On-line generation of third-order liquid chromatography–excitation-emission fluorescence matrix data. Quantitation of heavy-polycyclic aromatic hydrocarbons

Maira D. Carabajal, Juan A. Arancibia,* and Graciela M. Escandar*

Instituto de Química Rosario (CONICET-UNR), Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531 (2000), Rosario, Argentina

Corresponding Authors

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

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

Abstract

For the first time, third-order liquid chromatography with excitation-emission fluorescence matrix detection (LC-EEFM) data were generated on-line and chemometrically processed for the simultaneous quantitation of the heavy-polycyclic aromatic hydrocarbons fluoranthene, pyrene, benz[*a*]anthracene, chrysene, benzo[*b*]fluoranthene, benzo[*k*]fluoranthene, benzo[*a*]pyrene, and dibenz[*a*,*h*]anthracene. The applied experimental strategy is very simple, and is based on the reduction of the linear flow rate by fitting a larger diameter connecting-tube between the column outlet and the fluorimetric detector. In this way, EEFMs were successfully recorded on-line, without involving a large total analysis time. Because in the studied system quadrilinearity was fulfilled, four-way parallel factor (PARAFAC) analysis was applied for data processing. The second-order advantage, which is an intrinsic property of data of at least second-order, allowed the quantification of the analytes in interfering media. Moreover, resolution of the system with a high degree of collinearity was achieved thanks to the third-order advantage. In addition to a selectivity improvement, third-order/four-way calibration increased the sensitivity, with limits of detection in the range of 0.4–2.9 ng mL^{-1}. After a solid-phase extraction procedure with C18 membranes, considerably lower concentrations (between 0.033–2.70 ng mL^{-1}) were determined in real waters, with most recoveries in the range 90–106%.

Keywords: Liquid chromatography-excitation-emission fluorescence; Third-order/four-way calibration; Third-order advantage; Heavy-polycyclic aromatic hydrocarbons

1. Introduction

Third-order/four-way multivariate calibration is a very useful technique that is being increasingly applied for analytical purposes [1]. This fact can be justified considering that, in addition to the second-order advantage (quantification of analytes in the presence of uncalibrated sample constituents) [2], third-order/four way calibration allows the development of more sensitive and selective methods. Sensitivity is improved since the measurement of redundant data decreases the relative impact of the noise in the signal, while selectivity is also increased because each new instrumental mode contributes positively to the overall selectivity [3,4]. However, what clearly distinguishes third-order/four way calibration is its ability to deal with strong collinearity problems which cannot be solved by second-order calibration. This property is called the "third-order advantage" [5,6].

Although third-order data can be generated in different ways, the most commonly applied procedures include two-dimensional gas and liquid chromatography (GC-GC and LC-LC) with spectral detection, and excitation-emission fluorescence matrices (EEFMs) coupled to a kinetic reaction or as detecting system to unidimensional chromatography [7]. In the latter case, the approach consists in recording EEFMs as a function of the elution time; the main issue is to find adequate experimental and instrumental conditions that allow obtaining an adequate number of EEFMs containing satisfactory information in a short time. Very recently, Montemurro et al. described analytical methodologies for the generation of third-order LC-EEFM data [8]. The authors discussed the following suitable strategies: (1) injecting the sample several times and recording the emission wavelength–elution time matrix, each time at a different excitation wavelength [9,10], and (2) collecting elution fractions at the end of the chromatographic procedure every few seconds, and then measuring EEFMs for each collected aliquot [11,12].

A third option involves the direct measurement of EEFMs by using a traditional chromatograph-spectrofluorimeter hyphenated system. Although this is the most attractive one, since neither flow interruption nor fraction collection are required, it was deficient in the mode in which it was implemented. This is due to the presence of strong dependence between the phenomena corresponding to the excitation and time modes [8].

In the present work, a procedure which allows the on-line recording of third-order LC-EEFM data in a simpler experimental way is proposed. A connector tube between the column and the detector of a larger diameter allowed to slow the linear flow rate (LFR) and to measure substantially more EEFMs per chromatographic peak. The obtained LC-EEFM data could be arranged as a four-way array complying with the quadrilinearity condition and, therefore, four-way parallel factor analysis (four-way PARAFAC) was applied for data treatment [13,14].

The target analytes, namely fluoranthene (FL), pyrene (PYR), benz[*a*]anthracene (BaA), chrysene (CHR), benzo[*b*]fluoranthene (BbF), benzo[*k*]fluoranthene (BkF), benzo[*a*]pyrene (BaP), and dibenz[*a*,*h*]anthracene (DBA), were selected considering that heavy-PAHs human exposure is associated with serious diseases like cancer, and their quantification in the environment is of prime importance [15–17]. The selected system represents an actual chemometric challenge and demonstrates the ability of third-order/four way calibration to solve high collinearity issues in the spectral and chromatographic profiles.

2. Theory

2.1. Four-way PARAFAC

Four-way data are created by joining the third-order data arrays for the calibration samples and for each of the analyzed validation or test samples. Application of the PARAFAC model to the latter four-way data arrays requires fitting the following expression:

$$F_{ijkl} = \sum_{n=1}^{N} a_{in} b_{jn} c_{kn} d_{ln} + E_{ijkl} \quad (1)$$

where F_{ijkl} is an element of the four-way array of elution time-excitation emission fluorescence matrix signals, N is the total number of responsive components, a_{in} is the relative concentration of component n in sample i; b_{jn}, c_{kn}, and d_{ln} are the normalized intensities at the time channel j, emission wavelength k, and excitation wavelength l, respectively, and E_{ijkl} is an element of the array of errors not fitted by the model. The scores (relative concentrations) are collected in matrix **A**, of size $(I_{cal}+1)\times N$, where I_{cal} is the number of calibration samples. The loadings (normalized intensities) are collected into the profile matrices **B**, **C** and **D**, of size $J\times N$, $K\times N$ and $L\times N$, respectively, where J, K and L are the number of time channels, emission and excitation, respectively. The structure of the model Eq. (1) is called quadrilinear, and the unique decomposition is usually accomplished through alternating least-squares [13,18]. This constitutes the basis of the so-called second-order advantage, which should allow the analyst to obtain the concentration values of calibrated constituents in the presence of any number of uncalibrated components, such as the samples analyzed in the present report.

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

Initializing PARAFAC for the study of four-way arrays can be done using several options implemented in the PARAFAC package [13]: (1) singular value decomposition (SVD) vectors, (2) random orthogonalized values and (3) the best-fitting model of several models fitted using a few iterations. In our case, PARAFAC was initialized with the loadings giving the best fit after a small number of trial runs, selected from the comparison of the results provided by several random loadings [13]. Scores and loadings were restricted to be non-negative during the alternating least-squares fitting phase, and convergence was achieved when the relative change in fit was 1×10^{-6}. The number of components (N) was estimated by the analysis of residuals [13], considering the sum of squared errors (SSE), i.e., the sum of squared elements of the array $\boldsymbol{E}$ in Eq. (1):

$$\mathrm{SSE} = \sum_{i=1}^{I_{cal}+1}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l=1}^{L}(E_{ijkl})^2 \qquad (2)$$

This parameter decreases with increasing N, until it stabilizes at a value corresponding to the optimum number of components. In addition, the spectral profiles produced by the addition of subsequent components were evaluated. If a new component generated repeated profiles, suggesting overfitting, it was discarded and the previous number was selected.

Identification of the chemical constituent under investigation is done with the aid of the extracted profiles **B**, **C**, and **D**, in comparison with those for analyte standards.

Absolute analyte concentrations are obtained after calibration, because the four-way array decomposition only provides relative values (the scores contained in matrix **A**). Calibration is usually done by means of the set of standards with known analyte concentrations, a procedure which is repeated for each new test sample analyzed.

2.2. Stepwise description of external calibration mode for third-order/four-way data using PARAFAC

1) Build an $(I_{cal} + 1) \times J \times K \times L$ array with the first I EEFM-elution time for the training samples and the last one for the unknown.

2) Decompose the array and obtain **A**, **B**, **C**, and **D**.

3) Identify the nth analyte of interest from **B**, **C**, and **D** profiles.

4) Regress the first I_{cal} elements of column a_n against known standard concentrations c_{cal} of analyte n: $[a_{1,n}\ldots aI_{cal,n}] = k_1 \times c_{cal}$ (pseudo-univariate calibration).

5) Convert relative to absolute concentration of n in the unknown, starting from the last element of column a_n: $c_{unk} = a_{Ical+1,n}/k_1$

2.3. Software

The data were handled using the MATLAB computer environment (MATLAB R2012a), and were implemented using the graphical interface MVC3 [19], which is an integrated MATLAB toolbox for third-order calibration. It is freely available on the Internet [20].

3. Experimental

3.1. Reagents and materials

All PAHs were purchased from Sigma-Aldrich (Milwaukee, USA). Both acetonitrile and methanol were obtained from Merck (Darmstadt, Germany). All reagents were of high-purity grade and used as received. Stock solutions of all PAHs of about 500 μg mL^{-1} were prepared in acetonotrile. From these solutions, more diluted acetonitrile solutions of about 5 μg mL^{-1} were obtained. Working solutions were prepared immediately before their use by taking appropriate aliquots of solutions and diluting with acetonitrile and water (85:15 v/v) to the

185 desired concentrations. The PAHs were handled with extreme caution, using gloves and
186 protective clothing.
187
188 *3.2. Apparatus and procedure*
189
190 The chromatographic runs were performed on a Shimadzu Prominence HGE-UV liquid
191 chromatograph equipped with an oven column compartment, and the LabSolutions V 5.82
192 software package to control the instrument data acquisition and data analysis. A 100 µL loop
193 was employed to introduce each sample onto an Agilent Poroshell 120 EC-C18 column (2.7
194 µm average particle size, 50 mm×4.6 mm i.d.). The column temperature was controlled by
195 setting the oven temperature at 27 °C. A 50 cm PTFE tube of 3.17 mm i.d. was used to
196 connect the column with the detector. PTFE tubing of 0.76 mm i.d. was used for all the
197 remaining connections. The mobile phase was the same mixture of acetonitrile and water
198 (85:15 v/v) used to prepare the samples. Samples were filtered through 0.22 µm nylon
199 membranes before injection. The volumetric flow rate (VFR) was maintained at 0.4 mL
200 min^{-1}.
201 An Agilent Cary-Eclipse luminescence spectrometer (Agilent Technologies, Waldbronn,
202 Germany) was used as detector, employing an 8 µL quartz flow cell (Starna, CA, USA) of
203 1 mm optical path. The excitation and emission slit widths were 10 nm, photomultiplier
204 sensitivity was 800 V, and spectral scanning speed of 18,000 nm min^{-1}.
205 Chromatographic data were collected from 4 to 17 min each 0.28 min, and EEFMs were
206 recorded from 350 nm to 480 nm each 3.75 nm (emission) and from 240 to 300 nm each 5 nm
207 (excitation). The reading of each EEFM required a time of approximately 17 s, allowing to
208 register 45 EEFMs for each sample. In this way, data arrays of size 45×36×13 for temporal,
209 emission spectral and excitation spectral modes were respectively generated. For data

modeling, the spectral data points were reduced (emission range: 350–458 nm) in order to eliminate the Rayleigh contribution. The complete analysis, performed under isocratic conditions, was carried out in about 17 min.

3.3. Calibration, validation, and test samples

A calibration set of 22 samples was prepared (Table S1 of Supplementary data). Twelve of these samples corresponded to the concentrations provided by a Plackett–Burman design, one sample corresponded to a blank solution, another sample contained all the studied PAHs at average concentrations, and the remaining eight samples included each pure analyte, also at an average concentration. The tested concentrations for FL, BaA, CHR, BbF, BkF, BaP and DBA were in the ranges 0–100 ng mL^{-1}, and for PYR it was 0–200 ng mL^{-1}. A validation set was prepared employing concentrations different than those used for calibration and following a random design. Calibration and validation samples were prepared by measuring appropriate aliquots of standard solutions, placing them in 5.00 mL volumetric flasks to obtain the desired concentrations, and completing to the mark with mobile phase.

Test samples were prepared containing random concentrations of the 8 studied compounds and additional PAHs selected as potential interferences, namely azulene (AZU), phenanthrene (PHEN), indeno[1,2,3-*cd*]pyrene (IP), and benzo[*j*]fluoranthene (BjF). These additional PAHs, besides being potentially present in the same samples as the evaluated analytes and having similar toxicological effects, showed coelution and spectral overlapping with the calibrated analytes, representing a real challenge for the current research. The maximum concentrations of AZU, PHEN, IP, and BjF tested in these latter samples were 400 ng mL^{-1}.

3.4. Water sample procedure

Underground and stream water samples were prepared by spiking them with standard solutions of the studied PAHs, obtaining concentration levels in the range 0–3 ng mL^{-1}. These samples were prepared in duplicate and were filtered through 0.45 µm pore size nylon membranes. Samples were subjected to solid-phase extraction with C18 disks. Each disk was previously conditioned with 0.5 mL of methanol and 1 mL of ultrapure water. Aliquots of 20, 100 or 200 mL were passed through the disks under vacuum, with a flow rate of 10 mL min^{-1}. After elution of the retained organic constituents with 1 mL of acetonitrile, the solvent was evaporated with nitrogen, the residue was reconstituted with either 0.50 or 1.00 mL of mobile phase, and the obtained solutions were subjected to the same chromatographic analysis as the validation samples.

4. Results and discussion

4.1. Third-order LC-EEFM data on-line generation

Four-way PARAFAC has relevant advantages such as: (1) the resolution is unique, (2) in general, restrictions are not required [13] and, (3) the sensitivity is maximum [21]. Therefore, efforts were made to generate quadrilinear data and to take advantage of the benefits of using four-way PARAFAC.

There are two fundamental issues to consider when working with chromatographic data to be processed by multiway PARAFAC. One of them is the possible lack of repeatability in the elution times between successive runs, representing a limitation when data processing is performed with multi-linear algorithms, since they require tri- or quadrilinearity for second-

and third-order data, respectively [4]. In our system, replicate analysis showed that no significant changes in the elution time profiles ocurred between different runs.

Another issue inherent to the present multiway data is the signal measurement of a moving sample, which makes the local analyte concentration variable during the spectral data acquisition. This should in principle lead to a loss of multinearity. For second-order elution time-fluorescent emission wavelength data, it was demonstrated that emission spectra can be obtained in a very short time by using a fast-scanning spectrofluorimeter: each emission spectrum is recorded in a time appreciably smaller than the base width of a chromatographic peak [22,23]. However, for third-order LC-EEFM data, matrix measurements in an extremely short time are not feasible. Even with a modern fast-scanning spectrofluorimeter, the usual time required to measure a complete EEFM is long enough to produce significant variations of the local analyte concentration as a function of excitation wavelength. The excitation spectrum is deformed as a function of the measurement time, and this mutual dependence between excitation and time profiles leads to a significant loss of quadrilinearity [4].

EEFMs can be collected in an appropriate time during chromatographic elution by decreasing the linear flow rate (LFR) of the mobile phase, without increasing the total time of analysis. The volumetric flow rate (VFR) of a fluid is related to the LFR and the tube cross sectional area (A) through the following equation:

$$VFR = LFR \times A \quad (3)$$

It is evident that the LFR can be decreased at constant VFR by increasing the diameter of the flowing tube. Therefore, typical PTFE tubing of 0.76 mm i.d. was used for all chromatographic connections, except in the section that connects the column with the detector, where the diameter was 3.17 mm. Three lengths of this latter tube (25, 50 and 100 cm) and VFR values in the range 0.4–0.7 mL min^{-1} were probed. It was corroborated that a 50

283 cm tube length and a VFR of 0.4 mL min^{-1} produced better signals, allowing the effective
284 acquisition of the EEFMs in a total time of about 17 min.

285 As will be shown below, the above experimental considerations allowed us to conclude
286 that quadrilinearity was fulfilled and that four-way PARAFAC could be safely applied,
287 leading to satisfactory analytical results.

288

289 *4.2. PAHs system*

290

291 Fig. 1A shows the chromatogram and the emission and excitation spectra for FL, PYR,
292 BaA, CHR, BbF, BkF, BaP and DBA under the selected experimental conditions. For a
293 comparison with a classical chromatogram of the same analytes see Fig. S1 (Supplementary
294 data). Although the peaks are wider under the presently applied conditions, none of the
295 chromatograms show full peak resolution. In any case, this is not mandatory for successful
296 mathematical resolution using multi-way calibration.

297 Relevant challenges in relation to Fig. 1A are the total co-elution of BaA and CHR in the
298 time profile, and a marked spectral similarity in some signals: e.g. in the emission profiles of
299 the pairs FL/ BbF, and BaP/BkF, and in the excitation profiles of FL/BaA. Under these
300 conditions, second-order calibration may not be able to resolve such highly collinear systems.
301 However, this issue can be solved by third-order/four way calibration.

302 Sidiropoulos and Bro generalized Kruskal's result on the uniqueness of trilinear
303 decomposition of three-way arrays to the case of multilinear decomposition of higher-way
304 arrays. They showed that the four- way array can be successfully decomposed, even when two
305 profiles are identical in one mode [24]. The unique decomposition in the latter case, as
306 achieved by third-order/four way calibration, represents the third-order advantage.

The present calibration methodology provides additional benefits in the case of chromatographic measurements with fluorescence detection. In second-order analysis, the choice of the excitation wavelength is generally performed favoring the analyte with the weakest fluorescence signal. This is often not suitable for other sample analytes. In the presently proposed procedure, however, a wide range of excitation wavelengths is recorded. This increases the method sensitivity, since each PAH is irradiated at its optimal excitation wavelength, and the selectivity, because each PAH has its characteristic excitation profile.

4.3. Validation set

Four-way PARAFAC was firstly applied to the validation set of samples, as described in Section 2.1. Ten PARAFAC components, corresponding to the eight presently studied PAHs and two blank signals, were required to describe the variability in these data arrays, providing appropriate emission, excitation and chromatographic profiles (Fig. 1B).

The quality of the recovered profiles was evaluated through the similarity coefficient (r) between the reference (Fig. 1A) and the retrieved spectral profiles (Fig. 1B) [25]. The values of r found for FL, PYR, BaA, CHR, BbF, BkF. BaP and DBA were, respectively, 0.9987, 0.9979, 0.9989, 0.9994, 0.9950, 0.9832, 0.9980 and 0.9990 for the emission profiles, and 0.9993, 0.9974, 0.9406, 0.9992, 0.9987, 0.9953, 0.9997 and 0.9998 for the excitation profiles, indicating a satisfactory match between the resolved and pure spectra of the studied analytes.

Figs. 2A and 2B show the good prediction results for validation samples for individually evaluated PAH and for all analytes, respectively. For a statistical evaluation of the obtained results, the EJCR (elliptical joint confidence region) test for slopes and intercepts of the found vs nominal concentrations plots was also included in these figures [26]. Because all ellipses include the theoretically expected values of slope = 1 and intercept = 0, the accuracy of the

methodology can be stated. The statistical results for validation samples are completed with the parameters shown in Table 1.

Limits of detection (LODs) were estimated using the rigorous expression recommended by the International Union of Pure and Applied Chemistry (IUPAC), which take into account the so-called type 1 and 2 errors (false detects and false non-detects, respectively) [3,21]:

$$\mathrm{LOD} = 3.3(\mathrm{SEN}^{-2}\,\sigma_x^2 + h_0\mathrm{SEN}^{-2}\,\sigma_x^2 + h_0\sigma_{ycal}^2)^{1/2} \quad (4)$$

where the factor 3.3 is the sum of *t*-coefficients accounting for type I and II errors at 95% confidence level, SEN is the sensitivity, σ_x^2 is the instrumental variance, h_0 is the sample leverage at zero analyte concentration, σ_{ycal}^2 is the variance in calibration concentrations.

The sensitivity for a given analyte is defined as [21]:

$$\mathrm{SEN} = \{\mathbf{g}^{\mathrm{T}}[\mathbf{Z}_{\mathrm{exp}}{}^{\mathrm{T}}(\mathbf{I} - \mathbf{Z}_{\mathrm{unx}}\mathbf{Z}_{\mathrm{unx}}{}^{+})\mathbf{Z}_{\mathrm{exp}}]^{-1}\mathbf{g}\}^{-1/2} \quad (5)$$

In the latter equation, **I** is a unit matrix, **g** is a column vector of size $N_{\mathrm{cal}}\times 1$ (N_{cal} = number of analytes) with all zeros except a single one in the analyte position, $\mathbf{Z}_{\mathrm{exp}}$ is given by:

$$\mathbf{Z}_{\mathrm{exp}} = m\mathbf{D}_{\mathrm{exp}}\odot\mathbf{C}_{\mathrm{exp}}\odot\mathbf{B}_{\mathrm{exp}} \quad (6)$$

where *m* is the slope of the pseudounivariate calibration curve for the analyte, $\mathbf{D}_{\mathrm{exp}}$, $\mathbf{C}_{\mathrm{exp}}$ and $\mathbf{B}_{\mathrm{exp}}$ are the matrices of profiles in the three modes for all analytes, and "⊙" indicates the Khatri–Rao product operator. The matrix $\mathbf{Z}_{\mathrm{unx}}$, on the other hand, is:

$$\mathbf{Z}_{\mathrm{unx}} = [\mathbf{d}_1\otimes\mathbf{c}_1\otimes\mathbf{I}_b|\mathbf{d}_1\otimes\mathbf{I}_c\otimes\mathbf{b}_1|\mathbf{I}_d\otimes\mathbf{c}_1\otimes\mathbf{b}_1|\mathbf{d}_2\otimes\mathbf{c}_2\otimes\mathbf{I}_b|\mathbf{d}_2\otimes\mathbf{I}_c\otimes\mathbf{b}_2|\mathbf{I}_d\otimes\mathbf{c}_2\otimes\mathbf{b}_2|...] \quad (7)$$

where $\mathbf{b}_1$, $\mathbf{b}_2$, …, $\mathbf{c}_1$, $\mathbf{c}_2$, …, $\mathbf{d}_1$, $\mathbf{d}_2$, … are columns of the **B**, **C**, **D** matrices for the unexpected constituents, $\mathbf{I}_b$, $\mathbf{I}_c$ and $\mathbf{I}_d$ are appropriately dimensioned unit matrices, and "⊗" is the Kronecker product. The numbers 1, 2, ... run up to the total number of unexpected constituents.

The limit of quatitation was defined as [21]:

$$\mathrm{LOQ} = 3\times\mathrm{LOD} \quad (8)$$

In a previous work, where a similar PAHs system was resolved through second-order calibration using chromatographic data with fluorescence detection, LODs were 20, 26, 17, 10, 4, 1, 2, and 2 ng mL^{-1} for FL, PYR, BaA, CHR, BbF, BkF, BaP and DBA, respectively [22]. The comparison of these values with those here obtained (LODs between 0.4 and 2.9 ng mL^{-1}) demonstrates the positive influence of the third-order/four-way calibration in the sensitivity of the method.

The advantages of this type of calibration are also corroborated through the good obtained selectivities, all higher than 0.49, and reasonably large analytical sensitivities (Table 1).

The selectivity (SEL) for each analyte is given by [21]:

$$SEL = SEN/m \quad (9)$$

The analytical sensitivity (γ) has been proposed as a better indicator for comparison purposes [21], as the ratio between sensitivity and instrumental noise (σ_x):

$$\gamma = SEN/\sigma_x \quad (10)$$

Both the root-mean square errors and the relative errors of prediction, computed with respect to the mean calibration concentration of each analyte (below 7 ng mL^{-1} and 10 %, respectively) are acceptable considering the complexity of the evaluated system.

4.4. Test set

The power of the proposed method was also evaluated in the presence of AZU, PHEN, IP, and BjF, which showed coelution and spectral overlapping with the calibrated analytes. The chromatographic and spectral profiles for these potential interferents are shown in Fig. 3A, and their strong overlapping with the studied PAHs can be appreciated in Fig. 3B. Further, collinearity is detected between the chromatographic bands for BbF (analyte) and BjF (a non-calibrated compound). This situation is similar to that indicated above for the analytes CHR

and BaA in the validation samples. As was discussed in relation to the third-order advantage, it does not represent a serious problem when third-order/four-way calibration is applied, provided the spectral profiles are different.

Figs. 4A and 4B show the good individual and global four-way PARAFAC predictions, respectively, corresponding to the test samples containing potential interferences. Although the calculated values for BbF and DBA show a slight dispersion with respect to the perfect fit, the ellipses obtained when the EJCR analysis is applied imply accurate predictions and the ability of four-way PARAFAC to resolve highly overlapped system. This conclusion is corroborated by the statistical results shown in Table 1. While the presence of non-calibrated PAHs in the samples produces an increase in the REP values, LODs do not appear to be considerably affected by the presence of the studied interferences.

4.5. Water samples

An underground sample taken from a countryside zone of La Pampa province (Argentina) and different stream water samples collected near industrial and rural areas of south of the Santa Fe province (Argentina) were selected as real matrices for assaying the proposed method. Since these samples were found to be free from the studied PAHs, spiked samples were prepared and a recovery study was performed. Because the levels of PAHs that can be found in surface and underground waters vary from a few parts-per-trillion to parts-per-billion in contaminated areas [27,28], a wide range of concentrations was covered, applying a very simple pre-concentration treatment. It should be noted that US EPA (United States Environmental Protection Agency) has set in drinking water a maximum contaminant level (MCL) of 0.2 ng mL^{-1} for BaP, BbF, and BkF, while for BaA and DBA the MCL values are of 0.1 and 0.3 ng mL^{-1}, respectively [29].

Table 2 displays the obtained concentration recoveries, and the statistical EJCR test (Fig. 5) supports that there are no statistical differences between found and nominal concentrations, suggesting that foreign compounds present in the studied matrices do not produce a significant interference in our analysis.

5. Conclusions

The presently proposed strategy provides a useful and refreshingly easy way of measuring on-line chromatographic-EEFM third-order data. In contrast to the usual methodologies for the generation of this type of data, such as fraction collection and multiple chromatographic runs per sample, the fluorescence matrices are recorded in parallel with the chromatographic procedure, drastically decreasing the experimental time, the solvent consumption, the waste generation, and using an equipment of low complexity. Thus, the develop method is in accordance with the green analytical chemistry principles. The third-order advantage, evidenced through the appropriate resolution of a system with high degree of collinearity among the analytes themselves and non-calibrated constituents, is added to the well-known benefits of increasing the number of modes in multivariate calibration.

Acknowledgements

This work was financially supported by Universidad Nacional de Rosario (Project BIO 237), CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas, Project PIP 0163), and ANPCyT (Agencia Nacional de Promoción Científica y Tecnológica, Project PICT 2016-1122). M. D. C. thanks ANPCyT for a doctoral fellowship.

Appendix A. Supplementary data

Supplementary data related to this article is included

References

[1] A. Muñoz de la Peña, H.C. Goicoechea, G.M. Escandar, A.C. Olivieri (Eds), Data Handling in Science and Technology, Fundamental and Analytical Applications of Multiway Calibration, Elsevier, Amsterdam, 2015, Vol. 29.

[2] K.S. Booksh, B.R. Kowalski, Theory of Analytical Chemistry, Anal. Chem. 66 (1994) 782A–791A.

[3] C. Bauza, G.A. Ibañez, R. Tauler, A.C. Olivieri, Sensitivity equation for quantitative analysis with multivariate curve resolution-alternating least-squares: theoretical and experimental approach, Anal. Chem. 84 (2012) 8697–8706.

[4] A.C. Olivieri, G.M. Escandar, Practical Three-Way Calibration, Elsevier, Waltham, USA, 2014.

[5] C. Kang, H.L. Wu, L.X. Xie, S.X. Xiang, R.Q. Yu, Direct quantitative analysis of aromatic amino acids in human plasma by four-way calibration using intrinsic fluorescence: exploration of third-order advantages, Talanta 122 (2014) 293–301.

[6] M.D. Carabajal, J.A. Arancibia, G.M. Escandar, Excitation-emission fluorescence-kinetic data obtained by Fenton degradation. Determination of heavy-polycyclic aromatic hydrocarbons by four-way parallel factor analysis, Talanta 165 (2017) 52–63.

[7] G.M. Escandar, A.C. Olivieri, A road map for multi-way calibration models, Analyst 142 (2017) 2862–2873.

[8] M. Montemurro, G.G. Siano, M.R. Alcaráz, H.C. Goicoechea, Third order chromatographic-excitation-emission fluorescence data: advances, challenges and prospects in analytical applications, Trends Anal. Chem. 93 (2017) 119–133.

[9] V.A. Lozano, A. Muñoz de la Peña, I. Durán Merás, A. Espinosa Mansilla, G.M. Escandar, Four-way multivariate calibration using ultra-fast high-performance liquid chromatography with fluorescence excitation-emission detection. Application to the direct analysis of chlorophylls a and b and pheophytins a and b in olive oils. Chemom. Intell. Lab. Syst. 125 (2013) 121–131.

[10] M. Montemurro, L. Pinto, G. Véras, A. de Araújo Gomes, M.J. Culzoni, M.C. Ugulino de Araújo, H.C. Goicoechea, Highly sensitive quantitation of pesticides in fruit juice samples by modeling four-way data gathered with high-performance liquid chromatography with fluorescence excitation-emission detection, Talanta 154 (2016) 208–218.

[11] M.R. Alcaráz, G.G. Siano, M.J. Culzoni, A. Muñoz de la Peña, H.C. Goicoechea, Modeling four and three-way fast high-performance liquid chromatography with fluorescence detection data for quantitation of fluoroquinolones in water samples, Anal. Chim. Acta 809 (2014) 37–46.

[12] M.R. Alcaráz, S.A. Bortolato, H.C. Goicoechea, A.C. Olivieri, A new modeling strategy for third-order fast high-performance liquid chromatographic data with fluorescence detection. Quantitation of fluoroquinolones in water samples, Anal. Bioanal. Chem. 407 (2015) 1999–2011.

[13] R. Bro, PARAFAC. Tutorial and applications, Chemom. Intell. Lab. Syst. 38 (1997) 149–171.

[14] A.C. Olivieri, J.A. Arancibia, A. Muñoz de la Peña, I. Durán Merás, A. Espinosa Mansilla, Second-order advantage achieved with four-way fluorescence excitation-emission-

kinetic data processed by parallel factor analysis and trilinear least-squares. Determination of methotrexate and leucovorin in human urine, Anal. Chem. 76 (2004) 5657–5666.

[15] T. Wenzl, R. Simon, J. Kleiner, E. Anklam, Analytical methods for polycyclic aromatic hydrocarbons (PAHs) in food and the environment needed for new food legislation in the European Union, Trends Anal. Chem. 25 (2006) 716–725.

[16] http://monographs.iarc.fr/ENG/Classification/ (accessed August 2017).

[17] H.I. Abdel-Shafy, M.S.M. Mansour, A review on polycyclic aromatic hydrocarbons: source, environmental impact, effect on human health and remediation, Egypt. J. Petrol. 25 (2016) 107–123.

[18] P. Paatero, A weighted non-negative least squares algorithm for three-way 'PARAFAC' factor analysis, Chemom. Intell. Lab. Syst. 38 (1997) 223–242.

[19] A.C. Olivieri, H.L. Wu, R.Q. Yu, MVC3: A MATLAB graphical interface toolbox for third-order multivariate calibration, Chemom. Intell. Lab. Syst. 116 (2012) 9–16.

[20] www.iquir-conicet.gov.ar/descargas/mvc3.rar (accessed August 2017).

[21] A.C. Olivieri, Analytical figures of merit: from univariate to multiway calibration, Chem. Rev. 114 (2014) 5358–5378.

[22] S.A. Bortolato, J.A. Arancibia, G.M. Escandar, Non-trilinear chromatographic time retention–fluorescence emission data coupled to chemometric algorithms for the simultaneous determination of ten polycyclic aromatic hydrocarbons in the presence of interferences, Anal. Chem. 81 (2009) 8074–8084.

[23] J.A. Arancibia, G.M. Escandar, Second-order chromatographic photochemically-induced fluorescence emission data coupled to chemometric analysis for the simultaneous determination of urea herbicides in the presence of matrix co-eluting compounds, Anal. Methods 6 (2014) 5503–5511.

[24] N.D. Sidiropoulos, R. Bro, On the uniqueness of multilinear decomposition of *N*-way arrays, J. Chemometrics 14 (2000) 229–239.

[25] V. Gómez, M. Miró, M.P. Callao, V. Cerdá, Coupling of sequential injection chromatography with multivariate curve resolution-alternating least-squares for enhancement of peak capacity, Anal. Chem. 79 (2007) 7767–7774.

[26] A.G. González, M.A. Herrador, A.G. Asuero, Intra-laboratory testing of method accuracy from recovery assays, Talanta 48 (1999) 729–736.

[27] Md.J. Alam, D. Yuan, Y.J. Jiang, Y. Sun, Y. Li, X. Xu, Sources and transports of polycyclic aromatic hydrocarbons in the Nanshan Underground River, China, Environ. Earth Sci. 71 (2014) 1967–1976.

[28] R. Sarria-Villa, W. Ocampo-Duque, M. Páez, M. Schuhmacher, Presence of PAHs in water and sediments of the Colombian Cauca River during heavy rain episodes, and implications for risk assessment, Sci. Total Environ. 540 (2016) 455–465.

[29] https://www.atsdr.cdc.gov/csem/csem.asp?Csem=13 & po=8 (accessed August 2017).

Figure Caption

Fig 1. (A) Experimental chromatographic (top), emission (medium) and excitation (bottom) profiles for FL (red), PYR (blue), BaA (green), CHR (pink), BbF (light blue), BkF (gray), BaP (violet), and DBA (light green). (B) Profiles retrieved by four-way PARAFAC after processing a validation sample. Blanks were omitted for clarity. All intensities are normalized to unit length.

Fig. 2. (A) Individual plots of predicted concentrations as a function of the nominal values for FL (red), PYR (blue), BaA (green), CHR (pink), BbF (light blue), BkF (gray), BaP (violet), and DBA (light green) in validation samples, and the corresponding elliptical joint regions for the slopes and intercepts of the regressions for predictions. (B) Plot of predicted concentrations as a function of the nominal values for all evaluated PAHs and that corresponding to the global ellipse including all predictions. Black circles in the elliptical plots mark the theoretical (intercept = 0, slope = 1) point.

Fig. 3. (A) Chromatographic (top), emission (medium) and excitation (bottom) profiles for the potential interferences AZU (long-dashed line), PHEN (dashed-dotted line), IP (solid line), and BjF (short-dashed line). (B) Chromatographic (top), emission (medium) and excitation (bottom) profiles for the studied analytes (colour codes as in Fig. 1), superimposed with the potential interferences shown in (A). All intensities are normalized to unit length.

Fig. 4. (A) Individual plots of predicted concentrations as a function of the nominal values for FL (red), PYR (blue), BaA (green), CHR (pink), BbF (light blue), BkF (gray), BaP (violet), and DBA (light green) in test samples, and the corresponding elliptical joint regions for the slopes and intercepts of the regressions for predictions. (B) Plot of predicted concentrations as a function of the nominal values for all evaluated PAHs and that corresponding to the global ellipse including all predictions. Black circles in the elliptical plots mark the theoretical (intercept = 0, slope = 1) point.

Fig. 5. (A) Plot of PAHs predicted concentrations in water samples as a function of the nominal values (the solid line is the perfect fit). The inset shows the predictions in the low concentrations range. (B) Elliptical joint region at 95% confidence level for slope and intercept of the regression of four-way PARAFAC. Black circle marks the theoretical (intercept = 0, slope = 1) point.

Table 1
Statistical results for the analytes in validation samples and in samples with AZU, PHEN, IP, and BjF as potential interferences.

	FL	PYR	BaA	CHR	BbF	BkF	BaP	DBA
	Validation samples							
γ	1.6	1.2	4.2	3.6	2.1	9.2	6.5	3.3
SEL	0.90	0.75	0.78	0.77	0.55	0.49	0.56	0.82
LOD	2.1	2.9	0.8	1.0	1.6	0.4	0.5	1
LOQ	6.4	8.9	2.5	2.9	4.9	1.2	1.6	3.1
RMSEP	4	7	3	3	5	3	2	4
REP	8	7	6	6	10	6	4	8
	Samples with potential interferences							
γ	1.1	0.77	3.9	2.4	2.2	10	5.6	3.1
SEL	0.65	0.49	0.72	0.51	0.58	0.50	0.56	0.79
LOD	3.0	4.4	0.9	1.5	1.6	0.3	0.6	1.1
LOQ	9.2	13	2.7	4.4	4.7	1.0	1.9	3.4
RMSE	6	11	4	5	6	4	4	7
REP	12	11	8	10	12	8	8	14

γ (ng^{-1} mL), analytical sensitivity; SEL, selectivity; LOD (ng mL^{-1}), limit of detection, and LOQ (ng mL^{-1}), limit of quantitation, were calculated according to Ref. [21]; RMSE (ng mL^{-1}), root-mean square error; REP (%), relative error of prediction.

Table 2
Recovery study for the studied PAHs in spiked water samples.

	FL	PYR	BaA	CHR	BbF	BkF	BaP	DBA
Water I								
Taken	0.055	0.044	0.054	0.055	0.052	0.056	0.054	0.045
Found	0.052(1)	0.045(5)	0.049(1)	0.056(2)	0.053(2)	0.058(1)	0.054(1)	0.035(2)
Recovery	95	102	91	102	102	104	100	78
Water II								
Taken	0.183	0.880	0.400	0.160	0.360	0.300	0.140	0.250
Found	0.195(9)	0.80(5)	0.35(2)	0.12(1)	0.30(3)	0.26(1)	0.13(1)	0.26(2)
Recovery	106	91	88	75	83	87	93	104
Water III								
Taken	0.091	0.165	0.080	0.090	0.052	0.080	0.080	0.055
Found	0.075(7)	0.145(7)	0.060(5)	0.089(2)	0.054(2)	0.082(1)	0.059(8)	0.042(2)
Recovery	82	88	75	99	104	103	74	76
Water IV								
Taken	0.303	0.445	0.306	0.300	0.360	0.336	0.271	0.567
Found	0.315(2)	0.447(2)	0.223(6)	0.270(3)	0.31(3)	0.343(9)	0.23(2)	0.70(2)
Recovery	104	100	73	90	86	102	85	123
Water V								
Taken	0.153	0.033	0.150	0.180	0.180	0.195	0.160	0.193
Found	0.15(2)	0.04(1)	0.12(3)	0.16(1)	0.16(1)	0.16(1)	0.12(1)	0.200(3)
Recovery	98	131	80	89	89	82	75	104
Water VI								
Taken	0.620	0.800	0.500	0.520	0.460	0.560	0.540	0.540
Found	0.72(5)	0.80(1)	0.42(4)	0.48(4)	0.46(9)	0.56(3)	0.51(3)	0.563(1)
Recovery	116	100	84	92	100	100	94	104
Water VII								
Taken	0.552	0.288	0.250	0.330	0.260	0.233	0.300	0.270
Found	0.58(9)	0.33(2)	0.23(3)	0.26(1)	0.27(3)	0.28(1)	0.26(3)	0.27(1)
Recovery	105	114	92	79	104	120	87	100
Water VIII								
Taken	1.07	1.86	1.25	1.38	0.88	1.55	1.25	1.21
Found	1.00(1)	1.80(1)	1.25 (9)	1.45(6)	0.88(1)	1.55(2)	1.10(9)	1.16(5)
Recovery	93	97	100	105	100	100	88	96
Water IX								
Taken	0.850	2.65	0.750	0.900	0.900	0.800	0.850	0.730
Found	0.87(6)	2.70(8)	0.65(3)	0.850(7)	0.80(8)	0.76(3)	0.85(8)	0.757(3)
Recovery	102	102	87	94	89	95	100	104

Waters from the following sources: I and II, Salvat stream (Santa Fe, Argentina); III and IV, Ludueña stream (Santa Fe, Argentina); V and VI, Ibarlucea stream (Santa Fe, Argentina); VII, Santa Rosa underground (La Pampa, Argentina); VIII and IX, Andino stream (Santa Fe, Argentina). Concentrations are given in ng mL^{-1}, and recoveries are given in percentage. Experimental standard deviations of duplicates are given between parentheses and correspond to the last significant figure.

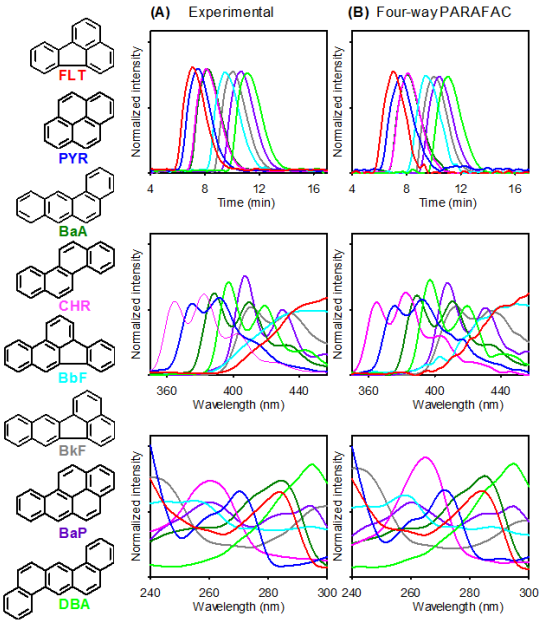
(A) Experimental
(B) Four-way PARAFAC
FLT
PYR
BaA
CHR
BbF
BkF
BaP
DBA
Normalized intensity
Time (min)
Wavelength (nm)

Fig. 1

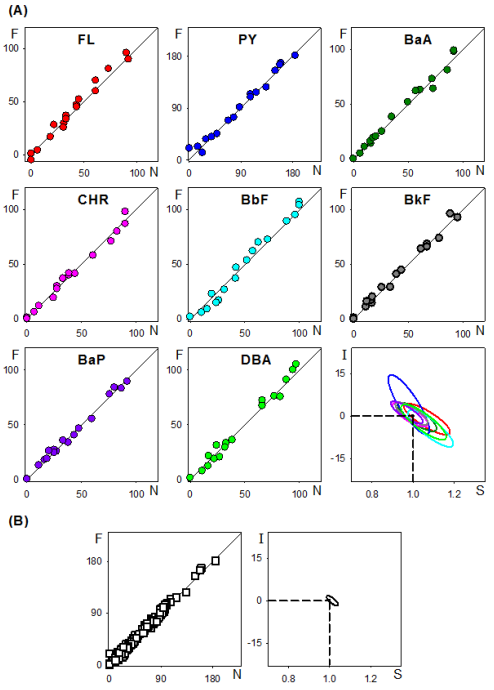
(A)
FL
PY
BaA
CHR
BbF
BkF
BaP
DBA
F
N
I
S
(B)

Fig. 2

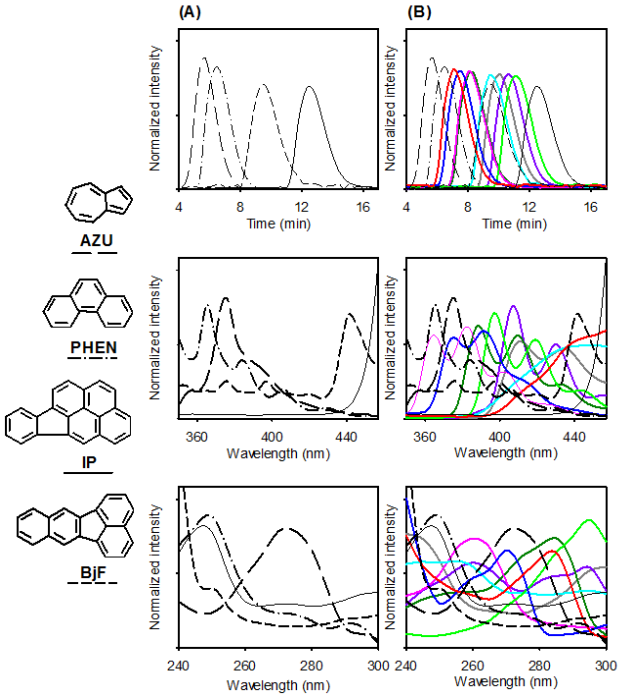
(A)
(B)
Normalized intensity
Time (min)
AZU
PHEN
IP
BjF
Wavelength (nm)

Fig. 3

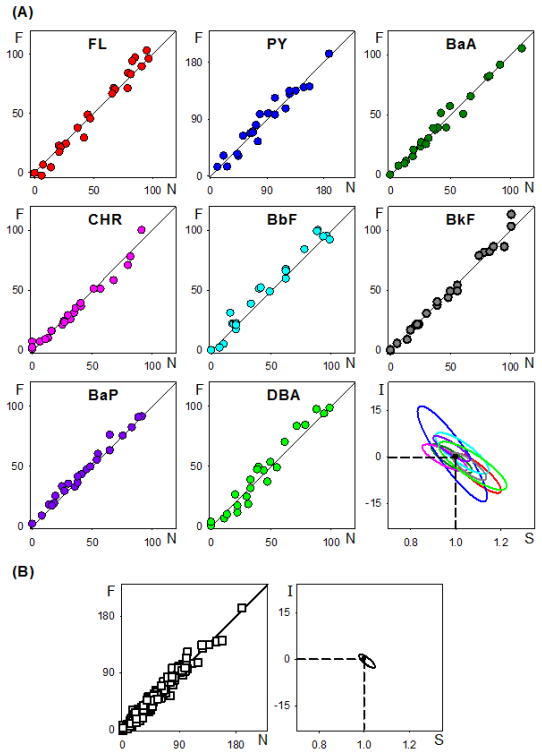
(A)
FL
PY
BaA
CHR
BbF
BkF
BaP
DBA
F
I
N
S
(B)

Fig. 4

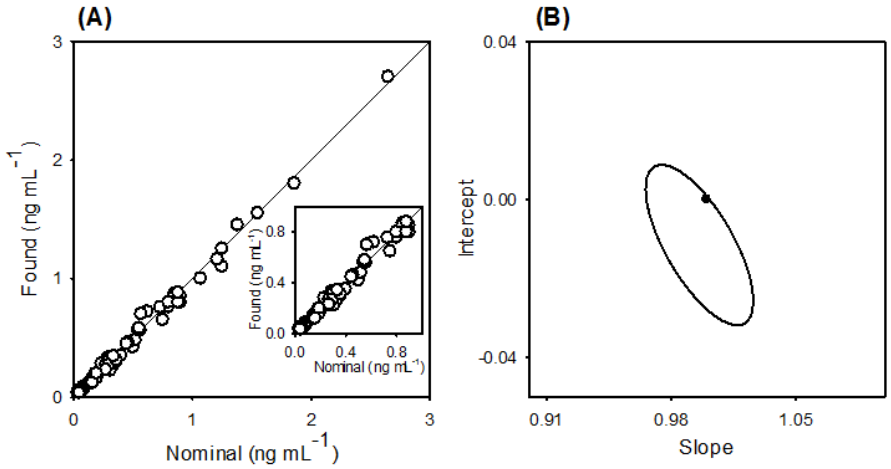
(A)
(B)
Found (ng mL-1)
Nominal (ng mL-1)
Found (ng mL-1)
Nominal (ng mL-1)
Intercept
Slope
0
1
2
3
0.0
0.4
0.8
-0.04
0.00
0.04
0.91
0.98
1.05

Fig. 5