Chemometrics and Intelligent Laboratory Systems 109 (2011) 178–185

Contents lists available at SciVerse ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

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

Rubén M. Maggio [a], Arsenio Muñoz de la Peña [b], Alejandro C. Olivieri [a,*]

[a] Departamento de Química Analítica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Instituto de Química de Rosario (IQUIR-CONICET), Suipacha 531, S2002LRK, Argentina
[b] Departamento de Química Analítica, Universidad de Extremadura, 06006, Badajoz, Spain

ARTICLE INFO

Article history:
Received 17 June 2011
Received in revised form 4 September 2011
Accepted 7 September 2011
Available online 16 September 2011

Keywords:
Five-way data
Partial least-squares
Residual quadrilinearization
Fluorescent analysis

ABSTRACT

Unfolded partial least-squares in combination with residual quadrilinearization (U-PLS/RQL), is developed as a new latent structured algorithm for the processing of fourth-order instrumental data. In order to check its analytical predictive ability, fluorescence excitation-emission-kinetic-pH data were measured and processed. The concentration of the fluorescent pesticide carbaryl was determined in the presence of the pesticides fuberidazole and thiabendazole as uncalibrated interferents, in the first example of fourth-order multivariate calibration. The hydrolysis of the analyte was followed at different pH values using a fast-scanning spectrofluorimeter, recording the excitation-emission fluorescence matrices during its evolution to produce 1-naphthol, which does also emit fluorescence. A set of test samples containing the above mentioned fluorescent contaminants was analyzed with the new model, comparing the results with those from parallel factor analysis (PARAFAC). The newly developed U-PLS/RQL model provides better figures of merit for analyte quantitation (average prediction error, 7 μg L^{-1}, relative prediction error, 5%, calibration range, 50–250 μg L^{-1}), and is considerably simpler than PARAFAC in its implementation. The latter, however, furnishes important physicochemical information regarding the chemical process under study, although this requires the data to be unfolded into an array of lower dimensions, due to the lack of quadrilinearity of the experimental data.

1. Introduction

The recent developments in multi-dimensional analytical instrumentation and data collection are producing data arrays of increasing complexity, which are particularly useful for quantitative analysis in complex mixtures. It is apparent that this progress towards multi-way data offers theoretical and practical advantages from an analytical point of view [1-3]. For example, whereas zeroth-order univariate calibration cannot detect sample components producing an interfering signal, first-order calibration, which operates using a vector of data per sample, may compensate for these potential interferents, provided they are included in the calibration set, a property known as the 'first-order advantage' [4]. Going one step further, second-order data lead to three-way arrays which can be uniquely decomposed, allowing relative concentrations and profiles of the individual components in the different domains to be extracted directly. In this way, analytes can be quantified even in the presence of unknown interferents which are not included in the calibration set. This property has been generally recognized as the 'second-order advantage', a term coined in 1994 [4], although the first experimental demonstration of this advantage was reported in 1978, when perylene was determined in mixtures with anthracene, by suitable processing of fluorescence excitation-emission matrix data [5]. Second-order data for a given sample can be easily produced in a variety of ways, either in a single instrument or by resorting to instrument hyphenation. One of the simplest examples is a fluorescence excitation-emission matrix (EEM). When recorded for a group of samples, they can be 'stacked' into a three-way array, which can be processed by any of the available second-order algorithms, whose relative advantages have been recently reviewed [1-3]. These characteristics of second-order data support the attractive predominance of second-order calibration and the numerous analytical applications reported in the subject in recent years.

Multi-dimensional calibration methods are not limited to second-order. Third-order, and, in general, *N*th-order multivariate calibration, can also be applied with similar purposes. Third-order data can be obtained in different ways; one common strategy is to record EEMs when following the kinetic evolution of a chemical reaction for a

* Corresponding author.
E-mail address: olivieri@iquir-conicet.gov.ar (A.C. Olivieri).

0169-7439/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemolab.2011.09.002

single sample [2]. Third-order calibration includes a similar second-order advantage, i.e., the components of interest can be quantified in the presence of uncalibrated interferents. It may also hold additional advantages, at least according to some authors. For example, the following candidate properties have been proposed for the 'third-order advantage': 1) decomposition of the third-order data array for a single sample, 2) improved algorithmic resolution of highly collinear data, or 3) increased sensitivity and selectivity [6]. However, no general consensus appears to exist on this matter, other than the fact that analytical advantages of any order, if they indeed exist, should be defined in terms of what is really gained in a strictly analytical sense [6].

Third-order data have been usually processed by resorting to the well-known parallel factor analysis (PARAFAC) [7]. The combination of trilinear least-squares with residual trilinearization (TLLS/RTL) has also been proposed as a new algorithm for four-way data processing, and shown to be useful for the analysis of complex samples [8,9]. Alternative methodologies based on the use of latent variables do also exist for processing third-order data, such as multi-dimensional or *N*-way partial least-squares (N-PLS) and the unfolded variant U-PLS, both of them lacking the second-order advantage. However, when U-PLS and N-PLS are coupled to the separate procedure known as RTL, they are also able to achieve the second-order advantage [9-12]. Matrix-based methodologies can also be applied to third-order data by first unfolding them into matrices, as has been done with multivariate curve resolution coupled to alternating least-squares (MCR-ALS) [13].

Surprisingly, only in a few cases have third-order data been recorded and used to construct quantitative calibration models and to develop analytical methodologies. This may be attributed to our lacking of a thorough understanding of their analytical advantages, or to the fact that the practical acquisition of these data arrays is still difficult to implement. Hence, although one can imagine a large number of possible forms of obtaining third-order data, those commonly used are the following: 1) with a single instrument, EEMs as a function of reaction time or decay time, and 2) with hyphenated instruments, bidimensional chromatography with time of flight mass spectrometry (TOFMS) or diode array detection (DAD), such as GC × GC-TOFMS or LC × LC-DAD, and LC-DAD as a function of reaction time (GC = gas chromatography, LC = liquid chromatography) [14].

The first third-order analytical example involved the modelling of EEMs as a function of fluorescence decay time for resolving a binary mixture using the generalized rank annihilation method (GRAM) as early as in 1990 [15]. Recently, similar data (excitation-emission phosphorescence matrices as a function of decay time, measured by laser excitation modulated emission wavelength delay-time spectrophosphorimetry in Spol'skii matrices at liquid helium temperature) were modelled by PARAFAC for analyzing 2,3,7,8-tetrachlorodibenzo-*p*-dioxin in complex environmental samples [16].

One of the first examples of the processing of third-order time-evolving EEMs was the modelling of the degradation of chlorophylls *a* and *b* extracted from spinach [17]. On the strictly analytical side, several successful third-order multivariate calibrations were developed using analogous strategies. Non-fluorescent cathecholamines were derivatized to the fluorescent 3,5,6-trihydroxyindole derivatives and analyzed by PARAFAC and N-PLS [18]. Based on analyte oxidation by permanganate, these determinations were reported: methotrexate and leucovorin in urine by PARAFAC [9], and by TLLS and unfolded-PLS (U-PLS), both coupled to RTL [8], folic acid and methotrexate in urine by PARAFAC and N-PLS [19], and in serum by U-PLS/RTL [10], and folic acid in the presence of its two main metabolites, 5-methyltetrahydrofolic acid and tetrahydrofolic acid by U-PLS/RTL and N-PLS/RTL [11]. Alkaline hydrolysis allowed similar four-way analyses: procaine and its metabolite *p*-amino benzoic acid in equine sera by N-PLS/RTL [12] and by a PARAFAC variant [20], carbaryl in effluents [21] and carbaryl and 1-naphthol in water samples in the presence of interferents [22]. Finally, the Hantzsch reaction was used for malonaldehyde quantitation in olive oils, using a non-linear variant of PARAFAC and a neural network coupled to RTL [23]. The kinetics of photochemically induced EEMs and of the degradation of EEMs were processed by four-way PARAFAC for the analysis of several non-fluorescent pesticides [24] and of ternary mixtures of polycyclic aromatic hydrocarbons respectively [25].

Examples of four-way GC × GC-TOFMS data processed by PARAFAC have been described for the resolution of four isomers (*iso*-butyl, *sec*-butyl, *tert*-butyl and *n*-butyl benzenes) [26], for the analysis of environmental samples containing fuel components, pesticides and natural products [27], and for the identification of chemical differences in metabolic extracts isolated from fermenting and respiring yeast cells [28]. Likewise, LC × LC-DAD data were used for the resolution of a sample composed of 26 indolic compounds in maize seedlings [29], to remove a background drift in samples of the traditional Chinese medicines *Rhizoma Chuanxiong* [30], and for metabolic studies applying MCR-ALS [31]. Finally, LC-DAD-kinetic data allowed to follow the hydrolysis of the Aly pesticide [32].

Fourth-order data would display the obvious advantage of providing richer analytical information, implying more stable methods towards interferences and matrix effects, and less prone to minor changes in reaction conditions. This should allow for an improvement in predictive ability. To the best of our knowledge, however, no analytical application of fourth-order data has been reported to date, in spite of the different possibilities for obtaining them. It is likely that in the near future more applications will be seen where multi-dimensional data of increasing complexity are generated. In this sense, there is a continuous requirement of algorithms which are able to extract the maximum information for a variety of purposes.

In this report, we describe a new fourth-order multivariate calibration model based on the combination of U-PLS and residual quadrilinearization (RQL), which is introduced for the first time in this work. In order to test its performance in a real case, it was applied to fourth-order data corresponding to aqueous mixtures of the pesticide carbaryl (1-naphthyl methyl carbamate) and other interfering fluorescent pesticides which are usually encountered in environmental samples, such as fuberidazole and thiabendazole. A fast-scanning spectrofluorimeter allowed to record a complete EEM in a short time, and thus excitation-emission-kinetic-pH fourth-order arrays have been easily measured for each experimental sample. Relative advantages of the newly proposed U-PLS/RQL algorithm and PARAFAC are discussed in light of the analytical and physicochemical results. The introduction of a novel methodological multi-way approach provides a new tool for the resolution of complex analytical problems.

2. Experimental

2.1. Apparatus

All fluorescence measurements were performed on a fast-scanning Varian Cary Eclipse spectrofluorometer (Melbourne, Australia) equipped with two Czerny-Turner monochromators, a xenon flash lamp, a quartz cell, and connected to a PC microcomputer via an IEEE 488 (GPIB) serial interface. Excitation-emission fluorescence matrices were recorded in the following ranges: excitation, 244–312 nm each 4 nm, emission, 311–491 nm each 2 nm, during a time of 11.2 min (2–13.2 min from the beginning of the reaction each 0.8 min). For each sample, five third-order arrays were recorded at the following pH values: 9.5, 9.8, 10.0, 10.2 and 10.8. The complete data were collected into a fourth-order array of size 18×91×15×5, making a total of 122,850 data points. The slit widths for the excitation and emission monochromators were fixed at 5 nm, and the detector voltage at 600 V. A wavelength scanning speed of 12,000 nm/min was employed, so that a complete excitation-emission fluorescence matrix was obtained in 36 seconds. The total

experimental time for a given sample was 60 min. The cell was thermostatized at 35 °C.

All pH measurements were performed on a Metrohm 713 pH Meter (Herisau, Switzerland), equipped with a glass membrane electrode, a reference electrode and a temperature sensor.

2.2. Reagents and Stock Solutions

Analytical reagents grade chemicals, pure solvents and Milli-Q (Millipore, Maryland, USA) water were used in all experiments. Carbaryl, fuberidazole and thiabendazole were purchased from Sigma-Aldrich Co. (St. Louis, MO, USA).

Stock standard solutions of carbaryl, fuberidazole and thiabendazole (1.000 g L^{-1}) were prepared in 10.00 mL volumetric flasks by dissolving accurately weighed amounts of each drug in methanol and completing to the mark with the same solvent. Working solutions of each compound (1.000 mg L^{-1}) were prepared by appropriate dilution of the corresponding stock standard solution in water, employing 10.00 mL volumetric flasks.

2.3. Calibration and Test Samples

A set of 5 calibration solutions containing the analyte in the range 50–250 μg L^{-1} for carbaryl were prepared adding 1 mL of borate buffer (0.025 mol L^{-1}, pH 9.5, 9.8, 10.0, 10.2, or 10.8) to the corresponding mixtures of working standard solutions, and completing to the mark with water in 2.00 mL volumetric flasks. Each sample was measured 2 min after its preparation, as reported above.

A set of 9 test spiked samples, containing the analyte in the range 100–250 μg L^{-1}, were prepared in 2.00 mL volumetric flasks by appropriate dilution of mixtures of the corresponding working standard solutions and the interfering agrochemicals fuberidazole and thiabendazole with water before 1 mL of borate buffer (0.025 mol L^{-1}, pH 9.5, 9.8, 10.0, 10.2, or 10.8) was added. The interferent concentrations in these samples were as follows: samples 1–5 contained thiabendazole 25 μg L^{-1}, and samples 6–9 contained fuberidazole 125 μg L^{-1}. Each sample was measured 2 min after its preparation, as reported above.

These test samples are intended to mimic truly unknown samples composed of uncalibrated substances, where a responsive background may occur. The inclusion of known chemical components in these samples had the purpose of checking whether the multivariate algorithm is able to successfully retrieve their corresponding profiles.

3. Theory

3.1. Nomenclature

In multivariate calibration the term 'order' is usually employed to denote the number of modes for the data array which is recorded for a *single sample*. The term 'way', on the other hand, is reserved for the number of modes of the mathematical object which is built by joining data arrays measured for a *group of samples*. In this sense, the classical univariate calibration, which operates using a single datum per sample, is a zeroth-order and also a one-way method. Correspondingly, first-order data per sample leads to two-way data sets, second-order data per sample to three-way data sets, third-order data per sample to four-way data sets, fourth-order data per sample to five-way data sets, etc. The analytical community seems to prefer 'order' for distinguishing the various calibration scenarios, focusing on the data dimensions collected for each sample. This is also linked to the expression 'second-order advantage', a popular expression in analytical chemistry works. On the other hand, in the context of multivariate data modelling, in applications outside the analytical arena, and in many basic chemometric works, the preferred expression is 'way'. However, this latter terminology should strictly be applied to truly multi-way algorithms such as PARAFAC. In multivariate calibration using U-PLS regression, for example, a multi-way data array is never built.

Concerning the nomenclature for the different variables described in the present manuscript, scalars are noted as italicized letters (e.g., x or X), vectors as boldface lowercase letters (e.g., $\mathbf{x}$), matrices as boldface capital letters (e.g., $\mathbf{X}$), and multi-way arrays of three or more dimensions as italicized boldface capital letters (e.g., $\boldsymbol{X}$).

3.2. U-PLS/RQL

Calibration with the U-PLS algorithm implies building a classical PLS model after unfolding the calibration data into vectors, without including data for the unknown sample [33]. The I_{cal} fourth-order calibration data arrays (size $J\times K\times L\times M$, where J, K, L and M are the number of data points in each measuring mode) are thus vectorized into $JKLM\times 1$ vectors, and a PLS model is built using these data together with the vector of calibration concentrations $\mathbf{y}$ (size $I_{cal}\times 1$). This provides a set of loadings $\mathbf{P}$ and weight loadings $\mathbf{W}$ (both of size $JKLM\times A$, 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].

If no unexpected components occurred in the test sample (i.e., the sample only contains calibrated components), $\mathbf{v}$ could be employed to estimate the analyte concentration according to:

$$y_u = \mathbf{t}_u{}^T\mathbf{v} + \bar{y}_{cal} \quad (1)$$

where $\mathbf{t}_u$ is the vector of test sample scores (size $A\times 1$) and $\bar{y}_{cal}$ is the mean calibration concentration. The second term in the right hand side of Eq. (1) is needed to de-center the predicted concentration if mean centered data are employed for PLS modeling.

The score vector $\mathbf{t}_u$ is obtained by projecting the vectorized data for the test sample vec($\boldsymbol{X}_u$) onto the space of the A latent factors:

$$\mathbf{t}_u = \left(\mathbf{W}^T\mathbf{P}\right)^{-1}\mathbf{W}^T\mathrm{vec}(\boldsymbol{X}_u) \quad (2)$$

where vec($\cdot$) is the vectorization operator.

When unexpected constituents occur in $\boldsymbol{X}_u$, the sample scores given by Eq. (2) are unsuitable for analyte prediction through Eq. (1). In this case, the residuals of the U-PLS prediction step (s_p, see Eq. (3) below) will be abnormally large in comparison with the typical instrumental noise level:

$$s_p = \|\mathbf{e}_p\|/(JKLM-A)^{1/2} = \|\,\mathrm{vec}(\boldsymbol{X}_u) - \mathbf{P}\,\mathbf{t}_u\|/(JKLM-A)^{1/2} \quad (3)$$

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

This situation can be handled by a separate procedure, introduced for the first time in the present report, and called residual quadrilinearization. The underlying philosophy of RQL is analogous to those for the already described bi- and trilinearization procedures RBL [35,36] and RTL [9], extended to the fourth dimension. RQL involves decomposing the overall signal into a part which can be modeled using the calibration data, and a remaining part due to the interferents:

$$\mathrm{vec}(X_u) = \text{Modeled signal} + \text{Interferent signal} + \mathbf{e}_u \quad (4)$$

where $\mathbf{e}_u$ is a vector collecting the unmodeled errors.

The contribution of the unexpected components is a fourth-dimensional data array, which is explained in RQL using a four-dimensional Tucker3 model. The latter can be considered as a multi-dimensional version of principal component analysis, and allows to model the interferent signal in a

flexible manner using a latent structure. Therefore, the Tucker3 loadings are in general combinations of the true spectra for the interferents.

In more specific terms, the aim of RQL is to minimize the residuals when fitting the sample data to the sum of the relevant contributions in Eq. (4):

$$\mathrm{vec}(\boldsymbol{X}_\mathrm{u}) = \mathbf{P}\mathbf{t}_\mathrm{u} + \mathrm{vec}\left[\mathrm{Tucker3}\left(\boldsymbol{E}_\mathrm{p}\right)\right] + \mathbf{e}_\mathrm{u} \tag{5}$$

$$\boldsymbol{E}_\mathrm{p} = \mathrm{reshape}\left(\mathbf{e}_\mathrm{p}\right) \tag{6}$$

where reshape(·) indicates transforming a $JKLM\times 1$ vector into a $J\times K\times L\times M$ four-way array and $\mathbf{e}_\mathrm{p}$ is from Eq. (3). The Tucker3 model in Eq. (5) is built with a number of components which should be tuned for each specific unknown sample under study (see below). This is due to the fact that RQL models the contribution from the unexpected interferents, and this might be different for each unknown sample studied. It may be noticed that the computer time required to complete the iterative Tucker3 model of Eq. (5) is comparable to that needed for PARAFAC modeling of the multi-way data array built with each test sample and the calibration samples.

During the RQL procedure, the loadings **P** are kept constant at the calibration values in Eq. (5), and $\mathbf{t}_\mathrm{u}$ is varied until the final RQL residual error s_u is minimized using a Gauss-Newton procedure, with s_u given by:

$$s_\mathrm{u} = ||\mathbf{e}_\mathrm{u}||/(JKLM)^{1/2} \tag{7}$$

where $\mathbf{e}_\mathrm{u}$ is from Eq. (4).

The overall problem can thus be formulated as the search of the vector $\mathbf{t}_\mathrm{u}$ which minimizes the residuals s_u:

$$\mathbf{t}_\mathrm{u} = \min ||\,\mathrm{vec}(\boldsymbol{X}_\mathrm{u}) - \mathbf{P}\,\mathbf{t}_\mathrm{u} - \mathrm{vec}\left[\mathrm{Tucker3}\left(\boldsymbol{E}_\mathrm{p}\right)\right]|| \tag{8}$$

When the minimization process finishes, the residuals are minimized to a level compatible with the instrumental noise, and the contribution of the potential interferents is accounted for by the loadings of Tucker3 model of Eq. (5). Therefore, the final scores $\mathbf{t}_\mathrm{u}$ are free from interferent effects, and effectively represent the contribution to the signal arising from the calibrated analytes. These scores can be confidently employed for prediction using the U-PLS predictive expression (1), multiplying the calibration vector of regression coefficients **v** by the final $\mathbf{t}_\mathrm{u}$ scores provided by the Gauss–Newton minimization.

In the present work, the Tucker3 model in Eq. (5) is constructed by restricting the loadings to be orthogonal, and with no special constraints on the core elements. For a single unexpected component, the Tucker3 model is built with a single component in all dimensions, which is straightforward and provides the four corresponding interferent profiles. For additional unexpected constituents, however, the retrieved profiles no longer resemble true spectral, time or pH profiles. Moreover, in this latter case, several different Tucker3 models could in principle be constructed, because the number of loadings may be different in each dimension. We notice that the aim which guides the RQL procedure is the minimization of the residual error term s_u of Eq. (7) to a level compatible with the degree of noise present in the measured signals. Therefore, if two unexpected components are considered, for example, one should explore the possible Tucker3 models having one or two loadings in each dimension, and select the simplest model giving a residual value of s_u which is not statistically different from the minimum one. For more unexpected components a similar procedure is recommended. The finally selected Tucker3 model is the simplest one which provides a value of s_u which is not statistically different than the noise level. This can be assessed by comparing the different squared s_u on the basis of the so-called generalized cross-validation error (GCVE), already employed in the framework of the parent RBL procedure [37]. The definition can be suitably adapted to RQL as:

$$\mathrm{GCVE} = s_\mathrm{u}\frac{(JKLM)^{1/2}}{\left[(J-N_\mathrm{RQL})(K-N_\mathrm{RQL})(L-N_\mathrm{RQL})(M-N_\mathrm{RQL})-A\right]^{1/2}} \tag{9}$$

where N_RQL is the number of RQL components. The GCVE parameter penalizes the residuals s_u for excess number of parameters, and decreases on adding RQL components, until a minimum is reached when the correct number of RQL components is achieved.

Two important properties of the U-PLS/RQL procedure must be noted in connection with the presently discussed analytical problem: 1) the predictive Eq. (1) is unique, regardless of the fact that the analyte is converted into another responsive species through a chemical reaction, and 2) the latent variable structures of both U-PLS and RQL strategies make the algorithm suitable for the analysis of data arrays which do not comply with the quadrilinearity condition. These two aspects of U-PLS/RQL constitute advantages over the more restricted PARAFAC model described in the next section.

3.3. PARAFAC

For data processing with the PARAFAC algorithm, each of the I_cal training arrays $\boldsymbol{X}_{i,\mathrm{cal}}$ are joined with the unknown sample array $\boldsymbol{X}_\mathrm{u}$ into a five-way data array $\boldsymbol{X}$, whose dimensions are $[(I_\mathrm{cal}+1)\times J\times K\times L\times M]$. Provided $\boldsymbol{X}$ follows a quadrilinear PARAFAC model, it can be written in terms of five vectors for each responsive component, designated as $\mathbf{a}_n$, $\mathbf{b}_n$, $\mathbf{c}_n$, $\mathbf{d}_n$ and $\mathbf{f}_n$, collecting the relative concentrations $[(I_\mathrm{cal}+1)\times 1]$ for component n, and the profiles in all modes ($J\times 1$, $K\times 1$, $L\times 1$ and $M\times 1$) respectively. The specific expression for a given element of $\boldsymbol{X}$ is [38]:

$$x_{ijklm} = \sum_{i=1}^{N} a_{in}b_{jn}c_{kn}d_{ln}f_{mn} + e_{ijklm} \tag{10}$$

where N is the total number of responsive components, a_{in} is the relative concentration (or score) of component n in the ith. sample, and b_{jn}, c_{kn}, d_{ln} and f_{mn} are the corresponding intensities at channels j, k, l and m, respectively. The values of $\boldsymbol{e}_{ijklm}$ are the elements of the array $\boldsymbol{E}$, which is a residual error term of the same dimensions as $\boldsymbol{X}$. The column vectors $\mathbf{a}_n$ are collected into the corresponding score matrix **A**, and the profiles $\mathbf{b}_n$, $\mathbf{c}_n$, $\mathbf{d}_n$ and $\mathbf{f}_n$ into the loading matrices **B**, **C**, **D** and **F**. A successful decomposition of $\boldsymbol{X}$, usually accomplished through an alternating least-squares minimization scheme [39,40], provides access to the relative concentrations (**A**) of individual components in the $(I_\mathrm{cal}+1)$ mixtures, whether they are chemically known or not. This constitutes the basis of the second-order advantage.

There are several relevant issues regarding the application of the PARAFAC model for the calibration of five-way data: 1) initializing the algorithm, 2) applying restrictions to the least-squares fit, 3) establishing the number of responsive components, 4) identifying specific components from the information provided by the model and 5) calibrating the model in order to obtain absolute concentrations for a particular component in an unknown sample.

Initializing PARAFAC for the study of five-way arrays can be done using several options implemented in the PARAFAC package [39]: 1) singular value decomposition (SVD) vectors, 2) random orthogonalized values and 3) the best-fitting model of several models fitted using a few iterations. The first of these alternatives was employed in the present work.

Restrictions during the PARAFAC fit are employed for retrieving physically recognizable profiles in all dimensions. In the present case, however, such restrictions were not necessary.

The number of responsive components (N) can be estimated by several methods. A useful technique is CORCONDIA, a diagnostic tool considering the PARAFAC internal parameter known as core

consistency [41]. Another useful technique is the consideration of the PARAFAC sum of squared errors (SSE), i.e., the sum of squared elements of the array $\boldsymbol{E}$ in Eq. (1)[41]:

$$\mathrm{SSE} = \sum_{i=1}^{I\mathrm{cal}+1} \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{l=1}^{L} \sum_{m=1}^{M} \left(e_{ijklm}\right)^2 \tag{11}$$

This parameter decreases with increasing N, until it stabilizes at a value corresponding to the optimum number of components.

Identification of the chemical constituents under investigation is done with the aid of the estimated profiles, mainly the emission and excitation spectra, and comparing them with those for a standard solution of the analyte of interest. This is required since the components obtained by decomposition of $\boldsymbol{X}$ 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 five-way array decomposition only provides relative values (**A**). Calibration is performed by means of the set of standards with known analyte concentrations (contained in an $I_{\mathrm{cal}} \times 1$ vector **y**), and regression of the first I_{cal} elements of column $\mathbf{a}_n$ (corresponding to the calibration samples) against **y**:

$$k = \mathbf{y}^{+} \times \mathbf{a}_n(1 \ldots I_{\mathrm{cal}}) \tag{12}$$

where '+' implies taking the generalized inverse. Conversion of relative to absolute concentration of n in the unknown proceeds by division of the last element of $\mathbf{a}_n$ (corresponding to the test sample) by the slope of the calibration graph k:

$$y_{\mathrm{u}} = a_n(I\mathrm{cal}+1)/k \tag{13}$$

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

It should be noticed that even when several interconverting species for a given analyte may occur, the values contained in the vector **y** are total analyte concentrations. In contrast, the scores $\mathbf{a}_n$ are specific for a given analyte *species*. Therefore, several pseudo-univariate graphs can in principle be envisaged, by regressing the scores for each analyte species against the nominal analyte concentrations. Usually this does not represent a problem, and analysts choose the most sensitive of these graphs to predict the analyte concentration.

In the presently studied case, where the data are not quadrilinear due to the fact that **D** and **F** profiles are mutually dependent, it might be necessary to unfold the five-way data array into a four-way array by concatenating the **D** and **F** dimensions. In such a case, the decomposition would be carried out through the following expression:

$$x_{ijkp} = \sum_{i=1}^{N} a_{in} b_{jn} c_{kn} g_{pn} + e_{ijkp} \tag{14}$$

where g_{pn} identifies an element of the **G** loading matrix (size $LM \times N$) with p running from 1 to $L \times M$. The remaining discussion concerning analyte prediction is analogous to that discussed above for the five-way array decomposition.

3.4. Software

MATLAB 7.10 was used for all calculations [42]. The PARAFAC algorithm was taken from the web page maintained by Bro (http://www.models.kvl.dk/algorithms, accessed March 2011). The U-PLS/RQL MATLAB code is available from the authors on request.

4. Results and discussion

4.1. Carbaryl hydrolysis

Carbaryl is a successful carbamate insecticide due to its broad-spectrum efficacy to control over 100 species of insects on citrus, fruit, cotton, forests and other crops. The pesticide is used indiscriminately, so its toxicity has raised the public concern about the ecosystem and human health. Carbaryl is rather stable in acidic conditions, but suffers a hydrolytic process at alkaline pH values, giving rise to 1-naphthol as the hydrolysis product (see Fig. 1) [20,21,43]. In this hydrolytic process, the fluorescence emission considerably increases, as carbaryl is significantly less fluorescent than 1-naphthol. It is also known that the pseudo first-order kinetics of the carbaryl hydrolysis is strongly pH-dependent [43]. Therefore, the presently studied hydrolysis of the analyte carbaryl introduces several challenges to multilinear algorithms such as PARAFAC: 1) strong linear dependencies, due to the fact that in the time dimension the reagent profile of the analyte is correlated to the profile for the reaction product, and 2) multilinearity losses in the data structure, due to the following reasons: a reaction which is progressing during the time required for registering a complete EEM, and the dependence of the time profiles with the pH, which precludes the obtainment of unique time and pH profiles.

4.2. U-PLS/RQL results

This model requires first to assess the number of U-PLS latent variables in order to model the calibration data. This was done with the aid of leave-one-out cross validation [34], which led to the conclusion that two PLS factors were enough to model the variability across the calibration set of samples, as expected because two responsive components (carbaryl and 1-naphthol) are present in the calibration set.

The next step was the processing of each test sample data, estimating the number of RQL components by gradually increasing the dimensionality of the Tucker3 model in Eq. (5) until the residual fit s_{u} (Eq. (7)) stabilized. A typical result in terms of retrieved profiles, i.e., the Tucker3 loadings obtained during the RQL procedure, is shown in Fig. 2. This corresponds to the test sample No. 4, which contained fuberidazole as a potential interferent, and required a single RQL component for modeling its signal. Notice that for proper comparison with the best four-way PARAFAC model (see below), the RQL-Tucker3 loadings in the pH and time modes were concatenated into a single loading. In any case, the RQL profiles are almost identical to those found by PARAFAC (Fig. 2, see the next section), and aids in isolating the interferent contribution to the total signal in order to achieve a successful analyte prediction in the test samples. The results for the remaining test samples were equally satisfactory. The complexity of the RQL-Tucker3 model needed for a successful modeling of the interfering signals was assessed by comparison of the fitting residuals s_{p} and s_{u} (Eqs. (3) and (7) respectively) for increasing number of components. In most of the test samples this number was one. The need of an additional RQL component in samples 3, 5 and 8 appears to be due to a small background signal found by RQL in the latter samples. In all cases, the initial fitting residuals s_{p} (Eq. (3)) were significantly larger than the final values of s_{u} (Eq. (7)), which were in the range from 1.3 to 2.3 fluorescence units, and comparable to the

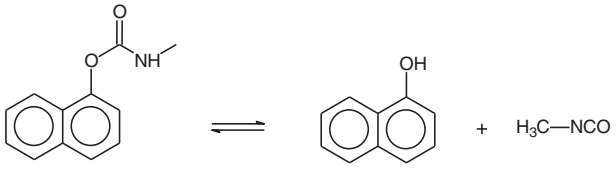

Fig. 1. Hydrolysis of carbaryl to 1-naphthol.

instrumental noise level (the latter was assessed by replicate blank measurements). It is important to note that the interferent profiles retrieved by U-PLS/RQL for a single RQL component are physically interpretable, as is shown in Fig. 2, where the profiles match those corresponding to the pure interferent. For additional RQL components, however, this qualitative information is partially lost.

It may be noted that carbaryl and its hydrolysis product 1-naphthol are mutually correlated, yet during the application of U-PLS/RQL the analyte prediction is made directly from the calibration model, without any algorithmic modification with respect to cases where the responsive constituents are uncorrelated. Table 1 also shows the U-PLS/RQL analytical results and statistical indicators for the set of test samples. Good analyte predictions are obtained, which can be attributed to the flexibility of each of the two latent-structured strategies involved in U-PLS/RQL, when handling data which are not quadrilinear. Moreover, the relative error of prediction (REP = 4.7%, Table 1) is considerably lower than that achieved in a previous third-order multivariate calibration for the same analyte (*ca.* 10%), which was carried out at a fixed pH value [22]. This indicates that a significant improvement in prediction performance is obtained by the present methodology.

4.3. PARAFAC Results

When processing five-way data for a set composed of signals for a group of samples, a requirement for a successful PARAFAC decomposition is that the data follow the quadrilinear condition, meaning that the profiles in all modes are independent of each other, and common to all samples for a given component. In principle, the excitation and emission profiles comply with this condition. However, the pH and time profiles are mutually dependent, since the kinetics of carbaryl hydrolysis is known to depend on the solution pH [43], which make the data non quadrilinear. In addition, the time profiles may introduce an additional source of quadrilinearity loss, as the practical recording of a complete EEM is a 36-second process, which takes place while the chemical reaction proceeds and modifies the constituent concentrations.

Nevertheless, a first attempt was made to process a five-way data set only containing the calibration samples, using a two-component PARAFAC model, in order to test the severity of the experimental loss of quadrilinearity. Initialization was made with the default SVD profiles included in the PARAFAC package, with no specific restrictions during the least-squares fitting phase. The results were successful in terms of excitation and emission profiles for the analyte and also for the reaction product 1-naphthol (see Fig. 3A and B). However, the retrieved time profiles are identical for all pH values, which is not the experimentally observed kinetic behavior. This should in principle lead to a lower modeling power of five-way PARAFAC towards the present data.

In order to successfully reproduce the kinetic and pH profiles using PARAFAC, it is necessary to reduce the dimensions of the five-way data set to four-way, unfolding the data by combining the pH and time modes into a single one. When this is done, and the calibration data set is processed by four-way PARAFAC, the excitation and emission profiles for the reagent (carbaryl) and for the reaction product (1-naphthol) are correctly retrieved (Fig. 3A and B). Moreover, the combined pH-time mode shows the expected increase in reaction rate in going to higher pH values. Fig. 3C shows the obtained profiles in this combined mode in comparison with those obtained by five-

Table 1
Prediction results for the test sample set.[a]

Sample	Nominal carbaryl[b]	U-PLS/RQL[c]		Four-way PARAFAC[d]	
		Predicted carbaryl	N_{RQL}	Predicted carbaryl	N
1	100	95	1	98	3
2	125	123	1	112	3
3	150	144	2	153	3
4	200	199	1	190	3
5	250	240	2	231	3
6	100	87	1	95	3
7	100	95	1	92	3
8	200	207	2	201	3
9	250	244	1	243	3
RMSE/μg L^{-1}		7.0		9.3	
REP/ %		4.7		6.2	

[a] All concentrations in μg L^{-1}. RMSE = root mean square error, REP = relative error of prediction.
[b] Besides the analyte carbaryl, samples 1–5 contained thiabendazole 25 μg L^{-1}, and samples 6–9 contained fuberidazole 125 μg L^{-1}.
[c] The number of RQL components is denoted as N_{RQL}.
[d] The number of PARAFAC components used to build the model is denoted as N.

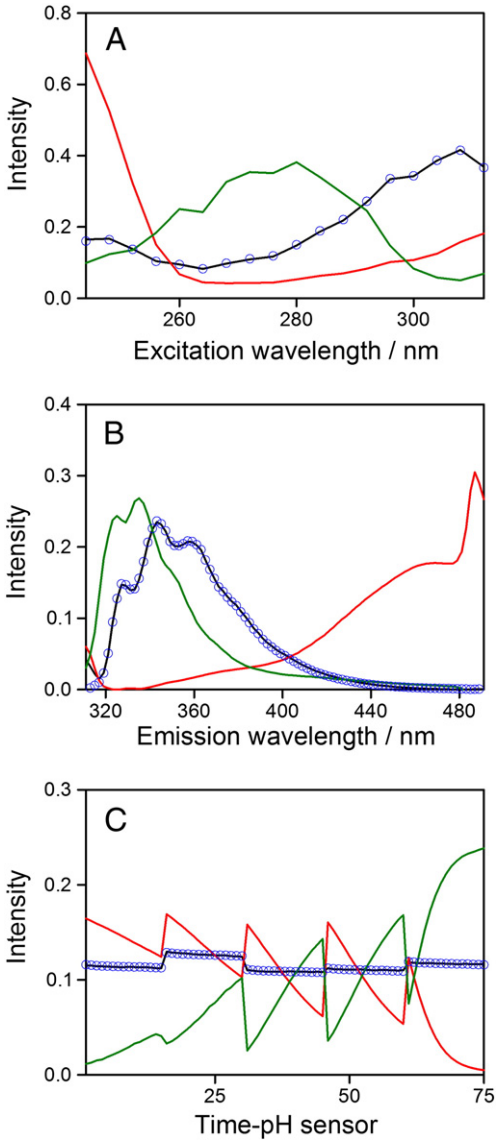

Fig. 2. Profiles recovered by the employed algorithms after processing five-way data composed of data for calibration and the test sample No. 4 containing the interferent fuberidazole. A) Excitation mode. B) Emission mode. C) Combined pH and reaction time mode. In all graphs, profiles obtained by four-way PARAFAC decomposition are as follows: green solid line, carbaryl, red solid line, 1-naphthol, black solid line, interferent (in this case matching the profiles for fuberidazole). The blue circles indicate the profiles obtained by U-PLS/RQL for the interferent, i.e., the Tucker3 loadings obtained during the RQL procedure (in plot C the pH and time profiles were concatenated for comparison with PARAFAC). All profiles are normalized to unit length, with the vertical axis in arbitrary units.

way PARAFAC (the latter were plotted by first concatenating the separate pH and time PARAFAC profiles into a single one).

It should be noticed that the SSE parameter in Eq. (2) for five-way PARAFAC (original data) was *ca.* 6×10^6 squared fluorescence, significantly larger than that computed for four-way PARAFAC unfolding the pH and time modes, which was about 3×10^6 squared units. This latter value corresponds to a residual fit of 1.6–1.9 fluorescence units, which is similar to the instrumental noise level, confirming that better modeling is obtained on lowering the number of dimensions to get a multi-linear data array.

From the individual time profiles at each pH value which are shown in Fig. 3C for the four-way PARAFAC decomposition, it is possible to estimate the hydrolytic first-order rate constant k at each pH. The corresponding values are linearly related to the pH, as can be observed in Fig. 4. This latter figure does also show the values of the rate

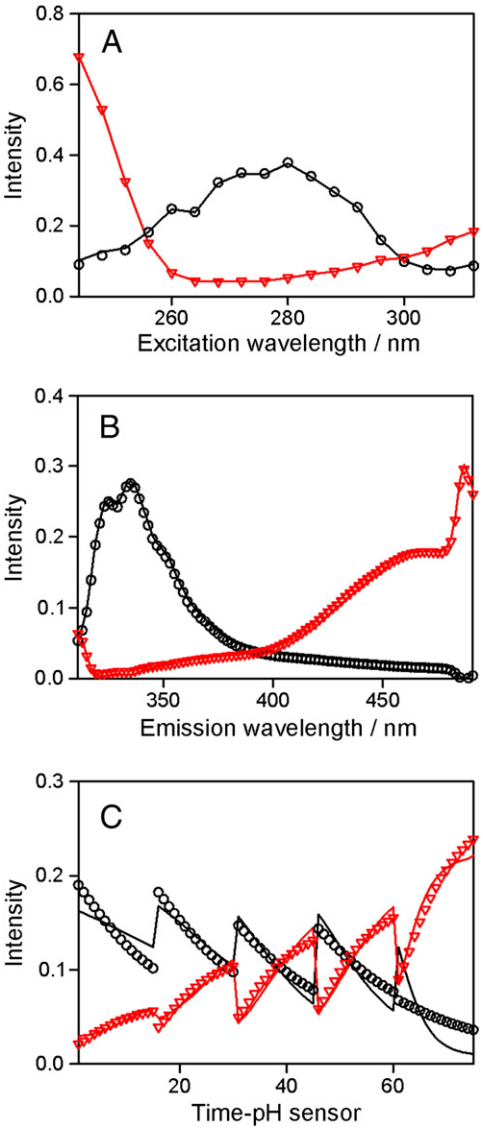

Fig. 3. Profiles recovered for the analyte carbaryl after four- and five-way PARAFAC decomposition of the calibration data set. A) Excitation mode. B) Emission mode. C) Combined pH-reaction time mode. Solid lines correspond to the four-way PARAFAC decomposition of a four-way data set obtained by combining the temporal and pH modes: black solid lines, reagent carbaryl, red solid lines, reaction product 1-naphthol. The symbols correspond to the five-way PARAFAC decomposition of the complete five-way data set (in plot C the pH and time profiles were concatenated for comparison): black circles, reagent carbaryl, red triangles, reaction product 1-naphthol. All profiles are normalized to unit length, with the vertical axis in arbitrary units.

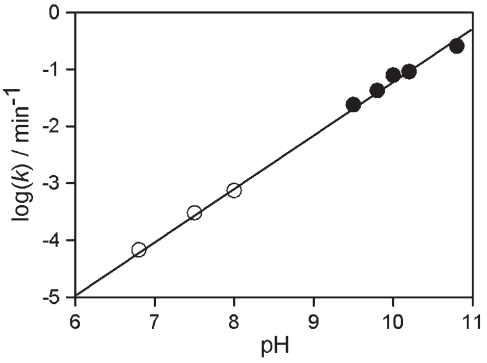

Fig. 4. Logarithmic plot of the pseudo first-order kinetic constant (k) for the hydrolysis of carbaryl as a function of pH. Black circles, this work, white circles, Ref. [43]. The solid line is for visually appreciating the correlation.

constants which were previously estimated at lower pH values [43]. A satisfactory correlation is obtained between the present results and those previously obtained (Fig. 4).

Since the best PARAFAC model is the one corresponding to the four-way array, this was applied to the analyte prediction in the test set of samples. Hence, third-order data for each test sample (obtained by concatenating the pH and time dimensions of the original fourth-order data) were joined in turn with those for the calibration samples, and the four-way arrays were submitted to PARAFAC decomposition and analyte prediction as explained above. In these cases the models were built with three components, i.e., one component in addition to the two required for the calibration samples (Table 1). The profiles for the additional component were successfully recovered by PARAFAC as corresponding to the interferent (either thiabendazole or fuberidazole). Fig. 2 shows the excitation, emission, and combined time-pH profiles recovered for the sample No. 4, where the interferent is correctly identified as the fungicide fuberidazole. Similarly satisfactory results were obtained for the remaining test samples.

As explained above, in analytical systems like the presently discussed one, two separate pseudo-univariate PARAFAC calibration graphs are possible for quantitating the analyte, i.e., one for carbaryl scores and the other one for 1-naphthol scores. This latter component displays a more intense fluorescence emission, which is less overlapped with the interferents, and hence 1-naphthol scores were selected for calibration, because they provided better sensitivity. The analytical results, in terms of recovery of the analyte concentration in the test samples and statistical indicators, are collected in Table 1. Although satisfactory results were obtained, the predictive ability of the new U-PLS/RQL model is better in this regard.

It should be noticed in this context that previous theoretical works indicate a lower statistical efficiency of multi-way decomposition when data sets are unfolded into arrays of lower dimensions [44]. This lower efficiency is expected to be translated into lower predictive ability towards the analyte. Hence in the present case the application of PARAFAC is limited by a trade-off between the need of fulfilling the quadrilinearity condition, which is achieved in four-way data sets, and the advantage of improving the predictive performance, which is possible in five-way data sets.

5. Conclusions

Fourth-order excitation-emission-kinetic-pH data have been applied for the first time for a practical analytical application: the determination of carbaryl in water samples containing other fluorescent pesticides as potential interferents. The kinetic transformation of carbaryl to 1-naphtol has been followed by recording the EEMs of the samples as a function of the reaction time and at different pH values. To process the five-way data, the classical PARAFAC model and a new algorithm developed for this purpose and introduced for the first time

in the present report, U-PLS/RQL, have been used and compared. The newly developed algorithm seems to achieve a superior performance towards the studied five-way data set in terms of predictive ability, because of its inherent latent-structured flexibility, which allows the handling of data that are not quadrilinear, as in the present case. The use of higher-order data, as the presently reported fourth-order data, opens new analytical strategies for resolving analytical situations in complex samples. The complete characterization of the advantages of using higher-order data is still necessary for a widespread popularization of this kind of data, whose detailed analytical properties are still difficult to grasp.

Acknowledgment

A.C.O. and R.M.M. gratefully acknowledge the financial support of Universidad Nacional de Rosario, CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas, Project No. PIP 1950) and Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT, PICT 2010–0084). A.M.P. acknowledges the Ministerio de Ciencia e Innovacion of Spain (Project CTQ2011-25388) and Junta de Extremadura (Consolidation Proyect GR10033 Research Group FQM003 co-financed by FEDER funds).

References

[1] R. Bro, Review on multiway analysis in chemistry — 2000–2005, Critical Reviews in Analytical Chemistry 36 (2006) 279–293.
[2] 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 in Analytical Chemistry 26 (2007) 752–765.
[3] 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 in Analytical Chemistry 30 (2011) 607–617.
[4] K.S. Booksh, B.R. Kowalski, Theory of analytical chemistry, Analytical Chemistry 66 (1994) 782A–791A.
[5] C.N. Ho, G.D. Christian, E.R. Davidson, Application of the method of rank annihilation to quantitative analyses of multicomponent fluorescence data from the video fluorometer, Analytical Chemistry 50 (1978) 1108–1113.
[6] A.C. Olivieri, Analytical advantages of multivariate data processing. One, two, three, infinity? Analytical Chemistry 80 (2008) 5713–5720.
[7] C.M. Andersen, R. Bro, Theory of net analyte signal vectors in inverse regression, Journal of Chemometrics 17 (2003) 200–215.
[8] 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, Analytical Chemistry 76 (2004) 5657–5666.
[9] 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 analysis of four-way instrumental data, Chemometrics and Intelligent Laboratory Systems 80 (2006) 77–86.
[10] A. Muñoz de la Peña, I. Durán Merás, A. Jiménez 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) 1261–1267.
[11] 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, Analytica Chimica Acta 622 (2008) 94–103.
[12] P.C. Damiani, I. Durán Merás, A.G. García Reiriz, A. Jiménez 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, Analytical Chemistry 76 (2007) 6949–6958.
[13] J. Jaumot, V. Marchán, R. Gargallo, A. Grandas, R. Tauler, Multivariate curve resolution applied to the analysis and resolution of two-dimensional [^{1}H,^{15}N] NMR reaction spectra, Analytical Chemistry 76 (2004) 7094–7101.
[14] A.E. Sinha, B.J. Prazen, R.E. Synovec, Trends in chemometric analysis of comprehensive two-dimensional separations, Analytical and Bioanalytical Chemistry 378 (2004) 1948–1951.
[15] 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, Journal of Chemometrics 4 (1990) 15–28.
[16] H.C. Goicoechea, S. Yu, A.C. Olivieri, A.D. Campiglia, 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 Spol'skii spectroscopy, Analytical Chemistry 77 (2005) 2608–2616.
[17] Y. Tan, J.H. Jiang, H.L. Wu, H. Cui, R.Q. Yu, Resolution of kinetic system of simultaneous degradations of chlorophyll a and b by PARAFAC, Analytica Chimica Acta 412 (2000) 195–202.
[18] R.P.H. Nikolajsen, K.S. Booksh, A.M. Hansen, R. Bro, Quantifying catecholamines using multi-way kinetic modelling, Analytica Chimica Acta 475 (2003) 137–150.
[19] 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, Analytical and Bioanalytical Chemistry 385 (2006) 1289–1297.
[20] 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, Journal of Chemometrics 21 (2007) 133–144.
[21] 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.
[22] 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-naphthol in water samples in the presence of fluorescent interferents, Analytica Chimica Acta 677 (2010) 97–107.
[23] 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, Analytical Chemistry 80 (2008) 7248–7256.
[24] 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 analyses, Analytical Chemistry 77 (2005) 7679–7686.
[25] M.L. Nahorniak, G.A. Cooper, Y.C. Kim, K.S. Booksh, Fiber-optic surface plasmon resonance for vapor phase analyses, Analyst 130 (2005) 85–93.
[26] 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, Journal of Chromatography. A 1056 (2004) 145–154.
[27] 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, Journal of Chromatography. A 1027 (2004) 269–277.
[28] 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, Analytical Chemistry 78 (2006) 2700–2709.
[29] 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, Analytical Chemistry 78 (2006) 5559–5569.
[30] 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, Journal of Chromatography. A 1167 (2007) 178–183.
[31] H.P. Bailey, S.C. Rutan, Chemometric resolution and quantification of four-way data arising from comprehensive 2D-LC-DAD analysis of human urine, Chemometrics and Intelligent Laboratory Systems 106 (2011) 131–141.
[32] 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, Chemometrics and Intelligent Laboratory Systems 81 (2006) 82–93.
[33] S. Wold, P. Geladi, K. Esbensen, J.J. Öhman, Multi-way principal components-and PLS-analysis, Journal of Chemometrics 1 (1987) 41–56.
[34] D.M. Haaland, E.V. Thomas, Partial least-squares methods for spectral analyses. 2. Application to simulated and glass spectral data, Analytical Chemistry 60 (1988) 1193–1202.
[35] J. Öhman, P. Geladi, S.J. Wold, Residual bilinearization. Part 1: Theory and algorithms, Journal of Chemometrics 4 (1990) 79–88.
[36] A.C. Olivieri, On a versatile second-order multivariate calibration method based on partial least-squares and residual bilinearization: Second-order advantage and precision properties, Journal of Chemometrics 19 (2005) 253–265.
[37] S.A. Bortolato, J.A. Arancibia, G.M. Escandar, Chemometrics-assisted excitation-emission fluorescence spectroscopy on nylon membranes. Simultaneous determination of benzo[a]pyrene and dibenz[a, h]anthracene at parts-per-trillion levels in the presence of the remaining EPA PAH priority pollutants as interferences, Analytical Chemistry 80 (2008) 8276–8286.
[38] S. Leurgans, R.T. Ross, Multilinear models: applications in spectroscopy, Statistical Sciences 7 (1992) 289–319.
[39] R. Bro, PARAFAC: tutorial and applications, Chemometrics and Intelligent Laboratory Systems 38 (1997) 149–171.
[40] P. Paatero, A weighted non-negative least squares algorithm for three-way PARAFAC factor analysis, Chemometrics and Intelligent Laboratory Systems 38 (1997) 223–242.
[41] R. Bro, H.A.L. Kiers, A new efficient method for determining the number of components in PARAFAC models, Journal of Chemometrics 17 (2003) 274–286.
[42] MATLAB 7.10, The MathWorks Inc., Natick, MA, 2010.
[43] J.M. Patel, D.E. Wurster, Catalysis of carbaryl hydrolysis in micellar solutions of cetyltrimethylammonium bromide, Pharmaceutical Research 8 (1991) 1155–1158.
[44] X. Liu, S.D. Sidiropoulos, Cramer–Rao lower bounds for low-rank decomposition of multidimensional arrays, IEEE Transactions on Signal Processing 49 (2001) 2074–2086.