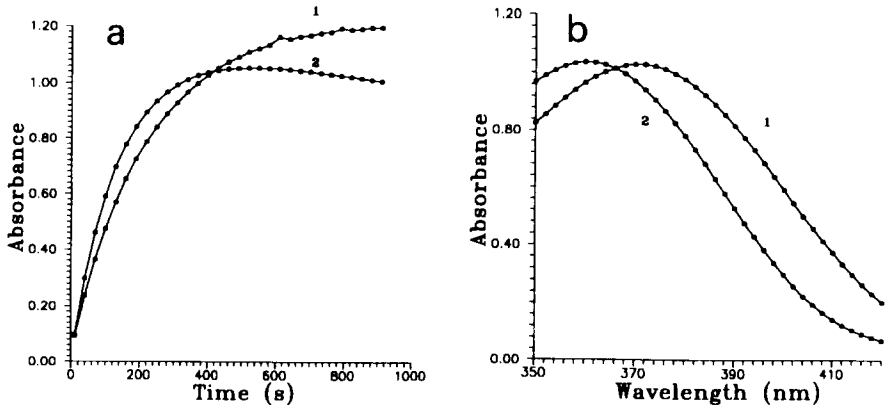

Fig. 3. Kinetic curves (left) and azodye spectra (right) for 5.06×10^{-5} M *o*-ABA (1) and *m*-ABA (2).

noise, the first spectrum ($t = 10$ s) was rejected and, therefore, each kinetic data set consisted of 30 spectra.

4. Results and discussion

4.1. Optimization of the reaction conditions

Since only the basic form of the substrates reacts, coupling with diazonium ions is largely affected by pH [34]. Thus, at pH values lower than 3.5, the reactions were too slow, and at pH values higher than 5.5, some of the analytes coupled in less than 1 min, which was not suitable for the manual mixing procedure used. Also, at pH values higher than 5.5, the absorbance of the reagent blank was large. This has been shown to be due to hydrolysis of the diazonium ion to yield a phenol which couples with the excess reagent [35]. Also, at higher pH values, the azo dyes were unstable, and the absorbance decreased rapidly after reaching a maximum value. It was observed that the addition of SDS alleviated this problem and, therefore, a final concentration of 2% SDS was used.

To evaluate the rate constants, three solutions of each substrate at increasing concentrations were prepared, the procedure given above was applied, and the

Table 2
Composition of the mixtures used for calibration (orthogonal design) and evaluation of the PLS procedures

Experiment No. [a]	*o*-ABA/ORC [b] ($\times 10^{-5}$ mol l^{-1})	*m*-ABA/*p*-ABA [b] ($\times 10^{-5}$ mol l^{-1})	*o*-ABA/*m*-ABA [b] ($\times 10^{-5}$ mol l^{-1})
1	2.759/3.456	1.839/1.945	2.299/2.299
2	4.139/3.456	3.219/1.945	2.299/3.679
3	5.518/3.456	4.598/1.945	2.299/5.058
4	2.759/5.184	1.839/3.403	3.679/2.299
5	4.139/5.184	3.219/3.403	3.679/3.679
6	5.518/5.184	4.598/3.403	3.679/5.058
7	2.759/6.912	1.839/4.861	5.058/2.299
8	4.139/6.912	3.219/4.861	5.058/3.679
9	5.518/6.912	4.598/4.861	5.058/5.058
10	3.679/4.320	2.759/3.889	3.219/2.759
11	3.679/6.048	3.679/2.917	4.139/4.599
12	4.599/4.320	2.299/3.403	4.139/2.759
13	4.599/6.048	3.219/2.431	3.219/4.599

[a] 1–9, calibration set; 10–13, evaluation set.
[b] ABA, amino benzoic acid; ORC, orciprenaline.

corresponding kinetic curves were fitted. The Powell nonlinear fitting algorithm was used [36]. The results are given in Table 1, and showed a large variation of the rate constants with pH. The *o*-ABA/ORC mixtures had the largest spectral and kinetic differences between the two analytes, whereas for the *o*-ABA/*m*-ABA mixtures, the ratio of the rate constants was the smallest, of about 1.6, and the spectra overlapped seriously. The kinetic curves and the spectra of the azo dyes obtained with two *o*-ABA and *m*-ABA solutions are shown in Fig. 3. It can be observed that the *m*-ABA azodye was not stable, which is not considered in Eq. (1).

4.2. Multivariate calibration

To establish the calibration sets, nine mixtures of each binary combination of substrates were prepared following an orthogonal design. To evaluate the predictive capacity of the PLS methods for each combination of substrates, the corresponding evaluation sets were prepared. The concentrations used were within the ranges embraced by the calibration sets (Table 2). For each binary combination, and on the basis of single wavelength kinetic curves, a two-way PLS calibration model was constructed. Three-way modelling was applied to series of three-way data arrays con-

Table 3
The estimated error AARE of the evaluation mixtures (experiments 10–13 listed in Table 2) by using two- and three-way PLS with different sizes

Substrates [a]	No. of wavelengths	Wavelength range used in the evaluation ($\Delta\lambda$) and the corresponding AARE (%) [b]									
o-ABA/ORC	1	$\Delta\lambda$:	350	362	374	386	398	410	422	434	422
		AARE:	8.5	6.5	4.0	2.5	2.5	2.9	3.6	5.0	6.1
	6	$\Delta\lambda$:	350–360	374–384	386–396	398–408	410–420	434–444			
		AARE:	7.5	2.6	2.4	2.8	3.1	3.9			
	12	$\Delta\lambda$:	350–372	374–396	398–420	422–444					
		AARE:	5.1	2.4	3.0	4.1					
	24	$\Delta\lambda$:	350–396	398–444							
		AARE:	2.4	3.1							
	48	$\Delta\lambda$:	350–444								
		AARE:	2.6								
m-ABA/*p*-ABA	1	$\Delta\lambda$:	350	362	378	394	406	414	422	430	
		AARE:	4.5	4.2	4.1	4.0	3.8	4.6	9.1	15	
	5	$\Delta\lambda$:	350–358	370–378	380–388	390–398	400–408	420–428			
		AARE:	4.4	4.1	4.0	4.0	3.1	5.1			
	10	$\Delta\lambda$:	350–368	370–388	390–408	410–428					
		AARE:	4.3	4.0	3.1	2.6					
	20	$\Delta\lambda$:	350–388	390–428							
		AARE:	4.1	3.1							
	41	$\Delta\lambda$:	350–430								
		AARE:	4.1								
o-ABA/*m*-ABA	1	$\Delta\lambda$:	350	358	366	374	382	390	398	406	414	420
		AARE:	13	11	14	17	17	19	21	23	25	15
	6	$\Delta\lambda$:	350–360	362–372	374–384	386–420	398–408	410–420				
		AARE:	5.3	3.5	4.5	3.4	4.4	9.1				
	12	$\Delta\lambda$:	350–372	374–396	398–420							
		AARE:	4.2	4.2	3.3							
	24	$\Delta\lambda$:	350–386	388–420								
		AARE:	5.9	4.1								
	36	$\Delta\lambda$:	350–420									
		AARE:	4.1									

[a] ABA, amino benzoic acid; ORC, orciprenaline.
[b] AARE is the average absolute relative error defined in Eq. (12).

sisting of a number of selected wavelengths each. The number of selected wavelengths was increased until the whole spectra was included in the calibration model. The average absolute relative error (AARE) used for the evaluation of the results was

$$\mathrm{AARE} = \frac{\sum_{n=1}^{N}\sum_{l=1}^{L}\left|\frac{\hat{c}_{nl}-c_{nl}}{c_{nl}}\right| \times 100}{N \times L} \qquad (12)$$

where $\hat{c}_{nl}$ is the estimate for the true concentration, c_{nl}.

The results of two-way PLS for the three combinations of substrates are given in Table 3. The results obtained by three-way PLS with an increasing number of wavelengths are also listed in the same table, and the errors yielded when the whole spectrum was used are given in Table 4. As shown in Table 3, the *o*-ABA/ORC mixtures gave satisfactory results with most of the single wavelength PLS. In contrast, for the *m*-ABA/*p*-ABA mixtures, the PLS model based on single wavelength information gave a much higher error. These mixtures presented a similar rate constant ratio but a more serious spectral overlap than the *o*-ABA/ORC mixtures which can explain the difference. With both combinations of substrates, the results were worse at the ends of the wavelength range, where the molar absorptivities were low, and therefore, where the absorbance change produced by the reaction was small. When the information taken at several wavelengths was simultaneously used by three-way PLS, the results improved. However, when the wavelengths associated with larger values of the molar absorptivity were used, no advantage of the three-way PLS was obtained with respect to the single wavelength PLS. On the contrary, the single wavelength PLS method did not provide acceptable results when applied to the *o*-ABA/*m*-ABA mixtures (AARE was higher than 10%). With three-way PLS, and when the number of selected wavelengths was increased, the results improved dramatically. The results given by at least 12 wavelengths were quite acceptable.

The optimal number of the latent variables used in the PLS modelling was different for the three combinations of substrates. For the *o*-ABA/ORC and *m*-ABA/*p*-ABA pairs, the optimal number of latent variables retained in the PLS model was always two to three, no matter how many wavelengths were used to construct the PLS model. In contrast, for the *o*-ABA/*m*-ABA mixtures, and when a single wavelength was used for modelling, usually seven to eight latent variables were required. When the number of wavelengths increased, the number of latent variables for optimal prediction decreased, and when the whole spectrum was used, two latent variables were

Table 4
Estimated concentrations and relative errors for the evaluation mixtures (Nos. 10–13 in Table 2) by three-way PLS based on the whole spectrum (48, 41 and 36 wavelengths for the *o*-ABA/ORC, *m*-ABA/*p*-ABA and *o*-ABA/*m*-ABA mixtures, respectively)

Experiment No.	*o*-ABA (mol $l^{-1} \times 10^{-5}$)	Deviation (%)	ORC (mol $l^{-1} \times 10^{-5}$)	Deviation (%)
10	3.596	−2.3	4.254	−1.5
11	3.675	−0.1	5.901	−2.4
12	4.732	2.9	4.268	−1.2
13	4.432	−3.6	5.638	−6.8
	m-ABA (mol $l^{-1} \times 10^{-5}$)	Deviation (%)	*p*-ABA (mol $l^{-1} \times 10^{-5}$)	Deviation (%)
10	2.743	−0.6	3.873	−0.4
11	3.736	1.5	2.782	−4.4
12	2.545	11.	3.551	4.3
13	3.301	2.6	2.238	−7.9
	o-ABA (mol $l^{-1} \times 10^{-5}$)	Deviation (%)	*m*-ABA (mol $l^{-1} \times 10^{-5}$)	Deviation (%)
10	3.312	−2.7	2.872	4.1
11	4.186	1.1	4.713	2.5
12	4.456	7.7	2.934	6.3
13	3.058	−5.0	4.737	3.0

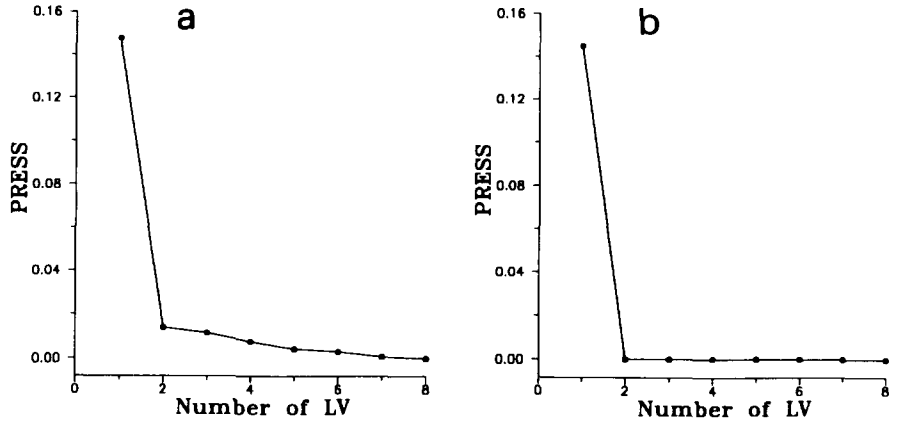

Fig. 4. Plot of predictive error sum of squares (PRESS) versus number of latent variables for the *o*-ABA/*m*-ABA mixtures when a single wavelength (384 nm, left) and the whole spectrum (right) were used.

enough to provide a good prediction, which is consistent with the fact that the system contained only two independent variables. The PRESS (prediction error sum of squares) versus the number of latent variables for a single wavelength and for the whole spectrum of the *o*-ABA/*m*-ABA mixtures is given in Fig. 4.

5. Conclusion

With three-way PLS the potential of the diode array detector to quasi-simultaneously provide kinetic and spectral data can be fully exploited. In comparison to PLS applied to single wavelength kinetic data, advantages of the three-way PLS are found in solving binary mixtures with both a low rate constant ratio and small spectral differences, or alternatively, with small relative changes in the signal. In addition, it has been shown that both PLS methods gave good results in the presence of the side reactions of decomposition of the azo dyes. This could be due to the linear nature of the effect produced by the side-reaction on the response, if the model of Eq. (3) was followed, but also to the potential of PLS and other linear soft-modelling methods to take care of moderate deviations from linearity. This would not be the case for any hard-modelling procedure, where the factors not considered in the model would degrade their performance very much. Application of hard-modelling techniques to industrial, environmental or biomedical samples is limited owing to the interference and matrix effects produced by the many factors which are ignored, or which would complicate the model excessively. However, as far as a linear behavior would be present, and in spite of the complexity of the matrix, PLS yields good results, which should stimulate the development of applications in these fields.

Acknowledgements

This work was supported by the DGICYT of Spain, Project PB93/355. Y.L.X. thanks the Ministry of Education and Science of Spain for a postdoctoral grant.

References

[1] A.K. Smilde, Three-way analysis, problems and prospects, *Chemometrics and Intelligent Laboratory Systems*, 15 (1992) 143–157.

[2] S. Wold, P. Geladi, Kim Esbensen and J. Ohman, Multi-way principal components and PLS-analysis, *Journal of Chemometrics*, 1 (1987) 41–56.

[3] E. Sanchez and B.R. Kowalski, Tensorial calibration. I. First-order calibration, *Journal of Chemometrics*, 2 (1988) 247.

[4] E. Sanchez and B.R. Kowalski, Tensorial calibration. II. Second order calibration, *Journal of Chemometrics*, 2 (1988) 265–280.

[5] E. Sanchez and B.R. Kowalski, Tensorial resolution: a direct trilinear decomposition, *Journal of Chemometrics*, 4 (1990) 29.

[6] D.S. Burdick, X.M. Tu, L.B. McGown and D.W. Millican, Resolution of multicomponent fluorescent mixtures by analysis of the excitation–emission frequency array, *Journal of Chemometrics*, 4 (1990) 15.

[7] S.S. Li and P.J. Gemperline, Eliminating complex eigenvectors and eigenvalues in multiway analysis, using the direct trilinear decomposition method, *Journal of Chemometrics*, 7 (1993) 77–88.

[8] A.K. Smilde and D.A. Doornbos, Three-way methods for the calibration of chromatographic systems: comparing PARAFAC and three-way PLS, *Journal of Chemometrics*, 5 (1991) 345–360.

[9] A.K. Smilde, P.H. van de Graaf, D.A. Doornbos, T. Steerneman and A. Sleurink, Multivariate calibration of reversed-phase chromatographic systems. Some designs based on three-way data analysis, *Analytica Chimica Acta*, 235 (1990) 41–51.

[10] A.K. Smilde and D.A. Doornbos, Simple validatory tools for judging the predictive performance of PARAFAC and three-way PLS, *Journal of Chemometrics*, 6 (1992) 11–28.

[11] Y. Zeng and P. Hopke, Methodological study applying three-mode factor analysis to three-way chemical data sets, *Chemometrics and Intelligent Laboratory Systems*, 7 (1990) 237–250.

[12] J. Ohman, P. Geladi and S. Wold, Residual bilinearization. Part 1. Theory and algorithm, *Journal of Chemometrics*, 4 (1990) 79–90.

[13] J. Ohman, P. Geladi and S. Wold, Residual bilinearization. Part II. Application to HPLC–DAD data and comparison with rank annihilation factor analysis, *Journal of Chemometrics*, 4 (1990) 135–146.

[14] Y.Z. Liang and O.M. Kvalheim, Constrained background bilinearization, *Chemometrics and Intelligent Laboratory Systems*, 12 (1992) 646–657.

[15] Y.L. Xie, Y.Z. Liang and R.Q. Yu, Constrained background bilinearization with generalized simulated annealing algorithm, *Journal of Chemometrics*, 7 (1993) 369–379.

[16] M. Silva, Recent strategies in automated reaction-rate based determination, *Analyst*, 118 (1993) 681–688.

[17] H.B. Mark Jr. and G.A. Rechnitz, *Kinetics in Analytical Chemistry*, Interscience, New York, 1968.

[18] H.A. Mottola, *Kinetic Aspects of Analytical Chemistry*, Wiley, New York, 1988.

[19] J. Havel, J.L. González and M.N. Moreno, Computation of kinetics. 2. Simultaneous regression estimation of rate constants and initial concentration, *Reaction Kinetics and Catalysis Letters*, 39 (1989) 41–48.

[20] J.J. Baeza-Baeza, G. Ramis-Ramos, F. Pérez-Plá and R. Valero-Molina, Multi-component analysis using OPKINE, a program for the non-linear treatment of kinetic problems, *Analyst*, 115 (1990) 721.

[21] F. Pérez-Plá, J.J. Baeza-Baeza, G. Ramis-Ramos and J. Palou, OPKINE, a multipurpose program for kinetics, *Journal of Computational Chemistry*, 12 (1991) 283.

[22] P.D. Wentzell, M.I. Kazagannis and S.R. Crouch, Simultaneous kinetic determinations with the Kalman filter, *Analytica Chimica Acta*, 224 (1989) 263–274.

[23] W.H. Lewis Jr. and S.C. Rutan, Guanidinium-induced differential kinetic denaturation of alkaline phosphatase isozymes, *Analytical Chemistry*, 63 (1991) 627–629.

[24] E. Forster, M. Silva, M. Otto and D. Pérez-Bendito, Enzymatic determination of alcohol mixtures at the nanogram level by the stopped flow technique, *Analytica Chimica Acta*, 274 (1993) 109–116.

[25] R. Xiong, A. Velasco, M. Silva and D. Pérez-Bendito, Performance of the Kalman filter algorithm in differential reaction-rate methods, *Analytica Chimica Acta*, 251 (1991) 313–319.

[26] A. Velasco, R. Xiong, M. Silva and D. Pérez-Bendito, Simultaneous kinetic determination of phenols by use of the Kalman filter, *Talanta*, 40 (1993) 1505–1510.

[27] S.C. Rutan and S.D. Brown, Estimation of first-order kinetic parameters by using the extended Kalman filter, *Analytica Chimica Acta*, 167 (1985) 23–27.

[28] B.M. Quencer and S.R. Crouch, Extended Kalman filter for multiwavelength, multicomponent kinetic determinations, *Analyst*, 118 (1993) 695–701.

[29] B.M. Quencer and S.R. Crouch, Multicomponent kinetic determination of Lanthanides with stopped-flow, diode array spectrophotometry and extended Kalman filters, *Analytical Chemistry*, 66 (1994) 458–463.

[30] R. Jiménez-Prieto, A. Velasco, M. Silva and D. Pérez-Bendito, Kalman filtering of data from first- and second-order kinetics, *Talanta*, 40 (1993) 1731–1739.

[31] J. Havel, F. Jiménez, R.D. Bautista and J.J. Arias León, Evaluation of multi-component kinetic analysis data by a partial least squares calibration method, *Analyst*, 118 (1993) 1355–1360.

[32] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, J. Riba and E. Rovira, Kinetic spectrophotometric determination of Ga(III)–Al(III) mixtures by stopped-flow injection analysis using principal component regression, *Talanta*, 40 (1993) 261–267.

[33] S. Wold, K. Esbensen and P. Geladi, Principal component analysis, *Chemometrics and Intelligent Laboratory Systems*, 2 (1987) 37–52.

[34] G. Ramis-Ramos, J.S. Esteve Romero and M.C. García Alvarez-Coque, Colorimetric determination of arylamines and sulphonamides by diazotization and coupling in a micellar solution, *Analytica Chimica Acta*, 223 (1989) 327–337.

[35] J.S. Esteve Romero, M.C. García Alvarez-Coque and G. Ramis-Ramos, Formation rates and protonation constants of azo dyes in a sodium dodecylsulphate micellar solution, *Talanta*, 38 (1991) 1285–1289.

[36] S.S. Rao, *Optimization: Theory and Applications*, Wiley Eastern Limited, New Delhi, 2nd edn., 1984.