Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Analytica Chimica Acta 651 (2009) 165–172

Contents lists available at ScienceDirect

Analytica Chimica Acta

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

A novel second-order standard addition analytical method based on data processing with multidimensional partial least-squares and residual bilinearization

Valeria A. Lozano, Gabriela A. Ibañez, Alejandro C. Olivieri *

Departamento de Química Analítica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario and Instituto de Química Rosario (IQUIR-CONICET), Suipacha 531, Rosario S2002LRK, Argentina

ARTICLE INFO

Article history:
Received 30 April 2009
Received in revised form 21 July 2009
Accepted 20 August 2009
Available online 26 August 2009

Keywords:
Standard addition
Second-order multivariate calibration
Second-order advantage
Partial least-squares
Residual bilinearization

ABSTRACT

In the presence of analyte–background interactions and a significant background signal, both second-order multivariate calibration and standard addition are required for successful analyte quantitation achieving the second-order advantage. This report discusses a modified second-order standard addition method, in which the test data matrix is subtracted from the standard addition matrices, and quantitation proceeds via the classical external calibration procedure. It is shown that this novel data processing method allows one to apply not only parallel factor analysis (PARAFAC) and multivariate curve resolution-alternating least-squares (MCR-ALS), but also the recently introduced and more flexible partial least-squares (PLS) models coupled to residual bilinearization (RBL). In particular, the multidimensional variant N-PLS/RBL is shown to produce the best analytical results. The comparison is carried out with the aid of a set of simulated data, as well as two experimental data sets: one aimed at the determination of salicylate in human serum in the presence of naproxen as an additional interferent, and the second one devoted to the analysis of danofloxacin in human serum in the presence of salicylate.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Calibration by standard addition is employed to cope with background effects, which are usually due to a change in analyte response brought about by interactions with the background, i.e., a change in the slope of the univariate signal–concentration relationship. Univariate standard addition is designed to circumvent this phenomenon [1]. More generally, an interfering background signal may be overlapped with that from the analyte. This problem can only be solved by univariate standard addition when the background response arises from the chemical treatment of the sample rather than from the sample itself. This allows one to adequately subtract it from the analyte signal (for example, by carrying out two standard additions on different sample amounts, or by combining standard addition with Youden calibration) [2]. This latter situation is not the most common one, however, and does not include the analysis of natural or biological samples containing a variety of responsive non-analytes. A background signal stemming from responsive non-analytes constitutes an interference in univariate analysis, and cannot be corrected by means of standard addition.

In the first-order multivariate calibration scenario, a generalized version of univariate standard addition method (the so-called GSAM) is available [3,4], which implies measuring first-order data (i.e., spectra) for various overlapping analytes embedded in a sample background. Generalized standard addition not only demands knowledge of the number and identity of the analytes, but also that standards of each of them are available, in order to be added in perfectly known amounts to each sample. The limitations of this method regarding the background effects are analogous to those for the univariate standard addition mode.

The presence of a responsive background, which does also affect the analyte response in a sample (for example, through inner filter effects or analyte–background interactions such as complex formation or protein binding) requires second-order standard addition for analyte quantitation [5]. This ubiquitous analytical problem can also be solved by second-order external calibration in the presence of background, provided the latter is available to be spiked with the analyte [6]. In general, however, this approach is not experimentally feasible.

Only a few references exist in the literature on this interesting standard addition multi-way research field [7–11]. The algorithm of choice for obtaining the second-order advantage from standard addition data is parallel factor analysis (PARAFAC) [12], although a recent report prefers the PARALIND variant [9] (a PARAFAC version adapted to linearly dependent systems, as described in Ref.

* Corresponding author. Tel.: +54 341 4372704; fax: +54 341 4372704.
E-mail addresses: olivieri@iquir-conicet.gov.ar, aolivier@fbioyf.unr.edu.ar (A.C. Olivieri).

0003-2670/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.aca.2009.08.027

Author's personal copy

[13]). This is because of the presence of linear dependencies in standard addition data when more than one interferent occurs in the test sample. In certain cases, standard addition PARAFAC could not be employed because of serious profile overlapping in one of the data dimensions, in which case multivariate curve resolution-alternating least-squares (MCR-ALS) [14] was successfully applied [11]. It should be noticed that linear dependency is a general phenomenon, which is not only present in standard addition data, but also in pH-gradient [14] or kinetic-modulated spectral experiments [15,16].

Recently, attention has been focused on alternative second-order multivariate calibration algorithms achieving the second-order advantage, which are based on powerful latent-structured methodologies. Pertinent examples are unfolded partial least-squares/residual bilinearization (U-PLS/RBL) [17] and multidimensional partial least-squares/residual bilinearization (N-PLS/RBL) [18]. These methods cannot be directly applied in the standard addition mode, because they include a calibration step in which nominal analyte concentrations are required, and these are neither available for test samples nor for those obtained after the addition of standards. This is somewhat deceptive, since PLS-based methods are more flexible and have been recently shown to provide better figures of merit than their competitors [6,19–22]. In some cases, they have even been found to be the only possible choice among the available second-order methodologies [19].

Interestingly, these second-order PLS/RBL methods can be applied to standard addition data, provided a recently discussed modification is incorporated, which consists of subtracting the test data matrix from the standard addition matrices, with quantitation proceeding by a classical external calibration procedure [11]. The purpose of the present work is to compare the performances of these new standard addition U-PLS/RBL and N-PLS/RBL algorithms with those based on PARAFAC, PARALIND and MCR-ALS analyses. Both simulated and experimental results indicate similar prediction abilities of the new models, suggesting that the methods herein described deserve to be added to the analyst resources for tacking complex samples with both a responsive background and analyte–background interactions.

2. Experimental

2.1. Equipment

Fluorescence excitation–emission matrices were measured with a PerkinElmer LS 55 luminescence spectrometer equipped with a xenon discharge lamp (equivalent to 20 kW for 8 μs duration) and connected to a PC microcomputer, using 1.00 cm quartz cells. Instrumental parameters were: excitation and emission slits, 5 nm, photomultiplier voltage 650, scan rate 1500 nm min^{-1}. For the experimental system 1 (analyte salicylate in serum in the presence of naproxen), excitation was scanned in the range 260–320 nm (each 0.5 nm), and emission in the range 330–494 nm (each 2 nm), producing matrices of size 121 × 83 data points. For the experimental system 2 (analyte danofloxacin in serum in the presence of salicylate), the corresponding ranges were 272–321 nm (each 0.5 nm) and 400–500 nm (each 2 nm) respectively, yielding matrices of size 99 × 51.

Data were saved in ASCII format, and transferred to a PC Sempron AMD microcomputer for subsequent manipulation by the multivariate programs.

2.2. Reagents

All chemicals used were of analytical reagent grade. For the experimental system 1, the following solutions were employed: NH_3 0.1 mol L^{-1}, prepared from commercial NH_3 (Merck, Darmstadt, Germany), stock solutions of sodium salicylate 1000 mg L^{-1} (Merck, Darmstadt, Germany) and of sodium naproxenate 1000 mg L^{-1} (Sigma, St. Louis, MO, USA), both prepared weighting the required amount of the corresponding compounds and dissolving it in doubly distilled water.

For the experimental system 2, a sodium acetate/acetic acid buffer (1.00 mol L^{-1}, pH 4.00) was used. Stock solutions of danofloxacin 100 mg L^{-1} (Riedel-de Haën, Sigma–Aldrich, Steinheim, Germany) in acetic acid 5×10^{-2} M, sodium salicylate 1000 mg L^{-1} (Merck, Darmstadt, Germany) were also prepared, weighting the required amount of the corresponding compound and dissolving it in doubly distilled water.

2.3. Procedure

For the determination of salicylate in serum in the presence of naproxen, appropriate aliquots of the corresponding stock solutions and 4.00 μL of serum were placed in a 2.00 mL volumetric flask and completion to the mark was achieved with NH_3 0.1 mol L^{-1}. The solution was placed in the measuring cell and the fluorescence excitation–emission matrix was measured. Three successive additions of analyte stock solution (1.4 μL) were then carried out, in such a way that the analyte concentrations were respectively increased by (1) 0.07, 0.14 and 0.21 mg L^{-1} for salicylate (concentration changes by dilution were considered negligible). After each addition, the samples were homogenized. The final concentration ranges for the analyzed drug was as follows (values refer to the measuring cell): salicylate, from 0.00 to 0.60 mg L^{-1}. We estimate the uncertainties in all these analyte concentrations to be of the order of ± 0.01 mg L^{-1}. The degree of serum dilution (1:500) was such that the maximum serum concentration of the studied drug was 300 mg L^{-1} for the salicylate, and ca. 100 mg L^{-1} for naproxen. All these concentration ranges are within the therapeutic values of the studied drugs in human serum.

For the determination of danofloxacin in serum in the presence of salicylate, appropriate aliquots of the corresponding stock solutions, 200 μL of acetic/acetate buffer and 13 μL of serum were placed in a 2.00 mL volumetric flask and completion to the mark was achieved with distilled water. The solution was placed in the cuvette and the matrix was measured. Three successive additions of analyte stock solution (1.0 μL) were then carried out, in such a way that the analyte concentrations were respectively increased by (1) 5.0, 10.0 and 15.0 ng L^{-1} for danofloxacin (concentration changes by dilution were considered negligible). After each addition, the samples were homogenized. The final concentration ranges for the analyzed drug was as follows (values refer to the measuring cell): danofloxacin, from 0.00 to 55.0 ng L^{-1}. We estimate the uncertainties in all these analyte concentrations to be of the order of ± 0.01 mg L^{-1}. The degree of serum dilution (1:150) was such that the maximum serum concentration of the studied drug was 5.00 mg L^{-1} for danofloxacin and ca. 200 mg L^{-1} for salicylate. All these concentration ranges are within the therapeutic values of the studied drugs in human serum.

3. Simulations

Data were simulated for multi-component mixtures having a single analyte and two potential interferents appearing in the test samples, and for the corresponding standard additions of pure analyte at known concentrations. Noiseless profiles for the analyte and for the potential interferents are shown in Fig. 1A and B in both data dimensions, leading to data matrices of size 50 × 40 data points. Using the analyte profiles shown in Fig. 1, 1000 test samples were created in which the analyte was considered to be present at concentrations which were taken at random from the range 0–1. These test samples did also contain both potential interferents, at con-

Author's personal copy

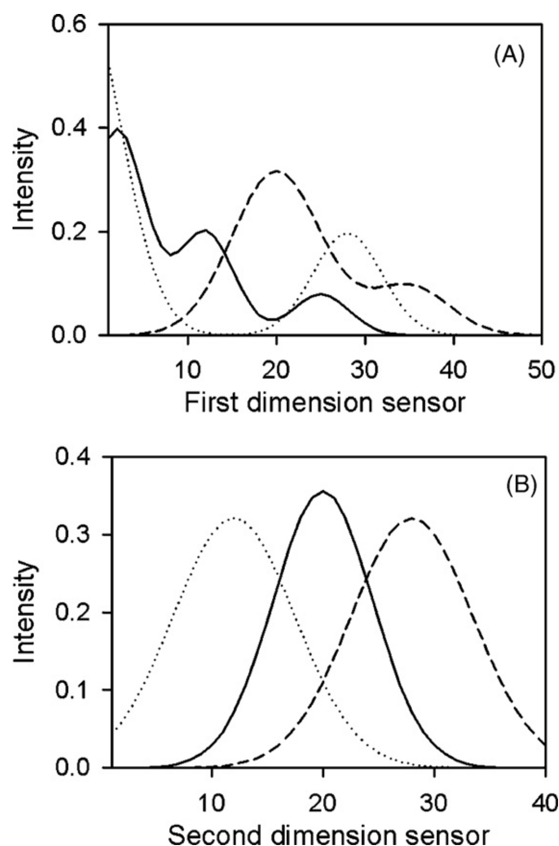

Fig. 1. Noiseless profiles employed for the simulations, in the first (A) and second dimension (B). In both cases, the solid line corresponds to the analyte, the dashed and dotted lines to the potential interferents.

centrations taken at random from the range 0.5–1.5 (to ensure that they always contain a significant amount of interferent). The matrix signal **X** for a typical test sample was given by:

$$\mathbf{X} = y_1\mathbf{b}_1\mathbf{c}_1^T + y_2\mathbf{b}_2\mathbf{c}_2^T + y_3\mathbf{b}_3\mathbf{c}_3^T \quad (1)$$

where the concentration of the analyte is y_1 and those for the interferents are y_2 and y_3, $\mathbf{b}_n$ and $\mathbf{c}_n$ (n = 1, 2, 3) are the ($J \times 1$) and ($K \times 1$) component profiles in dimensions 1 and 2 respectively, (J and K are the number of channels in each dimension) and the superscript 'T' indicates matrix transposition. The profiles $\mathbf{b}_n$ and $\mathbf{c}_n$ are shown in Fig. 1A and B, all normalized to unit length.

To give an idea of the spectral overlapping in this simulated three-component system, the selectivity parameter defined by Messick, Kalivas and Lang (MKL) [23] can be employed, as described in Ref. [24]. The corresponding value, calculated from the profiles shown in Fig. 1, is 0.89 (this can be compared below with both of the studied experimental systems).

For each of the test samples, three standard addition samples were built, having the analyte at concentrations which were 1, 2 and 3 units larger than the analyte concentration in the test samples. The interferent concentrations were kept constant in all these latter samples. Once the noiseless matrices were built, gaussian noise was added to all signals. The standard deviation of the added noise was 0.01 units, representing ca. 5% of the maximum signal for a typical test data matrix. Uncertainty in concentrations (with a standard deviation of 0.01 units) was also introduced both in the nominal analyte concentrations and in the nominal concentrations added to each sample.

The simulated data matrices, i.e., test data matrices and the corresponding standard additions matrices were processed with the following second-order multivariate calibration models: (1) PARAFAC, (2) PARALIND, (3) MCR-ALS, (4) U-PLS/RBL and (5) N-PLS/RBL, in a manner which will be described in detail below.

4. Theory

4.1. Algorithms

The theory of the second-order multivariate calibration algorithms applied in the present work is now well established and can be found in the relevant references: PARAFAC, Ref. [12], PARALIND, Ref. [13], MCR-ALS, Ref. [14], U-PLS/RBL, Ref. [17] and N-PLS/RBL, Ref. [18].

In the case of PARAFAC as applied in the usual standard addition mode 1, a three-way array is built with matrix data for a given test sample and the corresponding standard additions. Fig. 2A shows a scheme of the data matrices which are joined for PARAFAC analysis. The decomposition renders the so-called scores or relative concentrations (usually contained in the score matrix **A**, of size $I \times N$, where I is the number of samples and N the number of responsive components). The decomposition is accomplished by initializing the algorithm with scores and loadings provided by: (1) direct trilinear decomposition (DTLD) [25], (2) the best results of a set of a small number of runs, including DTLD results and vectors composed of random numbers. The possibility exists of applying non-negativity restrictions to all three modes during the least-squares fitting phase.

In the calibration mode 1, the analyte scores are employed to build a pseudo-univariate standard addition calibration graph against added analyte concentrations, predicting the concentration in the test samples in the usual univariate manner, i.e. [26]:

$$[a(1,n)|a(2,n)|\ldots|a(I,n)] = m_1[0|y^T] + n_1 \quad (2)$$

$$y_u = \frac{n_1}{m_1} \quad (3)$$

where n indicates the analyte, y_u is the predicted concentration, and **y** the vector [size $(I-1) \times 1$] of nominal concentrations added to the sample. It is assumed that sample no. 1 is the test sample and samples 2 ... I are the standard additions. As recently pointed out, in this case the scores for the interferents are constant in all samples, and hence the corresponding columns of the **A** matrix are linearly dependent [9]. One of the purposes of this simulated work was to investigate the limits in the applicability of standard PARAFAC to this analytical problem, probing different initialization conditions, and non-negativity restrictions during the least-squares fit. It should be noticed that PARALIND was recommended for this second-order standard addition mode [9]. We have applied this latter model to the simulated data, in a manner similar to PARAFAC, except that correlations between the scores of both interferents are taken into consideration, as described in Ref. [13]. The analytical phase which follows PARALIND decomposition is analogous to that discussed above for PARAFAC.

In the modified standard addition mode 2, the test data matrix is digitally subtracted from each of the standard addition matrices, creating a new data set comprised of the unknown matrix and data matrices representing the contribution of the pure analyte, embedded in the background of the test sample. The corresponding matrices are schematically shown in Fig. 2B. A new three-way array is created with these matrices, and subjected to PARAFAC decomposition. Quantitation is then possible using the standard calibration mode, i.e.:

$$[a(2,n)|a(3,n)|\ldots|a(I,n)] = m_2 y^T + n_2 \quad (4)$$

$$y_u = \frac{[a(1,n) - n_2]}{m_2} \quad (5)$$

Author's personal copy

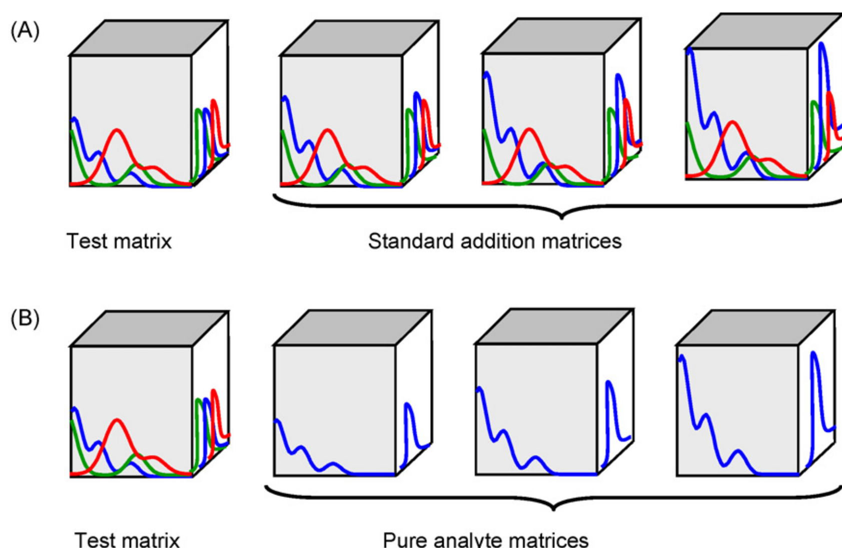

Fig. 2. Schematic representation of both standard addition modes. (A) The usual mode 1, where the matrices to be analyzed are the test sample data matrix and the standard addition matrices. The blue, red and green lines represent the profiles in both data dimensions of the analyte and two interferents, respectively. Thus, all matrices contain the same amount of interferents (red and green), whereas the analyte signal (blue) increases from left to right. (B) The modified mode 2, where the test sample data matrix is analyzed together with artificial pure analyte matrices. Here, only the test matrix contains all three components. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

In this mode, the scores for the interferents are no longer linearly dependent, since they are absent in the artificial pure analyte matrices. However, when a single test sample having more than one interferent occurs, the PARAFAC fitting may lead to linear combinations of the interferent scores and loadings. This does not imply that analyte quantitation is not possible, provided the analyte profiles are correctly retrieved, which can be achieved via proper initializing conditions and non-negativity restrictions in all three PARAFAC modes.

When applying MCR-ALS to this type of problems, the implementation of both standard addition modes is analogous to PARAFAC, as previously described in detail [11]. In the case of MCR-ALS, matrix augmentation was performed in the first dimension, initialization was made by resorting to the so-called purest variables found by SIMPLISMA (simple interactive self-modelling mixture analysis) [27], and restrictions were imposed during the least-squares phase (non-negativity in both concentration and spectral profiles). In the new standard addition mode 2, an additional restriction can be applied: the correspondence among components and samples, which involves information as to whether a given component is present or absent in certain samples. In this case, the interferent is only present in the test sample, information which is valuable during the decomposition of the augmented data matrix.

Finally, in applying the new standard addition U-PLS/RBL and N-PLS/RBL models, only the calibration mode 2 is possible. The implementation is thus analogous to that described in detail in the relevant references [17,18].

4.2. Software

All simulations were done using MATLAB [28]. PARAFAC, PARALIND and N-PLS were implemented with the codes provided by Bro in his webpage: http://www.models.kvl.dk/source/. The latter were incorporated into the useful MATLAB graphical interface MVC2 [29], available at http://www.chemometry.com/Index/Links%20and%20downloads/Programs.html. The MVC2 program does also implement the N-PLS/RBL and U-PLS/RBL combinations. MCR-ALS was applied with the graphical interface maintained by Tauler in http://www.ub.es/gesq/mcr/mcr.htm.

5. Results and discussion

5.1. Simulated data

The results of the analysis of a typical simulated test sample by different methods are now reported. PARAFAC was first applied considering three responsive components. In general, when the PARAFAC standard addition mode 1 is employed with no restrictions imposed, and starting the least-squares fit from the scores and profiles given by DTLD, the final profiles are not physically reasonable, i.e., some of the values are negative. However, as can be seen in Fig. 3, the analyte profiles and scores are correctly retrieved, permitting accurate quantitation of the analyte. The application of non-negativity restrictions leads to reasonable interferent profiles, rendering in this case reasonable scores not only for the analyte (i.e., values which increase with the addition of the analyte standard), but also for the interferents (approximately constant scores in all samples). It thus appears that non-negativity constrained PARAFAC is enough to allow for a qualitative analysis of the component profiles, and also to quantitate the analyte by standard addition, even in the presence of more than one interferent. Hence, especial variants such as PARALIND do not appear to be required.

However, when the above process was repeated for 1000 different samples, ca. 10% of the cases yielded significantly inaccurate results, even after applying non-negativity restrictions. Table 1 collects the results for the first 10 test samples, and Fig. 4 shows a box and whisker plot summarizing the complete 1000 results (algorithm 1). The results yielding predictions which are significantly different than the nominal values corresponded to PARAFAC solutions which are linear combinations of the known component profiles. The results were, however, correct when PARAFAC was initialized with scores and loadings provided by the best results of a set of 10 small runs, which included DTLD and vectors composed of random numbers (Fig. 4, algorithm 2). Table 1 collects the results for the first 10 samples using this strategy, which leaded to an overall root mean square error (RMSE) of 0.04 units (ca. 8% with respect to the mean test concentration for the analyte).

The same set of 1000 samples was studied using PARAFAC in the modified standard addition mode 2. The application of DTLD initialization and non-negativity constraints was enough to yield

Author's personal copy

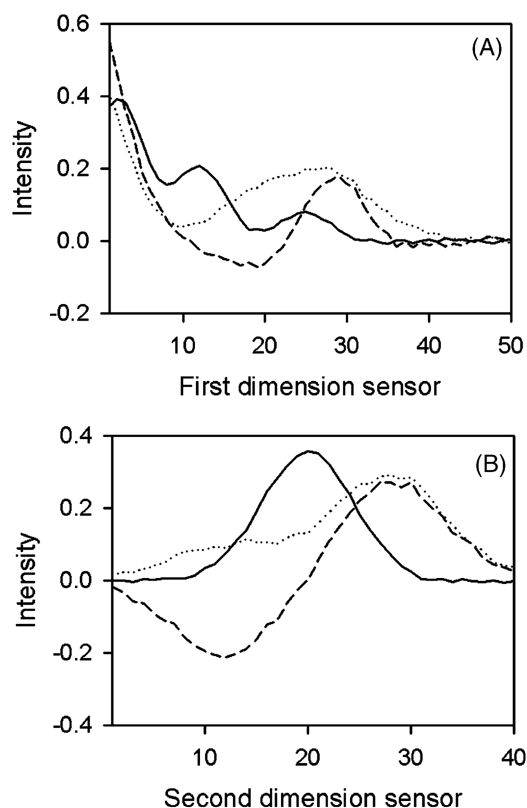

Fig. 3. Profiles retrieved when PARAFAC was applied with no restrictions, initialized with DTLD values, on a typical simulated sample using the classical standard addition mode 1. (A) Profiles in the first dimension, with the solid line corresponding to the analyte, and the dashed and dotted lines to the potential interferents. (B) Profiles in the second dimension. Lines are as in plot (A).

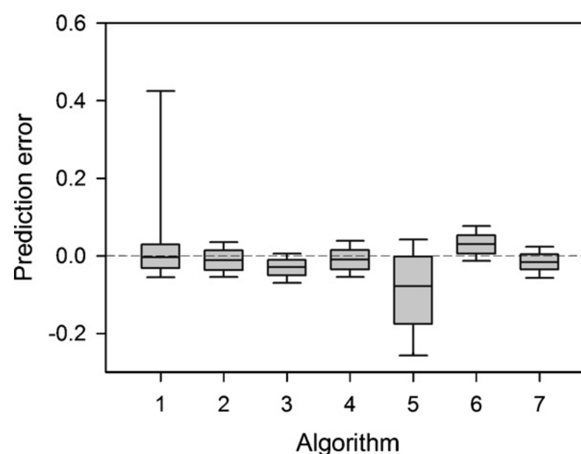

Fig. 4. Box and whisker plot of the complete 1000 prediction results corresponding to the simulated study. Algorithms are numbered in the horizontal axis as follows: (1) PARAFAC in mode 1, initialized with DTLD, (2) PARAFAC in mode 1, initialized with the best of 10 small runs, (3) PARAFAC in mode 2 initialized with DTLD, (4) PARALIND in mode 1, (5) MCR-ALS in mode 1, (6) MCR-ALS in mode 2, and (7) N-PLS/RBL in mode 2. For each algorithm, the gray boxes are bounded by the 25% and 75% quartiles with the median inside, whereas the extreme levels correspond to 5% and 95% quartiles. PARAFAC, PARALIND and MCR-ALS were all applied using non-negativity restrictions.

prediction results with an accuracy comparable to that reached by mode 1 (see Table 1 for specific results for the first 10 samples). However, Fig. 4 (algorithm 3) implies slightly poorer prediction results in comparison with the former alternative.

PARALIND was then applied considering three components. The model was initialized with the best profiles retrieved for a set of 20 calculations, run with a few iterations each. Non-negativity restrictions were imposed on all three modes. The results, in terms of predictive ability of the model, were very similar to those obtained for the best PARAFAC model, i.e., the one including non-negativity constraints and initialization with the best of 10 small runs (see Table 1 for the results corresponding to the first 10 samples, and Fig. 4 for the full predictions on the 1000 simulated test samples).

As recently discussed, MCR-ALS is another second-order multivariate calibration strategy which could in principle be applied to standard addition analytical problems in either of the presently discussed modes 1 or 2. Details are provided above and in Ref. [11]. As with PARAFAC, three components were considered to be present in all samples. The prediction results for the first 10 samples show RMSE which are comparable to the PARAFAC alternatives, as clearly displayed in Table 1. However, inspection of Fig. 4 reveals a bias in the complete results using mode 1, with a significant improvement on employment of mode 2 (in fact, the small remaining bias is comparable to the uncertainty in nominal concentrations, i.e., 0.01 units). The origin of the bias in the former case is unclear, but may be related to the strong correlations when mode 1 is used.

Finally, the latent variable structured models U-PLS/RBL and N-PLS/RBL models were applied to the 1000 sample test set, using the only possible standard addition strategy, i.e., mode 2. Calibration was performed using a single latent variable, while two components were included in the RBL phase. It should be noticed

Table 1
Prediction results in the first 10 samples of the simulated data set.

Nominal	PARAFAC[a]			PARALIND[a]	MCR-ALS[a]		N-PLS/RBL
	Mode 1[b]	Mode 1[c]	Mode 2[b]	Mode 1	Mode 1	Mode 2	Mode 2
0.49	0.48	0.43	0.45	0.45	0.53	0.48	0.46
0.15	0.18	0.22	0.20	0.21	0.19	0.14	0.21
0.78	0.68	0.71	0.73	0.73	0.82	0.77	0.74
0.07	0.03	0.04	0.03	0.04	0.11	0.06	0.06
0.11	0.46	0.08	0.09	0.09	0.15	0.10	0.10
0.76	1.20	0.75	0.75	0.74	0.80	0.75	0.77
0.50	0.52	0.51	0.50	0.50	0.54	0.49	0.52
0.55	1.09	0.47	0.49	0.48	0.59	0.54	0.51
0.05	0.70	0.04	0.05	0.05	0.09	0.04	0.07
0.13	0.12	0.11	0.12	0.12	0.17	0.18	0.13
RMSE[d]	0.32	0.05	0.04	0.04	0.04	0.02	0.03
REP%[d]	64	10.0	8.0	8.0	8.0	4.0	6.0

[a] Mode 1 is the usual standard addition mode, mode 2 is the modified standard addition strategy described in the present work.
[b] Initialized using the DTLD algorithm, and fitted by applying non-negativity restrictions during the least-squares phase.
[c] Initialized using the best of 10 small runs, and fitted by applying non-negativity restrictions during the least-squares phase.
[d] RMSE: root mean square error; REP%: relative error of prediction.

Author's personal copy

that in the case of two interferents in the test samples, the profiles retrieved by U-PLS/RBL and N-PLS/RBL do not resemble true component profiles, because they are obtained through principal component analysis [17]. The results using U-PLS/RBL were rather discouraging, since for some of the studied samples significantly inaccurate predictions were obtained. However, the use of the multidimensional variant N-PLS/RBL achieving the second-order advantage provided results with comparable accuracy to the PARAFAC and MCR-ALS alternatives. Table 1 demonstrates that the RMSE for N-PLS/RBL for the first 10 test samples is comparable to the remaining algorithms. Fig. 4 also shows that the complete results are of a quality comparable to the best PARAFAC/PARALIND results. It appears that the multidimensional variant of PLS, which maintains the matrix data structure rather than unfolding it as in the case of U-PLS, is able to better handling the analytical information in order to achieve the second-order advantage in the presently discussed systems.

5.2. Experimental data

5.2.1. Experimental system 1

The determination of salicylate in serum requires standard addition, due to changes in the analyte spectrum by interactions with the serum background [7]. When PARAFAC analysis of the different experimental data sets was attempted, the first step was the estimation of the number of responsive components. This can in principle be assessed using the diagnostic tool known as core consistency test [30], or the consideration of the residual fit of the PARAFAC model as the number of components is increased [20]. We applied this second-order method in the above discussed mode 2, with non-negativity restrictions and initialization with the best results for a set of 10 small runs. During the analysis of the experimental system 1, after PARAFAC processing of a three-way array composed of a typical spiked test serum sample and the set of standard addition samples, the progression of core consistency values was 100, 99.7, 99.5, 99.8 and 74.5 for 1–5 components respectively, while the residuals of the PARAFAC fit decreased as follows: 25, 3.5, 2.6, 2.5 and 2.4 arbitrary fluorescence units. This suggests that core consistency does not allow to firmly establish the number of responsive components, while three components appear to be optimum when analyzing the residual fit, which seemed to be a more reasonable choice.

It is interesting to inspect the three spectral profiles retrieved by PARAFAC when processing typical standard addition data in mode 2, which are shown in Fig. 5. The analyte salicylate shows excitation and emission maxima at 300 and 410 nm respectively, while the interferents can be identified as naproxen (excitation at 265 nm with a shoulder at 320 nm, and emission at 355 nm) and serum (excitation at 280 and emission at 340 nm, most probably corresponding to the fluorescence of tryptophan). When four components were extracted by PARAFAC, the fourth profile was similar to one of the first three, confirming that three components is a good choice for this type of samples. For comparison with the simulated system, the MKL selectivity in this experimental system is computed to be 0.92.

The prediction results for the set of spiked test samples are shown in Table 2, leading to a reasonably low root mean square error for the prediction of the analyte in this set of spiked samples, consistent with the high analyte selectivity. Application of PARALIND to this set of experimental data led to analogous results to those obtained from PARAFAC (Table 2).

When applying MCR-ALS, the processing of the data was similar to that already described in Ref. [11]. Three responsive components were included in the study, as confirmed by principal component analysis. The initial profiles employed to start the MCR-ALS fitting were estimated by SIMPLISMA, as discussed above in relation to

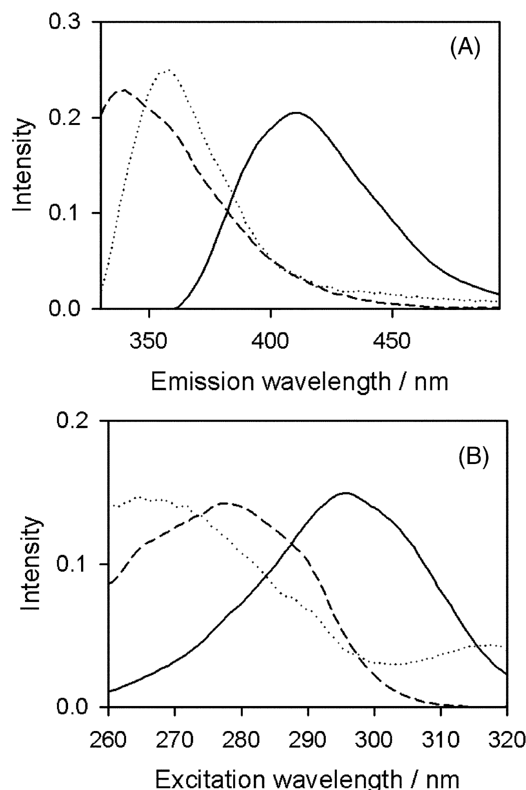

Fig. 5. Profiles retrieved when PARAFAC was applied in mode 2 with non-negativity restrictions, initialized with the best of 10 small runs, on a typical experimental sample of serum containing salicylate and naproxen. (A) Emission profiles, with the solid line corresponding to the analyte salicylate, and the dashed and dotted lines to the potential interferents. (B) Excitation profiles. Lines are as in plot (A).

the simulated data. The predictions concerning the experimental system 1 were of the same quality as PARAFAC (Table 2).

When applying U-PLS/RBL and N-PLS/RBL to this set of samples, the best result was obtained using the latter methodology. As in the case of the simulated data, highly inaccurate results were obtained for one of the samples, while N-PLS/RBL furnished an RMSE value comparable to those for PARAFAC, PARALIND and MCR-ALS (Table 2).

5.2.2. Experimental system 2

The determination of fluoroquinolone antibiotics in serum, such as danofloxacin, requires standard addition due to changes in the analyte spectrum by interactions with the serum background [11]. As in the previous system, PARAFAC was applied in mode 2, and was initialized with profiles obtained after a small set of trial runs, with non-negativity ensured in all three modes during the least-squares fit. In this case, decomposition of a three-way array composed of a typical spiked sample and the standard addition samples led to the following core consistency values: 100, 94.7, 89.9, 80.5 and 58.5 for 1–5 components respectively. The residuals of the PARAFAC fit, in turn, were estimated as 9.7, 4.7, 3.6, 3.5 and 3.5 arbitrary fluorescence units. Again, this suggests that three components appear to be optimum, a conclusion which can only be reached from the residual fit and not from core consistency diagnostics.

The three spectral profiles retrieved by PARAFAC are shown in Fig. 6 for a typical analysis. The one ascribed to the analyte danofloxacin shows excitation maxima at 280 and 320 nm and emission at 450 nm, while the interferents can be identified as salicylate (excitation at 300 nm emission at 410 nm) and serum (excitation at 280–290 nm and emission as a decreasing band which could be due to the one centred at 350 nm, which may correspond

Author's personal copy

Table 2
Prediction results in the experimental data sets.

Experimental system 1[a]					
Salicylate nominal	PARAFAC[b]	PARALIND[b]	MCR-ALS[b]	U-PLS/RBL[c]	N-PLS/RBL[c]
0.20	0.20	0.20	0.20	0.20	0.20
0.50	0.51	0.51	0.51	–[d]	0.50
0.05	0.06	0.06	0.06	0.06	0.06
0.00	0.02	0.02	0.01	0.01	0.01
0.03	0.03	0.03	0.04	0.04	0.04
RMSE[e]	0.01	0.01	0.01	0.01	0.01
Experimental system 2[f]					
Danofloxacin nominal	PARAFAC[b]	PARALIND[b]	MCR-ALS[b]	U-PLS/RBL[c]	N-PLS/RBL[c]
10	18	18	21	7	7
40	55	55	107	41	31
0	6	6	5	0	0
5	12	12	19	2	3
25	15	15	18	13	13
RMSE[e]	10	10	30	6	7

[a] Values in mg mL^{-1}. All samples contain serum and naproxen 0.20 mg mL^{-1}.
[b] PARAFAC, PARALIND and MCR-ALS were applied with three components in both data sets.
[c] U-PLS/RBL and N-PLS/RBL were applied with one latent variable for calibration and two interferents in both data sets.
[d] Highly inaccurate result. The RMSE is reported excluding this sample.
[e] RMSE: root mean square error.
[f] Values in ng mL^{-1}. All samples contain serum and salicylate (0.50 mg mL^{-1} in the first three samples and 0.30 mg mL^{-1} in the remaining two).

to tryptophan). In this case the MKL selectivity is lower than in the previous system, and is computed as 0.74.

Specific prediction results for the set of spiked test samples are shown in Table 2. In this case, where lower sensitivity towards the analyte is attained, and heavy spectral overlapping occurs in both data dimensions, the RMSE is rather high in comparison with the mean analyte concentration across the set of samples. As with the previous experimental system, the prediction results obtained from PARALIND were identical to those corresponding to PARAFAC (Table 2).

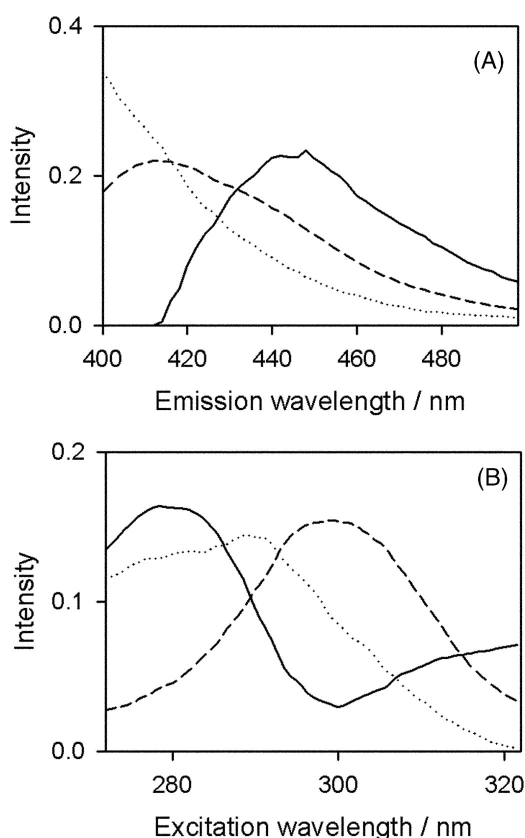

Fig. 6. Profiles retrieved when PARAFAC was applied in mode 2 with non-negativity restrictions, initialized with the best of 10 small runs, on a typical experimental sample of serum containing danofloxacin and salicylate. (A) Emission profiles, with the solid line corresponding to the analyte danofloxacin, and the dashed and dotted lines to the potential interferents. (B) Excitation profiles. Lines are as in plot (A).

When applying MCR-ALS, the predictions were clearly worse, indicating that the combination of low analyte signal and spectral overlapping have a stronger effect on this algorithm than on PARAFAC decomposition.

U-PLS/RBL and N-PLS/RBL, on the other hand, led to significantly better predictions (Table 2). This may be an indication that these latent structure methods may be better prepared to cope with the problems of severe spectral overlapping when analyzing standard addition data.

6. Conclusions

This work shows that partial least-squares models, both in the unfolded and multidimensional versions, can be applied to standard addition calibration of second-order instrumental data, provided the following operations are carried out: (1) they are coupled with residual bilinearization, which allows them to achieve the second-order advantage, and (2) the matrix data are modified, subtracting the test data matrix from all standard addition matrices. Both simulations and experiments indicate that multidimensional partial least-squares/residual bilinearization is able to render the best analytical results, comparable to those furnished by the classical algorithms parallel factor analysis (either in the usual version or by taking care of linear dependencies) and multivariate curve resolution-alternating least-squares.

References

[1] R.C. Castells, M.A. Castillo, Anal. Chim. Acta 423 (2000) 179–185.
[2] W.J. Youden, Anal. Chem. 19 (1947) 946–950.
[3] B.E.H. Saxberg, B.R. Kowalski, Anal. Chem. 51 (1979) 1031–1038.
[4] J.H. Kalivas, B.R. Kowalski, Anal. Chem. 53 (1981) 2207–2212.
[5] A.C. Olivieri, Anal. Chem. 80 (2008) 5713–5720.
[6] M.J. Culzoni, H.C. Goicoechea, A.P. Pagani, M.A. Cabezón, A.C. Olivieri, Analyst 131 (2006) 718–723.
[7] N.M. Sena, M.G. Trevisan, R.J. Poppi, Quim. Nova 28 (2005) 910–920.

Author's personal copy

[8] J.A. Arancibia, A.C. Olivieri, G.M. Escandar, Anal. Bioanal. Chem. 374 (2002) 451–459.
[9] M. Bahram, R. Bro, Anal. Chim. Acta 584 (2007) 397–402.
[10] V. Gómez, R. Cuadros, I. Ruisánchez, M.P. Callao, Anal. Chim. Acta 600 (2007) 233–239.
[11] V.A. Lozano, R. Tauler, G.A. Ibañez, A.C. Olivieri, Talanta 77 (2009) 1715–1723.
[12] R. Bro, Chemom. Intell. Lab. Syst. 38 (1997) 149–171.
[13] R. Bro, R.A. Harshman, N.D. Sidiropoulos, M.E. Lundy, J. Chemom. 23 (2009) 324–340.
[14] A. DeJuan, E. Casassas, R. Tauler, in: R.A. Myers (Ed.), Encyclopedia of Analytical Chemistry, vol. 11, Wiley, Chichester, 2002, pp. 9800–9837.
[15] M. Bahram, M. Mabhooti, J. Chemom. 23 (2009) 236–247.
[16] H. Abdollahi, S.M. Sajjadi, J. Chemom. 23 (2008) 139–148.
[17] A.C. Olivieri, J. Chemom. 19 (2005) 253–265.
[18] V.A. Lozano, G.A. Ibañez, A.C. Olivieri, Anal. Chim. Acta 610 (2008) 186–195.
[19] D.B. Gil, A. Muñoz de la Peña, J.A. Arancibia, G.M. Escandar, A.C. Olivieri, Anal. Chem. 78 (2006) 8051–8058.
[20] J.A. Arancibia, C.E. Boschetti, A.C. Olivieri, G.M. Escandar, Anal. Chem. 80 (2008) 2789–2798.
[21] P.C. Damiani, I. Durán-Merás, A. García-Reiriz, A. Jiménez-Girón, A. Muñoz de la Peña, A.C. Olivieri, Anal. Chem. 79 (2007) 6858–6949.
[22] A. García-Reiriz, P.C. Damiani, A.C. Olivieri, Talanta 71 (2007) 806–815.
[23] N.J. Messick, J.H. Kalivas, P.M. Lang, Anal. Chem. 68 (1996) 1572–1579.
[24] A.C. Olivieri, N.M. Faber, in: S. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive Chemometrics, vol. 3, Elsevier, Amsterdam, 2009, pp. 91–106.
[25] E. Sanchez, B.R. Kowalski, J. Chemom. 1 (1990) 29–45.
[26] J.N. Miller, J.C. Miller, Statistics and Chemometrics for Analytical Chemistry, Pearson-Prentice Hall, New York, 2005 (Chapter 2).
[27] W. Windig, J. Guilment, Anal. Chem. 63 (1991) 1342–1425.
[28] MATLAB 7.0, The Mathworks, Natick, Massachussets, USA, 2007.
[29] A.C. Olivieri, R.Q. Yu, H.L. Wu, Chemom. Intell. Lab. Syst. 96 (2009) 246–251.
[30] R. Bro, H.A.L. Kiers, J. Chemom. 17 (2003) 274–286.