Chemometrics and Intelligent Laboratory Systems 141 (2015) 1–11

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Novel augmented parallel factor model for four-way calibration of high-performance liquid chromatography–fluorescence excitation–emission data

Santiago A. Bortolato [a], Valeria A. Lozano [a], Arsenio Muñoz de la Peña [b], Alejandro C. Olivieri [a,*]

[a] Instituto de Química Rosario (CONICET-UNR), Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531, 2000, Rosario, Argentina
[b] Department of Analytical Chemistry, Faculty of Sciences, University of Extremadura, 06006 Badajoz, Spain

ARTICLE INFO

Article history:
Received 11 September 2014
Received in revised form 24 November 2014
Accepted 27 November 2014
Available online 5 December 2014

Keywords:
Liquid chromatography–excitation–emission fluorescence data
Multi-way calibration
Augmented parallel factor analysis

ABSTRACT

A new augmented parallel factor analysis model (Augmented PARAFAC) is presented, inspired by the useful augmentation concept employed in multivariate curve resolution-alternating least-squares (MCR-ALS), applicable to calibration based on non-quadrilinear four-way data, such as those produced by high-performance liquid chromatography with matrix excitation–emission fluorescence detection. The new model involves creating an augmented three-way array in the elution time direction, containing data for the calibration sample set and for each of the test samples, subsequently analyzed with an Augmented PARAFAC version. To test the properties of this approach, chromatographic data were simulated with different degrees of overlapping and misalignment among the chromatographic peaks. Additionally, experimental data from olive oil samples were tested with the new model, aimed at the quantitation of the level of chlorophylls and pheophytins. The results were compared with those obtained by data processing with MCR-ALS. Relative prediction errors (%) were: Augmented PARAFAC, 9.7, 21.0, 14.7 and 9.3, and MCR-ALS, 5.9, 14.5, 20.0 and 14.7 for Chl *a*, Chl *b*, Phe *a* Phe *b*, respectively, for concentrations in the range 0.00–1.00 μg mL^{-1}. Both MCR-ALS and Augmented PARAFAC allow one to obtain a detailed and realistic description of the analyzed samples, in terms of pure elution time, excitation and emission spectral profiles, which can be independently retrieved for every component.

1. Introduction

Chemical multi-way calibration has become an important frontier in chemometric research. The multivariate analysis of excitation–emission fluorescence matrix (EEM) data, constituting three-way arrays when joining data for a group of samples, has been summarized in two recent reviews and a tutorial, all reporting up-to-date applications in the biomedical, environmental and food analysis fields [1–3].

High-performance liquid chromatography (HPLC), when combined with spectroscopic techniques, such as UV–visible diode-array detection (DAD) or fast-scanning fluorescence detection (FSFD) is also able to yield spectral-elution time matrix data. The response can be arranged as a data matrix, where each column corresponds to a different wavelength and each row to a different elution time. When full selectivity in the chromatographic separation is not achieved, and even in the presence of unexpected components, multivariate calibration can be successfully applied to the three-way arrays obtained by joining data for a group of samples. Additional benefits are decreasing cost and times of analysis. Several recent reports deal with the advantages and drawbacks associated with the combination of multivariate calibration and chromatography, and pertinent references on the successful processing of spectroscopic–chromatographic data can be found [4–14].

Four-way data can be obtained by joining third-order data for a set of samples into a four-dimensional mathematical object. The latter data not only retain the 'second-order advantage' inherent to three-way/second-order calibration [15], but can also have additional advantages [16]. They would display the obvious advantage of providing richer analytical information, implying more stable methods towards interferences and matrix effects, and less sensitive to minor changes in reaction conditions, which should allow for an improvement in predictive ability. In addition, improvements in sensitivity and the resolution of collinearity problems have been reported [17]. However, it is interesting to note that only few experimental four-way data have been recorded and used to develop analytical methodologies to date, which can be attributed to the fact that the practical acquisition of these data arrays is still difficult to implement. In addition, a thorough understanding of their analytical advantages is still needed.

Four-way data can be collected with a single instrument; the most common approach is the recording of luminescence EEMs as a function of some factors such as reaction time, decay time or any additional variable affecting the analytical signal (pH, sample volume, quenching effects, etc.). These factors are introduced as independent analytical modes, to construct multi-way data and to exploit the additional

* Corresponding author. Tel./fax: +54 341 4672704.
E-mail address: olivieri@iquir-conicet.gov.ar (A.C. Olivieri).

http://dx.doi.org/10.1016/j.chemolab.2014.11.013
0169-7439/

information they provide. Several different experimental approaches, describing practical analytical applications, can be found in the literature [18–41].

In the case of chromatographic data, four-way arrays can be generated with comprehensive gas or liquid two-dimensional chromatographic systems equipped with detection based on time of flight mass spectrometry (TOFMS) or DAD, leading to GC × GC–TOFMS or LC × LC–DAD hyphenated data. This is a currently growing scenario, and examples of different analytical problems in samples of high complexity have been reported [42–47].

Unidimensional chromatography can also provide four-way data, and Bro was the first to explore this possibility by recording full fluorescence EEMs during detection [48]. An alternative are LC–DAD-kinetic data, collected while following the hydrolysis of the Aly pesticide [49]. Recently, two different approaches have been reported by recording four-way LC–EEM data, which were employed for the analysis of: chlorophylls and pheophytins in olive oil samples [50], and fluoroquinolones in water samples [51].

When component profiles change from sample to sample, as it usually happens during the collection of chromatographic data because of elution time shifts or peak shape changes, the four-way LC–EEM data are not quadrilinear (strictly speaking, low-rank quadrilinear) [52]. In this case, the elution time mode is considered to be the 'quadrilinearity breaking' mode, or the mode suspected of breaking the quadrilinearity of the data [51,53]. Only a few multivariate techniques have been duly tested for data processing of these data: multivariate curve resolution with alternating least-squares (MCR-ALS) is the most common one [51], although the variant parallel factor analysis-2 (PARAFAC2) has also been applied [48]. However, it should be noted that the use of MCR-ALS implies unfolding the four-way data set into a super-augmented matrix. The statistical efficiency of decomposing multiway arrays is higher (and consequently the sensitivity is larger) when the original data structure is maintained, in comparison with unfolding into arrays of lower dimensions [54]. PARAFAC2, on the other hand, has been shown to be less efficient than MCR-ALS when processing multiway data in the presence of unexpected interferents not included in the calibration phase [55].

To avoid the above mentioned potential disadvantages, we propose a new model based on three-way PARAFAC, taking advantage of the augmentation philosophy applied in MCR-ALS studies. In the present work, both simulated and experimental four-way liquid chromatographic data with EEM detection are analyzed using the new Augmented PARAFAC model which allows to model inter-sample chromatographic profile variations, with an exhaustive comparison with the MCR-ALS algorithm. In the case of experimental data, samples already analyzed in Ref. [50] were now processed by Augmented PARAFAC, MCR-ALS and PARAFAC2.

2. Theory

2.1. Simulations

Data have been synthesized for a system having two calibrated analytes and a single potential interferent in the test samples along with the analytes. All data arrays were built mimicking four-way chromatographic data with EEM detection (elution time-excitation wavelength-emission wavelength), similar to those recorded for the experimental systems. The simulated signal for component n at unit concentration ($\mathbf{m}_n$) is governed by the following equation:

$$\mathbf{m}_n = \mathbf{t}_n \otimes \mathbf{w}_n \otimes \mathbf{z}_n \quad (1)$$

where $\mathbf{m}_n$ is the $JKL \times 1$ vectorized signal at unit concentration (J, K and L are the number of channels in each mode – elution time, excitation and emission wavelength, respectively – and are equal to 30, 20 and 50), $\mathbf{t}_n$, $\mathbf{w}_n$ and $\mathbf{z}_n$ are the individual profiles in each data mode (of size $J \times 1$, $K \times 1$, and $L \times 1$, respectively), and the symbol '⊗' indicates the Kronecker product.

Representative Gaussian elution time profiles $\mathbf{t}_n$ ($n = 1, 2$ and 3), partially overlapped in the time mode, are shown in Fig. 1A, although they change randomly from sample to sample during the simulations. Various types of chromatographic shifts were introduced into these time profiles, in order to generate a comprehensive set of

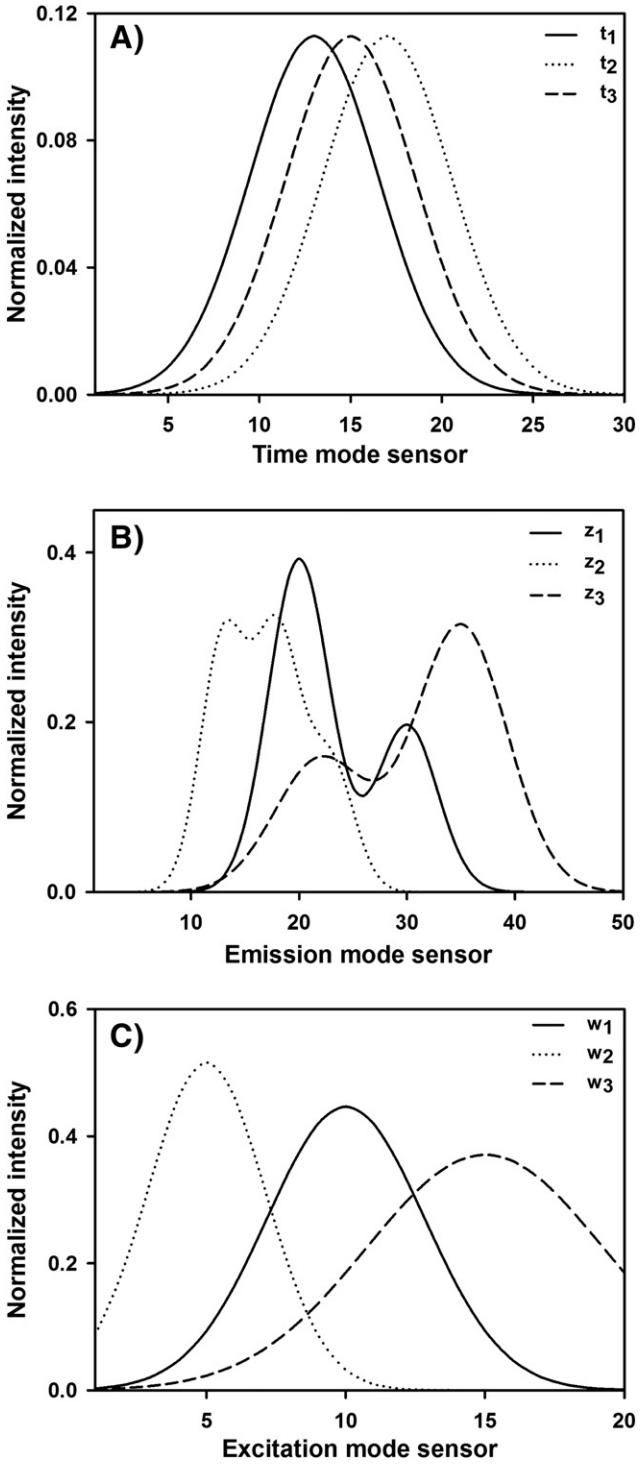

Fig. 1. Noiseless profiles employed for the simulations, in the elution time mode (A), in the emission mode (B) and in the excitation mode (C), for sample components at unit concentration. Solid line, analyte 1, dotted line, analyte 2, dashed line, potential interferent. The time profiles in (A) are scaled to unit area under each profile, while in (B) and (C) they are normalized to unit length.

cases to be studied. This was done to rigorously test the predictive ability of Augmented PARAFAC model and MCR-ALS towards analyte determination in the test sample sets. With regard to the spectral profiles ($\mathbf{z}_n$ and $\mathbf{w}_n$) for the sample components, they are shown in Fig. 1B and C, respectively, where considerable overlap can be observed among them. These profiles are common to all samples, and do not change from sample to sample, as is usual for fluorescence excitation and emission spectra.

To produce the calibration data, the unfold signal for a typical sample ($\mathbf{x}$) is given by the sum of the contributions of both analytes:

$$\mathbf{x} = y_1\mathbf{m}_1 + y_2\mathbf{m}_2 \tag{2}$$

with y_1 and y_2 being the component concentrations, and $\mathbf{m}_1$ and $\mathbf{m}_2$ given by an expression analogous to Eq. (1). In all the simulated data sets, calibration samples were created following a 9-sample central composite design with concentrations in the range 0.0–1.0. Later, fifty test samples containing the potential interferent along with the analytes, at concentrations taken at random from the range 0.2–1.5, were created. In this case, the test signals were given by the sum of three vectors given by ($y_n\,\mathbf{m}_n$), each of them provided by equations analogous to Eqs. (1) and (2). Once the noiseless calibration and test matrices were built, Gaussian noise was added to all signals. The standard deviation was 0.0015 units, representing 1% with respect to the maximum calibration signal of each analyte at unit concentration. The data sets were then submitted to third-order multivariate calibration for the determination of both calibrated analytes as described in the next sections.

All systems were processed by Augmented PARAFAC and MCR-ALS, which were applied in the usual way: (1) data for the unknown sample are joined with calibration data to create an augmented three-way array or a super-augmented matrix, respectively, (2) the latter array is adequately decomposed, (3) the calibration analyte scores are regressed against nominal concentrations, and (4) the test analyte score is interpolated in the calibration graph to estimate its concentration. The relevant details are discussed below for each of the implemented algorithms.

2.2. Augmented PARAFAC

In principle, each $\mathbf{x}$ vector described by Eq. (2) for a given sample can be rearranged to give a three-dimensional signal array, and a group of sample data can be joined into a four-way data array $\mathbf{X}$ of size $I \times J \times K \times L$. Provided the latter is quadrilinear (i.e., if chromatographic elution time, excitation and emission profiles were unique for all constituents and did not change from sample to sample), the well-known four-way PARAFAC model could be applied to process these data, as has been widely reported [56]. A four-way PARAFAC model can simply be written as follows, using the Khatri–Rao or column-wise Kronecker product indicated as '⊙' which makes the model description easier [57]:

$$\mathbf{X}^{(I\times JKL)} = \mathbf{A}_{4w}(\mathbf{D}_{4w}\odot\mathbf{C}_{4w}\odot\mathbf{B}_{4w})^T + \mathbf{E}^{(I\times JKL)} \tag{3}$$

where $\mathbf{X}^{(I\times JKL)}$ is a matrix of size $I \times JKL$ obtained by appropriately unfolding $\mathbf{X}$, and $\mathbf{A}_{4w}$, $\mathbf{B}_{4w}$, $\mathbf{C}_{4w}$ and $\mathbf{D}_{4w}$ are the four profile matrices in which $\mathbf{X}$ can be decomposed following a four-way PARAFAC model, of size $I \times N$, $J \times N$, $K \times N$ and $L \times N$, respectively. The model residuals are included in $\mathbf{E}^{(I\times JKL)}$.

However, in the presence of elution time profiles which change from sample to sample, either in peak position of profile shape, $\mathbf{X}$ will not be quadrilinear. One possibility to model this type of data is unfolding $\mathbf{X}$ into a super-augmented matrix and applying MCR-ALS [6], as pictorially depicted in Fig. 2A and detailed in the next section. A useful alternative is introduced in the present report, in the form of the Augmented PARAFAC model, where the four-way $\mathbf{X}$ array is unfolded into an augmented three-way array by unfolding $\mathbf{X}$ along the combined sample-elution time mode. Fig. 2B shows the generation of the three-way array for the Augmented PARAFAC model, contrasted with the super-augmented data matrix for MCR-ALS (Fig. 2A).

According to Fig. 2B, the augmented three-way array has dimensions $P \times KL$ ($P = IJ$), and the corresponding Augmented PARAFAC model can be represented by:

$$\mathbf{X}^{(P\times KL)} = \mathbf{AB}_{3w}(\mathbf{D}_{3w}\odot\mathbf{C}_{3w})^T + \mathbf{E}^{(P\times KL)} \tag{4}$$

where decomposition is accomplished into three loading matrices $\mathbf{AB}_{3w}$, $\mathbf{C}_{3w}$ and $\mathbf{D}_{3w}$, with sizes $P \times N$, $K \times N$ and $L \times N$, respectively. The model residuals are collected into $\mathbf{E}^{(P\times KL)}$, and the sum of its squared elements is minimized during data processing [56]. The matrix $\mathbf{AB}_{3w}$ collects the unfolded profiles along the combined sample-elution time modes, and carries concentration information regarding the

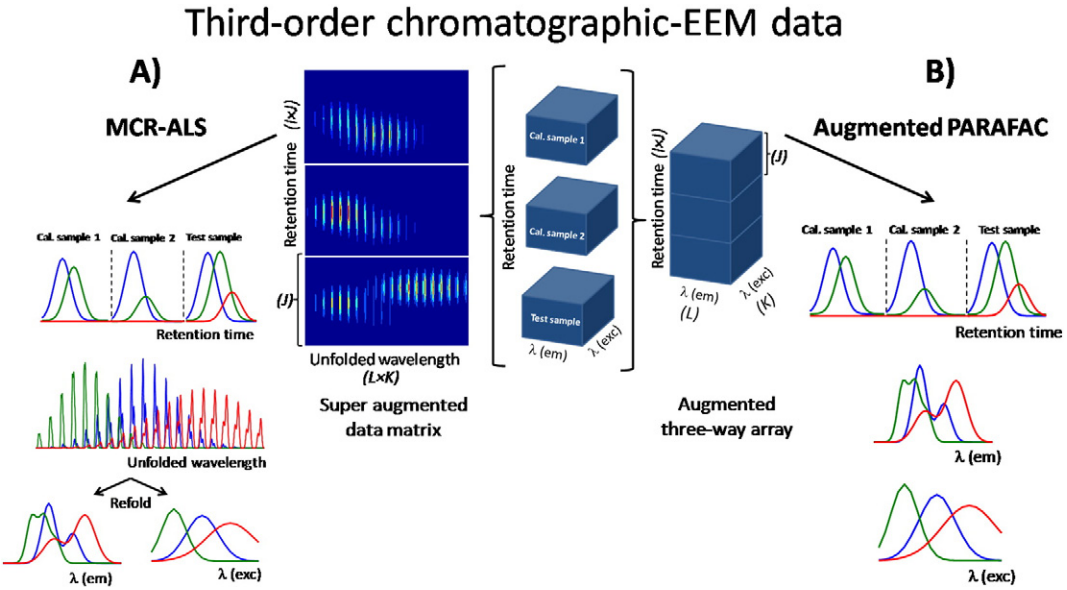

Fig. 2. Third-order chromatographic-EEM data. Flow chart to: A) MCR-ALS, B) Augmented PARAFAC.

responsive constituents. The central idea of the proposed model is thus to avoid the potential non-quadrilinearity problem.

Augmented PARAFAC can be implemented through a typical alternating least-squares process, which can be compactly described as follows:

(1) Randomly (or via known pure spectral profiles) initialize matrices $\mathbf{C}_{3w}$ and $\mathbf{D}_{3w}$, and find matrix $\mathbf{AB}_{3w}$ from the properly unfolded $\mathbf{X}$:

$$\mathbf{AB}_{3w} = \mathbf{X}^{(P\times KL)} \times \left[(\mathbf{D}_{3w} \otimes \mathbf{C}_{3w})^{+}\right]^{T} \quad (5)$$

where '+' denotes the Moore–Penrose generalized inverse.

(2) Update $\mathbf{C}_{3w}$ using the new $\mathbf{AB}_{3w}$:

$$\mathbf{C}_{3w} = \mathbf{X}^{(K\times PL)} \times \left[(\mathbf{AB}_{3w} \otimes \mathbf{D}_{3w})^{+}\right]^{T}. \quad (6)$$

(3) Update $\mathbf{D}_{3w}$ using the new $\mathbf{AB}_{3w}$ and $\mathbf{C}_{3w}$:

$$\mathbf{D}_{3w} = \mathbf{X}^{(L\times PK)} \times \left[(\mathbf{AB}_{3w} \otimes \mathbf{C}_{3w})^{+}\right]^{T}. \quad (7)$$

(4) Repeat steps 1–3 until certain stopping criterion is fulfilled, such as a convergence tolerance or maximum number of iterations, as is regularly done in PARAFAC studies.

The model can be initialized using spectral and chromatographic data which are known in advance for pure components, or automatically using direct trilinear decomposition (DTLD), as in regular PARAFAC [56]. The number of responsive components (N) can be estimated by consideration of the Augmented PARAFAC residual error, i.e., the standard deviation of the elements of the array $\mathbf{E}^{(P\times KL)}$ in Eq. (4) [56]. Usually this parameter decreases with increasing N, until it stabilizes at a value compatible with the instrumental noise. A reasonable choice for N is the smallest number of components for which the residual error is not statistically different than the instrumental noise.

Once the algorithm converges, the information gathered in the matrices $\mathbf{AB}_{3w}$, $\mathbf{C}_{3w}$ and $\mathbf{D}_{3w}$ serves to characterize both qualitatively and quantitatively the whole system. In our case, $\mathbf{C}_{3w}$ contains excitation spectra, and $\mathbf{D}_{3w}$ emission spectra of the chemical constituents under investigation. Identification is done with the aid of these two estimated profiles, and comparing them with those for a standard solution of the analyte of interest. This is required since the components obtained by decomposition of the three-way array are sorted according to their contribution to the overall spectral variance, and this order is not necessarily maintained when the unknown sample is changed.

Absolute analyte concentrations are obtained after calibration, because the three-way array decomposition only provides relative values ($\mathbf{AB}_{3w}$). However, since the resolved chromatographic profiles for each component in the matrix $\mathbf{AB}_{3w}$ consist of merged calibration and test profiles, once identification is done, caution must be taken when building the univariate calibration graph against analyte concentrations. Component scores are generated from the elements of $\mathbf{AB}_{3w}$ according to the following expression:

$$a_{3w}(i,n) = \sum_{p=1+(i-1)J}^{iJ} ab_{3w}(p,n) \quad (8)$$

in which $a_{3w}(i,n)$ is the score for the component n in the ith sample, extracted from the nth chromatographic profile of $\mathbf{AB}_{3w}$. Subsequently, calibration is done by means of the set of standards with known analyte concentrations (contained in an $I_{cal} \times 1$ vector $\mathbf{y}_{cal}$), and regression of the first I_{cal} elements against $\mathbf{y}_{cal}$:

$$k = \mathbf{y}_{cal}{}^{+}[a_{3w}(1,n) \mid \ldots \mid a_{3w}(I_{cal},n)]. \quad (9)$$

Conversion of relative to absolute concentration of n in the unknown (y_u) proceeds by division of the test sample element by the slope of the calibration graph k:

$$y_u = a_{3w}(I_{cal} + 1, n)/k. \quad (10)$$

The above procedure is repeated for each new test sample analyzed.

In principle, the three-way Augmented PARAFAC model of Eq. (4) is unique and does not require constraints for successful decomposition [53]. However, we have added the possibility of applying similar constraints to those regularly employed in the case of MCR-ALS, i.e., non-negativity and unimodality, which can be imposed on individual profiles along each mode and in each sub-array of the augmented data set, and correspondence among species and samples in the case of samples containing uncalibrated interferents. In fact, as we shall see, in the study of simulated cases, no restrictions were implemented, but, as discussed below, for the study of experimental data, some constraints were necessary for adequate resolution. Thereby, it is possible to impose non-negativity and/or unimodality constraints in each component profile, background or interferent independently, as in MCR-ALS. Later we will see the benefits that this option can provide.

2.3. MCR-ALS

The MCR-ALS model is illustrated in Fig. 2A, and requires the following activities: (1) for each sample data, unfolding the two spectral modes into a single one, transforming the third-order arrays into matrices, and (2) augmenting the latter matrices in the direction of the elution time, which is suspected of being quadrilinearity breaking, to create a super-augmented matrix $\mathbf{X}^{(P\times Q)}$ ($Q = KL$) [6,53]. The latter matrix is decomposed according to:

$$\mathbf{X}^{(P\times Q)} = \mathbf{AB}_{2w}\mathbf{CD}_{2w}{}^{T} + \mathbf{E}^{(P\times Q)} \quad (11)$$

where $\mathbf{AB}_{2w}$ is a matrix of the N pure profiles along the augmented unfolded mode for different samples and elution times, $\mathbf{CD}_{2w}$ is a matrix of the unfolded mode combining excitation and emission spectra, and $\mathbf{E}^{(P\times Q)}$ a matrix of residuals. Eq. (11) is identical to the matrix formulation of MCR-ALS analysis, which yields $\mathbf{AB}_{2w}$ and $\mathbf{CD}_{2w}^{T}$ after minimization of the residual matrix $\mathbf{E}^{(P\times Q)}$ through an alternating least-squares procedure.

As in the classical MCR-ALS representation, the $\mathbf{AB}_{2w}$ matrix in Eq. (11) contains component concentration information in all ($I_{cal} + 1$) samples for the N resolved components, similarly to the augmented $\mathbf{AB}_{3w}$ discussed above for the Augmented PARAFAC model. This concentration information is used to construct the univariate graph, by plotting the analyte concentration scores against the nominal analyte concentrations, with the scores defined as before:

$$a_{2w}(i,n) = \sum_{p=1+(i-1)J}^{iJ} ab_{2w}(p,n) \quad (12)$$

where $a_{2w}(i,n)$ is the score for the component n in the sample i. Analyte quantitation proceeds in a similar manner to that described above for the Augmented PARAFAC model. The pure-component spectrally-unfolded profile for each component, represented by a given row of $\mathbf{CD}_{2w}^{T}$ ($N \times LK$) can be reshaped into an EEM of size $L \times K$, and used for the identification of the resolved components (see below).

As is well-known, MCR-ALS requires initialization with system parameters as close as possible to the final results. Usually this requires

analyte and interferent spectra to be known or estimated by spectral selection based on the purest variables [58]. Moreover, convergence to the chemically desired solution can be achieved by imposing certain constraints. Non-negativity restrictions in the concentrations and spectra, as well as unimodality in the chromatographic profiles are highly useful, allowing the fit to converge to the minimum with physical meaning, while it is also very practical to use correspondence among species and samples in the case of samples containing interferents. When third-order data are processed by MCR-ALS, constraints can be applied on the different modes depending on the specific manner in which the unfolded modes are combined, including the possibility of imposing tri- or quadrilinearity in the data [59].

3. Materials and methods

3.1. Reagents and solutions

Chlorophylls *a* and *b* (Chl *a* and Chl *b*) were obtained from Sigma (St. Louis, MO, USA) and used as received. Stock solutions of chlorophylls *a* and *b* were prepared by dissolving the contents of ampules containing 1 mg of each chlorophyll in 25.00 mL of acetone and stored at −4 °C in darkness. Acetone was purchased from Merck (Darmstadt, Germany). Pheophytin (Phe) stock solutions of 40 μg mL^{-1} were prepared according to a previously described procedure [60]. Solutions of the four pigments of lower concentrations were prepared by appropriate dilution with acetone. These solutions were stored at −4 °C. All other chemicals utilized were of analytical reagent grade or better.

Methanol and 1-propanol, both HPLC-grade, were purchased from Sigma (Madrid, Spain). Ultrapure water provided by a Milli-Q purification system was used. Solvents and samples used to perform the chromatographic technique were filtered through 0.22 μm nylon filter membranes before each injection.

3.2. Apparatus and software

HPLC was carried out on an Agilent 1260 Infinity Series equipped with degasser, quaternary pump, column oven, manual six-way injection valve with a 10 μL fixed loop, a multi-scan fluorescent detector (G1321B FLD) and the ChemStation software package to control the instrument and data acquisition. The analytical column used was a Poroshell 120 EC-C18 column (4.6 × 50 mm, 2.7 μm, Agilent Technologies, Inc.). The Poroshell 120 packing has a solid core of 1.7 μm in size with a porous outer layer 0.5 μm thick and a total particle size of 2.7 μm. The column temperature was set at 20 °C. Data acquisition and instrument control were performed on the HPLC 1260 software package. The mobile phase consisted of a mixture of methanol and 1-propanol and these components were filtered through a 0.22 μm membrane nylon filter and degassed by ultra-sonication before use. The flow rate was 2.0 mL min^{-1}. The elution was made applying a gradient mode increasing the percentage of 1-propanol. The gradient program was the following: from 0 to 0.2 min the percentage of 1-propanol was 40%; from 0.2 to 0.3 min the percentage of 1-propanol was increased from 40 to 70% and then was maintained constant until 0.9 min; from 0.9 to 1.0 min the percentage of 1-propanol was decreased up to 40% and from 1 to 1.1 min, was maintained constant at 40%. The sequential mode was used and, for each sample, eight chromatograms were obtained exciting at eight different excitation wavelengths (from 350 to 490 nm in 20 nm steps) and recording the emission spectra from 620 nm to 680 nm every 1 nm, in 1.8 s time steps (the data points between 0 and 0.27 min were discarded because of a dead-time artifact). Each run was accomplished in 1.1 min and the complete analysis for a specific sample was carried out in 8.8 min. These matrices were then saved in ASCII format, and transferred to a computer for subsequent manipulation.

In terms of the models presented in Section 2, to facilitate the understanding of the equations used, we can consider that for each sample, *K* chromatographic runs were collected (excited from 350 to 490 nm in steps of 20 nm), in *J* time steps ($K = 8$ and $J = 37$). In each case, the fluorescence emission was read and saved as 61 different data points ($L = 61$). Thereby, for each *k* run an elution time–emission wavelength matrix (ETEM) of size $J \times L$ was obtained, in which *J* and *L* are the number of time sensors and emission wavelengths, respectively. Once the ETEMs were obtained, data were organized for their subsequent MCR-ALS analysis. The first stage was to unfold each ETEM generating, for each sample, a row vector of dimension ($1 \times LK$), and from the latter, a matrix of $J \times LK$. Subsequently, a super-augmented data matrix $\mathbf{X}^{(P \times Q)}$ of size $J(I_{cal} + 1) \times LK$ was built. The column-wise data arrangement $\mathbf{X}^{(P \times Q)}$ is very flexible and adequate to bilinear modeling requirements, because the number of wavelengths is equal for both all the chromatographic runs. Fig. 2A shows a flow chart of the complete MCR-ALS processing of the four-way data.

All calculations were made using in-house MATLAB 7.0 routines [61]. Augmented PARAFAC was implemented in-house and inspired in the freely available codes in Bro's webpage [62]. The routines used for MCR-ALS are freely available on the Internet [63]. All programs were run on an IBM-compatible microcomputer with an Intel Core(TM) i5-2310, 2.90 GHz microprocessor and 16.00 GB of RAM.

3.3. Calibration, validation and olive oil samples

For the experimental procedure corresponding to the four-way analysis for Chl *a*, Chl *b*, Phe *a* and Phe *b*, 25 samples of calibration and six samples of validation were prepared. The tested concentrations were established by the analysis of the linear fluorescence-concentration range, from 0.5 to 1.5 μg mL^{-1} for each analyte. The specific validation concentrations are provided in Table 1. Calibration and validation samples were prepared by measuring appropriate aliquots of standard solutions, placing them in 5.00 mL volumetric flasks in order to obtain the desired concentrations, and completing to the mark with the mobile phase. Injection into the chromatographic system was made in random order and in different days.

The olive oil samples (approximately 1 g) with and without addition of different amounts of each pigment were diluted to 10 mL with 1-propanol. Aliquots of 1 mL of these solutions were diluted with 3.0 mL of 1-propanol and with 6.0 mL of methanol up to a final volume of 10.0 mL. After that, aliquots (10 μL) of each sample were injected in the chromatographic system to obtain the chromatograms at the different selected excitation wavelengths. The four-way arrays obtained were

Table 1
Predicted concentrations (μg mL^{-1}) in the validation samples using MCR-ALS and Augmented PARAFAC.

Analyte	MCR-ALS	Augmented PARAFAC	Analyte	MCR-ALS	Augmented PARAFAC
Chl a			*Phe a*		
0.40	0.50	0.47	1.20	1.28	1.26
0.80	0.92	0.89	1.40	1.32	1.27
1.40	1.48	1.48	0.80	0.80	0.79
0.60	0.56	0.60	0.60	0.55	0.6
0.90	0.99	1.02	0.90	0.85	0.87
1.00	0.99	1.05	1.00	1.054	1.09
RMSEP [a]	0.08	0.07		0.06	0.06
REP [b]	11.2	10.4		7.6	9.3
Chl b			*Phe b*		
1.40	1.38	1.25	0.80	0.79	0.82
0.40	0.43	0.53	1.20	1.27	1.28
1.20	1.21	1.01	0.40	0.48	0.40
0.90	0.91	0.99	0.90	0.88	0.86
0.60	0.50	0.59	0.60	0.72	0.64
1.00	1.11	1.05	1.00	1.04	1.04
RMSEP[a]	0.06	0.11		0.07	0.04
REP[b]	8.6	15.0		9.2	5.9

[a] RMSEP: root mean square error of prediction.
[b] REP: relative error prediction in %.

used to measure the pigments in the olive oil samples. See additional details in Ref. [50].

4. Results and discussion

4.1. Simulated data

The generation of the simulated data has been described in detail in the relevant Section 2.1. To process the data, multivariate calibration was performed in order to predict the analyte concentrations in all test mixtures. The first applied model was the Augmented PARAFAC supplying the pure profiles in both spectral modes to start the iterations. In this case, it was not necessary to apply any constraints to achieve convergence, since the model is unique. Fig. 3A shows the model performance, with satisfactory results: the root mean square error predictions (RMSEPs) and the relative error predictions (REPs, in %) are equal to 0.01 and 2.5, respectively, for both analytes.

MCR-ALS was then applied to this simulated system (see Section 2.3.), considering three components, because the potential interferent is present in test samples along with the two analytes. The initialization of the algorithm was made with the unfolded pure excitation–emission profiles, and the applied constraints were unimodality in the elution time mode (in each sub-profile corresponding to each sample) and non-negativity in both modes, while in this case the correspondence constraint was not necessary. The results in terms of how well the predicted and nominal concentrations correlate are shown in Fig. 3B for all analyzed cases. It is apparent that the algorithm performance is also satisfactory, with RMSEP and REP values of 0.04 (7.2%) for analyte 1 and 0.06 (12.1%) for analyte 2. These results encouraged us to investigate the performance of these algorithms on experimental data. In this connection, the samples analyzed in Ref. [50] are used for such purpose, as we have already mentioned. Additionally, this is the first time MCR-ALS is used to resolve the systems present in the latter work, showing another example of the versatility of this algorithm, and contributes to the general knowledge of the chemometric processing of four-way data.

4.2. Experimental data

Fig. 4A shows the excitation and emission spectra for 1.00 μg mL^{-1} of the studied analytes in methanol:1-propanol 60:40 (v/v). As already mentioned in previous sections, for each sample, 61 emission spectra were recorded between 620 and 680 nm, with data interval of 1 nm, and at intervals of 1.8 s. The excitation wavelengths were ranged between 350 and 490 nm in steps of 20 nm. These eight matrices were then mathematically assembled to obtain a three-way array for each sample.

Fig. 4B shows the profile of the optimized elution gradient, and a chromatogram of an olive oil sample without addition of analytes and the same sample with addition of 0.45 μg mL^{-1} of each one eluted under the optimized conditions. As can be seen, the selected olive oil sample (and also the remaining ones) shows an important peak in the elution time where Chl *a* appears. The fluorescence intensity of each of the analytes increases at different elution times, and highlights the fact that a significant overlapping occurs between chlorophylls and pheophytins. At 0.30 min the first peak corresponding to Chl *b* begins to elute, finishing at 0.45 min, and Chl *a* starts appearing reaching its maximum at 0.45 min. Phe *b* appears at 0.65 min and continues until 0.81 min, and Phe *a* comes out at 0.85 min.

In order to quantitate Chl *a*, Chl *b*, Phe *a* and Phe *b* simultaneously, a set of six validation samples with the analytes (Table 1) was investigated with the aid of Augmented PARAFAC and MCR-ALS. In a first phase, the new model based on Augmented PARAFAC was used. An augmented data array was constructed as described in Section 2.2, placing calibration and test third-order arrays on top of each other. Unlike in the simulated systems, during optimization the following constraints were applied to get physically meaningful solutions: (1) the concentration and spectral profiles were constrained to be non-negative, and (2) individual concentration sub-profiles along the augmented elution time mode were constrained to be unimodal. The initialization in this instance was done using pure experimental spectra, and once the optimization converged to the reasonable tolerance, the information gathered in the matrix $\mathbf{AB}_{3w}$ was used to quantify the analytes under study, according to Section 2.2.

For MCR-ALS data processing, a super-augmented matrix was built by appending each unknown sample to the twenty-five calibration standard sample data for submission to MCR-ALS analysis. During optimization, the following constraints were applied to get physically meaningful solutions: (1) non-negativity in all concentration and spectral profiles and (2) unimodality in all concentration sub-profiles. In addition, both the number of components (four in the case of validation samples, and five to the olive oil samples) and their initial spectra were estimated according to Section 2.3.

As can be seen, the results of the models used in this work do not significantly differ from each other: the REP values (in %) obtained from MCR-ALS are equal to 11.2, 8.6, 7.6 and 9.2 for Chl *a*, Chl *b*, Phe *a* and Phe *b*, respectively, while for Augmented PARAFAC are 10.4, 15.0, 9.3 and 5.9 for Chl *a*, Chl *b*, Phe *a* and Phe *b*, respectively (see Table 1). Further, these results are similar to those obtained in Ref. [50] using the unfolded partial least-squares (U-PLS) algorithm, with reported REPs of 5, 9, 5 and 8, for Chl *a*, Chl *b*, Phe *a* and Phe *b*, respectively [50]. Indeed, the RMSEP values were statistically compared to those rendered by MCR-ALS or Augmented PARAFAC using the randomization test proposed by Van der Voet to compare prediction errors [64]. The result indicates that the RMSEPs found by U-PLS are not significantly smaller than the ones by de model assayed in this work, since the probability values obtained for the four analytes is much smaller than the critical level of

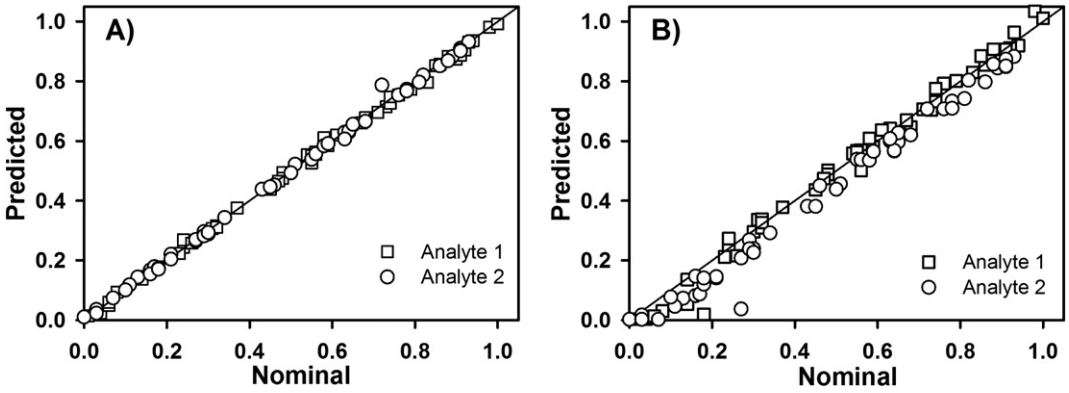

Fig. 3. Plots of predicted concentrations of the studied analytes in simulated system as a function of the nominal values, in fifty test samples with interferences: A) Augmented PARAFAC, and B) MCR-ALS.

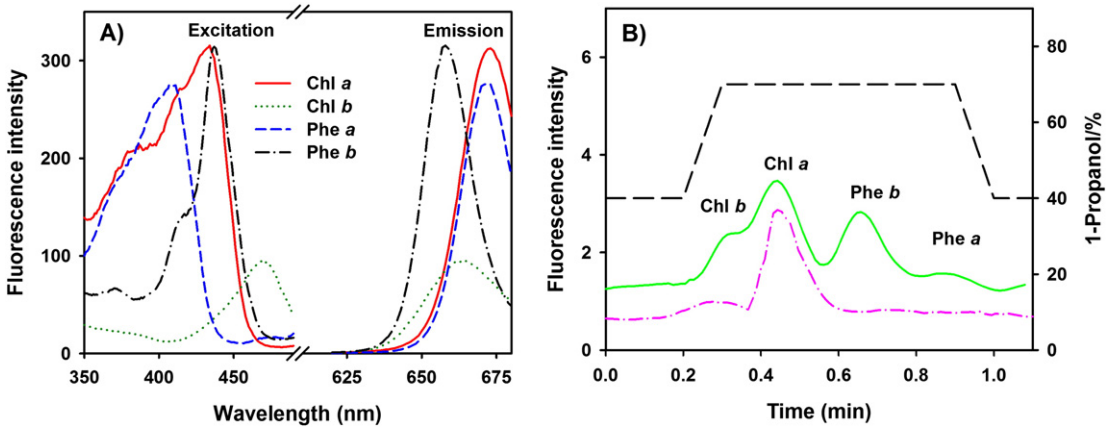

Fig. 4. (A) Excitation and emission spectra of Chl *a*, Chl *b*, Phe *a* and Phe *b* in methanol-1-propanol (60:40, v/v). All concentrations were 1.00 μg mL^{-1}. (B) Liquid chromatogram for an olive oil sample (dash-dot pink line). Spiked olive oil sample with 0.45 μg mL^{-1} of Chl *a*, Chl *b*, Phe *a* and Phe *b* (solid green line). Mobile phase: gradient program as function of 1-propanol (%) during the analysis time (dashed black line).

0.05. Besides, unlike the U-PLS approach, through both MCR-ALS and Augmented PARAFAC it is possible to obtain a more detailed description of the analyzed samples. In each case, the pure profiles (elution time, and excitation and emission fluorescence spectra) are achieved; however, while MCR-ALS gives the unfolded emission–excitation spectral profile along with the time profile for each analyte, Augmented PARAFAC gives the emission and excitation spectral profiles, independently for every component. For complex samples, this advantage allows a better identification of the analytes under study and perhaps increasing model versatility.

The above data were also processed using the PARAFAC2 model, with reasonable results regarding the validation samples, but discouraging in the case of the real samples: relative prediction errors for the two chlorophylls were ca. 25 and 35% respectively, while for the two pheophytins they were 12 and 17%. These results are consistent with previous findings concerning PARAFAC2 analyte quantitation in the presence of uncalibrated interferents [55]. The reasons why PARAFAC2 renders worse results than MCR-ALS for chromatographic-spectral data have already been explained [55].

Table 2 shows the figures of merit corresponding to the calibration curves obtained from both MCR-ALS and Augmented PARAFAC models for different samples. They were calculated according to Refs. [65] and [66] for MCR-ALS and Augmented PARAFAC. Concerning the latter, the expressions are in principle rather overoptimistic, since they were developed for quadrilinear four-way PARAFAC. However, to date there are no available expressions that compute figures of merit for Augmented PARAFAC, and therefore the present results should be employed only in order to make qualitative comparisons. Regarding the validation samples, the sensitivities (SENs), inverse of analytical sensitivities (ASENs), limits of detection (LODs) and limits of quantification (LOQs) are similar in both methods, and are close to those for the U-PLS approach [50], except for the sensitivity computed for PARAFAC in the case of Phe *a*, which is apparently low. It is well-known that figures of merit can be improved by using the data in their originally structured form [54], so better values would be expected for the model based on PARAFAC in comparison to either MCR-ALS or U-PLS. Unfortunately, this trend is not strictly observed; only the values of SEL are slightly better than the one corresponding to MCR-ALS, but the remaining results are inconclusive. Nevertheless, further research would be needed for a rigorous comparison of figures of merit.

In the case of the analyzed olive oil samples, the number of components for all cases was five, i.e., in addition to the analytes, an additional species that is not present in the calibration set is necessary. The profiles retrieved by MCR-ALS after decomposing the augmented $\mathbf{X}^{(P \times Q)}$ matrix from time-unfolded EEMs for Cornicabra olive oil sample are shown in Fig. 5. The decomposition of $\mathbf{X}^{(P \times Q)}$ furnishes a temporal profile for

Table 2
Figures of merit computed from validation and olive oil samples.[a]

Figure of merit	Chl *a*		Chl *b*		Phe *a*		Phe *b*	
	MCR-ALS	Augm. PARAFAC	MCR-ALS	Augm. PARAFAC	MCR-ALS	Augm. PARAFAC	MCR-ALS	Augm. PARAFAC
Validation samples								
SEN[b]	240	280	340	230	270	30	600	600
$[ASEN]^{-1} \times 10^{4}$[c]	4	4	1	6	7	10	2	2
SEL[d]	0.31	0.90	0.66	0.92	0.42	0.91	0.45	0.95
LOD (ng mL^{-1})[e]	1	1	4	1	2	2	1	1
LOQ (ng mL^{-1})[f]	4	4	12	4	7	7	4	4
Olive oil samples								
SEN[b]	260	270	260	190	270	100	520	405
$[ASEN]^{-1}$[c]	0.04	0.004	0.01	0.013	0.02	0.01	0.02	0.005
SEL[d]	0.29	0.84	0.51	0.71	0.40	0.65	0.43	0.80
LOD (ng mL^{-1})[e]	1	1	5	1	2	1	1	1
LOQ (ng mL^{-1})[f]	4	4	18	4	8	4	4	4

[a] The average of the figures of merit calculated for all validation and olive oil samples is reported.
[b] SEN: sensitivity, calculated according to Refs. [65] and [66] for MCR-ALS and Augmented PARAFAC, respectively.
[c] ASEN: analytical sensitivity calculated as sensitivity/sd_{test}; sd_{test}: standard deviation of replicate samples.
[d] SEL: calculated according to Refs. [65] and [66] for MCR-ALS and Augmented PARAFAC, respectively.
[e] LOD: limit of detection, calculated according to Refs. [65] and [66] for MCR-ALS and Augmented PARAFAC, respectively.
[f] LOQ: limit of quantitation, calculated according to Refs. [65] and [66] for MCR-ALS and Augmented PARAFAC, respectively, as LOD × (10/3.3).

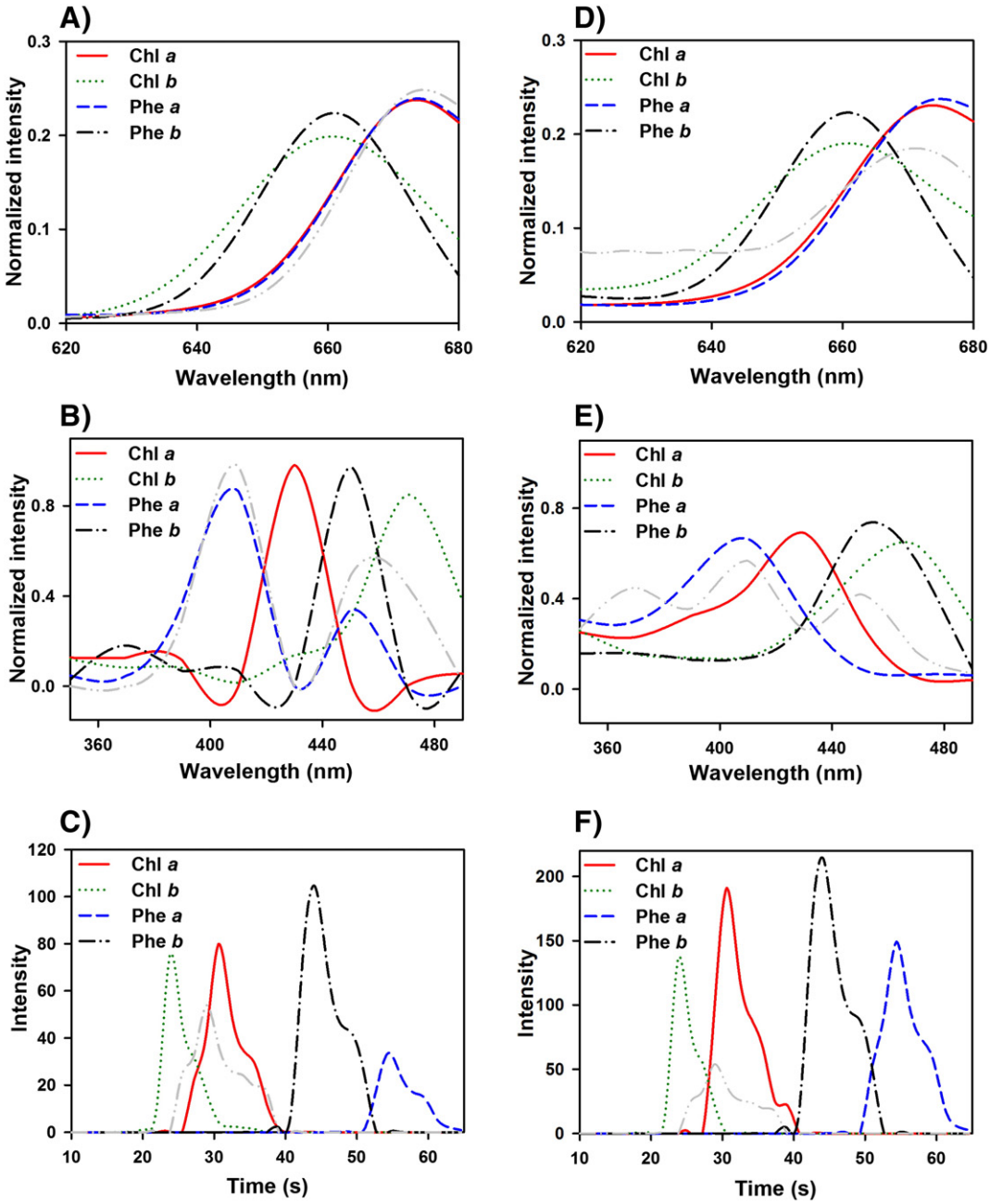

Fig. 5. Profiles retrieved by Augmented PARAFAC (A: emission, B: excitation and C: time) and MCR-ALS (D: emission, E: excitation and F: time) when the four-order data array corresponding to Cornicabra olive oil sample (1.10 μg mL^{-1} for Chl *a*, Chl *b*, Phe *a* and Phe *b*) was modeled. Chl *a*: red solid line, Chl *b*: green dotted line, Phe *a*: blue dashed line, Phe *b*: black dash-dotted line. Gray dash-dot-dotted lines correspond to the species present in the olive oil sample, not included in the calibration step (see details in text).

Chl *a*, Chl *b*, Phe *a* and Phe *b* (Fig. 5D) and an unfolded EEM profile for every component. The latter can be rearranged to obtain the excitation and emission profiles (Fig. 5E and F, respectively). This rearrangement was made by creating a suitable data matrix, submitted to singular value decomposition, and taking the first left and right eigenvectors as estimations of the true component profiles in both data modes.

Fig. 5C shows the time profiles retrieved by Augmented PARAFAC modeling of the third-order data for the same olive oil sample processed by MCR-ALS. A comparison of the spectral profiles retrieved of Augmented PARAFAC (Fig. 5A and B) with the pure spectra (Fig. 4A) enables us to conclude that the model performance is as satisfactory as for MCR-ALS. The predictive ability of the algorithms used to process the olive oil samples is summarized in Table 3. Because these samples did not contain Chl *a*, Chl *b*, Phe *a* and Phe *b* at levels higher than the attained detection limits, a recovery study was carried out by spiking them with different amounts of the four analytes. The average recoveries and REPs obtained suggest that the proposed methods are appropriate for the determination of the studied compound, besides being just as good as those obtained in Ref. [50].

The plots to be shown in Fig. 6 are the elliptical joint confidence regions (EJCR) of the slope and intercept of the linear regressions of predicted vs. nominal analyte concentrations for olive oil samples. The EJCR test is applied to gain insight into the accuracy and precision of analytical methods. If the ideal (1,0) point, i.e., unit slope and zero intercept, is contained within the ellipse, the method is considered to be accurate. On the other hand, the comparison of the relative sizes of the elliptical regions allows one to assess the relative precision of two methodologies, because the smaller the size the greater the precision.

Table 3
Predicted concentrations (μg mL^{-1}) in real samples using MCR-ALS and augmented PARAFAC models.

Olive oil[b]	Added [a]	Chl *a*		Chl *b*		Phe *a*		Phe *b*	
		MCR-ALS	Augm. PARAFAC	MCR-ALS	Augm. PARAFAC	MCR-ALS	Augm. PARAFAC	MCR-ALS	Augm. PARAFAC
Cornicabra									
	0.00	0.00	0.12	0.10	0.13	0.05	0.17	0.04	0.10
	0.55	0.48	0.55	0.71	0.72	0.54	0.58	0.67	0.58
	0.75	0.79	0.78	0.90	0.88	0.88	0.82	0.91	0.77
	1.10	1.08	1.04	1.08	1.02	1.16	1.11	1.15	1.09
Picual									
	0.00	0.01	0.04	0.08	0.23	0.03	0.21	0.11	0.08
	0.40	0.37	0.42	0.45	0.57	0.37	0.43	0.43	0.40
	0.65	0.62	0.73	0.76	0.72	0.69	0.77	0.78	0.66
	1.00	1.02	1.03	1.02	1.08	1.06	0.98	0.91	1.00
Manzanilla Cacereña									
	0.00	0.00	0.05	0.08	0.16	0.02	0.13	0.04	0.08
	0.45	0.43	0.47	0.54	0.68	0.50	0.54	0.49	0.49
	0.80	0.88	0.90	1.01	1.00	0.91	0.89	0.88	0.89
	0.95	1.00	0.98	1.20	0.94	1.07	1.01	0.99	1.05
Hojiblanca									
	0.00	0.01	0.04	0.02	0.20	0.13	0.08	0.00	0.08
	0.35	0.25	0.45	0.33	0.50	0.44	0.39	0.35	0.34
	0.70	0.68	0.87	0.67	0.88	0.96	0.92	0.68	0.66
	0.90	0.94	0.95	0.83	1.04	1.00	1.05	0.87	0.71
RMSEP[c]		0.044	0.073	0.11	0.16	0.15	0.11	0.08	0.07
REP[d]		5.9	9.7	14.5	21.0	20.0	14.7	10.7	9.3
Rec.[e]		97	108	111	122	111	111	107	101

[a] The added value is the same for all analytes in each sample.
[b] $N = 5$ for the all analyzed samples (four analytes and one interferent). The number of components was determined as explained in Sections 2.2 and 2.3.
[c] RMSEP: root mean square error of prediction.
[d] REP: relative error of prediction in %.
[e] Rec.: average recovery in %.

In the present case, as can be clearly appreciated, the sizes of the ellipses are very similar for both methods, confirming the equivalent results obtained using Augmented PARAFAC or MCR-ALS, which incidentally are all accurate because they contain the ideal (1,0) point within the ellipses.

5. Conclusions

A new PARAFAC model inspired by the augmentation concept used by MCR-ALS was developed for multivariate calibration of four-way data, based on the combination of high-performance liquid chromatography with excitation–emission fluorescence detection. The data were obtained in a very short time with an ultrafast chromatographic system.

The Augmented PARAFAC had a similar performance to the most widely used model for high-order chromatographic data (MCR-ALS), when simulated samples were processed. Moreover, these satisfactory results are achieved without use any restrictions. For the determination of Chl *a*, Chl *b*, Phe *a* and Phe *b* in different olive oils samples, the performance of Augmented PARAFAC is also satisfactory, although the presence of interferent species affects its efficiency in a similar way to MCR-ALS. It is likely that this fact made the expected improvement in figures of merit not to become fully evident. In any case, it is necessary to analyze a larger number of four-way systems to be conclusive about it.

Finally, the global results furnished by Augmented PARAFAC do not significantly differ from those already reported for the U-PLS algorithm, but, unlike the latter approach, through the new model it is possible to obtain a more detailed and realistic description of the analyzed samples, in terms of pure elution time, excitation and emission spectral profiles, which can be independently retrieved for every component. In addition, the Augmented PARAFAC approach allows applying different constraints, as non-negativity and unimodality on the spectral modes. For more complex systems, this advantage increases the potential versatility of the method. The present study is an introduction to the new algorithm and new studies are under way in our laboratory to test the advantages and/or potential shortcomings of the new algorithm when dealing with different classes of non-quadrilinear data.

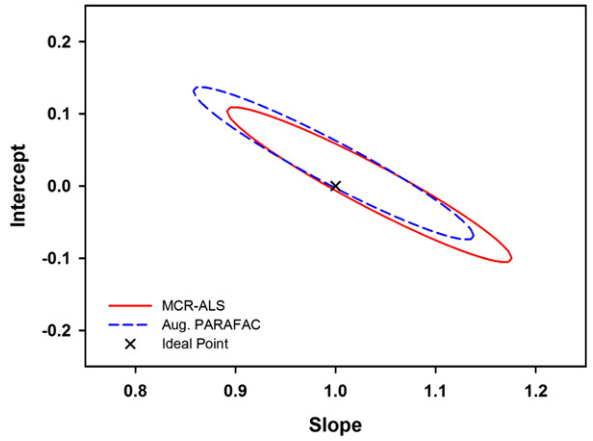

Fig. 6. Elliptic joint confidence region (EJCR) tests for the prediction results of Chl *a*, Chl *b*, Phe *a* and Phe *b* in olive oil samples for MCR-ALS (red solid line) and augmented PARAFAC (blue dash line).

Acknowledgments

The authors are grateful to the Ministerio de Economía y Competitividad of Spain (Project CTQ2011-25388) co-financed by the European FEDER funds (Projects CTQ2011-25388 and CTQ2014-

52309), the Gobierno de Extremadura and European FEDER Funds (Consolidation Project of Research Group FQM003, Project GR1003), Universidad Nacional de Rosario (Project 19B/357), Consejo Nacional de Investigaciones Científicas y Técnicas (Project PIP-0163) and Agencia Nacional de Promoción Científica y Tecnológica (Project PICT 2013-0136), for financially supporting this work.

References

[1] G.M. Escandar, N.M. Faber, H.C. Goicoechea, A. Muñoz de la Peña, A.C. Olivieri, R.J. Poppi, Second and third-order multivariate calibration: data, algorithms and applications, Trends Anal. Chem. 26 (2007) 752–765.
[2] A.C. Olivieri, G.M. Escandar, A. Muñoz de la Peña, Second-order and higher-order multivariate calibration methods applied to non-multilinear data using different algorithms, Trends Anal. Chem. 30 (2011) 607–617.
[3] G.M. Escandar, H.C. Goicoechea, A. Muñoz de la Peña, A.C. Olivieri, Second- and higher order data generation and calibration: a tutorial, Anal. Chim. Acta 806 (2014) 8–26.
[4] M. Daszykowsky, B.T. Walczak, Use and abuse of chemometrics in chromatography, Trends Anal. Chem. 25 (2006) 1081–1096.
[5] A. de Juan, R. Tauler, Chemometrics applied to unravel multicomponent processes and mixtures. Revisiting latest trends in multivariate resolution, Anal. Chim. Acta 500 (2003) 195–210.
[6] H. Parastar, R. Tauler, Multivariate curve resolution and hyphenated and multidimensional chromatographic measurements: a new insight to address current chromatographic challenges, Anal. Chem. 86 (2014) 286–297.
[7] C.J. Appellof, E.R. Davidson, Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents, Anal. Chem. 53 (1981) 2053–2056.
[8] R. Ferrer, J. Guiteras, J.L. Beltran, Development of fast-scanning fluorescence spectra as a detection system for high-performance liquid chromatography determination of polycyclic aromatic hydrocarbons in water samples, J. Chromatogr. A 779 (1997) 123–130.
[9] J.L. Beltran, J. Guiteras, R. Ferrer, Three-way multivariate calibration procedures applied to high-performance liquid chromatography coupled with fast-scanning fluorescence spectrometry detection. Determination of polycyclic aromatic hydrocarbons in water samples, Anal. Chem. 70 (1998) 1949–1955.
[10] S.A. Bortolato, J.A. Arancibia, G.M. Escandar, Non-trilinear chromatographic time retention — fluorescence emission data coupled to chemometric algorithms for the simultaneous determination of 10 polycyclic aromatic hydrocarbons in the presence of interferences, Anal. Chem. 81 (2009) 8074–8084.
[11] R.A. Gimeno, J.L. Beltran, R.M. Marce, F. Borrull, Determination of naphthalenesulfonates in water by on-line ion-pair solid-phase extraction and ion-pair liquid chromatography with fast-scanning fluorescence detection, J. Chromatogr. A 890 (2000) 289–294.
[12] F. Cañada Cañada, J.A. Arancibia, G.M. Escandar, G.A. Ibañez, A. Espinosa Mansilla, A. Muñoz de la Peña, A.C. Olivieri, Second-order multivariate calibration procedures applied to high-performance liquid chromatography coupled to fast-scanning fluorescence detection for the determination of fluoroquinolones, J. Chromatogr. A 1216 (2009) 4868–4876.
[13] A. Mancha de Llanos, M.M. de Zan, M.J. Culzoni, A. Espinosa Mansilla, F. Cañada Cañada, A. Muñoz de la Peña, H.C. Goicoechea, Determination of marker pteridines in urine by HPLC with fluorimetric detection and second-order multivariate calibration using MCR-ALS, Anal. Bioanal. Chem. 399 (2011) 2123–2135.
[14] M.J. Culzoni, A. Mancha de Llanos, M.M. de Zan, A. Espinosa Mansilla, F. Cañada Cañada, A. Muñoz de la Peña, H.C. Goicoechea, Enhanced MCR-ALS modeling of HPLC with fast scan fluorimetric detection second-order data for quantitation of metabolic disorder marker pteridines in urine, Talanta 85 (2011) 2368–2374.
[15] K.S. Booksh, B.R. Kowalski, Theory of analytical chemistry, Anal. Chem. 66 (1994) 782A–791A.
[16] A.C. Olivieri, Analytical advantages of multivariate data processing. One, two, three, infinity? Anal. Chem. 80 (2008) 5713–5720.
[17] H.L. Wu, Y. Li, R.Q. Yu, Recent developments of chemical multiway calibration methodologies with second-order or higher-order advantages, J. Chemometr. 28 (2014) 476–486.
[18] D.S. Burdick, X.M. Tu, L.B. McGown, D.W. Millican, Resolution of multicomponent fluorescent mixtures by analysis of the excitation–emission-frequency array, J. Chemometr. 4 (1990) 15–28.
[19] H.C. Goicoechea, S. Yu, A.F.T. Moore, A.D. Campligia, Four-way modeling of 4.2 time-resolved excitation–emission fluorescence data for the quantitation of polycyclic aromatic hydrocarbons in soil samples, Talanta 101 (2012) 330–336.
[20] H.C. Goicoechea, S. Yu, A.C. Olivieri, A.D. Campligia, Four-way data coupled to parallel factor model applied to environmental analysis: determination of 2,3,7,8-tetrachloro-dibenzo-para-dioxin in highly contaminated waters by solid–liquid extraction laser-excited time resolved Shpol'kii spectroscopy, Anal. Chem. 77 (2005) 2608–2616.
[21] X.D. Qing, H.L. Wu, X.F. Yan, Y. Li, L.Q. Ouyang, C.C. Nie, R.Q. Yu, Development of a novel alternating quadrilinear decomposition algorithm for the kinetic analysis of four-way room-temperature phosphorescence data, Chemom. Intell. Lab. Syst. 132 (2014) 8–17.
[22] Y. Tan, J.H. Jiang, H.L. Wu, H. Cui, R.Q. Yu, Resolution of kinetic system of simultaneous degradations of chlorophylls a and b by PARAFAC, Anal. Chim. Acta 412 (2000) 195–202.
[23] R.P.H. Nikolajsen, K.S. Booksh, A.M. Hansen, R. Bro, Quantifying catecholamines using multi-way kinetic modelling, Anal. Chim. Acta 475 (2003) 137–150.
[24] A.C. Olivieri, J.A. Arancibia, A. Muñoz de la Peña, I. Durán Merás, A. Espinosa Mansilla, Second-order advantage achieved with four-way fluorescence excitation–emission-kinetic data processed by parallel factor analysis and trilinear least-squares. Determination of methotrexate and leucovorin in human urine, Anal. Chem. 76 (2004) 5657–5666.
[25] J.A. Arancibia, A.C. Olivieri, D. Bohoyo Gil, A. Espinosa Mansilla, I. Durán Merás, A. Muñoz de la Peña, Trilinear least-squares and unfolded-PLS coupled to residual trilinearization: new chemometric tools for the data analysis of four-way instrumental data, Chemom. Intell. Lab. Syst. 80 (2006) 77–86.
[26] A. Muñoz de la Peña, I. Durán Merás, A. Jiménez Girón, Four-way calibration applied to the simultaneous determination of folic acid and methotrexate in urine samples, Anal. Bioanal. Chem. 385 (2006) 1289–1297.
[27] A. Muñoz de la Peña, I. Durán Merás, A. Jimenez Girón, H.C. Goicoechea, Evaluation of unfolded-partial least-squares coupled to residual trilinearization for four-way calibration of folic acid and methotrexate in human serum samples, Talanta 72 (2007) 1261–1268.
[28] A. Jiménez Girón, I. Durán Merás, A. Espinosa Mansilla, A. Muñoz de la Peña, F. Cañada Cañada, A.C. Olivieri, On line photochemically induced excitation–emission-kinetic four-way data. Analytical application for the determination of folic acid and its two main metabolites in serum by U-PLS and N-PLS/residual trilinearization (RTL) calibration, Anal. Chim. Acta 622 (2008) 94–103.
[29] P.C. Damiani, I. Durán-Merás, A. Garcia-Reiriz, A. Jimenez Girón, A. Muñoz de la Peña, A.C. Olivieri, Multiway partial least-squares coupled to residual trilinearization: a genuine multidimensional tool for the study of third-order data. Simultaneous analysis of procaine and its metabolite p-aminobenzoic acid in equine serum, Anal. Chem. 79 (2007) 6949–6958.
[30] S.H. Zhu, H.L. Wu, A.L. Xia, J.F. Nie, Y.C. Bian, C.B. Cai, R.Q. Yu, Excitation–emission-kinetic fluorescence coupled with third-order calibration for quantifying carbaryl and investigating the hydrolysis in effluent water, Talanta 77 (2009) 1640–1646.
[31] R.M. Maggio, P.C. Damiani, A.C. Olivieri, Four-way kinetic-excitation–emission fluorescence data processed by multi-way algorithms. Determination of carbaryl and 1-naphtol in water samples in the presence of fluorescent interferents, Anal. Chim. Acta 677 (2010) 97–107.
[32] A.G. García Reiriz, P.C. Damiani, A.C. Olivieri, F. Cañada Cañada, A. Muñoz de la Peña, Nonlinear four-way kinetic-excitation–emission fluorescence data processed by a variant of parallel factor analysis and by a neural network model achieving the second-order advantage: malonaldehyde determination in olive oil samples, Anal. Chem. 80 (2008) 7248–7256.
[33] Y.C. Kim, J.A. Jordan, M.L. Nahorniak, K.S. Booksh, Photocatalytic degradation excitation–emission matrix fluorescence for increasing the selectivity of polycyclic aromatic hydrocarbon analysis, Anal. Chem. 77 (2005) 7679–7686.
[34] M.L. Nahorniak, G.A. Cooper, Y.C. Kim, K.S. Booksh, Three and four-way parallel factor (PARAFAC) analysis of photochemically induced excitation–emission kinetic fluorescence data, Analyst 130 (2005) 85–93.
[35] A.L. Xia, H.L. Wu, S.F. Li, S.H. Zhu, L.Q. Hu, R.Q. Yu, Alternating penalty quadrilinear decomposition algorithm for an analysis of four-way data arrays, J. Chemometr. 21 (2007) 133–144.
[36] C. Kang, H.L. Wu, Y.J. Yu, Y.J. Liu, S.R. Zhang, X.H. Zhang, An alternative quadrilinear decomposition algorithm for four-way calibration to analysis of four-way fluorescence excitation–emission-pH data array, Anal. Chim. Acta 758 (2013) 45–47.
[37] Y.J. Liu, H.L. Wu, C. Kang, H.W. Gu, J.F. Nie, S.S. Li, Four-way self-weighted alternating normalized residue fitting algorithm with application for the analysis of serotonin in human plasma, Anal. Sci. 28 (2012) 1097–1104.
[38] C. Kang, H.L. Wu, L.X. Xie, S.X. Xiang, R.Q. Yu, Directly quantitative analysis of aromatic amino acids in human plasma by four-way calibration using intrinsic fluorescence: exploration of third-order advantages, Talanta 122 (2014) 293–301.
[39] N. Rodríguez, M.C. Ortiz, L.A. Sarabia, Fluorescence quantification of tetracycline in the presence of quenching matrix effect by means of a four-way model, Talanta 77 (2009) 1129–1136.
[40] H.Y. Fu, H.L. Wu, Y.J. Yu, L.L. Yu, S.R. Zhang, J.F. Nie, S.F. Li, R.Q. Yu, A new third-order calibration method with application for analysis of four-way data arrays, J. Chemometr. 25 (2011) 408–429.
[41] R.M. Maggio, A. Muñoz de la Peña, A.C. Olivieri, Unfolded partial least-squares with residual quadrilinearization: a new multivariate algorithm for processing five-way data achieving the second-order advantage. Application to fourth-order excitation–emission-kinetic-pH fluorescence analytical data, Chemom. Intell. Lab. Syst. 109 (2011) 178–185.
[42] A.E. Sinha, J.L. Hope, B.J. Prazen, C.G. Fraga, E.J. Nilsson, R.E. Synovec, Multivariate selectivity as a metric for evaluating comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry subjected to chemometric peak deconvolution, J. Chromatogr. A 1056 (2004) 145–154.
[43] A.E. Sinha, C.G. Fraga, B.J. Prazen, R.E. Synovec, Trilinear chemometric analysis of two dimensional comprehensive gas chromatography–time-of-flight mass spectrometry data, J. Chromatogr. A 1027 (2004) 269–277.
[44] R.E. Molher, K.M. Dombek, J.C. Hoggar, E.T. Young, R.E. Synovec, Comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry analysis of metabolites in fermenting and respiring yeast cells, Anal. Chem. 78 (2006) 2700–2709.
[45] S.E.G. Porter, D.R. Stoll, S.C. Rutan, P.W. Carr, J.D. Cohen, Analysis of four-way two-dimensional liquid chromatography-diode array data: application to metabolomics, Anal. Chem. 78 (2006) 5559–5569.
[46] Y. Zhang, H.L. Wu, A.L. Xiu, H.L. Hu, H.F. Zon, R.Q. Yu, Trilinear decomposition method applied to removal of three-dimensional background drift in comprehensive two-dimensional separation data, J. Chromatogr. A 1167 (2007) 178–183.

[47] H.P. Bailey, S.C. Rutan, Chemometric resolution and quantitation of four-way data arising from comprehensive 2D-LC–DAD analysis of human urine, Chemom. Intell. Lab. Syst. 106 (2011) 131–141.
[48] R. Bro, Multi-Way Analysis in the Food Industry. Models, Algorithms and Applications(Doctoral Thesis) University of Amsterdam, Netherlands, 1998. 253.
[49] E. Bezemer, S.C. Rutan, Analysis of three- and four-way data using multivariate curve resolution-alternating least squares with global multi-way kinetic fitting, Chemom. Intell. Lab. Syst. 81 (2006) 82–93.
[50] V.A. Lozano, A. Muñoz de la Peña, I. Durán-Merás, A. Espinosa Mansilla, G.M. Escandar, Four-way multivariate calibration using ultra-fast high-performance liquid chromatography with fluorescence excitation–emission detection. Application to the direct analysis of chlorophylls a and b and pheophytins a and b in olive oils, Chemom. Intell. Lab. Syst. 125 (2013) 121–131.
[51] M.R. Alcaráz, G.G. Siano, M.J. Culzoni, A. Muñoz de la Peña, H.C. Goicoechea, Modelling four and three-way fast high-performance liquid chromatography with fluorescence detection data for quantitation of fluoroquinolones in water samples, Anal. Chim. Acta 809 (2014) 37–46.
[52] A. Smilde, R. Bro, P. Geladi, Multi-Way Analysis: Applications in the Chemical Sciences, Wiley, Chichester, 2004.
[53] A.C. Olivieri, G.M. Escandar, Practical Three-Way Calibration, Elsevier, Waltham (MA), USA, 2014.
[54] X. Liu, S.D. Sidiropoulos, Cramer–Rao lower bounds for low-rank decomposition of multidimensional arrays, IEEE Trans. Signal Process. 49 (2001) 2074–2086.
[55] S.A. Bortolato, A.C. Olivieri, Chemometric processing of second-order liquid chromatographic data with UV–vis and fluorescence detection. A comparison of multivariate curve resolution and parallel factor analysis 2, Anal. Chim. Acta 842 (2014) 11–19.
[56] R. Bro, PARAFAC: tutorial & applications, Chemom. Intell. Lab. Syst. 38 (1997) 149–171.
[57] C.R. Rao, S. Mitra, Generalized Inverse of Matrices and its Applications, Wiley, New York, 1971.
[58] W. Windig, J. Guilment, Interactive self-modeling mixture analysis, Anal. Chem. 63 (1991) 1425–1432.
[59] A. Malik, R. Tauler, Extension and application of multivariate curve resolution-alternating least squares to four-way quadrilinear data-obtained in the investigation of pollution patterns on Yamuna River, India — a case study, Anal. Chim. Acta 794 (2013) 20–28.
[60] T. Galeano Díaz, I. Durán Merás, C. Arturo Correa, B. Roldan, M.I. Rodríguez Cáceres, Simultaneous fluorometric determination of chlorophylls a and b and pheophytins a and b in olive oil by partial least-squares calibration, J. Agric. Food Chem. 51 (2003) 6934–6940.
[61] MATLAB, Version 7.0, The Mathworks Inc., Natick, Massachusetts, 2010.
[62] http://www.models.life.ku.dk/.
[63] http://www.mcrals.info/.
[64] H. van der Voet, Comparing the predictive accuracy of models using a simple randomisation test, Chemom. Intell. Lab. Syst. 25 (1994) 313–323.
[65] C. Bauza, G.A. Ibañez, R. Tauler, A.C. Olivieri, Sensitivity equation for quantitative analysis with multivariate curve resolution-alternating least-squares: theoretical and experimental approach, Anal. Chem. 84 (2012) 8697–8706.
[66] A.C. Olivier, K. Faber, New developments for the sensitivity estimation in four-way calibration with the quadrilinear parallel factor model, Anal. Chem. 84 (2012) 186–193.