

Sensitivity Equation for Quantitative Analysis with Multivariate Curve Resolution-Alternating Least-Squares: Theoretical and Experimental Approach

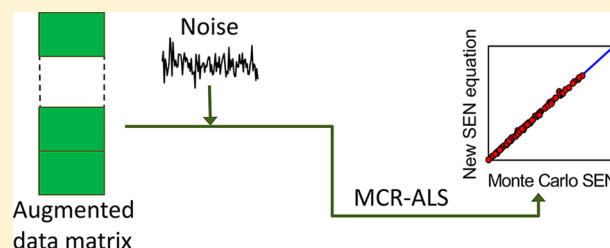
María C. Bauza,[†] Gabriela A. Ibañez,[†] Romà Tauler,[‡] and Alejandro C. Olivieri^{*,†}

[†]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

[‡]Institut de Diagnosi Ambiental i Estudis de l'Aigua (IDAEA-CSIC), Jordi Girona 18, Barcelona (08034) España

Supporting Information

ABSTRACT: A new equation is derived for estimating the sensitivity when the multivariate curve resolution-alternating least-squares (MCR-ALS) method is applied to second-order multivariate calibration data. The validity of the expression is substantiated by extensive Monte Carlo noise addition simulations. The multivariate selectivity can be derived from the new sensitivity expression. Other important figures of merit, such as limit of detection, limit of quantitation, and concentration uncertainty of MCR-ALS quantitative estimations can be easily estimated from the proposed sensitivity expression and the instrumental noise. An experimental example involving the determination of an analyte in the presence of uncalibrated interfering agents is described in detail, involving second-order time-decaying sensitized lanthanide luminescence excitation spectra. The estimated figures of merit are reasonably correlated with the analytical features of the analyzed experimental system.



Important analytical advantages are derived from multiway analysis, such as increased sensitivity and selectivity, and quantitative analyte determinations by processing calibration samples together with samples carrying potential interferences, i.e., the so-called second-order advantage.^{1–7} In the case of analyzing three-way data, it is imperative to assess if the data comply with the property of trilinearity, which in analytical chemistry terms basically implies that (1) the constituent signals are proportional to their concentrations, and (2) the component profiles along both instrumental modes (or ways) are unique and equal for all samples. Excitation–emission luminescence data, for example, are usually trilinear, because the signal intensity is directly proportional to concentration (within reasonable ranges), excitation/emission profiles of a given component are the same in all samples, and luminescence profiles for different components are linearly independent. The most common cause of trilinearity loss is the variation of the profiles for a given component in one data mode, e.g., from sample to sample.⁵ This usually occurs in chromatographic three-way data, because of the presence of run-to-run changes in both retention times and profile shapes.^{8,9}

A popular method based on the trilinear model is parallel factor analysis (PARAFAC)¹⁰ and its variants.^{11,12} Small deviations from the trilinear model can be taken into account by latent-variable based methodologies, such as unfolded and N-way partial least-squares combined with residual bilinearization (U-PLS/RBL and N-PLS/RBL, respectively).^{13,14} However, when sample-to-sample profile changes occur, the latter algorithms cannot in general be applied and an alternative in

such situations is the multivariate curve resolution-alternating least-squares (MCR-ALS)¹⁵ method. MCR-ALS is usually employed in works required to cope with deviations of the trilinear model, while the bilinear model still holds. Figure 1 illustrates the difference between a trilinear and a nontrilinear case when matrix chromatographic-spectral data are processed with MCR-ALS.

A variant of PARAFAC named PARAFAC2¹⁶ is also useful for cases where the deviations of trilinearity are not large, i.e., mostly caused by factor profiles shifting in one of the data modes.¹⁷ An additional situation where the trilinear model breaks is produced when component profiles are identical in one of the data modes. In this case, MCR-ALS still constitutes a valid alternative,^{18,19} as well as the application of methods based on the Tucker model are,²⁰ or the recently discussed version of U-PLS/RBL, which incorporates MCR-ALS to aid RBL in resolving the interfering agent profiles.²¹

In one- and two-way calibration problems the estimation of figures of merit is now firmly established, as documented in IUPAC's Technical Reports.^{22,23} The most relevant figure of merit is the sensitivity, because it is the crucial element in most expressions for estimating other figures, such as selectivity, limit of detection, limit of quantitation, and uncertainty in predicted concentrations.²⁴ Sensitivity may be defined as the change in

Received: July 10, 2012

Accepted: September 11, 2012

Published: September 11, 2012

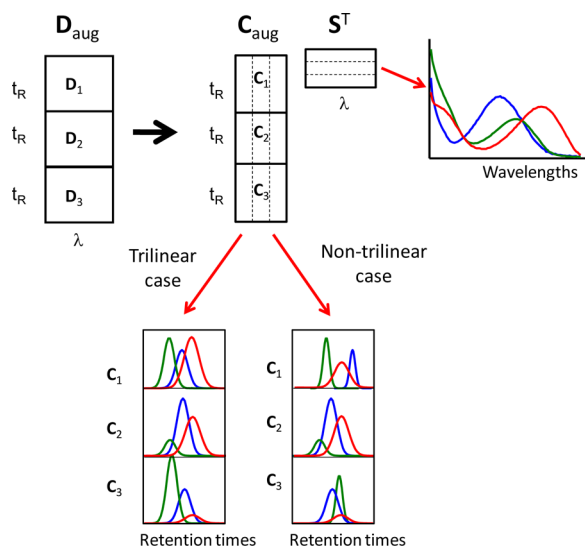


Figure 1. Scheme illustrating the difference between a trilinear data set and a nontrilinear data set, when chromatographic-spectral matrices are processed with MCR-ALS. This algorithm decomposes the augmented data matrix D_{aug} leading to the retrieval of spectral (S^T) and elution (C_{aug}) profiles for each component. In the trilinear case, the retention time profiles for each constituent are identical in all samples (submatrices C_1 , C_2 , and C_3), only differing in the relative scale (bottom-left plot). However, in the non-trilinear case, the elution profiles for a given component may differ in position and/or shape from sample to sample, as is evident in the figure (bottom-right plot).

(net) response for a given change in analyte concentration. While in univariate calibration, this is calculated as the slope of the calibration curve,²² in two-way multivariate calibration, this can be calculated as the slope of a pseudounivariate calibration curve from the value of the net analyte signal (NAS), which is estimated as the portion of the total signal uniquely ascribed to the analyte of interest.²⁵ An alternative definition, according to the so-called sensitivity analysis,²⁶ is based on the impact in the output of a model caused by the uncertainty of the input, which in analytical chemistry terms measures the sensitivity as the inverse of the ratio of predicted concentration uncertainty to signal uncertainty.²⁷

In three-way data analysis, different sensitivity definitions exist. In the framework of the PARAFAC model, for example, two alternative expressions were developed, one by Messick, Kalivas and Lang (MKL)²⁸ and another one by Ho, Christian, and Davidson (HCD).²⁹ Recently, however, it was shown that both of them are special cases of a general sensitivity definition by Faber and Olivieri (FO).³⁰ This somehow reflects the difficulties found by the NAS concept in this field.³¹ The extension of the latter approach to four-way data analysis and beyond has been troublesome, although an improved closed-form expression has recently been developed.³² In the RBL methodologies, some advances have taken place, although the picture is still far from being definite.¹⁴

In the case of MCR-ALS, some attempts have been made for defining figures of merit, based on resampling techniques or Monte Carlo noise addition approaches.³³ An experimental approach to figures of merit estimation can also be undertaken, considering the pseudounivariate scores-concentration calibration graph generated by MCR-ALS and defining parameters analogous to their univariate calibration counterparts.³⁴ This latter strategy allows one to estimate the limit of detection and quantitation. However, although analytical concentrations are

proportional to the scores recovered by the MCR-ALS algorithm, the latter have arbitrary units and do not reflect, in general, the effect of the overlapping of the analyte profiles with those for other samples components. Hence, the sensitivity cannot be defined as the slope of the MCR-ALS pseudounivariate calibration graph.

In the present report, we employ a sensitivity definition which is based on error propagation analysis as the inverse of the ratio of predicted concentration uncertainty to signal uncertainty.^{30–32} An expression suitable for MCR-ALS analytical studies is derived and supported by Monte Carlo noise addition simulations and experimental data. This may help in assessing other important figures of merit for this popular multivariate resolution methodology when used for quantitative purposes and also in putting the MCR-ALS sensitivity in perspective with the remaining second-order calibration algorithms.

THEORY

MCR-ALS. A brief description of the so-called extended MCR-ALS methodology is provided in the Supporting Information, and a derivation of the expression to estimate the sensitivity (SEN_{MCR}) is given in the Appendix:

$$SEN_{MCR} = m_n [J(S^T S)^{-1}]^{-1/2} \quad (1)$$

where n is the index for the analyte of interest in a multicomponent mixture, m_n is the slope of the MCR pseudounivariate calibration graph for this analyte, S^T is a matrix containing the profiles for all sample components in the nonaugmented MCR direction, and J is the number of channels in the test sample data matrix in the augmented MCR direction. For full details concerning these parameters see the Supporting Information.

Software. All calculations were implemented with MATLAB 7.10 routines,³⁵ available from the authors on request. The MCR-ALS algorithm was obtained from the Internet page <http://www.mcrals.info>.

Data Simulations. The first step in the simulations of the various data sets consists in creating data matrices for a number of simulated samples. In these simulations, for each sample a data matrix is created with dimensions $J \times K$, where J is the number of rows and K is the number of columns. The total number of data samples to be simultaneously analyzed using the extended MCR-ALS method is I (this includes the calibration data matrices and a single test data matrix). Details on the construction of the data sets and values of I , J , and K are given below.

Notice that, as explained in the Supporting Information, the discussed extended MCR-ALS model can be used to analyze different types of analytical scenarios. Typically the augmented data matrix is resolved into profiles contained in the C and S^T MCR matrices (see the Supporting Information) which may adopt different interpretations. For example, in the trilinear data sets B1, B2, T1, T2, T3, and Q1 to be discussed below, the C matrix represents successive excitation fluorescence profiles along the augmented mode and S^T the emission spectra along the nonaugmented mode. In the nontrilinear data set B3 (and also in the experimental data set for the determination of the analyte furosemide in the presence of uncalibrated flufenamic acid), C represents successive time decay profiles and S^T luminescence excitation spectra. Finally, in the nontrilinear

data set B4, **C** represents successive chromatographic retention time profiles, while S^T the detection absorption spectra.

To represent the two instrumental measurement modes for each sample, noiseless Gaussian-shaped profiles for four different constituents (1, 2, 3, and 4) were defined for the two instrumental data modes. All pure component profiles were normalized in both modes, so that their area under the profile (the total signal for each pure constituent) was one (Figure 2).

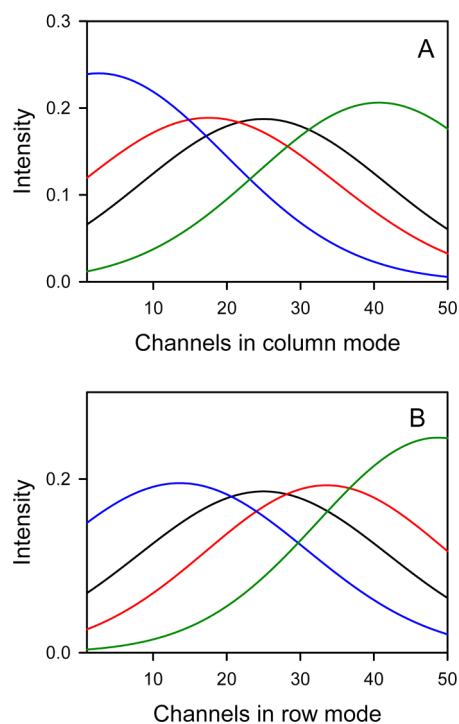


Figure 2. Noiseless component profiles used to build up the different simulated data sets: (A) column component profiles and (B) row component profiles. Black lines identify the analyte of interest (component 1); remaining colors correspond to other sample components: blue, component 2; green, component 3; red, component 4.

Simulated data have 50 values in the column vector space (Figure 2A) and 50 in the row vector space (Figure 2B), so that $J = K = 50$. In all cases, the peak maxima for the Gaussian profiles of constituent 1 (the analyte of interest) were fixed at the center of each of the data ranges, i.e., at position 25 in the column mode and at position 25 in the row mode (Figure 2A,B). The peak maxima for the remaining constituents were placed in both modes at 100 different random positions (except in system B3 in the column mode, see below). Figure 2 shows a particular situation for the four possible constituents.

In all simulated data sets, 10 calibration samples were created, with component concentrations taken at random and uniformly distributed in the range 0–1. A single test sample was produced, having component concentrations taken randomly from the range 0.4–0.6. In the present simulations, therefore, $I = 11$ (the 10 calibration samples plus the test sample).

A first group of simulations corresponded to six data sets fulfilling the trilinear model, where each component was characterized by spectral profiles in each data mode which were invariant from sample to sample. These data sets contain the following constituents: data set B1 with analytes 1 and 2; data

set B2 with analyte 1 and interfering agent 2; data set T1 with analytes 1, 2, and 3; data set T2 with analytes 1 and 2 and interfering agent 3; data set T3 with analyte 1 and interfering agents 2 and 3; and data set Q1 with analytes 1, 2, 3, and 4. These simulated data are typical of excitation–emission fluorescence data because usually the fluorescence excitation and emission spectra are invariant throughout the samples. For MCR-ALS analysis, the augmented mode was the column mode, which represents the excitation spectral profiles, and the non augmented mode was the row mode, which represents the emission spectral profiles.

Two additional nontrilinear data sets were produced. In the binary data set B3, analyte 1 occurs in all samples, while interfering agent 2 appears only in the test samples. The profiles for both components are invariant in all samples, as in the trilinear models, but they are identical for the two components in the column vector space. Specifically, the peak maximum for component 2 was placed at position 25 in the profile of the column mode in all samples, coincident with the analyte, and at random positions in the profile of the row mode. Since the profiles in the column mode are identical, the B3 data set should not be considered strictly a trilinear model but a degeneration of it.²⁰ This simulation tries to mimic some experimental systems, such as those based on time decay–luminescence data matrices to be described below. For MCR-ALS analysis, the augmented mode was the column mode, in which the analyte and interfering agent profiles are identical, and the nonaugmented mode was the row mode, where the analyte and interfering agent profiles are different.

In the data set B4, the peak maximum for component 2 was placed at 100 random positions in both modes, but in the column mode the positions and peak widths varied from sample to sample in each simulation, by random values within the range ± 5 units in peak position and ± 10 units (full peakwidth at half-maximum). In this way, the profiles for the same component are not invariant in the augmented column (concentration) mode, and thus it is not possible to define a unique component profile for it in this mode because they change along this mode, both in bandwidth and in position of the maximum. This is similar to what usually happens in chromatographic separations with spectroscopic detection, where the elution peak retention times and shapes of the same eluted component can change from run to run (in the augmented column mode). For MCR-ALS analysis, the augmented mode was the column mode, in which the component profiles change from sample to sample, and the non augmented mode was the row mode, where the profiles do not change from sample to sample.

All data sets were analyzed by extended MCR-ALS, which was applied with following the steps: (1) the whole data set including the unknown sample data matrix were joined with the calibration data matrices to create a new column-wise augmented data matrix, (2) the column-wise augmented data matrix was decomposed using a bilinear model, initializing the algorithm with profiles in the row vector space estimated from the purest variables in this mode³⁶ and imposing suitable constraints (see below), (3) MCR-ALS resolved scores for component 1 (analyte of interest) in the calibration samples were regressed against its nominal concentrations, and (4) the test analyte score was interpolated in the calibration graph to estimate its concentration. The following constraints were applied during the ALS phase: non-negativity constraints to all resolved profiles for all components; correspondence constraint

between profiles of the same component in the case of data sets B2, T2, and T3 having uncalibrated interfering agents and the trilinearity constraint³⁷ in the case of data sets B1, B2, T1, T2, T3, and Q1.

The sensitivity parameter is estimated in this report as the ratio of the noise introduced in the signal and the uncertainty obtained in the predicted concentrations (see below).^{30–32} For this purpose, instrumental uncertainty was only added to the test sample data matrices while keeping the calibration data precise, as in previous studies.^{30–32} This is usually done when estimating the sensitivity by Monte Carlo noise addition, because in this way the concentration uncertainty depends only on the sensitivity and on the signal uncertainty.^{30–32} If noise were also added to the calibration signals, the concentration uncertainty would also be a function of the sample leverage, a dimensionless parameter which positions the test sample in the calibration space.^{27,30} It would then be difficult to separate the effect propagated by the test sample from those propagated by the calibration samples in order to estimate the sensitivity. In each of the analyzed data sets, the value of the signal uncertainty (i.e., the value of $s_{d_{test}}$ see the Appendix) was 0.01% of the mean of all values of the elements of the corresponding calibration three-way data array.

This calibration/prediction procedure was repeated 1 000 times using different random seeds for the signal noise. The variance in the estimated concentrations of the analyte of interest (the constituent 1 in all cases) was calculated for the test sample in each case after the 1 000 Monte Carlo cycles as

$$\text{var}(c_1) = \sum_{i=1}^{1000} \frac{(c_{i1} - \bar{c}_1)^2}{999} \quad (2)$$

where c_{i1} is the estimated concentration of analyte 1 at each Monte Carlo cycle, and \bar{c}_1 is the mean estimated concentration of this analyte across all Monte Carlo cycles. This leads to Monte Carlo sensitivities SEN_{MC} toward analyte 1 through

$$\text{SEN}_{\text{MC}} = s_{d_{test}}/s_{c1} \quad (3)$$

where $s_{d_{test}}$ is the standard deviation of the Gaussian noise introduced in the test sample signals, and s_{c1} is the standard deviation in the predicted analyte concentrations, i.e., $[\text{var}(c_1)]^{1/2}$ or the square root of the value obtained in eq 2.

In all these calculations, it was assumed that rotation ambiguities associated to the MCR-ALS bilinear decomposition are removed or that they are low because of the applied constraints, especially those related with trilinearity, species correspondence, and local rank regions (see refs 15, 33, 37–39).

EXPERIMENTAL DATA

Lanthanide-sensitized luminescence signals were recorded on an Aminco Bowman series 2 spectrofluorimeter (Urbana, IL) equipped with a 7 W pulsed xenon lamp and a thermostatted (20.0 °C) quartz cell, through the AB2 software operating under Windows. Instrumental parameters were excitation and emission slits, 8 nm; delay time, 300 μs ; gate time, 4000 μs ; minimum flash period, 10 ms; photomultiplier sensitivity (PMT), 650 mV; masked detector. Matrix data were registered as a function of decay time and excitation wavelength, at a fixed emission wavelength of 545 nm. The time decay range was from 300 to 4000 μs each 100 μs (38 data points), and the excitation range was from 230 to 380 nm each 5 nm (31 data points). Each measured matrix was organized and saved as a

data matrix, in which the columns were the time decay profiles at each wavelength and the rows were the excitation spectra at each decay time.

All reagents were of analytical grade. Furosemide (Merck, Darmstadt, Germany) and flufenamic acid (Fluka, Buchs, Switzerland) were employed to prepare concentrated solutions in methanol (Merck, Darmstadt, Germany) and by adequate dilution with water for furosemide and with methanol for flufenamic acid, 10.0 mg L⁻¹ solutions. Terbium(III) (1.50 $\times 10^{-2}$ mol L⁻¹) was prepared by dissolving terbium chloride hexahydrate (Fluka, Buchs, Switzerland) in distilled water. Trioctylphosphine oxide (TOPO) (2.00 $\times 10^{-2}$ mol L⁻¹) was made by dissolving the reagent (Aldrich, Gillingham, Dorset, U.K.) in ethanol (Merck, Darmstadt, Germany). Triton X-100 (0.9% w/v) was prepared by direct dilution of the surfactant (J.T. Baker) in water. A tris-(hydroxymethyl)-aminomethane (Tris) buffer solution (0.05 M, pH = 6) was also prepared by dissolving solid Tris (Merck, Darmstadt, Germany) in water, adding diluted HCl to reach the desired pH value.

A total of 14 duplicate calibration samples were prepared containing the analyte furosemide at 7 different concentration levels, i.e., 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6 mg L⁻¹. Two duplicate blank samples were added to the analysis. In total, 20 test samples were also prepared in duplicate, with analyte concentrations within the calibration range, but different than those employed for calibration, and flufenamic acid in the range 0.05–0.65 mg L⁻¹. All samples contained Tb(III) 7.50 $\times 10^{-4}$ M, TOPO 1.60 $\times 10^{-4}$ M, triton X-100 0.060% w/v, and Tris buffer 0.05 M (final pH value = 6).

Lanthanide-sensitized time decay-excitation matrices were measured in random order for all samples. In these data matrices, the time decay profiles for both constituents are almost identical but the excitation profiles differ. Hence, selectivity is only present in the spectral mode. Therefore, this system can be analyzed with a few second-order multivariate algorithms, the most usual being MCR-ALS. An additional strategy has been recently described, based on a modified version of U-PLS/RBL (which in fact incorporates MCR-ALS within the RBL procedure).²¹

RESULTS AND DISCUSSION

Simulated Data. Simulated trilinear data sets B1, T1, Q1, B2, T2, and T3 were analyzed using MCR-ALS, as described in the corresponding section. In the first three cases, 2, 3, and 4 components occur both in calibration and test samples, respectively. Constraints applied during the ALS phase were non-negativity of profiles in both modes and trilinearity. After finishing the calibration/prediction process, the concentration of analyte 1 in the simulated test sample of 1 000 different Monte Carlo replicates, and considering 100 different overlapping situations among the profiles in both data modes, was estimated and the Monte Carlo sensitivity values, SEN_{MC} , calculated according to eq 3. For each of these cases, analyte sensitivities for component 1 were also estimated using eq 1. Both sets of results are compared in Figure 3A–C. As it can be seen, good agreement is found for these three investigated cases, i.e., B1, T1, and Q1. It is interesting to note that the sensitivity decreases when the number of overlapping components increases, as expected. This is already observed in Figure 3A–C but also numerically substantiated, since the average of Monte Carlo sensitivity values were 0.39, 0.16, and 0.02 for data sets B1, T1, and Q1, with 2, 3, and 4 sample components, respectively. It should be noticed that only eq 1 is

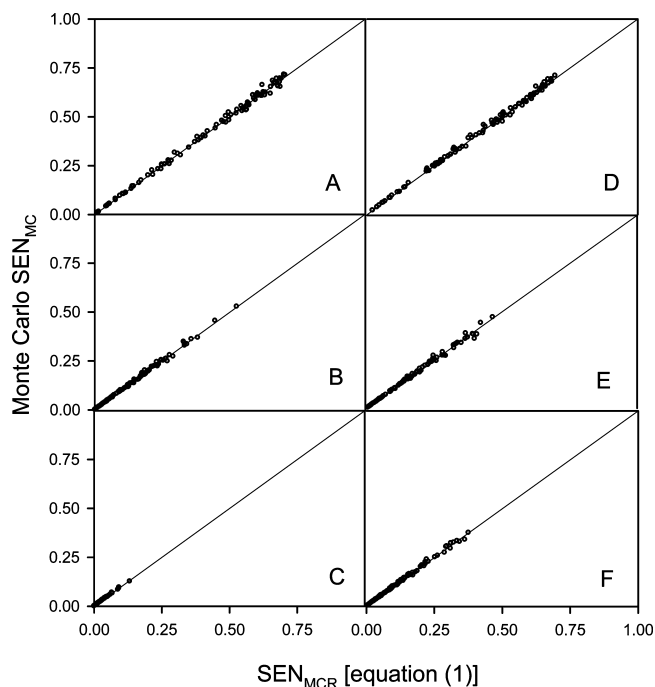


Figure 3. Comparison of sensitivities for analyte 1 using Monte Carlo MCR-ALS results (eq 3) and sensitivities obtained using eq 1 in the analysis of the trilinear data sets. Results for the (A) two-component system B1, (B) three-component system T1, (C) four-component system Q1, (D) two-component system (B2), (E) three-component system T2, and (F) three-component system T3 (see the section Data Simulations for the explanation of these systems).

able to successfully reproduce the values of sensitivity estimated by the present Monte Carlo simulations. If the slope of the MCR-ALS calibration graph m_n were considered as a measure of sensitivity, in the present simulations all sensitivities would be equal and independent of the profile overlapping in the

nonaugmented mode, in contrast to the values furnished by eq 1.

For B2, T2, and T3 data sets, the test samples contained interfering agent components which were not included in the calibration set of samples, meaning that the second-order advantage is needed for successful analyte prediction. In these cases, not only the constraints mentioned above were applied during MCR-ALS decomposition but also the one referred to species correspondence. This is because some components are known to be absent in the calibration samples in these systems, and this (local) rank information is very useful for a successful decomposition of the augmented data matrix without ambiguities.^{38,39} The results, presented in Figure 3D–F, do also show agreement between Monte Carlo simulations and the new eq 1. As with the previous investigated systems B1, T1, and Q1, the sensitivity decreases on increasing the total number of components. For binary system B2, the average value is 0.38, whereas for the ternary systems T2 and T3, they are both ~ 0.16 .

Concerning nontrilinear data sets B3 and B4, it should be noticed that MCR-ALS can be successfully applied also to model them as it has been shown before.^{15,38,39} In rank deficient system B3, two components occur, with identical column profiles and different (but partially overlapped) row profiles. The analyte (component 1) is present in both the calibration and test samples, but the interfering agent is only present in the test samples. This creates severe complications for some second-order multivariate algorithms, such as PARAFAC and U-PLS/RBL.²¹ In data set B3, for example, the PARAFAC sensitivity toward analyte 1 is known to be given by the well-known HCD expression:

$$SEN_{HCD} = m_n [(\mathbf{S}^T \mathbf{S})_{nn}^{-1} (\mathbf{C}_{test}^T \mathbf{C}_{test})_{mm}^{-1}]^{-1/2} \quad (4)$$

where \mathbf{S} and \mathbf{C}_{test} are the profile matrices retrieved by PARAFAC in both data modes and m_n the slope of the pseudounivariate calibration graph (recall that in the trilinear

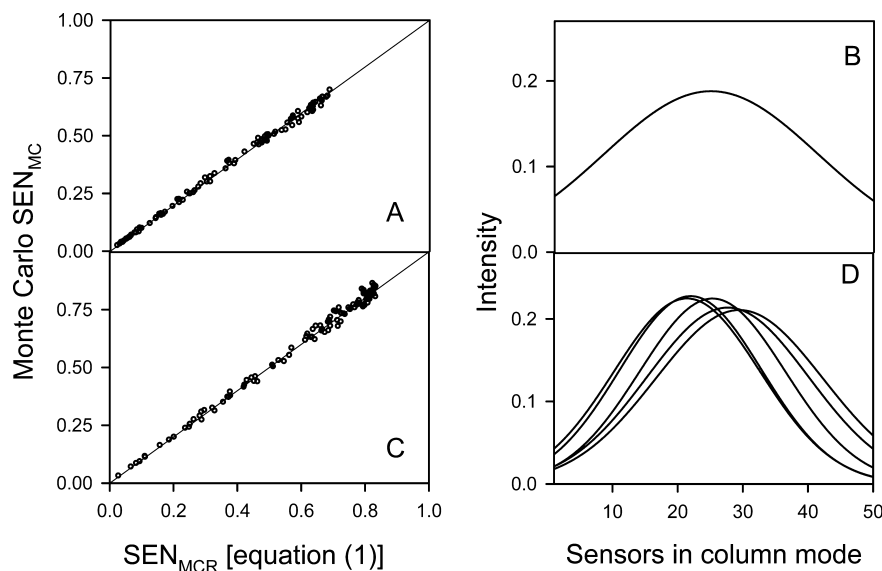


Figure 4. Comparison of sensitivities obtained using Monte Carlo MCR-ALS results and of sensitivities obtained using eqs 1 and 3 in the analysis of the nontrilinear data sets. (A) Comparison of sensitivities of analyte 1 in system B3. (B) MCR-ALS resolved column profile for component 2 in system B3 for the test sample, which is identical to the theoretical one for analyte 1 (see Figure 1). (C) Comparison of sensitivities of analyte 1 in binary system B4. (D) MCR-ALS resolved column profiles for analyte 1 in the five different samples of system B4, illustrating their changes in band shapes and peak positions from sample to sample (see the section Data Simulations for the explanation of these systems).

PARAFAC model both the S^T and C_{test} profiles are identical for all samples). From eq 4, it is clear that rank deficiency in the column vector space leads to $\text{SEN}_{\text{HCD}} = 0$, because the matrix ($C_{\text{test}}^T C_{\text{test}}$) is singular and cannot be inverted. This means that these types of data sets cannot be properly analyzed by regular PARAFAC.^{18,19,40} An analogous result is obtained for the RBL procedure, which cannot distinguish between the analyte and the interfering agent signal, because the former is based on principal component analysis of a rank one residual data matrix having two contributions (analyte and interfering agent) with an equal profile in one of the modes.²¹

Using MCR-ALS, successful analyte prediction is possible in these cases because it takes advantage of the rank augmentation obtained by the matrix augmentation procedure. Although the selectivity in the column vector space for a single data matrix is zero, a nonzero selectivity occurs in the augmented matrix, which is built by augmentation along the column vector space. The results from the present Monte Carlo simulations are collected in Figure 4A, where the agreement between results from simulations and from eq 1 is confirmed. Figure 4B shows the MCR-ALS resolved profile for component 2 in the column vector space in the test sample, which is identical to the analyte profile shown in Figure 2B.

Finally, the simulated nontrilinear binary data set B4 poses another challenge to second-order calibration algorithms, i.e., profiles in one data mode (the column vector space) varying for different data matrices (from sample to sample), both in peak shape and in peak position of its maximum. This is frequently encountered in the chemical type of data (chromatography, reaction based systems, etc.), and it is easily accounted for by the bilinear matrix augmentation models like those used by MCR-ALS, because the augmented concentration matrix retrieved by this algorithm allows component profiles in successive data submatrices to be different and specific for each data submatrix corresponding to each sample. To compute the sensitivity in this case, however, only the S profiles (row profiles) are required according to eq 1. Figure 4C confirms the adequacy of the presently proposed approach in reproducing Monte Carlo sensitivities for system B4. In Figure 4D, in turn, the different column profiles for component 1 in the five different data matrices (samples) were correctly resolved by MCR-ALS.

In summary, Monte Carlo simulations give support to the adequacy of eq 1 for the estimation of analyte sensitivities when MCR-ALS is employed for data processing in scenarios of different complexity. For a variety of systems with different degrees of overlapping, and also different number of calibrated analytes and uncalibrated interfering agents, good agreement has been found between the Monte Carlo estimated sensitivity and the values provided by eq 1. This confirms that the slope of the pseudounivariate MCR-ALS calibration graph (m_n) is not the correct measure of the sensitivity.

Experimental Data. In the experimental data set, a single analyte occurs (the diuretic furosemide) while all test samples contain an interfering agent (the anti-inflammatory flufenamic acid). The data structure is close to that of the simulated system B3, i.e., the profiles of both components in the time decay domain (the column mode of each data matrix) are almost identical, while one of the components only exists in the test samples. Selectivity is present in the excitation spectral mode, which corresponds to the row vector space of each data matrix. A typical calibration set employed in quantitative analytical studies was prepared, with all calibration samples as duplicates

and spanning the range of analyte concentrations with seven different concentration levels. Two duplicate blank samples were also added to the analysis. Test samples were prepared in duplicate, with analyte concentrations at random values within the calibration range.

Test samples were processed individually, by joining each of them in turn with the calibration samples, creating the new column-wise augmented data matrix, and analyzing it by MCR-ALS. The direction of matrix augmentation was the unselective time decay mode (the column space), and the more selective excitation spectral mode (the row space) was the non-augmented mode in all data matrices. In this way, the chemical rank of the augmented data matrix was two, because matrix augmentation with different data matrices provided a full chemical rank (mathematical rank in absence of noise) of two and selectivity in the augmented time decay mode. If augmentation had been made in the row mode, i.e., in the excitation spectral mode, the augmented data matrix would have been rank deficient with a chemical rank of one, because the time decay profile for both sample components are practically identical.

During the ALS optimization, non-negativity constraints were applied to both spectral and concentration profiles, taking into account the correspondence between the different components of the samples. For instance, the interfering agent (flufenamic acid) was only present in the test samples; therefore, it was absent in the calibration samples. This decreases considerably the rotation ambiguity for the analyte and provides more efficient bilinear decompositions of all augmented data matrices. Figure 5A,B shows the MCR-ALS results after the processing of a typical test sample. The resolved excitation spectra of the analyte (in blue) and of the interfering agent (in red) in the column vector space are shown in Figure 5A. On the other hand, Figure 5B shows the successive time profiles in the augmented time decay data mode for the test sample, a blank sample, and seven calibration samples (as indicated). Important issues to be noticed in this latter figure are (1) time decay profiles for analyte and interfering agent (the latter one is only present in the test sample) were almost identical and they required appropriate matrix augmentation in the time decay direction to avoid rank deficiency, (2) absence of the interfering agent in the calibration samples (red line), which allowed to use the component correspondence constraint during the ALS optimization, and (3) successive time decay profiles for the analyte in calibration samples (blue line) showed a proper linear relationship with increasing nominal analyte concentration in calibration samples.

Once decomposition was accomplished, a pseudounivariate calibration graph was plotted with analyte scores as a function of nominal analyte concentrations, as shown in Figure 6. Notice the good linear correlation and the low dispersion of values for duplicate samples. From the calibration graph in Figure 6, some figures of merit were estimated, in the same way as proposed in ref 34. For example, the limits of detection and quantitation were assessed using modern IUPAC's definition,⁴¹ as 0.04 and 0.11 mg L⁻¹, respectively. However, sensitivity cannot be obtained from Figure 6, because the vertical scale is based on scores in arbitrary units, which do not represent analytes signals and do not take into account the effect of the profile overlapping between analyte and other sample components and interfering agents. In fact, the slope of the calibration graph in Figure 6 is numerically equal to m_n , but this is only one of

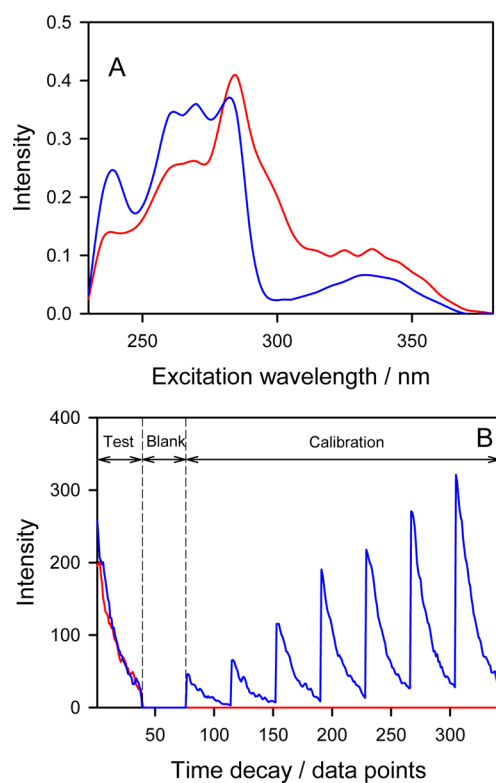


Figure 5. Excitation (A) and time decay (B) profiles resolved by MCR-ALS in the simultaneous analysis of the experimental system constituted by one typical test sample, a blank sample, and seven calibration samples (as shown in part B). Blue lines, analyte (furosemide); red lines, interfering agent (flufenamic acid).

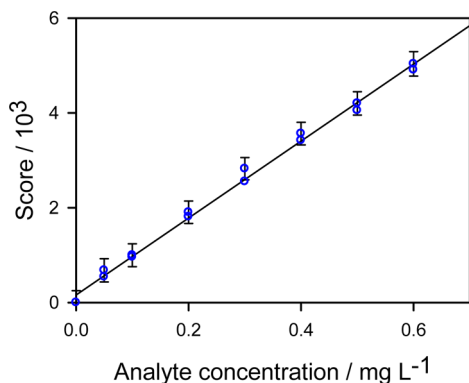


Figure 6. Pseudounivariate calibration plot of analyte scores vs nominal concentrations for calibration samples in the analysis of the experimental system involving time-decaying excitation luminescence data matrices for the determination of furosemide in the presence of flufenamic acid. Error bars correspond to $t_{\alpha,\nu} \times s$, where t is the Student coefficient at a confidence level of $100 \times (1 - \alpha)\%$ ($\alpha = 0.05$) with ν degrees of freedom ($\nu = 16 - 2$), and s is the average standard error at each calibration point. The best least-squares regression line is also given as a solid line.

the parameters required in eq 1 to give a more reliable estimation of this figure of merit.

Interpolation of the analyte score for the test sample in this calibration graph leads to analyte prediction for all test samples, as collected in Table 1. As it can be seen, reasonable recovery results were obtained, with a mean prediction error of 0.04 mg L^{-1} and a good correlation coefficient between nominal and

Table 1. Analytical Results for the Experimental Example Using MCR-ALS

sample ^a	furosemide/ mg L^{-1}		flufenamic acid/ mg L^{-1}
	nominal	predicted ^b	
1	0.00	0.05(1)	0.15
2	0.00	0.06(1)	0.30
3	0.08	0.14(1)	0.15
4	0.10	0.14(1)	0.10
5	0.10	0.12(1)	0.15
6	0.10	0.13(1)	0.20
7	0.15	0.22(1)	0.05
8	0.15	0.19(1)	0.15
9	0.24	0.26(1)	0.10
10	0.24	0.26(1)	0.20
11	0.30	0.37(1)	0.14
12	0.30	0.37(2)	0.65
13	0.34	0.34(1)	0.20
14	0.35	0.38(4)	0.35
15	0.45	0.44(2)	0.20
16	0.45	0.43(1)	0.60
17	0.50	0.54(1)	0.25
18	0.50	0.52(1)	0.60
19	0.55	0.52(1)	0.15
20	0.55	0.51(1)	0.35
RMSE		0.042	
REP %		14.0	
R^2		0.9883	

^aRMSE = root-mean-square error, REP % = relative error of prediction based on mean calibration concentration, R^2 = correlation coefficient. ^bExperimental standard deviation in parentheses.

predicted values for the 20-sample test set. Duplicate analysis for all test samples did also provide an experimental estimation of the uncertainty in the predicted analyte concentration, which was $\sim 0.01 \text{ mg L}^{-1}$. These values are also quoted in Table 1 for each concentration level in the test samples.

Equation 1 implies that the sensitivity depends on the number of repeated measurements in a given data matrix in the time-decay mode where augmentation is performed (the value of J in eq 1), and on the degree of spectral overlapping in the excitation spectral mode (the row vector space which is the nonaugmented mode), measured by $[(S^T S)_{nn}^{-1}]^{-1/2}$ (n identifies the component number corresponding to the analyte). Both terms tend to decrease the value of the final sensitivity from the initial theoretical value in absence of interfering agents evaluated from the slope value m_n . Inserting the appropriate parameters in eq 1, the sensitivity is computed as $600 \text{ luminescence intensity units L mg}^{-1}$.

Using the latter value of SEN_{MCR} , crude estimations of the limit of detection (LOD), limit of quantitation (LOQ), and concentration uncertainty can be made (these estimations ignore the effect of the calibration uncertainty) as²³

$$\text{LOD} = 3.3s_{d\text{test}}/\text{SEN}_{\text{MCR}} = 0.03 \text{ mg L}^{-1} \quad (5)$$

$$\text{LOQ} = 10s_{d\text{test}}/\text{SEN}_{\text{MCR}} = 0.09 \text{ mg L}^{-1} \quad (6)$$

$$s_c = S_{d\text{test}}/\text{SEN}_{\text{MCR}} = 0.01 \text{ mg L}^{-1} \quad (7)$$

where the parameter 3.3 follows from 5% probability assigned to both Type I and Type II errors.⁴¹ The value of $s_{d\text{test}}$ was estimated as 5 arbitrary luminescence units, from statistical analysis of samples replication. The resulting LOD, LOQ, and s_c

are reasonable and consistent with those obtained from the pseudounivariate calibration, indicating that the value of SEN_{MCR} provided by eq 1 is an appropriate approximation to the analyte sensitivity in this system.

Finally, the selectivity can be assessed using eq 23 (see the Appendix), which renders a value of 0.41 (the maximum selectivity value would be 1.0 or 100%). Notice that this parameter cannot be computed from the pseudounivariate calibration graph, which assumes the analyte signal has been totally isolated from the interfering agents. In any case, a selectivity of ~40% is consistent with the large degree of spectral overlapping occurring in the spectral mode (see Figure 5A).

CONCLUSIONS

The presently proposed approach provides a useful and reliable way of estimating the sensitivity in MCR-ALS, when quantitative estimations are made in cases where rotation ambiguities have been practically suppressed by a proper application of constraints like trilinearity, species correspondence, and local rank constraints. The estimated sensitivity allows the calculation of other figures of merit such as detection capabilities, from the knowledge of the sample component properties and instrumental noise, as an alternative to other empirical approaches.

APPENDIX

Calculation of the Sensitivity

In extended MCR-ALS, the calibration scores are employed to build a pseudo-univariate calibration line, leading to an estimation of the corresponding slope (m_n) and offset (n_n) (see the Supporting Information). The analyte score $a_{test,n}$ in the test sample is then interpolated in the calibration line to yield the predicted analyte concentration c_n :

$$c_n = (a_{test,n} - n_n)/m_n \quad (8)$$

The sensitivity of the analyte determination in this case cannot be defined simply as the slope of the pseudo-univariate calibration graph, because the vertical scale of the latter is arbitrary and the scores are neither true analyte signals nor net analyte signals (because they do not take into account the overlapping among sample components). We therefore employ the more general definition of sensitivity, which is given by the ratio of predicted concentration uncertainty to signal uncertainty, as shown in eq 3. This definition is fully consistent with those successfully employed in first-order multivariate calibration,²⁷ and in other second- and higher-order multivariate algorithms.^{30–32}

The estimation of the uncertainty in the predicted concentration assumes that the calibration is precise. This is usual in sensitivity studies from error propagation theory, where only the test sample is considered to carry instrumental noise, unlike the data for the calibration samples. This is done in order to leave the concentration uncertainty as only depending on the sensitivity and on the signal to noise ratio.³⁰ Thus in eq 8, the uncertainty in the predicted concentration c_n comes mainly from the test sample score $a_{test,n}$:

$$\text{var}(c_n) = (m_n)^{-2} \text{var}(a_{test,n}) \quad (9)$$

where $\text{var}(\)$ indicates variance. The score $a_{test,n}$ is equal to the sum of the elements of the analyte profile in the test sample in the augmented mode (see the Supporting Information). Hence,

the variance in $a_{test,n}$ will depend on the variance of these elements. The latter are estimated by least-squares during MCR-ALS data processing, and thus they may be mutually correlated. Therefore, to estimate $\text{var}(a_{test,n})$, not only the variances of each summed element should be taken into account but also the covariances between all individual row values of the resolved profile for component n . From eq 9:

$$\text{var}(c_n) = (m_n)^{-2} \sum_{j=1}^J \sum_{j'=1}^J V_n(j, j') \quad (10)$$

where $V_n(j, j')$ are elements of the $J \times J$ variance-covariance V_n matrix, which is a sub-matrix of the full V matrix for all values of the elements of all the resolved component profiles. The latter matrix can be estimated from the Jacobian matrix as follows. First assume a test data matrix D_{test} for a test sample composed of the analyte of interest and an additional component, defined as a function of the component profiles as

$$D_{test} = C_{test} S^T = c_1 s_1^T + c_2 s_2^T \quad (11)$$

where $C_{test} = [c_1 \ c_2]$ and

$$S^T = \begin{bmatrix} s_1^T \\ s_2^T \end{bmatrix}$$

Unfolding D_{test} into a vector leads to

$$d_{test} = \text{vec}(D_{test}) = s_1 \otimes c_1 + s_2 \otimes c_2 \quad (12)$$

where the \otimes symbol is the Kronecker product (see the Supporting Information for a detailed definition of the latter operation).

It can be shown that the required Jacobian J matrix for the elements of c_1 and c_2 in eq 12 contains a first block of J columns with the derivatives dd_{test}/dc_1 , followed by a second block of J columns with the derivatives dd_{test}/dc_2 :

$$J = [s_1 \otimes I_c \quad s_2 \otimes I_c] \quad (13)$$

where I_c is a $J \times J$ identity matrix. From eq 13, the $(2J \times 2J)$ variance-covariance matrix for all estimated parameters is

$$V = \text{var}(d_{test})(J^T J)^{-1} \quad (14)$$

where $\text{var}(d_{test})$ is the variance in the measured data (provided the noise is identically and independently distributed). A convenient way of obtaining the $J \times J$ sub-matrix V_1 of V , corresponding to analyte 1, involves three steps. First a projection matrix is defined, orthogonal to the second block of J , i.e., $(s_2 \otimes I_c)$ in eq 13. This matrix is orthogonal to the space spanned by component 2 in the column mode:³²

$$P_{s_2} = I - (s_2 \otimes I_c)(s_2 \otimes I_c)^+ \quad (15)$$

Second, the first block of J is projected onto P_{s_2} to yield J_1^+ :³²

$$J_1^+ = P_{s_2}(s_1 \otimes I_c) \quad (16)$$

Finally, the required V_1 matrix is given by³²

$$V_1 = \text{var}(d_{test})(J_1^+)^T J_1^+ \quad (17)$$

Inserting this result in eq 10 leads to

$$\begin{aligned} \text{var}(c_1) &= (m_1)^{-2} \sum_{j=1}^J \sum_{j'=1}^J V(j, j') \\ &= (m_1)^{-2} \text{var}(d_{\text{test}}) J \| \mathbf{P}_{s_2} \mathbf{s}_1 \|^2 \end{aligned} \quad (18)$$

where $\| \cdot \|$ indicates the Euclidean norm. Equation 18 can be conveniently re-written by noting that the product $(\mathbf{P}_{s_2} \mathbf{s}_1)$ is the projection of \mathbf{s}_1 orthogonal to \mathbf{s}_2 , i.e., the first row of the generalized inverse matrix \mathbf{S}^+ , leading to

$$\begin{aligned} \text{var}(c_1) &= (m_1)^{-2} \text{var}(d_{\text{test}}) J \| \text{first row of } \mathbf{S}^+ \|^2 \\ &= (m_1)^{-2} \text{var}(d_{\text{test}}) J (\mathbf{S}^T \mathbf{S})_{11}^{-1} \end{aligned} \quad (19)$$

where the shorthand notation " \mathbf{A}_{nn}^{-1} " indicates the (n, n) element of the inverse of the square matrix \mathbf{A} (in this case $n = 1$). From eq 19, the uncertainty in predicted concentration can be estimated simply by taking the square root of both sides:

$$s_{c1} = (s_{d\text{test}}/m_1) [J(\mathbf{S}^T \mathbf{S})_{11}^{-1}]^{1/2} \quad (20)$$

The sensitivity towards analyte 1 is the ratio of signal uncertainty to concentration uncertainty, and thus

$$\text{SEN}_{\text{MCR}} = (s_{d\text{test}}/s_{c1}) = m_1 [J(\mathbf{S}^T \mathbf{S})_{11}^{-1}]^{-1/2} \quad (21)$$

An interesting conclusion to be drawn from this result is that the sensitivity is lower than the slope of the pseudo-univariate calibration graph (m_1) and decreases with increasing overlap between the profiles of the sample components (profiles in the nonaugmented mode vector space or \mathbf{S}^T spectral vector space in this case), as measured by the overlapping factor $[(\mathbf{S}^T \mathbf{S})_{11}^{-1}]^{-1/2}$. Equation 21 can be easily generalized to any number of sample components, leading to

$$\text{SEN}_{\text{MCR}} = m_n [J(\mathbf{S}^T \mathbf{S})_{nn}^{-1}]^{-1/2} \quad (22)$$

where n indicates the analyte of interest, m_n is the slope of the pseudo-univariate calibration graph for analyte n , \mathbf{S}^T contains the profiles for all sample components in the non-augmented direction, and J is the number of channels in the sample data matrix in the augmented direction.

Finally, from eq 22 it is easy to define the selectivity (SEL) as the ratio between actual sensitivity and the value in the absence of interfering agents:

$$\text{SEL} = [(S^T \mathbf{S})_{nn}^{-1}]^{1/2} \quad (23)$$

■ ASSOCIATED CONTENT

Supporting Information

Additional information as noted in text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: olivieri@iquir-conicet.gov.ar.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Universidad Nacional de Rosario, CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas, Project No. PIP 1950), ANPCyT (Agencia Nacional de Promoción Científica y Tecnológica, Project No. PICT-2010-0084), and the Spanish Ministry of Science and Innovation (Project No.

CTQ2009-11572) are gratefully acknowledged for financial support.

■ REFERENCES

- (1) Booksh, K. S.; Kowalski, B. R. *Anal. Chem.* **1994**, *66*, 782A–791A.
- (2) Escandar, G. M.; Faber, N. M.; Goicoechea, H. C.; Muñoz de la Peña, A.; Olivieri, A. C.; Poppi, R. J. *Trends Anal. Chem.* **2007**, *26*, 752–765.
- (3) Escandar, G. M.; Damiani, P. C.; Goicoechea, H. C.; Olivieri, A. C. *Microchem. J.* **2006**, *82*, 29–42.
- (4) Olivieri, A. C. *Anal. Chem.* **2008**, *80*, 5713–5720.
- (5) Olivieri, A. C.; Escandar, G. M.; Muñoz de la Peña, A. *Trends Anal. Chem.* **2011**, *30*, 607–617.
- (6) Bro, R. *Crit. Rev. Anal. Chem.* **2006**, *36*, 279–293.
- (7) Gómez, V.; Callao, M. P. *Anal. Chim. Acta* **2008**, *627*, 169–183.
- (8) Arancibia, J. A.; Damiani, P. C.; Escandar, G. M.; Ibañez, G. A.; Olivieri, A. C. *J. Chromatogr., B* **2012**, DOI: 10.1016/j.jchromb.2012.02.004.
- (9) Amigo, J. M.; Skov, T.; Bro, R. *Chem. Rev.* **2010**, *110*, 4582–4605.
- (10) Bro, R. *Chemom. Intell. Lab. Syst.* **1997**, *38*, 149–171.
- (11) Chen, Z. P.; Wu, H. L.; Jiang, J. H.; Li, Y.; Yu, R. Q. *Chemom. Intell. Lab. Syst.* **2000**, *52*, 75–86.
- (12) Xia, A. L.; Wu, H. L.; Fang, D. M.; Ding, Y. J.; Hu, L. Q.; Yu, R. Q. *J. Chemom.* **2005**, *19*, 65–76.
- (13) Öhman, J.; Geladi, P.; Wold, S. *J. Chemom.* **1990**, *4*, 79–90.
- (14) Olivieri, A. C. *J. Chemom.* **2005**, *19*, 253–265.
- (15) Tauler, R. *Chemom. Intell. Lab. Syst.* **1995**, *30*, 133–146.
- (16) Kiers, H. A. L.; Ten Berge, J. M. F.; Bro, R. *J. Chemom.* **1999**, *13*, 275–294.
- (17) Bro, R.; Anderson, C. A.; Kiers, H. A. L. *J. Chemom.* **1999**, *13*, 295–309.
- (18) Izquierdo-Ridora, A.; Saurina, J.; Hernández-Cassou, S.; Tauler, R. *Chemom. Intell. Lab. Syst.* **1997**, *38*, 183–196.
- (19) Saurina, J.; Hernández-Cassou, S.; Tauler, R.; Izquierdo-Ridora, A. *J. Chemom.* **1998**, *12*, 183–203.
- (20) Smilde, A.; Bro, R.; Geladi, P. *Multivariate Analysis with Applications in the Chemical Sciences*; John Wiley & Sons: West Sussex, England, 2004.
- (21) Olivieri, A. C.; Ibañez, G. A.; Lozano, V. A. *Anal. Chem.* **2010**, *82*, 4510–4519.
- (22) Danzer, K.; Currie, L. A. *Pure Appl. Chem.* **1998**, *70*, 993–1014.
- (23) Olivieri, A. C.; Faber, N. M.; Ferré, J.; Boqué, R.; Kalivas, J. H.; Mark, H. *Pure Appl. Chem.* **2006**, *78*, 633–661.
- (24) Olivieri, A. C.; Faber, N. M. Validation and error. In *Comprehensive Chemometrics*; Brown, S., Tauler, R., Walczak, B., Eds.; Elsevier: Amsterdam, The Netherlands, 2009; Vol. 3, pp 91–120.
- (25) Lorber, A. *Anal. Chem.* **1986**, *58*, 1167–1172.
- (26) Saltelli, A.; Ratto, M.; Tarantola, S.; Campolongo, F. *Chem. Rev.* **2005**, *105*, 2811–2828.
- (27) Olivieri, A. C. *J. Chemom.* **2002**, *16*, 207–217.
- (28) Messick, N. J.; Kalivas, J. H.; Lang, P. M. *Anal. Chem.* **1996**, *68*, 1572–1579.
- (29) Ho, C. N.; Christian, G. D.; Davidson, E. R. *Anal. Chem.* **1980**, *52*, 1071–1079.
- (30) Olivieri, A. C.; Faber, N. M. *J. Chemom.* **2005**, *19*, 583–592.
- (31) Olivieri, A. C. *Anal. Chem.* **2005**, *77*, 4936–4946.
- (32) Olivieri, A. C.; Faber, N. M. *Anal. Chem.* **2012**, *84*, 186–193.
- (33) Jaumot, J.; Gargallo, R.; Tauler, R. *J. Chemom.* **2004**, *18*, 327–340.
- (34) Saurina, J.; Leal, C.; Compañó, R.; Granados, M.; Dolors Prat, M.; Tauler, R. *Anal. Chim. Acta* **2001**, *432*, 241–251.
- (35) *MATLAB 7.10*; The MathWorks Inc.: Natick, MA, 2010.
- (36) Windig, W.; Guilment, J. *Anal. Chem.* **1991**, *63*, 1425–1432.
- (37) Tauler, R.; Marqués, I.; Casassas, E. *J. Chemom.* **1998**, *12*, 55–75.
- (38) Tauler, R.; Smilde, A.; Kowalski, B. R. *J. Chemom.* **1995**, *9*, 31–58.

- (39) Tauler, R.; Maeder, M.; de Juan, A. Multiset Data Analysis: Extended Multivariate Curve Resolution. In *Comprehensive Chemometrics*; Brown, S., Tauler, R., Walczak, B., Eds.; Elsevier: Amsterdam, The Netherlands, 2009; Vol. 2, pp 473–505.
- (40) Smilde, A.; Tauler, R.; Saurina, J.; Bro, R. *Anal. Chim. Acta* **1999**, *398*, 237–251.
- (41) Currie, L. A. *Pure Appl. Chem.* **1995**, *67*, 1699–1723.