Journal of Chromatography A, 1216 (2009) 7063–7070

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

# Fast chromatographic method for the determination of dyes in beverages by using high performance liquid chromatography—Diode array detection data and second order algorithms

María J. Culzoni[a], Agustina V. Schenone[a], Natalia E. Llamas[b], Mariano Garrido[b], Maria S. Di Nezio[b], Beatriz S. Fernández Band[b,*], Héctor C. Goicoechea[a,*]

[a] Laboratorio de Desarrollo Analítico y Quimiometría (LADAQ), Cátedra de Química Analítica I, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral – CONICET, Ciudad Universitaria, Santa Fe (S3000ZAA), Argentina
[b] FIA Laboratory, INQUISUR – CONICET, Universidad Nacional del Sur, Av. Alem 1253, (B8000CPB) – Bahía Blanca, Buenos Aires, Argentina

ARTICLE INFO

*Article history:*
Received 24 June 2009
Received in revised form 25 August 2009
Accepted 28 August 2009
Available online 1 September 2009

*Keywords:*
Second order advantage
Dyes
Beverages samples
HPLC-DAD
Multivariate curve resolution alternating least squares
Residual bilinearization

ABSTRACT

A fast chromatographic methodology is presented for the analysis of three synthetic dyes in non-alcoholic beverages: amaranth (E123), sunset yellow FCF (E110) and tartrazine (E102). Seven soft drinks (purchased from a local supermarket) were homogenized, filtered and injected into the chromatographic system. Second order data were obtained by a rapid LC separation and DAD detection. A comparative study of the performance of two second order algorithms (MCR-ALS and U-PLS/RBL) applied to model the data, is presented. Interestingly, the data present time shift between different chromatograms and cannot be conveniently corrected to determine the above-mentioned dyes in beverage samples. This fact originates the lack of trilinearity that cannot be conveniently pre-processed and can hardly be modelled by using U-PLS/RBL algorithm. On the contrary, MCR-ALS has shown to be an excellent tool for modelling this kind of data allowing to reach acceptable figures of merit. Recovery values ranged between 97% and 105% when analyzing artificial and real samples were indicative of the good performance of the method. In contrast with the complete separation, which consumes 10 mL of methanol and 3 mL of 0.08 mol $L^{-1}$ ammonium acetate, the proposed fast chromatography method requires only 0.46 mL of methanol and 1.54 mL of 0.08 mol $L^{-1}$ ammonium acetate. Consequently, analysis time could be reduced up to 14.2% of the necessary time to perform the complete separation allowing saving both solvents and time, which are related to a reduction of both the costs per analysis and environmental impact.



## 1. Introduction

The use of second order multivariate algorithms has been shown to play a critical role in several analytical fields, as can be gathered from a literature survey in relevant analytical, chemometrics and applied journals [1,2]. One of the ways to obtain second order data is by coupling two "hyphenated" first order instruments, as in tandem high performance liquid chromatography-diode array detector (HPLC-DAD), gas chromatography–mass spectrometry (GC–MS), GC–GC, MS–MS, etc. measurements. Specifically, an important number of reports have been presented focusing on the resolution of really complex samples by using liquid chromatography [3] and exploiting the second order advantage [4]. In this field, important issues such as reduction in the time of analysis and consequently costs and amount of contaminant solvents should be considered.

A large number of algorithms can be cited among the approaches involving the second order advantage: generalized rank annihilation (GRAM) [5], direct trilinear decomposition (DTLD) [6], self-weighted alternating trilinear decomposition (SWATLD) [7], alternating penalty trilinear decomposition (APTLD) [8], parallel factor analysis (PARAFAC) [9], multivariate curve resolution alternating least squares (MCR-ALS) [10–12], and the most recently implemented bilinear least squares (BLLS) [13], unfolded partial least squares/residual bilinearization (U-PLS/RBL) [14,15] and artificial neural networks followed by residual bilinearization (ANN/RBL) [16].

In chromatography, the retention factor ($k$) is the degree of retention of the sample component in the column. In most chromatographic analysis, analytes elute with retention factors between 1 and 20 allowing their complete separation. A peak with $k$ equal to 0 is a component that does not interact with the stationary phase and elutes in the void volume [17]. If analytes are in

* Corresponding authors. Tel.: +54 342 4575205.
*E-mail addresses:* usband@criba.edu.ar (B.S. Fernández Band), hgoico@fbcb.unl.edu.ar (H.C. Goicoechea).

0021-9673/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2009.08.077

their ionized form, $k$ will be significantly lower than the one that would be obtained if they were in their neutral form [17]. A value of $k$ lower than 1 does not result in a differential migration of the component and originates dissimilarities in both time of elution and peak shapes, leading to data without the property of trilinearity, and making necessary the use MCR-ALS (which can solve this type of problems by resorting to the mathematical resource of matrix augmentation) or alternatives such as PARAFAC2, a variant of PARAFAC allowing for distinct time profiles in each experimental sample [18]. However, when data are conveniently pre-treated in order to alleviate the above-mentioned problems, good results can be obtained by using PARAFAC or RBL based algorithms [19,20].

Synthetic dyes are usually added to foodstuffs and soft drinks not only to improve appearance, colour and texture but also to maintain the natural colour during process or storage. Synthetic dyes show several advantages compared with natural dyes such as high stability to light, oxygen and pH, colour uniformity, low microbiological contamination and relatively lower production costs [21]. However, many of them may exhibit adverse health effects (allergy, respiratory problems, thyroid tumours, chromosomal damage, hyperactivity, abdominal pain, etc.) [22].

On the other hand, in some cases the use of food dyes is also indicative of foodstuff adulteration such as in their addition to fruit juices. Thus, the use of synthetic dyes is strictly controlled by laws, regulations and acceptable daily intake (ADI) values [23]. These regulations frame the role of the analytical chemist who has to test for the levels of dyes added to food. Amaranth (E123), sunset yellow FCF (E110) and tartrazine (E102) are among the synthetic dyes mainly used in non-alcoholic beverages and the ADI values are between 0 and 0.5 mg kg$^{-1}$ for amaranth (A), 0–2.5 mg kg$^{-1}$ for sunset yellow (SY) and 0–7.5 mg kg$^{-1}$ for tartrazine (T) [23].

Some problems found in artificial colour determination are related to the variety of dyes mixtures and the potential interferences present in the commercial samples. Therefore, the analyses have traditionally been focused on separation methods. The analytical methods frequently used for the determination of amaranth, sunset yellow and tartrazine include thin layer chromatography (TLC) [24,25], capillary electrophoresis (CE) [26,27], and mainly high performance liquid chromatography (HPLC) [28,29]. However, some disadvantages arise from these methods, such as usage of toxic solvents, spending of time, and the need of sample pretreatments. Chemometric modelling was used to overcome the lack of specificity due to spectra overlapping for the direct UV–vis spectrophotometric determination, representing a rapid, simple, and cheap method for the determination of these colourants [30,31]. Very recently, a spectrophotometric method employing the multivariate curve resolution alternating least squares algorithm was presented by our group [32].

In this work we present a comparative study of the performance of two algorithms: MCR-ALS and U-PLS/RBL on data obtained by a rapid LC separation and DAD detection, which present time shift between different chromatograms and cannot be conveniently corrected, to determine the above-mentioned dyes in beverage samples.

## 2. Theory

### 2.1. MCR-ALS

A data set can be considered trilinear when the resolved profiles of the same component in different data matrices for a particular direction ($\mathbf{C}$ or $\mathbf{S}^{\mathrm{T}}$) have the same shape [33]. MCR-ALS is capable of dealing with data sets deviating from trilinearity. Instead of forming a three-dimensional data array, the latter is unfolded along the mode which is suspected of breaking the trilinear structure, i.e. if a matrix-to-matrix variation of profiles occurs along the column direction, a column-wise augmented matrix is created. The bilinear decomposition of the augmented matrix $\mathbf{D}$ is performed according to the expression:

$$\mathbf{D} = \mathbf{C} \times \mathbf{S}^{\mathrm{T}} + \mathbf{E} \tag{1}$$

where the rows of $\mathbf{D}$ contain the absorption spectra measured as a function of time, the columns of $\mathbf{C}$ contain the time profiles of the compounds involved in the process, the columns of $\mathbf{S}$ their related spectra, and $\mathbf{E}$ is a matrix of residuals not fitted by the model. Appropriate dimensions of $\mathbf{D}$, $\mathbf{C}$, $\mathbf{S}$ and $\mathbf{E}$ are thus $K \times (1+I) \times J$, $K \times (1+I) \times F$, $J \times F$ and $K \times (1+I) \times J$, respectively ($I$ = number of training samples, $J$ = number of digitized wavelengths, $F$ = number of factors extracted and $K$ = number of elution times). Decomposition of $\mathbf{D}$ is achieved by iterative least squares minimization of $||\mathbf{E}||$ under suitable constraining conditions, i.e. nonnegativity in spectral profiles, unimodality and nonnegativity in concentration profiles.

In the present context, it is necessary to point out that MCR-ALS requires initialization with system parameters as close as possible to the final results. In the column-wise augmentation mode, the species spectra are required, as obtained from either pure analyte standards or from the analysis of the purest spectra.

### 2.2. U-PLS/RBL

The U-PLS algorithm has a calibration step, which employs concentration information (without including data for the unknown sample). First, the $I$ calibration data matrices $\mathbf{X}_{c,i}$ (size $J \times K$, where $J$ and $K$ are the number of channels in each dimension) are vectorized (unfolded) and a usual U-PLS model is calibrated with these data and the vector of calibration concentrations $\mathbf{y}$ ($N_c \times 1$, where $N_c$ is the number of calibration samples). This provides a set of loadings $\mathbf{P}$ and weight loadings $\mathbf{W}$ (both of size $JK \times A$, where $A$ is the number of latent factors), as well as regression coefficients $\mathbf{v}$ (size $A \times 1$). The parameter $A$ can be selected by techniques such as leave-one-out cross-validation [34] or by the randomization test proposed by Farber and co-workers [35]. If no unexpected interferences occur in the test sample, $\mathbf{v}$ can be employed to estimate the analyte concentration:

$$y_u = \mathbf{t}_u^{\mathrm{T}} \mathbf{v} \tag{2}$$

where $\mathbf{t}_u$ is the test sample score, obtained by projection of the unfolded data for the test sample $\mathbf{X}_u$ onto the space of the $A$ latent factors:

$$\mathbf{t}_u = (\mathbf{W}^{\mathrm{T}}\mathbf{P})^{-1}\mathbf{W}^{\mathrm{T}}\mathrm{vec}(\mathbf{X}_u) \tag{3}$$

When unexpected constituents occur in $\mathbf{X}_u$, then the sample scores given by Eq. (3) are not suitable for analyte prediction using Eq. (2). In this case, the residuals of the U-PLS prediction step, represented by $s_p$ in Eq. (4), will be abnormally large in comparison with the typical instrumental noise, which can be easily assessed by replicate measurements:

$$\begin{aligned} s_p &= ||\mathbf{e}_p||/(JK - A)^{1/2} \\ &= ||\mathrm{vec}(\mathbf{X}_u) - \mathbf{P}(\mathbf{W}^{\mathrm{T}}\mathbf{P})^{-1}\mathbf{W}^{\mathrm{T}}\mathrm{vec}(\mathbf{X}_u)||/(JK - A)^{1/2} \\ &= ||\mathrm{vec}(\mathbf{X}_u) - \mathbf{P}\mathbf{t}_u||/(JK - A)^{1/2} \end{aligned} \tag{4}$$

where $||\cdot||$ indicates the Euclidean norm.

If the test sample contains unexpected components, the situation can be handled by a separate procedure called residual bilinearization, which is based on singular value decomposition (SVD) modelling of the interferent effects [14,15]. RBL aims at minimizing the norm of the residual vector $\mathbf{e}_u$, computed while fitting

**Table 1**
Composition and prediction of the ternary mixtures used for calibration and validation.

| Sample | Tartrazine (mg L$^{-1}$) | | | Amaranth (mg L$^{-1}$) | | | Sunset yellow (mg L$^{-1}$) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Nom. | MCR[a] | RBL[b] | Nom. | MCR[a] | RBL[b] | Nom. | MCR[a] | RBL[b] |
| 1 | 5.89 | 6.13 | 5.85 | 2.20 | 2.20 | 2.05 | 12.09 | 11.95 | 12.31 |
| 2 | 2.61 | 2.26 | 2.80 | 1.05 | 1.15 | 1.07 | 32.91 | 33.80 | 34.43 |
| 3 | 5.89 | 5.81 | 5.44 | 1.05 | 1.11 | 1.05 | 12.09 | 12.50 | 12.89 |
| 4 | 2.61 | 2.56 | 2.48 | 2.20 | 2.24 | 2.25 | 32.91 | 33.87 | 34.31 |
| 5 | 4.25 | 4.09 | 4.53 | 0.65 | 0.74 | 0.67 | 22.50 | 22.80 | 22.75 |
| 6 | 2.61 | 2.72 | 2.38 | 2.20 | 2.20 | 2.14 | 12.09 | 12.00 | 11.72 |
| 7 | 1.50 | 1.39 | 1.58 | 1.63 | 1.67 | 1.56 | 22.50 | 23.08 | 23.94 |
| 8 | 4.25 | 4.46 | 4.43 | 1.63 | 1.64 | 1.62 | 5.00 | 5.58 | 5.84 |
| 9 | 4.25 | 4.47 | 4.32 | 2.60 | 2.70 | 2.78 | 22.50 | 23.37 | 23.75 |
| 10 | 5.89 | 5.92 | 6.26 | 2.20 | 2.22 | 2.17 | 32.91 | 34.09 | 34.45 |
| 11 | 5.89 | 5.64 | 5.96 | 1.05 | 1.16 | 1.03 | 32.91 | 33.87 | 34.38 |
| 12 | 2.61 | 2.60 | 2.66 | 1.05 | 1.12 | 1.13 | 12.09 | 12.27 | 13.78 |
| 13 | 7.00 | 7.11 | 6.96 | 1.63 | 1.68 | 1.66 | 22.50 | 23.17 | 24.43 |
| 14 | 4.25 | 4.21 | 4.20 | 1.63 | 1.69 | 1.66 | 22.50 | 22.66 | 24.36 |
| 15 | 4.25 | 4.00 | 3.98 | 1.63 | 1.69 | 1.61 | 40.00 | 41.09 | 44.29 |
| Mean recovery (%)[c] | – | 98.8(5.1) | 100.3(5.3) | – | 104.4(4.1) | 100.4(3.9) | – | 102.7(2.9) | 106.3(5.0) |
| REP (%)[d] | – | 4.1 | 4.9 | – | 3.9 | 4.4 | – | 3.1 | 7.5 |

[a] MCR-ALS: three factors (regions: see the text).
[b] U-PLS/RBL: five latent variables for each analyte (regions: see the text).
[c] Between parenthesis the standard deviation.
[d] Relative error of prediction, $\mathrm{REP} = \frac{100}{\bar{c}}\left[\frac{1}{I}\sum_{1}^{I}(c_{\mathrm{act}} - c_{\mathrm{pred}})^2\right]^{1/2}$, where $I$ is the number of samples, $c_{\mathrm{act}}$ and $c_{\mathrm{pred}}$ are the actual and predicted concentrations, and $\bar{c}$ is the mean concentration.

the sample data to the sum of the relevant contributions to the sample signal. For a single unexpected component:

$$\mathrm{vec}(\mathbf{X}_u) = \mathbf{P}\mathbf{t}_u + \mathrm{vec}[g_{unx}\mathbf{b}_{unx}(\mathbf{c}_{unx})^T] + \mathbf{e}_u \quad (5)$$

where $\mathbf{b}_{unx}$ and $\mathbf{c}_{unx}$ are the left and right eigenvectors of $\mathbf{E}_p$ and $g_{unx}$ is a scaling factor:

$$(g_{unx}, \mathbf{b}_{unx}, \mathbf{c}_{unx}) = \mathrm{SVD}_1(\mathbf{E}_p) \quad (6)$$

where $\mathbf{E}_p$ is the $J \times K$ matrix obtained after reshaping the $JK \times 1$ $\mathbf{e}_p$ vector of Eq. (4), and $\mathrm{SVD}_1$ indicates taking the first principal component.

During this RBL procedure, $\mathbf{P}$ is kept constant at the calibration values and $\mathbf{t}_u$ is varied until $||\mathbf{e}_u||$ is minimized. The minimization can be carried out using either a Gauss–Newton (GN) procedure or an alternating algorithm, in both cases starting with $\mathbf{t}_u$ from Eq. (2). Once $||\mathbf{e}_u||$ is minimized in Eq. (5), the analyte concentrations are provided by Eq. (2), by introducing the final $\mathbf{t}_u$ vector found by the RBL procedure.

The number of interferents $N_{unx}$ can be assessed by comparing the final residuals $s_u$ with the instrumental noise level:

$$s_u = ||\mathbf{e}_u||/[JK - (N_c + N_{unx})]^{1/2} \quad (7)$$

where $\mathbf{e}_u$ is from Eq. (5). Typically, a plot of $s_u$ computed for trial number of components will show decreasing values, starting at $s_p$ when the number of components is equal to $A$ (the number of latent variables used to described the calibration data), until it stabilizes at a value compatible with the experimental noise, allowing to locate the correct number of components. It should be noticed that for $N_{unx} > 1$, the profiles provided by the SVD analysis of $\mathbf{E}_p$ unfortunately no longer resemble the true interferent profiles, due to the fact that the principal components are restricted to be orthonormal.

It should be taken into account that adding more latent variables than the number of chemical compounds when applying RBL methods could somehow compensate for the lack of trilinearity [19].

## 3. Experimental

### 3.1. Reagents and solutions

All solutions were prepared daily. Analytical reagent-grade chemicals and milli-Q water were used. Amaranth, sunset yellow FCF and tartrazine 0.10 mol L$^{-1}$ stock solutions (all from Aldrich) were prepared in milli-Q water. Standard solutions and mixtures of dyes were freshly prepared by appropriated dilution of stock solutions with milli-Q water. Methanol and ammonium acetate were obtained from Sintorgan (Buenos Aires, Argentina) and Cicarelli (San Lorenzo, Argentina), respectively.

### 3.2. Apparatus and software

Both chromatographic procedures were carried out using five modules (degasser, pump, injection valve, autosampler and DAD detector) of an Agilent 1100 Series instrument (Agilent Technologies, Waldbronn, Germany). The measurements were done on a 5 μm ZORBAX Eclipse XDB-C18 column (4.6 mm × 150 mm).

The MCR-ALS algorithm was downloaded from the multivariate curve resolution web page: http://www.ub.edu/mcr/welcome.html. A useful interface for data input and parameter setting was employed for U-PLS/RBL implementation written by Olivieri et al. [36]. Both algorithms were implemented in MATLAB 7.1 [37].

### 3.3. Procedure

#### 3.3.1. Calibration standards and mixtures of dyes

Standard calibration solutions of the analytes (five standards for each dye) were prepared in the concentration ranges between 0.05 and 3.00 mg L$^{-1}$ for amaranth, 1.00 and 45.00 mg L$^{-1}$ for sunset yellow and 0.50 and 8.00 mg L$^{-1}$ for tartrazine. On the other hand, 15 ternary mixtures of the three analytes were prepared according to a central composite design in order to evaluate the prediction error of the chemometric algorithms (see Table 1).

**Table 2**
Prediction on real samples by application of the reference HPLC method and the proposed fast chromatography coupled to both MCR-ALS and U-PLS/RBL.

| Sample | Tartrazine (mg L$^{-1}$) | | | Amaranth (mg L$^{-1}$) | | | Sunset yellow (mg L$^{-1}$) | | |
|---|---|---|---|---|---|---|---|---|---|
| | HPLC[a] | MCR[b] | RBL[c] | HPLC[a] | MCR[b] | RBL[c] | HPLC[a] | MCR[b] | RBL[c] |
| 1 | – | – | – | 0.22 | 0.23 | 0.28 | 30.65 | 30.46 | 30.81 |
| 2 | – | – | – | 0.14 | 0.16 | 0.18 | 15.61 | 16.05 | 15.70 |
| 3 | 3.12 | 2.87 | 2.60 | 0.13 | 0.13 | 0.14 | 6.02 | 5.81 | 5.12 |
| 4 | 2.29 | 2.02 | 2.01 | 0.77 | 0.70 | 0.60 | 6.33 | 6.80 | 4.90 |
| 5 | 0.81 | 0.75 | 0.84 | – | – | – | 14.55 | 14.75 | 13.20 |
| 6 | 0.94 | 0.87 | 1.17 | – | – | – | 14.37 | 14.80 | 12.50 |
| 7 | 1.40 | 1.69 | 1.62 | 0.11 | 0.11 | 0.12 | 2.34 | 2.93 | 2.98 |
| Mean recovery (%)[d] | – | 97.2(13.3) | 103.0(16.7) | – | 101.9(8.5) | 110.1(20.5) | – | 105.1(9.5) | 95.5(16.3) |
| REP (%)[e] | – | 10.5 | 14.8 | – | 10.1 | 25.5 | – | 3.0 | 8.6 |

[a] HPLC method taken as reference (see Ref. [21]).
[b] MCR-ALS: factors and regions depending on the sample.
[c] U-PLS/RBL: five latent variables for each analyte (regions: see the text).
[d] Between parenthesis the standard deviation.
[e] Relative error of prediction, $\mathrm{REP} = \frac{100}{\bar{c}}\left[\frac{1}{I}\sum_{1}^{I}(c_{act} - c_{pred})^2\right]^{1/2}$, where $I$ is the number of samples, $c_{act}$ and $c_{pred}$ are the actual and predicted concentrations, and $\bar{c}$ is the mean concentration.

### 3.3.2. Beverage samples

The analyzed samples were seven soft drinks that were purchased from a local supermarket. Samples were homogenized, filtered through 0.45 μm membranes and injected into the chromatographic system.

### 3.3.3. Reference method

The reference procedure was adapted for the one proposed by Pereira Alves et al. [21]. Chromatograms were recorded at room temperature, using a mixture of (methanol:ammonium acetate 0.08 mol L$^{-1}$) (23:77) as mobile phase flowing at 1 mL min$^{-1}$ with ultraviolet detection at 454, 484 and 550 nm for tartrazine, sunset yellow and amaranth, respectively. In these conditions, the total analysis time for each chromatogram was 13 min. The results obtained for all the analyzed samples are depicted in Table 2.

### 3.3.4. Fast chromatographic method

With the aim of developing a faster methodology than that proposed in the literature [21], the composition of the mobile phase was totally inverted, i.e. (methanol:ammonium acetate 0.08 mol L$^{-1}$) (77:23). All other chromatographic conditions were maintained as in the reference methodology.

## 4. Results and discussion

### 4.1. General concerns

Fig. 1 shows the complete chromatographic separation of the three dyes and other components in a beverage sample (sample number 4 in Table 2) by using the method proposed by Pereira Alves et al. [21]. As can be seen, the complete separation between the analytes and interferents is achieved in 13 min. At least three interferents appear at elution time ranges of 1.0–1.5 min and 3.5–4.5 min.

A faster chromatographic run would be preferred because solvents and time savings are related to a reduction of both the costs per analysis and environmental impact. The analysis time can be significantly reduced by changing the composition of the mobile phase. This fact produces overlapping peaks, originating data which can be conveniently processed by using multivariate algorithms in order to achieve selectivity by mathematical means.

Fig. 2 shows the landscape obtained for a ternary mixture (number 1 of Table 1) when the chromatographic separation is performed in 1.85 min (each 0.41 s) and recorded with a diode array detector in the region of 440–570 nm (each 2 nm), i.e. a matrix of 100 × 66 points per sample. This figure also shows both the time and spectra profiles of the three dyes. Time elution profiles were recorded at the corresponding maximum wavelength for each compound (i.e. T: 450 nm, SY: 490 nm and A: 530 nm). As can be observed, a severe overlapping exists for the three compounds making impossible the use of univariate calibration. Interestingly, the elution profile corresponding to sunset yellow is extremely misshapen. This situation can be understood by the fact that besides the retention factors for all the dyes are zero; SY is also present in the samples in the highest concentration levels.

Another item that should be considered is the presence of unexpected components in real beverage samples as was commented above when analyzing Fig. 1, making necessary to exploit the second order advantage [4]. Fig. 3 shows the elution profiles recorded at λ = 440 nm, corresponding to a ternary mixture (number 1 in Table 1) and to a real sample (number 5 in Table 2). As can be appreciated in this figure, unexpected compounds appear, especially between 1.4 and 1.9 min.

On the other hand, a visual inspection of Fig. 4, in which the chromatograms corresponding to the 15 ternary mixtures (see Table 1) recorded at λ = 440 nm are showed, could lead to the conclusion

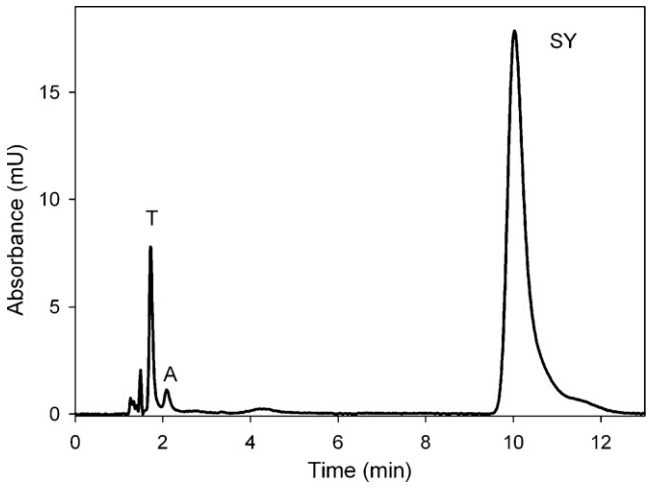


**Fig. 1.** Complete chromatographic separation (13 min) of the three dyes and other components in a beverage sample (sample number 4 in Table 2) by using the method proposed by Pereira Alves et al. (see Ref. [21]).

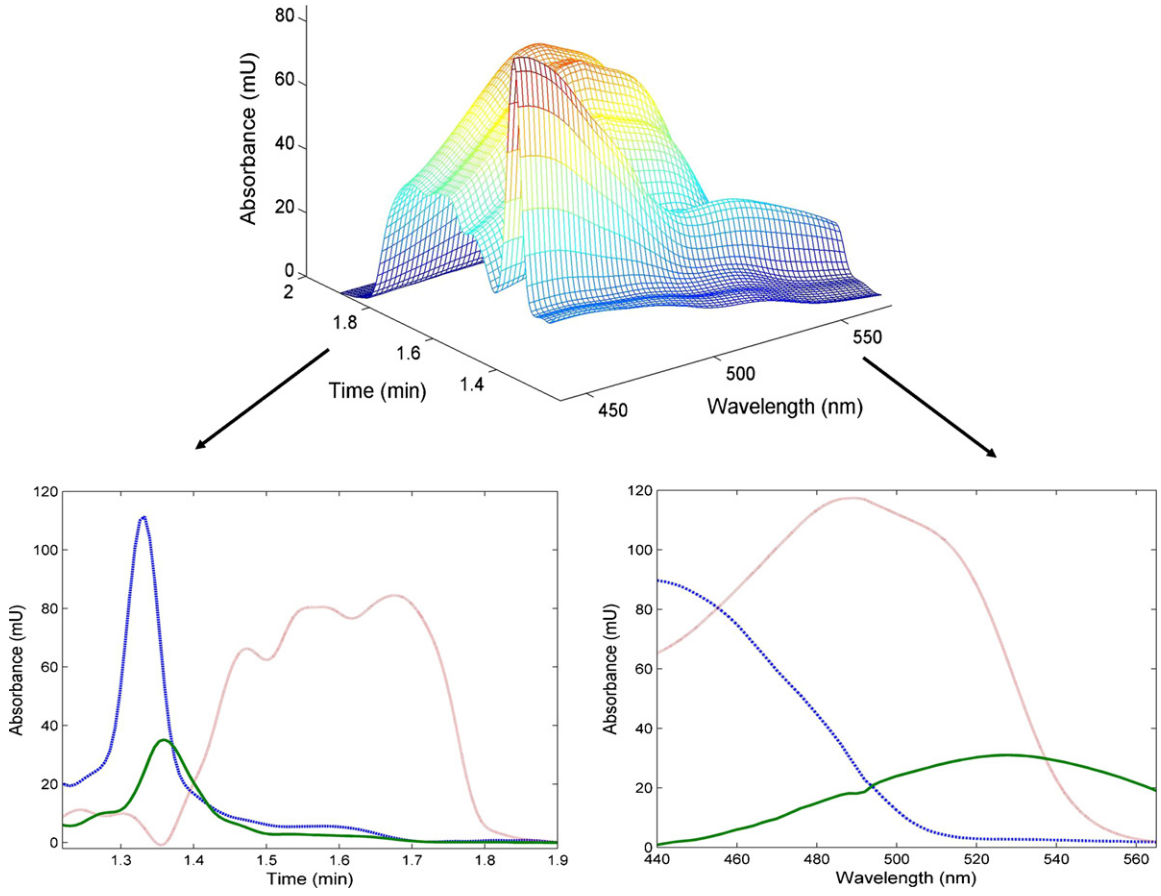


**Fig. 2.** Landscape obtained for a ternary mixture (number 1 of Table 1) when the chromatographic separation is performed in 1.85 min and recorded with a diode array detector in the region of 440–570 nm. Time and spectral profiles of the three dyes (T: blue line, A: green line and SY: red line). Time elution profiles were recorded at the corresponding wavelength maximum for each compound (i.e. T: 450 nm, SY: 490 nm and A: 530 nm). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

that a remarkable time shift effect is present. This fact is evident especially in the range of 1.5–1.6 min. On the contrary, differences in the region of 1.3–1.4 min cannot be only be ascribed to shift time effect, but also to changes in the relative concentration of T and A, whose peaks are strongly overlapped. This fact can be better appreciated in Fig. 5. This figure shows the chromatograms (recorded at $\lambda = 500$ nm) corresponding to standard solutions of T (8.00 mg $L^{-1}$), A (1.53 mg $L^{-1}$) and a mixture of both dyes at those concentrations. Thus, it would result impossible to align peaks without introducing error. Consequently, the obvious challenge in this work was the

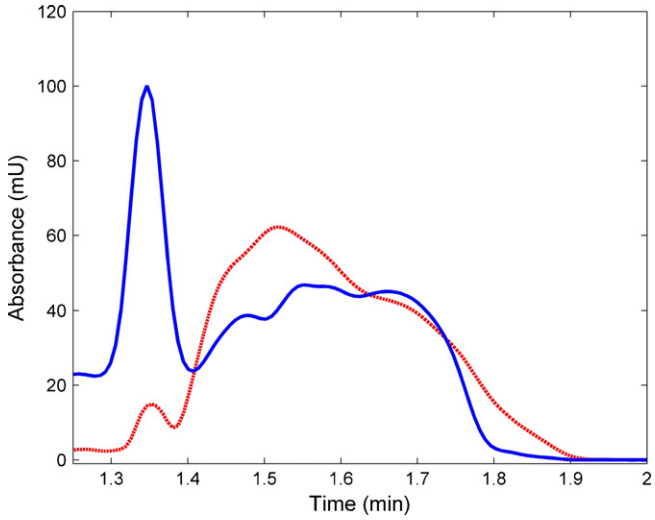


**Fig. 3.** Chromatograms recorded at $\lambda = 440$ nm, corresponding to a ternary mixture in blue solid line (number 1 in Table 1) and to a real sample in dashed red line (number 5 in Table 2). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

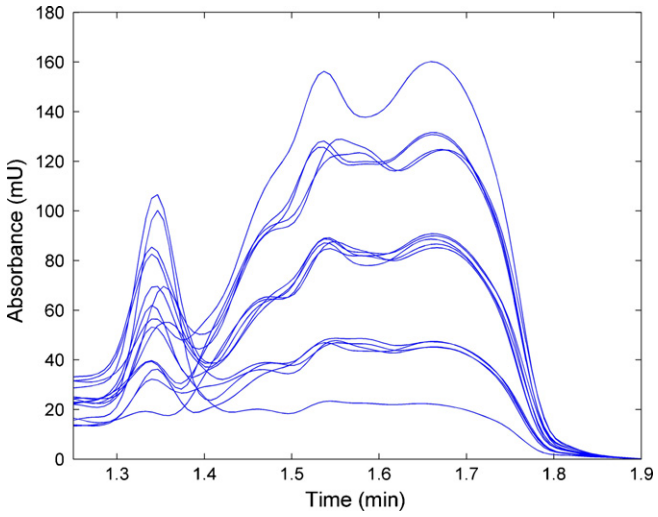


**Fig. 4.** Chromatograms corresponding to the 15 ternary mixtures (see Table 1) recorded at $\lambda = 440$ nm.

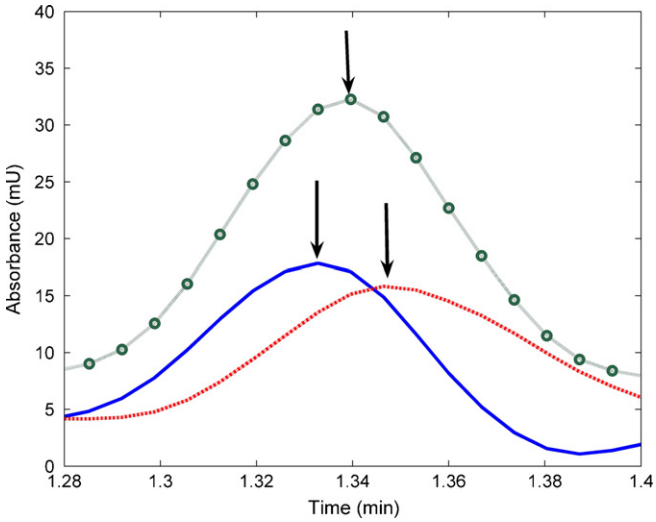


**Fig. 5.** Chromatograms (recorded at $\lambda = 500$ nm) corresponding to standard solutions of T (8.00 mg $L^{-1}$) in blue solid line, A (1.53 mg $L^{-1}$) in red dashed line and a mixture of both dyes at those concentrations in green circle line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

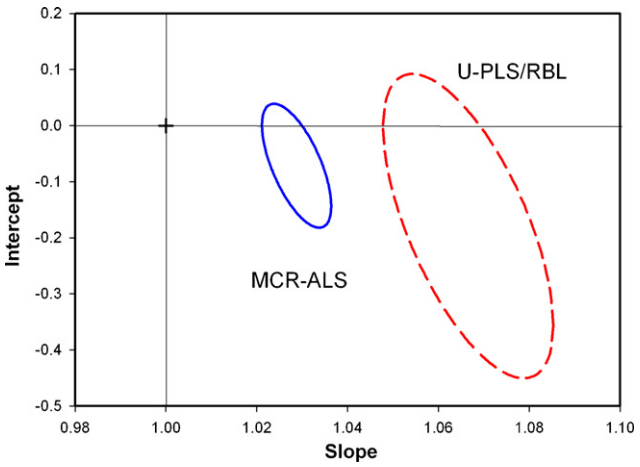


**Fig. 6.** Elliptical regions for the global data set for predictions using U-PLS/RBL and MCR-ALS algorithms on the 15 validation mixtures of Table 1.

quantitation of the three dyes in beverages, mainly without signal pre-treatments.

### 4.2. Application of second order algorithms

#### 4.2.1. Artificial ternary mixtures

The MCR-ALS algorithm was applied to the simultaneous analysis of the 15 ternary samples (Table 1) by using the pure standards described in Section 3. Column-wise augmented data matrices (**D**) were generated arranging $\mathbf{D}_i$ matrices corresponding to spectra recorded during the chromatographic process for standards and samples.

In the present work, nonnegativity in spectral and concentration profiles for all analytes and unimodality in concentration profiles only for tartrazine and amaranth were applied. The most important issue is that the pure spectra of the compounds should be the same in all experiments, but the profiles in the different **C** sub-matrices need not to share a common shape, as when the trilinearity is lost (the present case).

Before starting resolution, the determination of the number of contributions to each **D** data matrix was carried out by applying singular value decomposition. After that, the $\mathbf{S}^T$-type initial estimates were built by the selection of purest spectra based on SIMPLISMA [38]. In some cases, resolution results obtained using these initial estimations were unsatisfactory. Despite that fact, the optimized spectral profiles gathered by MCR were stored and used as initial estimations for subsequent MCR analysis until successful MCR quality parameters were reached. This strategy proved to be very effective in cases in which SIMPLISMA was not able to provide suitable initial information. Results for the validation samples are presented in Table 1, and as can be appreciated reasonable figures of merit are obtained.

On the other hand, the number of calibration latent variables when applying U-PLS/RBL for each analyte was set at 1 because at the beginning pure standards were used for calibration purposes. The presence of unexpected components (in this case the other two dyes) had to be considered to decrease the prediction residuals for the test samples until they stabilized at a value compatible with the instrumental noise. The prediction residuals were monitored as a function of trial values of $N_{unx}$ for all the samples. When $N_{unx}$ was equal to 3, the residual value was comparable with the instrumental noise (ca. 0.06 absorbance units in this system). Thus, $N_{unx} = 3$ was the corresponding correct choice. Evidently, this system would only need two unexpected components due to the fact that the mixtures are composed by three components. But, a third component is needed probably due to the lack of trilinearity. Results were unacceptable, with relative error of prediction of ca. 40% for the three dyes.

Consequently, a new strategy was followed for U-PLS/RBL modelling. Ternary samples were used instead of pure standards. In a recently published work, Cañada-Cañada et al. obtained good results applying a RBL based method (N-PLS/RBL) to LC fluorescence data by using a designed mixture of the two analytes for calibration [19]. In the present work, 15 calibration models were performed with 14 mixture samples described in Table 1, and predicting the rest of the samples. This procedure was repeated till the prediction value for the whole test samples was obtained. In this case, the number of latent variables estimated by the leave-one-sample out cross-validation method was equal to 3, and the number of unexpected components was zero, i.e. no second order advantage was required. As can be appreciated in Table 1, results present comparable figures of merit with those obtained by MCR-ALS.

To assess the accuracy of the models, the obtained values by applying both algorithms were compared with the nominal ones corresponding to the three analytes. For this purpose, the joint statistical test for the slope and the intercept of the linear regression between the measured concentration values versus those predicted was applied. The multivariate model is regarded as being accurate if the theoretical values of intercept and slope (zero and unity, respectively) are included within the ellipse, which describes the mutual confidence region. As has been previously suggested, when multivariate analysis is performed, it is highly convenient to include experimental data corresponding to all analytes, in order to better estimate the variance corresponding to the regression discussed above. This avoids the oversizing of the joint confidence region due to large experimental random errors and thus the probability of not detecting the presence of bias [39]. Fig. 6 shows these regions for predictions of the global data sets using U-PLS/RBL and MCR-ALS algorithms. As can be seen, both ellipses contain the theoretically expected value (0) for the intercept (at a confidence level of 95%). On the other hand, they do not contain the theoretically expected value (1) for the slope. This fact is indicative of the presence of a proportional error. In addition, the smaller size and the closeness to the expected value (1) of the ellipse corresponding to MCR-ALS allows one to conclude about both a higher precision and a lower proportional error of this algorithm when is compared with U-PLS/RBL

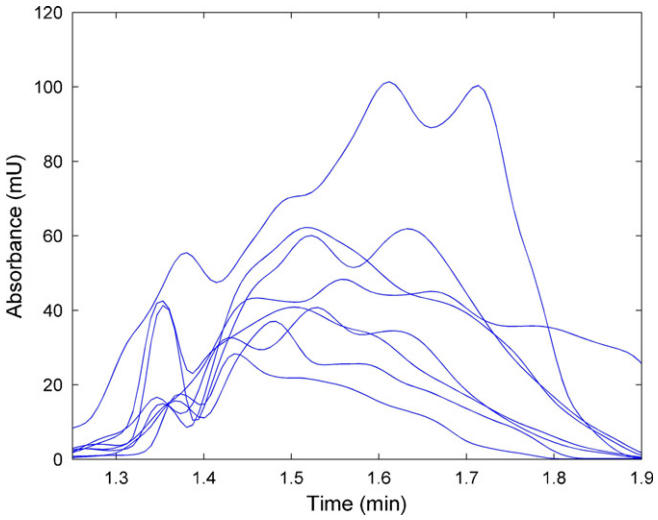


**Fig. 7.** Chromatograms corresponding to the eight real samples recorded at λ = 440 nm.

[40]. It should be taken into account that the MCR-ALS models were built only with pure standards, while for U-PLS/RBL it was necessary to build the models with ternary mixtures. These facts show once again a higher predictive ability of MCR-ALS in those cases in which data are not trilinear.

*4.2.2. Real samples*

The complexity of the real samples can be appreciated in Fig. 7. These samples were analyzed as was indicated for the 15 validation samples and the obtained results are displayed in Table 2. As can be seen, results rendered by U-PLS/RBL (the 15 ternary samples were used for building the calibration models) are very poor when the predictions are compared with results obtained by the HPLC method proposed in the literature, which guarantee the complete separation of the analytes and interferents (see Fig. 1). This fact suggests that these data cannot be correctly modelled likely due to the strong lack of trilinearity of these data. Interestingly, in most of the samples two or three interferences should be considered to reach residuals values comparable with the instrumental noise.

On the other hand, Table 2 shows that results proportioned by MCR-ALS modelling can be considered acceptable when they are matched with those obtained by the reference method. As an example of how MCR-ALS was implemented, Fig. 8 A and B shows both the time and the spectral profiles extracted by the algorithm when analyzing sample 7 (Table 2), which contains the three dyes plus an unexpected interference. As can be appreciated, this compound presents a spectrum similar to the one corresponding to tartrazine, while its time profile coelutes with sunset yellow. The excellent statistical parameters (percent of lack of fit = 3.31 and percent of variance explained, $r^2$ = 99.8904) and the reasonable figures of merit obtained when comparing MCR-ALS results with those obtained by the HPLC reference method, are indicative of the acceptable performance of this algorithm when processing such a complex instrumental data.

Finally, a consideration about solvent and cost saving should be considered. Firstly, the analysis time can be reduced up to 14.2% of the necessary time to perform the complete separation. On the other hand, a complete separation consumes 10.01 mL of methanol and 2.99 mL of 0.08 mol $L^{-1}$ ammonium acetate, while the proposed fast chromatography method requires only 0.46 mL of methanol and 1.54 mL of 0.08 mol $L^{-1}$ ammonium acetate. The facts support the use of this kind of methodology, even more if green chemistry is seriously taken into account.

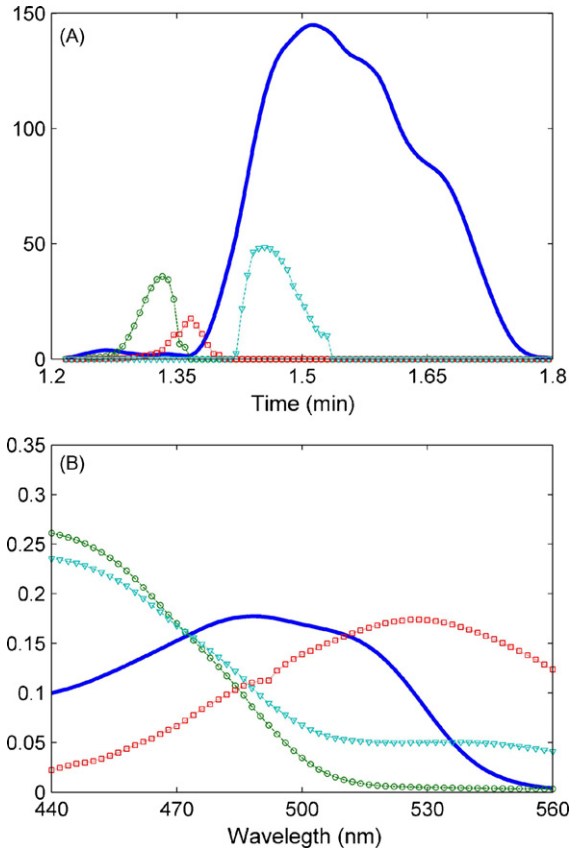


**Fig. 8.** (A) Time profiles extracted by the MCR-ALS algorithm when analyzing sample 7 (Table 2), which contains the three dyes (T: green circle line, A: red square line and SY: blue solid line) plus an unexpected interference (triangle cyan line). (B) Spectral profiles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

## 5. Conclusions

Reduction in the analysis time for a chromatographic method should be seriously taken into account considering that solvent saving is an issue with a strong impact in environment.

LC-DAD data with lack of trilinearity which cannot be conveniently pre-processed can hardly be modelled by using U-PLS/RBL algorithm. On the contrary, MCR-ALS became an excellent tool for modelling this kind of data allowing to reach acceptable figures of merit which are indicative of a good performance of the proposed methodology.

## Acknowledgements

Financial support from Universidad Nacional del Sur and Universidad Nacional del Litoral (Project CAI+D PI No. 12–65) is gratefully acknowledged. M.J. Culzoni, A.V. Schenone, M. Garrido, N. Llamas, B.S. Fernández Band and H.C. Goicoechea wish to thank Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

## References

[1] G.M. Escandar, N.M. Faber, H.C. Goicoechea, A. Muñoz de la Peña, A.C. Olivieri, R.J. Poppi, TrAC Trends Anal. Chem. 26 (2007) 752.
[2] V. Gómez, M.P. Callao, Anal. Chim. Acta 627 (2008) 169.
[3] M. Martínez Galera, M.D. Gilgarcía, H.C. Goicoechea, TrAC Trends Anal. Chem. 26 (2007) 1032.
[4] K.S. Booksh, B.R. Kowalski, Anal. Chem. 66 (1994) 782A.
[5] E. Sánchez, B.R. Kowalski, Anal. Chem. 58 (1986) 496.
[6] E. Sánchez, B.R. Kowalski, J. Chemometrics 1 (1990) 29.

[7] Z.P. Chen, H.L. Wu, J.H. Jiang, Y. Li, R.Q. Yu, Chemom. Intell. Lab. Syst. 52 (2000) 75.
[8] A.L. Xia, J.L. Wu, D.M. Fang, Y.J. Ping, L.Q. Hu, R.Q. Yu, J. Chemometrics 19 (2005) 65.
[9] R. Bro, Chemom. Intell. Lab. Syst. 38 (1997) 149.
[10] R. Tauler, Chemom. Intell. Lab. Syst. 30 (1995) 133.
[11] E. Peré-Trepat, S. Lacorte, R. Tauler, Anal. Chim. Acta 595 (2007) 228.
[12] A. De Juan, R. Tauler, J. Chromatogr. A 1158 (2007) 184.
[13] M. Linder, R. Sundberg, Chemom. Intell. Lab. Syst. 42 (1998) 159.
[14] S. Wold, G.P.K. Esbenesen, J. Öhman, J. Chemometrics 1 (1987) 41.
[15] A.C. Olivieri, J. Chemom. 19 (2005) 253.
[16] A.C. Olivieri, J. Chemom. 19 (2005) 615.
[17] M.W. Dong, Modern HPLC for Practicing Scientists, Wiley Synomics Pharmaceutical Services, LLC, Wareham, Massachusetts, 2006.
[18] R. Bro, Doctoral Thesis, University of Amsterdam, The Netherlands, 1998.
[19] F. cañada-Cañada, J.A. Arancibia, G.M. Escandar, G.A. Ibañez, A. Espinosa Mansilla, A. Muñoz de la Peña, A.C. Olivieri, J. Chromatogr. A 1216 (2009) 4868.
[20] M.D. Gil García, M.J. Culzoni, M.M. De Zan, R. Santiago Valverde, M. Martinez Galera, H.C. Goicoechea, J. Chromatogr. A 1179 (2008) 115.
[21] S. Pereira Alves, D. Mares Brum, E. Castello Branco de Andrade, A. Duarte Pereira Netto, Food Chem. 107 (2008) 489.
[22] M.L.S. Silva, B.Q. García, J.L.F.C. Lima, E. Barrado, Talanta 72 (2007) 282.
[23] Joint Expert Committee on Food Additives depending on Food and Agriculture Organization (FAO) and World Health Organization (WHO).
[24] D.M. Milojovic-Opsenica, K. Lazarevic, V. Ivackovic, Z.L. Tesic, J. Planar. Chromatogr. 16 (2003) 276.
[25] S. Dixit, S.K. Khanna, M. Das, J. AOAC Int. 91 (2008) 1387.
[26] A.V. Jager, F.G. Tonin, M.F.M. Tavares, J. Sep. Sci. 28 (2005) 957.
[27] K.S. Lee, M.J.A. Shiddiky, S.H. Park, D.S. Park, Y.B. Shim, Electrophoresis 29 (2008) (1910).
[28] M. Ma, X. Luo, B. Chen, S. Su, S. Yao, J. Chromatogr. A 1103 (2006) 170.
[29] N. Yoshioka, K. Ichihashi, Talanta 74 (2008) 1408.
[30] Y. Ni, Y. Wang, S. Kokot, Talanta 78 (2009) 432.
[31] D.W. Lachenmeier, W. Kessler, J. Agric. Food Chem. 56 (2008) 5463.
[32] N.E. Llamas, M. Garrido, M.S. Di Nezio, B.S. Fernández Band, Anal. Chim. Acta, submitted for publication.
[33] J. Jaumot, R. Gargallo, A. de Juan, R. Tauler, Chemom. Intell. Lab. Syst. 76 (2005) 101.
[34] D.M. Haaland, E.V. Thomas, Anal. Chem. 60 (1988) 1193.
[35] S. Wiklund, D. Nilsson, L. Eriksson, M. Sjvstrvm, S. Wold, K. Faber, J. Chemom. 21 (2007) 427.
[36] A.C. Olivieri, H.-L. Wu, R.-Q. Yu, Chemom. Intell. Lab. Syst. 96 (2009) 246.
[37] MATLAB 7.1, The MathWorks Inc., Natick, Massachussets, USA, 2005.
[38] W. Windig, J. Guilment, Anal. Chem. 63 (1991) 1425.
[39] A. Martínez, J. Riu, O. Busto, J. Guasch, F.X. Rius, Anal. Chim. Acta 406 (2000) 257.
[40] A.G. González, M.A. Herrador, A.G. Asuero, Talanta 48 (1999) 729.