OPEN ACCESS DOCUMENT

Information of the Journal in which the present paper is published:

- **Analytica Chimica Acta**, 2020, 1113, 52-65
- <https://doi.org/10.1016/j.aca.2020.03.057>

Application of the area correlation constraint in the MCR-ALS quantitative
analysis of complex mixture samples

Mehrnoosh Bayat^{a, *}, Marc Marín-García^b, Jahan B. Ghasemi^a and Roma Tauler^{b, *}

^a Department of Chemistry, Faculty of Sciences, University of Tehran, P.O. Box 14155–6455, Tehran, Iran

^b Department of Environmental Chemistry, IDAEA-CSIC, Jordi Girona 18–26, 08034 Barcelona, Catalonia, Spain

* Corresponding author E-mail: E-mail address: Roma.Tauler@idaea.csic.es (R. Tauler)

Abstract

The application of the recently developed area correlation constraint in Multivariate Curve Resolution–Alternating Least Squares (MCR–ALS) for the quantitative determination of analyte mixtures is shown. The feasibility of the proposed constraint is tested firstly for the calibration and quantitation of PAHs mixtures in their synthetic mixtures (validation samples) and in river water samples dissolved organic matter (DOM) using EEM fluorescent three-way data. In this case, MCR–ALS results obtained with the proposed area correlation constraint are comparable with the results obtained with methods based on the fulfillment of the trilinear model, like PARAFAC and MCR–ALS with the trilinearity constraint. Secondly, the possibility of applying this new area correlation constraint is extended to the analytical determination of lipid mixtures in synthetic and cell culture samples by LC–MS, where the trilinear model does not hold. The applicability of the proposed area correlation constraint is assessed, and it is proposed as a general tool for the quantitative determination of unknown mixtures of analytes in complex natural samples with severe profile overlapping and unknown composition, whatever the data structure is.

Keywords: MCR–ALS; area correlation constraint; excitation–emission data; LC–MS data

1. Introduction

Quantitative determination of multiple analytes in complex environmental and biological mixture samples is still a challenge in analytical chemistry [1, 2]. In recent years, different multivariate data analysis chemometric methods have been proposed to solve the mixture analysis type of problems [3–5]. Chemometric methods are powerful data analysis tools which are increasingly used in the study, analysis and resolution of complex multi-component mixture systems [6–8]. The application of multivariate calibration methods for second-order data, where the instrumental analysis of one single sample gives a data matrix, allows the analysis of complex natural samples in the presence of unknown interferences [9, 10]. The main advantage of these second-order calibration methods (called second-order advantage) is that the qualitative and quantitative determination of the considered analyte in the sample (or samples) can be achieved in the presence of unknown uncalibrated interferences, without the need of using costly separation nor sample pre-treatment methods, and that the selectivity and sensitivity of the analysis can be significantly improved [11]. Among these multivariate data analysis methods, the multivariate curve resolution–alternating least-squares (MCR–ALS) method is a widespread methodology used in a great variety of situations and application fields [14, 15].

MCR–ALS is an iterative soft modeling resolution method which allows for the resolution of the concentration and spectra profiles of constituents of a mixture, when no or little prior information is available for them. It has remarkable flexibility for the simultaneous analysis of multiple samples and of their associated two-way, three-way or multiway data sets [16, 17]. This method allows for the recovery of the multivariate responses (profiles) of the constituents as well as for the extraction of their quantitative composition in the analyzed samples. When MCR–ALS is applied to three-way data sets and the trilinear model holds, the rotation ambiguities associated with the MCR bilinear model can be eliminated by the application of the trilinearity constraint [18, 19]. The selection of suitable constraints is one of the main critical aspects of ALS optimization to reduce rotation ambiguities and to obtain optimal and reliable solutions [20]. Non-negativity constraints are the most general natural constraints that apply in most of the cases due to the natural non-negative nature of the measurements and responses of chemical systems. Apart from this constraint, other constraints and mathematical

conditions that the sought solutions should fulfill can be considered. When multiple data sets are simultaneously analyzed by the MCR-ALS method, the correspondence among the different chemical species in the different data sets can be used also as a constraint to decide about the presence or absence of a particular component and consequently to reduce the ambiguity in their simultaneous resolution [21, 22]. This constraint is especially helpful in the simultaneous analysis of calibration and test samples, due to the known composition of the analytes in the calibration samples usually prepared from standards and without interferences. In contrast, in the test samples, unknown interferences and matrix effects are present. When there is the possibility to model the data sets using the trilinear model [23], this can be implemented as a constraint during ALS optimization, and the obtained results are comparable to those obtained by other proposed multilinear based methods like the PARAFAC method. A more detailed description of the MCR-ALS method using the trilinearity as a constraint and of the PARAFAC method is given elsewhere [24, 25].

Two different multivariate calibration situations are frequently encountered in practice. One first situation is when the analysis of one sample gives a one-dimensional response, e.g. one single spectrum or one data vector per sample. This is usually referred to as the first-order calibration. In this case, a correlation constraint has been proposed for the MCR-ALS quantitative calibration of analytes in the presence of interferences [26–29]. Another situation is when the analysis of one sample gives a two-dimensional response or data table/matrix, like for instance the excitation–emission (EEM) fluorescent spectra of one sample. In this case, **a new different constraint named “area correlation constraint” has been proposed for the MCR-ALS** quantitative calibration of the analytes in the presence of interferences during the ALS optimization [30]. In this way, not only the concentrations of the analytes can be recovered in real concentration units, but also more accurate quantifications can be achieved, and rotation ambiguities associated with the bilinear model can be diminished [31, 32].

In the present work, the application of the recently developed area correlation constraint in MCR-ALS is shown for the improvement of quantitative results in two data examples [30–32]. Firstly, MCR-ALS with the area correlation constraint is tested for the analysis of PAH mixtures using Excitation–Emission (EEM) fluorescence second-order data in laboratory and river water samples and the results are compared with those obtained

using the trilinear decomposition model. Secondly, MCR-ALS is applied to the analysis of non trilinear second-order data obtained in the LC-MS analysis of lipid mixtures in synthetic and cell culture samples and the results are compared with those obtained by MCR-ALS with only non-negativity constraints. In this second case, the area correlation constraint is shown to be useful in the analysis of three-way (second-order) data where the trilinearity model does not hold, like frequently is the case of liquid chromatography hyphenated with spectroscopic analysis methods.

2. Theory and background

In this work, the Multivariate Curve Resolution (MCR-ALS, [14, 17, 33, 34]) method is proposed for the resolution and quantitative analysis of complex mixture samples. This method is based on a bilinear data decomposition model which individually in an element-wise way can be written according to the following equation

$$d_{ij} = c_{in} \cdot s_{jn} + e_{ij} \quad \text{Equation. 1}$$

where d_{ij} represents the data value at sample i ($i=1, \dots, I$) and variable j ($j=1, \dots, J$), c_{in} , s_{jn} are respectively the two factor terms related with the concentration of the component n ($n=1, \dots, N$) in sample i and with the spectral or signal contribution at variable (wavelength, channel) j , and e_{ij} is the residual term (part of the data not explained by the model). This set of linear equations described by equation 1 can be generalized and summarized using linear algebra data matrix notation as:

$$D(I, J) = C(I, N) \cdot S^T(N, J) + E(I, J) \quad \text{Equation 2}$$

In this equation, the set of experimental data, d_{ij} , are the elements of the data matrix, $D(I, J)$, which according to the bilinear model of Equation 2, is decomposed into the product of two factor matrices, $C(I, N)$ and $S^T(N, J)$, giving respectively the concentration and spectra profiles of the constituents (components) of the analyzed samples. Matrix $E(I, J)$, the residual matrix not explained by the bilinear model, has the same dimensions of matrix D , (I, J) .

To solve Equation 1, a number of components and initial estimates of either C or S^T should be proposed and an iterative optimization using an alternating least squares (ALS) algorithm is performed to estimate the concentration (in C), and spectra (in S^T)

profiles of the constituents (components) which better describe the analyzed data matrix, D . During the ALS optimization, different constraints like non-negativity, unimodality, closure, or local rank and selectivity can be applied to decrease the rotation ambiguities associated to the bilinear model and facilitate the convergence to chemically meaningful solutions. See references [19, 21, 24, 35] for more details of the MCR-ALS procedure.

The application of MCR-ALS can be easily extended to the simultaneous analysis of multiple data sets or data matrices coming from multiple samples or experiments. In these situations, the data can be arranged in a column-wise augmented data matrix, D_{aug} , and the MCR bilinear model can be applied as shown in the following equation (Equation2):

$$D_{aug} = \begin{pmatrix} D_1 \\ D_2 \\ D_3 \\ \dots \\ D_K \end{pmatrix} = \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ \dots \\ C_K \end{pmatrix} S^T + \begin{pmatrix} E_1 \\ E_2 \\ E_3 \\ \dots \\ E_K \end{pmatrix} = C_{aug} S^T + E_{aug} \quad \text{Equation 3}$$

where D_{aug} is the augmented data matrix with the dimension of $(K \times J, J)$ and S^T is the spectra matrix of the different resolved components which are assumed to be the same in the different data sets analyzed, D_k , $k=1, \dots, K$. C_{aug} ($K \times J, N$) is the augmented concentration matrix giving the concentration profiles of every component in every individual concentration matrix C_k ($k = 1, 2, \dots, K$), which for every resolved component can be different (in position and shape) in every C_k . [24, 31, 36]. In the case of EEM data, every data matrix D_k , has the EEM fluorescent spectra measured in a sample, and C_k will have the excitation spectra of the resolved components in sample k , and S^T the emission spectra of the resolved components common for all samples. In the case of LC-MS data, every data matrix D_k has the spectra measured at every retention time and chromatograms at every m/z value for every sample k . C_k has the elution (concentration) profiles of the resolved components in the sample k , and S^T the common mass spectra of them.

One important case of multiple data sets analysis (multiset data) is when the data can be structured in more than two ways (multiway data). In this case, the decomposition of a

data set arranged for instance in a three-way (cube) data set can be performed using a trilinear model (also called parallel factor analysis or PARAFAC model))[25], as given in an element-wise way by Equation 4 as:

$$x_{ijk} = \sum_{n=1}^N c_{in} s_{jn} z_{kn} + e_{ijk} \quad \text{Equation 4}$$

where x_{ijk} represents the ijk^{th} element in the three-way data set ($i=1,\dots,I$, $j=1,\dots,J$, and $k=1,\dots,K$), N is the number of components (chemical rank) common to the three modes ($n=1,\dots,N$), c_{in} , s_{jn} and z_{kn} are the elements of C , S^T , and Z factor matrices (loadings or component profiles in the three modes) and e_{ijk} is the residual term (part of the data not explained by the model).

For each slice of the data cube Equation 4 is described as follows (Equation 5):

$$\text{Equation 5}$$

As in Equation 1, C is the matrix of loadings in the first mode or scores (concentration matrix in many chemistry examples) and S^T is the matrix of loadings in the second mode (spectra matrix in many chemistry examples). The new factor matrix of the trilinear model, is a diagonal matrix giving the relative amounts of every component in each considered data matrix. In fact, in trilinear model, column profiles, C , and row profiles, S^T , are normalized and all scaling information about components are described by D_k . The set of D_k diagonal matrices can be arranged in the matrix Z of loadings in the third mode. The important aspect to remark here about the trilinear model is that the set of D_k matrices are decomposed with the same common C and S^T matrices and that they only differ in the D_k matrices, in the relative amounts of the components described by their diagonal elements. Whenever a unique concentration and spectrum profile for each individual component in all simultaneously analyzed matrices (is appropriate for the description of the observed changes, the trilinear model is appropriate for the simultaneous decomposition of the whole data set. Trilinear model enforces decompositions to give unique solutions for the three factor matrices (apart from scale and trivial permutation rotation ambiguities) and avoids the presence of rotational ambiguities associated with bilinear models [37]. The implementation of the trilinearity constraint and its application to different data systems has been described in detail elsewhere [38, 39]. When MCR-ALS with the trilinearity constraint or PARAFAC are

applied to the same system, obtained results will be practically the same for both algorithms. However, MCR-ALS with the trilinearity constraint has the possibility to apply the trilinearity constraint independently and optionally to each component of the data set. Therefore, it can be used for full trilinear and for only partially trilinear models, giving more flexibility to the whole data analysis.

In previous works in which MCR-ALS was applied to second-order data for quantitative analysis, a post-processing calibration step was performed [35]. In the post-processing calibration step first, the peak areas or heights of the resolved concentration profiles of the analytes in the calibration samples were regressed against their known analyte concentrations. Then, the obtained calibration parameters (offset and slope of the calibration equation) were used for the estimation of the concentration of the analytes in the unknown samples. However, this post-processing calibration step did not guarantee the optimal recovery of the quantitative information. Therefore, this procedure has been recently proposed to be implemented as a new constraint (area correlation constraint) during the ALS optimization. In this way, the results resulted to be optimal from a least-squares criterion, and the concentration profiles of the analytes were recovered in their proper quantitative units [30–32].

Fig. 1 near here

Applying area correlation means, at each iteration step of the ALS process, the areas or heights of the concentration profiles of the analyte in every sample analyzed are regressed against the known analyte concentrations. This strategy is schematically shown in Fig.1 As it can be seen in this figure, the area or height of the concentration profiles of each analyte are calibrated against their known concentration in the different samples, submatrices C_k , (c) using a linear regression equation (Equation 6):

Equation 6

where a_m represents the area of the concentration profile for the considered component (analyte), in the sample i (number of calibration samples $m=1, \dots, M$). c_m is the known concentration of this component (analyte) in the calibration sample m , and b_0 and b_1 are the intercept and slope of the calibration line. Slope and intercept of this

calibration model are then used to recalculate the concentration values in the profile of this component (analyte) in the calibration samples using Equations 7 and 8

Equation 7

Equation 8

Where C_m and C_m' are the concentration profiles of the analyte in the calibration samples ($m=1,\dots,M$) after and before rescaling with the parameters of the regression line and A_m and A_m' are the areas (sums) or heights of these analyte concentration profiles, respectively.

In the samples where the concentrations of the analytes are unknown, the same parameters of the regression are used according to the following equations (Equations 9 and 10)

Equation 9

Equation 10

where A_l is the area of the resolved analyte concentration profile in the unknown samples ($l=1,\dots,L$). C_l is the concentration the analyte in the unknown sample l which is obtained using the parameters of the calibration model, b_0 and b_1 , obtained previously (b_0 and b_1).

As shown in Fig.1, this procedure is implemented at every ALS step. When the convergence is achieved, estimations of the analyte concentrations in real concentration units are obtained, at the same time as all the other ALS optimized concentration and spectral profiles of all the components in the system. In the results section, this procedure is described in more detail for the two data sets analyzed in this work, for the EEM fluorescence determination of PAHs in laboratory and river water samples and for the LC-MS determination of lipids in synthetic and cell culture samples.

3. Materials and procedures

3.1. Preparation of polyaromatic hydrocarbons (PAH) mixtures in MiliQ water and in river water samples

Table 1 near here

Stock solutions (100 mg L^{-1}) of each PAH were prepared by dissolving appropriate amounts of the PAH standards in HPLC-grade methanol. Mixtures of the six PAHs (Fluorene, Acenaphthene, Naphthalene Phenanthrene, Pyrene, and Fluoranthene) were added to miliQ water and to different river water samples. A set of 18 calibration samples at five different concentrations of these PHAs were prepared using a random design. Final concentrations of the PAHs in these samples were in the following concentration ranges: $5\text{--}25 \text{ ng mL}^{-1}$ for Fluorene, Phenanthrene and Pyrene, $10\text{--}50 \text{ ng mL}^{-1}$ for Acenaphthene and $30\text{--}150 \text{ ng mL}^{-1}$ for Naphthalene and Fluoranthene. The final concentration of each PAH at each calibration sample is given in Table 1. Six validation samples were also prepared in the same range of concentrations for each PAH. All calibration and validation samples were prepared in miliQ water.

In addition to the previously described PAH synthetic mixtures in miliQ water, different river water samples, located at different sampling sites of the Llobregat river course (one of the main rivers located in Catalonia, at the North east of Iberian Peninsula), close to its origin, in the middle course and at the end (mouth) of the river were collected and used to validate the proposed method in more complex natural water samples. It is worth mentioning here that the river samples contained amounts of dissolved organic matters (DOMs) whose natures are unknown which also produce strong EEM signals, and therefore resulted to be strong interferences for the analysis of the PAHs in these samples. River samples were filtered through a $0.45 \mu\text{m}$ membrane and spiked with two different mixtures of the six PAH. The compositions of these spiked river samples are given in the results section.

3.2. Preparation of lipid mixtures and cell culture samples

The second set of samples consisted of 12 mixtures of nine lipids from three different families of glycerols (MG, DG, and TG), phospholipids (PC, PS, and PE) and lyso phospholipids (LysoPC, LysoPS, and LysoPG) prepared firstly, in synthetic mixture samples and secondly, in cell culture (biological) samples.

1 First, the lipid synthetic mixture samples were prepared from a stock solution that
2 contained all of them at 1000 ppm, which was then diluted for obtaining four new lipid
3 calibration mixtures at 2.5, 5, 10 and 20 ppm concentrations. In each calibration set all
4 the nine lipids had the same concentrations, and at each concentration three replicates
5 were analyzed. Therefore, a set of 12 different calibration samples was prepared at these
6 four concentrations with three different replicates.

11 Secondly, different volumes of the synthetic lipid mixtures were spiked to cell culture
12 samples [40] at the concentrations of 2.5, 5, 10 and 20 mg L⁻¹, with three replicates at each
13 concentration level. In these cell culture samples, a large number of unknown cell
14 culture sample constituents can be present and they are possible interferences in the
15 determination of the lipids. Matrix effects can also be present.

22 3.3. EEM experimental determinations of the PAH mixtures

25 Fluorescence measurements of PAH mixtures were performed using a Cary Eclipse
26 fluorescence spectrophotometer (Agilent Technologies, CA, USA). The slit widths for
27 both excitation and emission monochromators were set to 5 nm. Analyses were done in
28 a quartz cell of 1 cm pathlength at room temperature. EEM spectra of each calibration
29 sample were measured on the excitation spectral range from 230 to 300 nm (every 2
30 nm) and on the emission range from 290 to 550 nm (every 2 nm). For each sample and
31 EEM measurement, one data matrix of dimensions of 36 rows (excitation wavelengths)
32 × 241 columns (emission wavelengths) was obtained. The EEM data were preprocessed
33 to eliminate the effects of Rayleigh and Raman scatterings using the EEM filtering
34 preprocessing algorithm from the PLS toolbox (<http://www.eigenvecctor.com>) under
35 the MATLAB computer and visualization environment (The MathWorks, Inc., MA,
36 USA). In order to remove these undesirable non-linear effects, this algorithm removes
37 Raman specific bands and first and second-order Rayleigh scattering signals by
38 replacing them with interpolated values. Individual data EEM data sets were then
39 arranged in a column-wise augmented data matrix, with the emission wavelengths in the
40 common columns of the new augmented column-wise EEM data matrix.

56 Table 2 near here

In Table 2, the correlation coefficients between emission (lower triangle, in black) and excitation (upper triangle, in red italic) spectra of the six PAH are given respectively. The correlation coefficient between the emission spectra of naphthalene with Acenaphthene is 0.98, which means that they are extremely similar and their estimation will be more difficult, especially in the case of MCR-ALS with only non-negativity constraints.

3.4. LC-MS analysis of lipid mixtures

LC-MS analysis of lipid mixtures used a Waters Acquity UHPLC system connected to a Waters LCT Premier orthogonal accelerated time of flight (ToF) mass spectrometer (Waters), operated in positive and negative electrospray ionization modes (ESI + and ESI -). Full scan spectra between 50 and 1500 Da were acquired, and individual spectra were summed to produce data points every 0.2 s. Mass accuracy and reproducibility were maintained by using an independent reference spray via the LockSpray interference. The analytical column was a 100 × 2.1 mm inner diameter, 1.7 mm C8 Acquity UPLC bridged ethylene hybrid (Waters). The two mobile phases were, phase A: MeOH, 1mM ammonium formate and phase B: H₂O, 2mM ammonium formate, 0.2% formic acid for both solutions. The flow rate was 0.3 ml min⁻¹ and the gradient of A/B solvents started at 80:20 and changed to 99:1 in 6 min until min 15; remained 99:1 until minute 18; finally returned to the initial conditions until minute 20. The column was held at 30 °C.

In the LC-MS analysis of the lipid mixtures, a total number of 512 m/z values were selected after the application of the regions of interest using the ROIMCR procedure described in [40, 41]. The dimensions of the analyzed LC-MS data matrices for each calibration sample were 638 rows (elution times) × 512 columns (m/z values), and the augmented data matrix had the dimensions of 7656 × 512, for the 12 calibration data matrices. A total number of 30 was proposed and tested for this augmented data matrix and satisfactory results were obtained since all considered lipids were found and identified. The number of components proposed for MCR-ALS resolution is usually

higher than the number of constituents in the sample because there are many other features in the data apart from the sought ones, either from the solvent, background or instrumental.

In the case of the LC-MS analysis of the cell culture samples, a total number of 712 m/z ROI values were selected. The final MSROI augmented data matrix had 7356 rows and 712 (m/z ROI) columns, also for the 12 cell culture data matrices simultaneously analyzed. In this case, as a multiple sample interferences and features do exist, we increased the number of components and finally 96 components were proposed and all 9 considered lipids were properly resolved and identified. This number of components is high enough to include all the considered lipids and other data features and contributions from the LC-MS instrument, background and solvent. A first data set was used to build a calibration model with all three replicates of the cell culture samples at the four different lipid concentrations. A second data set was used to test the calibration model for the prediction of samples not considered in the calibration set. In this case, the lipid concentrations in two of the sample replicates were used to build the calibration curves and the concentrations of the lipids in the third sample replicate were used to validate the calibration model.

3.5. Software

All the calculations and programs were performed using the MATLAB computer and visualization environment. MCR-ALS toolbox with a graphical user-friendly (<http://www.mcrals.info/>) [21, 42] and the PLS-toolbox (<http://www.eigenvector.com>), both for MATLAB (The MathWorks, Inc., MA, USA) were used.

4. Results and discussion

Results obtained in the application of the area correlation constraint to the two previously described experimental data sets, the determination of PAH mixtures by EEM and the determination of lipid mixtures by LC-MS, are given below separately. These results are also compared with those obtained using other data analysis strategies.

4.1. Determination of PAH mixtures in miliQ and river waters by EEM

1 First, the results of the EEM determination of six PAHs in miliQ and in river water
2 samples using four different implementations of the MCR-ALS method have been
3 compared in this case: a) MCR-ALS with only non-negativity constraints (nn); b)
4 MCR-ALS with non-negativity and species (component) correspondence constraints
5 (nn and isp); c) MCR-ALS with non-negativity and trilinearity constraints (nn and trl);
6 and d) MCR-ALS with non-negativity and area correlation constraints (nn and acc). In
7 addition, these results are compared with those obtained by the application of the
8 PARAFAC method.
9

10
11
12
13
14
15
16 Table 3 near here
17

18
19 Data fitting results obtained by application of the MCR-ALS different approaches are
20 listed in Table 3. Convergence is achieved in all cases with low lack of fit (LOF) values
21 and high explained variances (R^2 close to 100%). Generally, the lack of fit will increase
22 somewhat when more constraints are applied during the ALS optimization due to the
23 decrease in the number of degrees of freedom, for instance when the trilinearity
24 constraint is applied. If the data system is not trilinear but it is forced to behave in a
25 trilinear way, the data fitting will be considerably worse than when it is modeled
26 assuming only the bilinear model. But this is not the case for the EEM data here
27 analyzed, where the trilinear model holds and the data fitting is similar to the bilinear
28 modeling case. When other constraints like local rank/selectivity, species
29 correspondence and area correlation constraints are applied, fitting results resulted to be
30 also good. Moreover, in Table 3, the results of the application of the area correlation
31 constraint are given separately for the determination of each individual PAH
32 component. Lack of fit and explained variance values using non-negativity and area
33 correlation constraints were like those obtained using only non-negativity constraints,
34 meaning that the application of this constraint explained well the experimental data.
35
36
37
38
39
40
41
42
43
44
45
46
47
48

49 Fig. 2 near here
50

51
52 In Fig. 2, PAHs emission spectra (from 290 to 530 nm) recovered by MCR-ALS with
53 the trilinearity constraint (black line), MCR-ALS with the area correlation constraint
54 (red dotted line), and MCR-ALS with only the non-negativity constraint (blue dotted
55 line) are given. Larger differences among the recovered emission spectra using these
56 three algorithms are observed in the case of naphthalene when MCR-ALS is applied
57
58
59
60
61
62
63
64
65

with only non-negativity constraints. In contrast, the good agreement between the profiles obtained using the area correlation constraint and those obtained using MCR with the trilinearity constraint showed that the application of the area correlation constraint was working well in this case of EEM data.

Table 4 near here

In Table 4, more detailed results about the quality of the resolved profiles and about the EEM quantitative determinations of the six PAHs using MCR-ALS in its different variants and using PARAFAC are compared. Optimal recoveries of the component profiles are achieved when the slope, offset and correlation coefficients of the regression lines obtained between predicted and actual analyte concentrations are close to 1, 0 and 1 respectively. The quantitative determination of Fluorene, Phenanthrene, Pyrene, Fluoranthene, gave very good regression lines, with slopes higher than 0.95, and low offset values close to zero. Correlation coefficients of these regression lines were also high in all cases (higher than 0.97). In contrast, for Naphthalene and Acenaphthene, these parameters were not so good when only non-negativity constraints were applied, with worse slopes and offsets. Relative quantitative errors (RE% and RMEC) in the estimation of the concentrations of the PAHS using the different approaches are also given in Table 4. For Fluorene, Phenanthrene, Pyrene, and Fluoranthene, relative prediction errors in all cases were low, which confirmed that all approaches were adequate for their calibration. However, the relative quantitative errors for the other two PAHs, Naphthalene and Acenaphthene were worse, especially in the case of only applying non-negativity constraints. In this case, the severe overlapping of their emission spectra (see previously in Table 2) and the corresponding non-resolved rotation ambiguity effects, implied worse recoveries of their concentrations. This is also corroborated from the spectra pairwise correlations and angles between the MCR-ALS calculated and reference (previously known) individual spectra for each PAH (also given also in Table 4). The correlation coefficients and spectra angles obtained by PARAFAC and MCR-ALS with the trilinearity constraint were the best ones, close to 1 and 0, respectively for all PAHs. MCR-ALS using the area correlation constraint also resolved correctly the PAHs spectra with good correlation coefficients and spectra angles, close to those obtained with the MCR-ALS trilinear and PARAFAC methods. In the other cases however, results were not so good, especially when MCR-ALS was

applied only with non-negativity constraints. The relative errors, RE%, in the prediction of the concentrations using the species correspondence constraint (isp), trilinearity and area correlation constraint, were however better than when only non-negativity constraints were applied. We can, therefore, assert that the performance of the proposed area correlation constraint was accurate and reliable.

Fig. 3 near here

The applicability of the area correlation constraint in the MCR-ALS algorithm was further tested to predict the concentrations of PAHs in external validation samples (not used to build the calibration model) and in new natural river water samples. In Fig. 3(a), the contour plots of the individual 2D-EEM spectra of the different PAHs are given in the same plot. The contour plots of the 2D-EEM spectra of the different river water samples, located at different sampling sites of the river course, are given in Fig. 3 (b), (c) and (d), i.e for sampling sites 1 (close to its origin), 2 (in the middle course), and 3 (at the end of course of the river, close to its mouth). The contour plot of the 2D-EEM spectrum of the mixture of the six PAHs in miliQ water is given in Fig. 3 (e). Finally, the contour plots of the 2D-EEM spectra of river water samples 1, 2 and 3 spiked with the PAH mixture are given in Fig. 3(f), 3(g) and 3(h). As it can be observed, the 2D-EEM contour plot spectra of the six PAHs are, overlapped significantly among them, as well as with the unknown DOM fluorescent interferences from river water samples. All these river water samples had strong EEM interferences from the dissolved organic matter (DOMs) as shown in Fig. 3(b), 3(c), 3(d), 3(f), 3(g) and 3(h).

Fig. 4 near here

In Fig. 4, results of the MCR-ALS analysis of the different data sets, including the miliQ and river water samples spiked with PAHs (augmented data matrix), are given. In addition to the excitation and emission spectra of the six PAHs, four new unknown DOM interferences (see Fig.4) were resolved. The spectra of these river water interferences can be assigned to different chemical groups characteristic of DOMs, like those from humic acids, aromatic proteins, fulvic acids and proteins [43].

Table 5 near here

In Table 5, the reliability and accuracy of the results obtained by the application of area correlation constraint in MCR-ALS are further shown for the analysis of all these water samples. Predicted concentrations of the six PAHs using the area correlation constraint are compared with their theoretical concentrations (in parenthesis) in the validation PAHs mixture samples (V_1 – V_6) and in river water samples (R_1 – R_6). The agreement between true and predicted concentrations were good in all cases, with good values of MR (mean recovery), RMSEP (root mean square of prediction errors) and %RE (relative prediction error). Relative quantitative error values in the validation MiliQ water samples were lower compared to those obtained in the river water samples having strong DOM interferences not present in the calibration samples. These good recovery values obtained in general confirmed the accuracy of the proposed method (see Table 5).

As a final conclusion, the application of the area correlation constraint was giving good results in all the different cases investigated here for the EEM analysis of PAHs in both synthetic mixture water samples and river water samples with interfering dissolved organic matter (Table 5). These satisfactory results suggest that accurate quantitation can be obtained also for non-trilinear data, as for instance in the case of LC-MS data, where it is not possible to apply the trilinearity constraint in MCR-ALS nor the PARAFAC method, as it is shown and discussed in the next section.

4.2. Determination of lipid mixtures by LC-MS

MCR-ALS with the area correlation constraint has been then applied to the analysis of mixtures of lipids in synthetic and cell culture samples (see [40]), In this case, the data do not fulfill the requirements of the trilinear model and therefore this constraint cannot be applied and rotation ambiguities associated to the bilinear model data decomposition can be present in case of strong overlap in their elution and/or spectra profiles [39]. The application of the area correlation constraint is tested in these cases as a way to improve quantitative estimations and to remove rotation ambiguities [31]. The experimental details of the LC-MS analysis of these mixtures are described in the experimental section and in previous work [40]. In fact, in previous work, the data were analyzed by the ROIMCR method [40, 41] and lipids were identified. Here, in the present work, the application of area correlation constraint in MCR-ALS was tested for quantification of

all considered lipids and it was then compared with the results of applying only non-negativity. See below in Fig. 5 for more details.

Fig.5 near here

As shown in Fig 5, three different families of lipids were resolved by the MCR-ALS method. Whereas the phospholipids (PC, PS, PE) and lyso phospholipids (LysoPC, LysoPS, LysoPG) families were coeluting strongly, glycerols family (MG, DG, and TG) was well separated in the chromatographic system. As an example, the elution profiles of the nine lipid constituents in one sample replicate at one of the concentrations (20 mg L^{-1}) are given in Fig.5.

Fig. 6 near here

Calibration curves using the area correlation constraint for the three sample replicates at four different concentrations for lyso phospholipids family (LysoPE, LysoPG and LysoPC) are shown at the top of Fig.6. The MCR-ALS resolved elution profiles of these lipids in all these samples (three replicates at four different concentrations) are also given in the middle of Fig. 6, as well as their MCR-ALS, resolved mass spectra (at the bottom).

In general, data from LC-MS analysis do not conform with the trilinear model and its application is not adequate [39, 44, 45]. As it has been shown in previous works, the application of the trilinear model data decomposition to LC-MS data would imply that the different lipids would produce an elution profile in the LC system having exactly the same retention time and peak shape for all the simultaneously analyzed samples. As shown in previous works [39, 44, 45] this situation is not achieved in practice in most cases. In fact, ubiquitous coelution problems modify the shape of the elution profiles and their retention times change among different chromatographic runs. This situation is encountered also in the analysis of the lipid mixtures investigated in this work, and therefore the results of the application of PARAFAC and MCR-ALS with the trilinearity constraint are not giving here for brevity. Due to this reason, the area correlation constraint was tested in the MCR-ALS analysis of the LC-MS data to see whether a significant improvement in the quantitative determination of the lipids could be achieved or not. The quantitative results obtained for the nine lipids using both,

1 none-negativity and area correlation constraints, were then compared with those
2 obtained with the application of only non-negativity constraints.
3
4

5 Table 6 near here
6

7 All nine lipids were correctly identified in their mixtures, both in synthetic and in cell
8 culture samples. Chromatographic retention times of the lipids at the peak maxima of
9 the MCR-ALS resolved elution profiles, and m/z values at the highest intensity of their
10 corresponding mass spectrum, are given in Table 6. The quantitative results obtained in
11 their calibration and prediction are also given in Table 6. Calibration MCR-ALS results
12 with only non-negativity constraints (nn) and with non-negativity and area correlation
13 constraints (nn and acc), are compared (first and the second group of results). As
14 described in the experimental section, in the first calibration analysis, all three replicates
15 at the four different concentrations were used. The slope, offset and correlation
16 coefficient (r^2) of the regression lines comparing predicted and actual lipid
17 concentration values, and the RMSEC and the %RE (relative Error) values are given in
18 this Table 6. The quantitative determination of the well-separated glycerols family (MG,
19 DG and TG), gave similar RE% and RMSEC results for MCR-ALS either using nn and
20 acc constraints and also for MCR-ALS using only nn constraints. No improvement was
21 achieved in this case after applying the area correlation constraint. In contrast, for the
22 lyso phospholipids and phospholipids, the values of %RE and RMSEC improved
23 significantly when MCR-ALS was applied using the area correlation constraint. This
24 corroborates that the application of this constraint produces better recoveries of the
25 elution profiles of those lipids more strongly coeluted and consequently better
26 quantitative estimations of their concentrations when the area correlation constraints
27 were used than when only non-negativity constraints were applied.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

47 The applicability of the area correlation constraint was further tested for the prediction
48 of lipids in those synthetic and cell culture samples not used in the calibration step. To
49 perform this comparison, as mentioned in the experimental section, two of the sample
50 replicates were used to build the calibration curve and then, the concentration of the
51 lipids in the third sample replicate was estimated. The results obtained are given also in
52 Table 6 (third group of results). As expected, their %RE and RMSEP values were a
53 little worse than in the previous calibration case, but the predictions of the lipid
54 concentrations obtained now were still reasonably good for both type of samples,
55
56
57
58
59
60
61
62
63
64
65

1 synthetic mixtures and cell culture samples. All these results confirmed again that the
2 area correlation constraint in MCR-ALS was working reasonably well also for non-
3 trilinear data too, both in calibration and in prediction. This is especially important for
4 those data sets where the trilinearity constraint cannot be applied. The application of the
5 area correlation constraint in these cases will assure better recoveries of the
6 concentration profiles and improved quantitative predictions.
7
8
9

10 11 12 5. Conclusions

13
14 In this work, the MCR-ALS method with the recently introduced area correlation
15 constraint is proposed to improve the quantitative determination of the constituents of
16 unresolved complex mixtures with highly overlapped responses in the analysis of two
17 different data sets. In the first data set, a mixture of PAHS was analyzed using EEM
18 fluorescence giving highly overlapped emission and excitation spectra. Results
19 confirmed that the area correlation constraint produced similar results to those obtained
20 with methods based on the application of the trilinear model, like PARAFAC or MCR-
21 ALS with the trilinearity constraint. Either by application of the area correlation
22 constraint or by application of the trilinear model, the quantitative results improved
23 compared with those obtained using only non-negativity constraints. In the second data
24 set, mixtures of lipids in their synthetic mixture and in cell cultures samples were
25 analyzed by LC-MS. Although in this case the experimental data were non-trilinear and
26 there was no possibility to apply the trilinearity constraint, MCR-ALS with the area
27 correlation constraint still provided satisfactory quantitative results even for the cases in
28 which the lipids were strongly coeluting. As a final conclusion, the proposed application
29 of the MCR-ALS method with the area correlation constraint is confirmed to be a
30 powerful tool for improved recovery of the quantitative information of the constituents
31 in complex unknown unresolved mixtures in general. The application of the area
32 correlation constraint allows for the direct recovery of the quantitative information of
33 the mixture constituents in concentration units, and also reduces the extension of the
34 rotation ambiguities associated with the solutions obtained by MCR-ALS when only
35 non-negativity constraints are used.
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58

59 Acknowledgments

60
61
62
63
64
65

Research grants from the Spanish Ministry of Science, Innovation, and Universities (MICINN) and the Government of Catalonia for projects CTQ2015-66254-C2-1-P and 2017-SGR-753, respectively, are acknowledged.

References:

- [1] M.C. Antunes, J.E. Simao, A.C. Duarteb, R. Tauler, Multivariate curve resolution of overlapping voltammetric peaks: quantitative analysis of binary and quaternary metal mixtures, *The Analyst*, 127 (2002) 809-817.
- [2] A. de Juan, R. Tauler, Chemometrics applied to unravel multicomponent processes and mixtures: Revisiting latest trends in multivariate resolution, *Anal. Chim. Acta*, 500 (2003) 195-210.
- [3] A. Malik, A. de Juan, R. Tauler, Multivariate Curve Resolution: A Different Way To Examine Chemical Data, 40 Years of Chemometrics – From Bruce Kowalski to the Future, *American Chem. Soc.* (2015) 95-128.
- [4] S. Mas, A. de Juan, R. Tauler, A.C. Olivieri, G.M. Escandar, Application of chemometric methods to environmental analysis of organic pollutants: A review, *Talanta*, 80 (2010) 1052-1067.
- [5] A.K. Smilde, R. Tauler, J. Saurina, R. Bro, Calibration methods for complex second-order data, *Anal. Chim. Acta*, 398 (1999) 237-251.
- [6] A.C. Olivieri, G.M. Escandar, A.M.d.I. Peña, Second-order and higher-order multivariate calibration methods applied to non-multilinear data using different algorithms, *TrAC Trends Anal. Chem.*, 30 (2011) 607-617.
- [7] J.A. Arancibia, A.C. Olivieri, G.M. Escandar, First- and second-order multivariate calibration applied to biological samples: determination of anti-inflammatories in serum and urine, *Anal. Bioanal. Chem.*, 374 (2002) 451-459.
- [8] R. Tauler, B. Kowalski, S. Fleming, Multivariate curve resolution applied to spectral data from multiple runs of an industrial process, *Anal. Chem.*, 65 (1993) 2040-2047.
- [9] H.-L. Wu, Y. Li, R.-Q. Yu, Recent developments of chemical multiway calibration methodologies with second-order or higher-order advantages, *J. Chemometrics*, 28 (2014) 476-489.
- [10] Y.-Z. Liang, H.-L. Wu, G. Shen, J. Jiang, S. Liang, R. Yu, Aspects of recent developments in analytical chemometrics, *Sci. China Series B Chem.*, 49 (2006) 193-203.
- [11] G.M. Escandar, H.C. Goicoechea, A. Muñoz de la Peña, A.C. Olivieri, Second- and higher-order data generation and calibration: A tutorial, *Anal. Chim. Acta*, 806 (2014) 8-26.
- [12] M.C. Bauza, G.A. Ibanez, 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.
- [13] A.C. Olivieri, Practical guidelines for reporting results in single- and multi-component analytical calibration: A tutorial, *Anal. Chim. Acta*, 868 (2015) 10-22.
- [14] C. Ruckebusch, L. Blanchet, Multivariate curve resolution: a review of advanced and tailored applications and challenges, *Anal. Chim. Acta*, 765 (2013) 28-36.

- [15] A. de Juan, J. Jaumot, R. Tauler, Multivariate Curve Resolution (MCR). Solving the mixture analysis problem, *Anal. Methods*, 6 (2014) 4964–4976.
- [16] A. de Juan, R. Tauler, Multivariate Curve Resolution (MCR) from 2000: Progress in Concepts and Applications, *Critical Rev. Anal. Chem.*, 36 (2006) 163–176.
- [17] R. Tauler, Multivariate curve resolution applied to second order data, *Chemom. Intell. Lab. Sys.*, 30 (1995) 133–146.
- [18] A. Malik, R. Tauler, Performance and Validation of MCR-ALS with Quadrilinear Constraint in the Analysis of Noisy Datasets, *Chemom. Intell. Lab. Sys.* 135 (2014) 223–234.
- [19] M. Alier, M. Felipe, I. Hernandez, R. Tauler, Trilinearity and component interaction constraints in the multivariate curve resolution investigation of NO and O3 pollution in Barcelona, *Anal. Bioanal. Chem.*, 399 (2011) 2015–2029.
- [20] R. Tauler, M. Maeder, A. de Juan, 2.24 –Multiset Data Analysis: Extended Multivariate Curve Resolution, in: S.D. Brown, R. Tauler, B. Walczak (Eds.) *Comprehensive Chemom.*, Elsevier, Oxford, (2009) 473–505.
- [21] J. Jaumot, R. Gargallo, A. de Juan, R. Tauler, A graphical user-friendly interface for MCR-ALS: a new tool for multivariate curve resolution in MATLAB, *Chemom. Intell. Lab. Sys.*, 76 (2005) 101–110.
- [22] R. Tauler, A. Smilde, B. Kowalski, Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution, *J. Chemom.*, 9 (1995) 31–58
- [23] R. Leardi, Multi-way analysis with applications in the chemical sciences, age smilde, Rasmus Bro and Paul Geladi, *J. Chemom.*, 19 (2005) 119–120.
- [24] R. Tauler, A. De Juan, Multivariate Curve Resolution for Quantitative Analysis, *Data Handling Sci.Tech.*, (2015) 247–292.
- [25] R. Bro, PARAFAC. Tutorial and applications, *Chemom. Intell. Lab. Sys.*, 38 (1997) 149–171.
- [26] G. Ahmadi, R. Tauler, H. Abdollahi, Multivariate calibration of first-order data with the correlation constrained MCR-ALS method, *Chemom. Intell. Lab. Sys.*, 142 (2015) 143–150.
- [27] S. Richards, E. Becker, R. Tauler, A. D Walmsley, A novel approach to the quantification of industrial mixtures from the Vinyl Acetate Monomer (VAM) process using Near Infrared spectroscopic data and a Quantitative Self Modeling Curve Resolution (SMCR) methodology, *Chemom. Intell. Lab. Sys.*, 94 (2008) 9–18.
- [28] T. Azzouz, R. Tauler, Application of multivariate curve resolution alternating least squares (MCR-ALS) to the quantitative analysis of pharmaceutical and agricultural samples, *Talanta*, 74 (2008) 1201–1210.
- [29] R.R. de Oliveira, K.M.G. de Lima, R. Tauler, A. de Juan, Application of correlation constrained multivariate curve resolution alternating least-squares methods for

determination of compounds of interest in biodiesel blends using NIR and UV–visible spectroscopic data, *Talanta*, 125 (2014) 233–241.

- [30] A.C. Neves, R. Tauler, K.M. de Lima, Area correlation constraint for the MCR-ALS quantification of cholesterol using EEM fluorescence data: A new approach, *Anal. Chim. Acta*, 937 (2016) 21–28.
- [31] M. Ghaffari, H. Abdollahi, A conceptual view to the area correlation constraint in multivariate curve resolution, *Chemom. Intell. Lab. Sys.*, 189 (2019) 121–129.
- [32] M. Ghaffari, A.C. Olivieri, H. Abdollahi, Strategy To Obtain Accurate Analytical Solutions in Second-Order Multivariate Calibration with Curve Resolution Methods, *Anal. Chem.*, 90 (2018) 9725–9733.
- [33] M. Esteban, C. Ariño, J.M. **Díaz-Cruz**, **M.S. Díaz-Cruz**, R. Tauler, Multivariate curve resolution with alternating least squares optimisation: a soft-modelling approach to metal complexation studies by voltammetric techniques, *TrAC Trends Anal. Chem.*, 19 (2000) 49–61.
- [34] N.E. Llamas, M. Garrido, M.S.D. Nezio, B.S.F. Band, Second order advantage in the determination of amaranth, sunset yellow FCF and tartrazine by UV–vis and multivariate curve resolution–alternating least squares, *Anal. Chim. Acta*, 655 (2009) 38–42.
- [35] N. Omidikia, S. Beyramysoltan, J. Mohammad Jafari, E. Tavakkoli, M. Akbari Lakeh, M. Alinaghi, M. Ghaffari, S. Khodadadi Karimvand, R. Rajkó, H. Abdollahi, Closure constraint in multivariate curve resolution, *J. Chemom.*, 32 (2018) 1–15.
- [36] E. Peré-Trepát, S. Lacorte, R. Tauler, Alternative calibration approaches for LC–MS quantitative determination of coeluted compounds in complex environmental mixtures using multivariate curve resolution, *Anal. Chim. Acta*, 595 (2007) 228–237.
- [37] R. Tauler, Calculation of maximum and minimum band boundaries of feasible solutions for species profiles obtained by multivariate curve resolution, *J. Chemom.*, 15 (2001) 627–646.
- [38] R. Tauler, I. Marqués, E. Casassas, Multivariate curve resolution applied to three-way trilinear data: Study of a spectrofluorimetric acid–base titration of salicylic acid at three excitation wavelengths, *J. Chemom.*, 12 (1998) 55–75.
- [39] A. de Juan, R. Tauler, Comparison of three-way resolution methods for non-trilinear chemical data sets, *J. Chemom.*, 15 (2001) 749–771.
- [40] N. Dalmau, C. Bedia, R. Tauler, Validation of the Regions of Interest Multivariate Curve Resolution (ROIMCR) procedure for untargeted LC–MS lipidomic analysis, *Anal. Chim. Acta*, 1025 (2018) 80–91.
- [41] E. Gorrochategui, J. Jaumot, R. Tauler, ROIMCR: a powerful analysis strategy for LC–MS metabolomic datasets, *BMC bioinformatics*, 20 (2019) 256.
- [42] J. Jaumot, A. de Juan, R. Tauler, MCR-ALS GUI 2.0: New features and applications, *Chemom. Intell. Lab. Sys.*, 140 (2015) 1–12.

- 1 [43] W. Chen, P. Westerhoff, J.A. Leenheer, K. Booksh, Fluorescence Excitation–Emission
2 Matrix Regional Integration to Quantify Spectra for Dissolved Organic Matter, Environ.
3 Sci. Technol., 37 (2003) 5701–5710.
- 4 [44] S.A. Bortolato, A.C. Olivieri, Chemometric processing of second-order liquid
5 chromatographic data with UV–vis and fluorescence detection. A comparison of
6 multivariate curve resolution and parallel factor analysis 2, Anal. Chim. Acta, 842 (2014)
7 11–19.
- 8 [45] M.B. Anzardi, J.A. Arancibia, A.C. Olivieri, Interpretation of matrix chromatographic–
9 spectral data modeling with parallel factor analysis 2 and multivariate curve resolution, J.
10 Chromatogr. A, 1604 (2019) 460502.
- 11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Table 1. Composition of PAH sample mixtures in ng mL ⁻¹						
Calibration samples	Fluorene	Acenaphthene	Naphthalene	Phenanthrene	Pyrene	Fluoranthene
C ₁	5.00	30.0	120	10.0	0.00	60.0
C ₂	20.0	0.00	150	20.0	25.0	150
C ₃	0.00	50.0	60.0	25.0	25.0	120
C ₄	25.0	40.0	30.0	15.0	0.00	30.0
C ₅	15.0	20.0	90.0	5.00	20.0	0.00
C ₆	15.0	50.0	0.00	10.0	5.00	60.0
C ₇	10.0	0.00	120	15.0	10.0	90.0
C ₈	10.0	10.0	0.00	5.00	15.0	30.0
C ₉	5.00	30.0	60.0	0.00	10.0	150
C ₁₀	25.0	30.0	90.0	25.0	5.00	0.00
C ₁₁	0.00	40.0	150	20.0	10.0	150
C ₁₂	25.0	10.0	60.0	0.00	15.0	30.0
C ₁₃	10.0	0.00	150	15.0	20.0	30.0
C ₁₄	20.0	40.0	120	5.00	5.00	0.00
C ₁₅	0.00	20.0	90.0	10.0	25.0	60.0
C ₁₆	15.0	50.0	0.00	20.0	15.0	120
C ₁₇	20.0	10.0	30.0	25.0	0.00	90.0
C ₁₈	5.00	20.0	30.0	0.00	20.0	90.0

Table 2. Pairwise correlations (similarities) between the emission and the <i>excitation (in cursive) spectra</i> of the different PAHs investigated in this work						
	Fluorene	Acenaphthene	Naphthalene	Phenanthrene	Pyrene	Fluoranthene
Fluorene	1.00	-0.42	0.46	0.43	-0.06	-0.51
Acenaphthene	0.32	1.00	0.14	-0.36	0.21	0.24
Naphthalene	0.41	0.98	1.00	-0.27	-0.22	-0.10
Phenanthrene	-0.20	0.19	0.14	1.00	0.13	-0.21
Pyrene	-0.26	-0.19	-0.21	0.43	1.00	0.58
Fluoranthene	-0.46	-0.49	-0.51	-0.49	-0.33	1.00

Table 4. Results about the quality of the resolved profiles of six PAHs using MCR-ALS in its different variants and PARAFAC						
PAH list		MCR (nn ^a)	MCR (nn and isp ^b)	MCR (nn and trl ^c)	PARAFAC	MCR (nn and acc ^d)
Fluorene	Slope	0.957	0.957	0.958	0.958	1.01
	offset	0.577	0.576	0.564	0.559	-0.231
	Line r ²	0.978	0.978	0.971	0.979	0.979
	Spectra angle	1.09	4.83	1.64	1.58	1.12
	Spectra r ²	0.999	0.996	0.999	0.999	0.999
	^e Rel. Error (%)	10.9	10.9	10.8	10.8	10.9
	^f RMSEC	1.74	1.74	1.72	1.71	1.74
Acenaphthene	Slope	0.880	0.999	0.972	0.972	0.998
	offset	3.13	-0.016	0.739	0.745	0.125
	Line r ²	0.938	0.9811	0.986	0.986	0.987
	Spectra angle	0.881	1.29	0.831	0.860	0.699
	Spectra r ²	0.999	0.999	0.999	0.999	0.999
	Rel. Error (%)	19.45	11.1	9.44	9.49	9.21
	RMSEC	6.15	3.51	2.98	3.00	2.91
Naphthalene	Slope	0.870	0.959	0.968	0.989	0.955
	offset	9.28	2.94	2.24	0.805	4.58
	Line r ²	0.938	0.979	0.984	0.994	0.975
	Spectra angle	4.91	5.21	3.83	3.08	3.52
	Spectra r ²	0.996	0.996	0.998	0.999	0.998
	Rel. Error (%)	20.8	11.7	10.2	9.53	12.1
	RMSEC	18.3	10.3	8.96	5.37	9.99
Phenanthrene	Slope	0.964	0.976	0.963	0.962	0.999
	offset	0.437	0.298	0.451	0.470	-0.001
	Line r ²	0.964	0.988	0.981	0.980	0.982
	Spectra angle	3.59	2.99	3.26	3.35	3.09
	Spectra r ²	0.998	0.998	0.998	0.998	0.998
	Rel. Error (%)	10.5	8.66	10.6	10.8	10.7
	RMSEC	1.55	1.29	1.57	1.65	1.58
Pyrene	Slope	0.973	0.982	0.977	0.980	0.999
	offset	0.371	0.242	0.301	0.268	0.001
	Line r ²	0.973	0.991	0.989	0.990	0.982
	Spectra angle	1.51	3.24	1.61	1.59	1.82
	Spectra r ²	0.999	0.991	0.999	0.999	0.998
	Rel. Error (%)	9.69	7.83	8.73	8.25	9.44
	RMSEC	1.60	1.29	1.44	1.37	1.56
Fluoranthene	Slope	0.990	0.990	0.990	0.990	0.980
	offset	0.639	0.648	0.638	0.640	1.89
	Line r ²	0.990	0.995	0.990	0.995	0.995
	Spectra angle	1.33	2.83	1.46	1.36	1.19
	Spectra r ²	0.999	0.999	0.999	0.999	0.999
	Rel. Error (%)	5.71	5.75	5.70	5.72	5.87
	RMSEC	4.51	4.54	4.51	4.51	4.64
^a non-negativity constraints						
^b corespondance of species constraints						
^c trilinearity constraints						
^d area correlation constraints						
^e Relative Error of calibration, RE (%) = $\frac{\sum_{i=1}^n c_{i,est} - c_{i,known} }{\sum_{i=1}^n c_{i,known}}$ where $c_{i,est}$ and $c_{i,known}$ are the MCR-ALS estimated and known concentration in calibration samples, respectively.						
^f Root Mean Square Error of Calibration, RMSEC= $\sqrt{\frac{\sum_{i=1}^n (c_{i,est} - c_{i,known})^2}{n}}$ where $c_{i,est}$ and $c_{i,known}$ are the MCR-ALS estimated and known concentration in calibration samples, respectively.						

Table 3. MCR-ALS fitting results			
Data set 1: EEM determination of PAH mixtures			
Methods	Analytes	lof ^a	R ² , ^b
MCR-ALS (nn) ^c	All	2.09	99.96
MCR-ALS (nn and isp) ^d	All	3.09	99.91
MCR-ALS (nn and trl) ^e	All	3.99	99.85
MCR-ALS (nn and acc) ^f	Fluorene	2.23	99.95
	Acenaphthene	2.15	99.95
	Naphtalene	2.51	99.93
	Phenanthrene	2.09	99.96
	Pyrene	2.09	99.96
	Fluoranthene	2.31	99.95
Data set 2: Chromatographic determination of the lipids in their sybthetic mixtures			
Methods	Analyte	lof ^a	R ²
MCR-ALS (nn) ^c	All	6.21	99.61
MCR-ALS (nn and acc) ^f	LysoPE	6.12	99.63
	LysoPG	6.14	99.62
	LysoPC	6.16	99.62
	MG	6.11	99.63
	PS	6.11	99.63
	PE	6.15	99.62
	PC	6.17	99.62
	DG	6.12	99.63
	TG	6.11	99.63
^a MCR-ALS lack of fit, lof $\frac{\sum_{i=1}^n \text{res}_i^2}{\sum_{i=1}^n \text{data}_i^2}$, where res_i is an element of input data and data_i is the corresponding MCR-ALS residual (not explained part).			
^b R2 MCR-ALS explained variances, $\frac{\sum_{i=1}^n \text{data}_i^2 - \sum_{i=1}^n \text{res}_i^2}{\sum_{i=1}^n \text{data}_i^2}$, where data_i is an element of input experimental data and res_i is the corresponding MCR-ALS residual(not explained part)			
^c MCR-ALS with non-negativity constraints			
^d MCR-ALS with non-negativity and correspondence of species constraints			
^e MCR-ALS with non-negativity and trilinearity constraints			
^f MCR-ALS with non-negativity and area correlation constraints			

Table 4. Results about the quality of the resolved profiles of six PAHs using MCR-ALS in its different variants and PARAFAC						
PAH list		MCR (nn ^a)	MCR (nn and isp ^b)	MCR (nn and trl ^c)	PARAFAC	MCR (nn and acc ^d)
Fluorene	Slope	0.957	0.957	0.958	0.958	1.01
	offset	0.577	0.576	0.564	0.559	-0.231
	Line r ²	0.978	0.978	0.971	0.979	0.979
	Spectra angle	1.09	4.83	1.64	1.58	1.12
	Spectra r ²	0.999	0.996	0.999	0.999	0.999
	^e Rel. Error (%)	10.9	10.9	10.8	10.8	10.9
	^f RMSEC	1.74	1.74	1.72	1.71	1.74
Acenaphthene	Slope	0.880	0.999	0.972	0.972	0.998
	offset	3.13	-0.016	0.739	0.745	0.125
	Line r ²	0.938	0.9811	0.986	0.986	0.987
	Spectra angle	0.881	1.29	0.831	0.860	0.699
	Spectra r ²	0.999	0.999	0.999	0.999	0.999
	Rel. Error (%)	19.45	11.1	9.44	9.49	9.21
	RMSEC	6.15	3.51	2.98	3.00	2.91
Naphthalene	Slope	0.870	0.959	0.968	0.989	0.955
	offset	9.28	2.94	2.24	0.805	4.58
	Line r ²	0.938	0.979	0.984	0.994	0.975
	Spectra angle	4.91	5.21	3.83	3.08	3.52
	Spectra r ²	0.996	0.996	0.998	0.999	0.998
	Rel. Error (%)	20.8	11.7	10.2	9.53	12.1
	RMSEC	18.3	10.3	8.96	5.37	9.99
Phenanthrene	Slope	0.964	0.976	0.963	0.962	0.999
	offset	0.437	0.298	0.451	0.470	-0.001
	Line r ²	0.964	0.988	0.981	0.980	0.982
	Spectra angle	3.59	2.99	3.26	3.35	3.09
	Spectra r ²	0.998	0.998	0.998	0.998	0.998
	Rel. Error (%)	10.5	8.66	10.6	10.8	10.7
	RMSEC	1.55	1.29	1.57	1.65	1.58
Pyrene	Slope	0.973	0.982	0.977	0.980	0.999
	offset	0.371	0.242	0.301	0.268	0.001
	Line r ²	0.973	0.991	0.989	0.990	0.982
	Spectra angle	1.51	3.24	1.61	1.59	1.82
	Spectra r ²	0.999	0.991	0.999	0.999	0.998
	Rel. Error (%)	9.69	7.83	8.73	8.25	9.44
	RMSEC	1.60	1.29	1.44	1.37	1.56
Fluoranthene	Slope	0.990	0.990	0.990	0.990	0.980
	offset	0.639	0.648	0.638	0.640	1.89
	Line r ²	0.990	0.995	0.990	0.995	0.995
	Spectra angle	1.33	2.83	1.46	1.36	1.19
	Spectra r ²	0.999	0.999	0.999	0.999	0.999
	Rel. Error (%)	5.71	5.75	5.70	5.72	5.87
	RMSEC	4.51	4.54	4.51	4.51	4.64
^a non-negativity constraints						
^b corespondance of species constraints						
^c trilinearity constraints						
^d area correlation constraints						
^e Relative Error of calibration, RE (%) = $\frac{\sum_{i=1}^n c_{i,est} - c_{i,known} }{\sum_{i=1}^n c_{i,known}}$ where $c_{i,est}$ and $c_{i,known}$ are the MCR-ALS estimated and known concentration in calibration samples, respectively.						
^f Root Mean Square Error of Calibration, RMSEC= $\sqrt{\frac{\sum_{i=1}^n (c_{i,est} - c_{i,known})^2}{n}}$ where $c_{i,est}$ and $c_{i,known}$ are the MCR-ALS estimated and known concentration in calibration samples, respectively.						

Table 5. Prediction of the concentration of PAHs in the validation (V₁-V₆) miliQ water samples and in the river water samples (R₁-R₆) using MCR-ALS with the area correlation constraint

Mixture	PAH					
	Fluorene	Acenaphthene	Naphthalene	Phenanthrene	Pyrene	Fluoranthene
V ₁	14.9 (14.0) ^a	46.7 (45.0)	71.7 (70.0)	12.2 (12.0)	16.4 (15.0)	94.4 (100)
V ₂	17.4 (16.0)	27.1 (30.0)	87.6 (80.0)	17.7 (18.0)	31.7 (35.0)	63.7 (60.0)
V ₃	5.59 (6.00)	21.9 (20.0)	114 (100)	5.34 (6.00)	26.1 (30.0)	85.9 (90.0)
V ₄	19.4 (18.0)	3.96 (5.00)	137 (130)	11.3 (10.0)	18.3 (20.0)	55.9 (50.0)
V ₅	10.2 (12.0)	20.4 (15.0)	49.4 (50.0)	13.0 (12.0)	20.4 (20.0)	68.4 (70.0)
V ₆	9.35 (8.00)	38.5 (40.0)	129 (110)	20.1 (18.0)	11.6 (10.0)	123 (130)
Rel error ^b	10.1	9.56	11.5	8.41	10.2	5.70
RMSEP ^c	1.31	2.81	10.8	1.15	2.37	4.99
Mean Recovery	103	102	108	101	104	100
R ₁ -water A ^d	9.72 (12.0)	14.1 (15.0)	43.2 (50.0)	10.4 (12.0)	17.1 (20.0)	67.5 (70.0)
R ₂ -water A ^e	5.08 (6.00)	16.4 (20.0)	87.4 (100)	5.40 (6.00)	29.9 (30.0)	99.0 (100)
R ₃ -water B ^f	11.9 (12.0)	14.8 (15.0)	49.4 (50.0)	11.7 (12.0)	16.6 (20.0)	70.1 (70.0)
R ₄ -water B	5.01 (6.00)	18.9 (20.0)	112 (100)	6.11 (6.00)	24.6 (30.0)	81.9 (100)
R ₅ -water C	11.1 (12.0)	12.4 (15.0)	41.3 (50.0)	10.3 (12.0)	22.4 (20.0)	76.6 (70.0)
R ₆ -water C	6.04 (6.00)	17.7 (20.0)	115 (100)	5.62 (6.00)	32.4 (30.0)	109 (100)
Rel error	12.2	12.0	13.2	12.5	12.4	10.3
RMSEP	1.15	2.12	10.4	1.18	3.15	8.82
Mean Recovery	90.3	90.1	97.1	92.6	95.2	99.4

^a Actual concentration value in parentheses

^b Relative error of prediction RE (%) = $\frac{|C_{est} - C_{known}|}{C_{known}} \times 100$ where C_{est} and C_{known} are the MCR-ALS estimated and known concentration in prediction samples, respectively

^c Root mean square error of prediction RMSEP = $\sqrt{\frac{1}{n} \sum_{i=1}^n (C_{est,i} - C_{known,i})^2}$ where $C_{est,i}$ and $C_{known,i}$ are the MCR-ALS estimated and known concentration in prediction samples, respectively.

^d R₁, R₂: river water samples from sites near the river spring spiked with a mixture of the six PAHs similar to water validation samples V₅ and V₃, respectively. (Actual concentration values are given in parentheses)

^e R₃, R₄: river water samples from sites in the middle course of the river spiked with the mixture of the six PAHs similar to water validation samples V₅ and V₃, respectively. (Actual concentration values are given in parentheses)

^f R₅, R₆: river water samples from sites near the mouth of the river spiked with the mixture of the six PAHs similar to water validation V₅ and V₃, respectively. (Actual concentration values are given in parentheses)

Table 6. MCR-ALS calibration and prediction results in the LC-MS analysis of nine lipids in their mixtures																			
			Calibration ^a of lipids in synthetic mixture samples										Prediction ^b of lipids in synthetic mixture samples						
Lipids list	m/z ROI value	rt (min)	MCR_ALS (nn ^c)						MCR_ALS (nn and acc ^d)						MCR_ALS (nn and acc ^d)				
			Slope	offset	Line r ²	RMSEC	Rel. Error (%)		Slope	offset	Line r ²	RMSEC	Rel. Error (%)		Slope	offset	Line r ²	RMSEP	Rel. Error (%)
LysoPE	466.2917	2.45	0.974	0.240	0.987	0.876	7.61		0.917	1.14	0.988	0.729	6.33		0.961	0.260	0.988	1.02	8.84
LysoPG	497.2869	2.52	0.973	0.249	0.987	1.04	9.01		1.028	-0.477	0.988	0.533	4.63		0.960	0.074	0.988	1.03	8.97
LysoPC	510.3546	2.89	1.00	0.00	0.941	1.74	15.1		0.947	0.748	0.990	0.797	6.92		0.954	0.345	0.989	0.963	8.35
PS	793.7368	8.35	0.964	0.339	0.982	1.06	9.25		0.997	0.044	0.981	0.902	7.83		0.952	0.189	0.982	1.28	11.1
PE	791.7797	8.45	0.896	0.969	0.947	2.28	19.7		0.991	0.101	0.973	1.60	13.9		0.928	0.721	0.968	2.40	20.8
PC	749.7122	8.61	0.916	0.788	0.957	2.10	17.6		1.01	-0.107	0.986	1.14	9.91		0.957	0.204	0.983	1.38	11.9
MG	362.2805	3.29	0.934	0.919	0.986	0.685	5.95		0.977	0.215	0.988	0.977	5.94		0.965	0.225	0.987	1.14	9.86
DG	619.6030	10.9	0.944	0.526	0.971	0.865	7.50		0.937	0.906	0.971	0.863	7.49		0.945	0.211	0.972	1.61	13.9
TG	866.8206	17.6	0.923	0.376	0.997	0.509	4.41		1.03	-0.727	0.997	0.505	4.38		1.06	-1.04	0.997	0.874	7.58
			Calibration ^a of lipids in cell culture samples)										Prediction ^b of lipids in cell culture samples						
			MCR_ALS (nn ^c)						MCR_ALS (nn and acc ^d)						MCR_ALS (nn and acc ^d)				
LysoPE	466.2934	2.30	1.00	-0.207	0.984	0.840	7.29		0.907	1.09	0.994	0.722	6.26		0.948	0.662	0.990	0.973	8.45
LysoPG	497.2859	2.64	0.954	0.433	0.976	1.57	13.6		0.970	0.448	0.979	1.39	12.1		0.932	0.764	0.974	1.69	14.7
LysoPC	510.3584	2.89	0.979	0.192	0.989	0.474	4.11		0.986	0.207	0.999	0.303	2.63		0.946	0.326	0.989	0.790	6.72
PS	793.7429	8.33	0.976	0.365	0.966	1.79	15.5		0.976	0.366	0.966	1.59	13.5		1.38	-3.46	0.990	2.56	21.2
PE	791.7760	8.54	0.897	0.968	0.947	1.82	15.7		1.038	-0.522	0.992	1.29	11.3		0.837	1.46	0.981	1.55	13.5
PC	749.7160	8.61	1.07	-0.233	0.977	1.44	12.6		0.983	0.157	0.991	0.876	7.59		1.22	-3.00	0.986	1.16	10.1
MG	362.3410	3.28	0.954	0.426	0.977	0.737	6.40		1.01	0.00	0.988	0.514	4.45		0.904	0.803	0.973	1.58	8.75
DG	619.6046	10.9	0.945	0.519	0.972	1.63	14.1		1.04	-0.541	0.974	1.54	13.4		0.805	0.475	0.999	1.89	16.4
TG	866.8210	17.6	1.00	-0.845	0.998	0.679	5.29		1.01	-0.845	0.998	0.677	5.27		0.975	-0.868	0.998	1.07	9.15
^a In the calibration set, the three replicates at each of the four concentration levels were used for calibration.																			
^b In the prediction set only two of the three replicates at each of the four concentration levels were used for calibration and then, the concentration of the lipids in the third replicate was predicted..																			
^c non-negativity constraints																			
^d non-negativity constraint and area correlation constraints																			

Table 6. MCR-ALS calibration and prediction results in the LC-MS analysis of nine lipids in their mixtures																			
			Calibration ^a of lipids in synthetic mixture samples										Prediction ^b of lipids in synthetic mixture samples						
Lipids list	m/z ROI value	rt (min)	MCR_ALS (nn ^c)						MCR_ALS (nn and acc ^d)						MCR_ALS (nn and acc ^d)				
			Slope	offset	Line r ²	RMSEC	Rel. Error (%)		Slope	offset	Line r ²	RMSEC	Rel. Error (%)		Slope	offset	Line r ²	RMSEP	Rel. Error (%)
LysoPE	466.2917	2.45	0.974	0.240	0.987	0.876	7.61		0.917	1.14	0.988	0.729	6.33		0.961	0.260	0.988	1.02	8.84
LysoPG	497.2869	2.52	0.973	0.249	0.987	1.04	9.01		1.028	-0.477	0.988	0.533	4.63		0.960	0.074	0.988	1.03	8.97
LysoPC	510.3546	2.89	1.00	0.00	0.941	1.74	15.1		0.947	0.748	0.990	0.797	6.92		0.954	0.345	0.989	0.963	8.35
PS	793.7368	8.35	0.964	0.339	0.982	1.06	9.25		0.997	0.044	0.981	0.902	7.83		0.952	0.189	0.982	1.28	11.1
PE	791.7797	8.45	0.896	0.969	0.947	2.28	19.7		0.991	0.101	0.973	1.60	13.9		0.928	0.721	0.968	2.40	20.8
PC	749.7122	8.61	0.916	0.788	0.957	2.10	17.6		1.01	-0.107	0.986	1.14	9.91		0.957	0.204	0.983	1.38	11.9
MG	362.2805	3.29	0.934	0.919	0.986	0.685	5.95		0.977	0.215	0.988	0.977	5.94		0.965	0.225	0.987	1.14	9.86
DG	619.6030	10.9	0.944	0.526	0.971	0.865	7.50		0.937	0.906	0.971	0.863	7.49		0.945	0.211	0.972	1.61	13.9
TG	866.8206	17.6	0.923	0.376	0.997	0.509	4.41		1.03	-0.727	0.997	0.505	4.38		1.06	-1.04	0.997	0.874	7.58
			Calibration ^a of lipids in cell culture samples)										Prediction ^b of lipids in cell culture samples						
			MCR_ALS (nn ^c)						MCR_ALS (nn and acc ^d)						MCR_ALS (nn and acc ^d)				
LysoPE	466.2934	2.30	1.00	-0.207	0.984	0.840	7.29		0.907	1.09	0.994	0.722	6.26		0.948	0.662	0.990	0.973	8.45
LysoPG	497.2859	2.64	0.954	0.433	0.976	1.57	13.6		0.970	0.448	0.979	1.39	12.1		0.932	0.764	0.974	1.69	14.7
LysoPC	510.3584	2.89	0.979	0.192	0.989	0.474	4.11		0.986	0.207	0.999	0.303	2.63		0.946	0.326	0.989	0.790	6.72
PS	793.7429	8.33	0.976	0.365	0.966	1.79	15.5		0.976	0.366	0.966	1.59	13.5		1.38	-3.46	0.990	2.56	21.2
PE	791.7760	8.54	0.897	0.968	0.947	1.82	15.7		1.038	-0.522	0.992	1.29	11.3		0.837	1.46	0.981	1.55	13.5
PC	749.7160	8.61	1.07	-0.233	0.977	1.44	12.6		0.983	0.157	0.991	0.876	7.59		1.22	-3.00	0.986	1.16	10.1
MG	362.3410	3.28	0.954	0.426	0.977	0.737	6.40		1.01	0.00	0.988	0.514	4.45		0.904	0.803	0.973	1.58	8.75
DG	619.6046	10.9	0.945	0.519	0.972	1.63	14.1		1.04	-0.541	0.974	1.54	13.4		0.805	0.475	0.999	1.89	16.4
TG	866.8210	17.6	1.00	-0.845	0.998	0.679	5.29		1.01	-0.845	0.998	0.677	5.27		0.975	-0.868	0.998	1.07	9.15
^a In the calibration set, the three replicates at each of the four concentration levels were used for calibration.																			
^b In the prediction set only two of the three replicates at each of the four concentration levels were used for calibration and then, the concentration of the lipids in the third replicate was predicted..																			
^c non-negativity constraints																			
^d non-negativity constraint and area correlation constraints																			

Figure Captions:

Fig.1 Schematic implementation of the area correlation constraint in the MCR-ALS analysis of multiple data matrices using a bilinear decomposition model (see Part 2 Theory).

Fig.2 MCR-ALS resolved emission spectra of the six individual PAHs in their mixtures in miliQ water using non-negativity (blue dotted line), non-negativity and trilinearity (black line) and non-negativity and area correlation constraints (red dotted line).

Fig.3 Contour plots of the experimental 2D-EEM spectra of (a) individual PAHs in miliQ water; (b) river water sample 1; (c) river water sample 2; (d) river water sample 3; (e) the mixture of PAHs in miliQ water; (f) the mixture of PAHs in river water sample 1; (g) the mixture of PAHs in river water sample 2; and (h) mixture of PAHs in river water sample 3.

Fig.4 MCR-ALS resolved excitation (a) and emission (b) spectra of the six PAH and of the six DOM interferences detected in the river water samples.

Fig.5 MCR-ALS resolved elution profiles of one of the sample replicates of the mixture of nine lipids analyzed by LC-MS at the concentration of 20 mg L^{-1} .

Fig.6 (a) Plot of the MCR-ALS predicted lipid concentrations versus actual lipid concentrations using the area correlation constraint. (b) MCR-ALS resolved elution profiles of three of the lipids in the mixture samples at four different concentrations and three replicates. (c) MCR-ALS resolved pure mass spectra of these three lipids allowing their identification (given in the left corner of the subplot).

Fig.1 Schematic implementation of the area correlation constraint in the MCR-ALS analysis of multiple data matrices using a bilinear decomposition model (see Part 2 Theory).

Fig.2 MCR-ALS resolved emission spectra of the six individual PAHs in their mixtures in miliQ water using non-negativity (blue dotted line), non-negativity and trilinearity (black line) and non-negativity and area correlation constraints (red dotted line).

Fig.3. Contour plots of the experimental 2D-EEM spectra of (a) individual PAHs in miliQ water; (b) river water sample 1; (c) river water sample 2; (d) river water sample 3; (e) the mixture of PAHs in mili Q water; (f) the mixture of PAHs in river water sample 1; (g) the mixture of PAHs in river water sample 2; and (h) mixture of PAHs in river water sample 3.

Fig.4 MCR-ALS resolved excitation (a) and emission (b) spectra of the six PAH and of the six DOM interferences detected in the river water samples.

Fig. 5 MCR-ALS resolved elution profiles of one of the sample replicates of the mixture of nine lipids analyzed by LC-MS at the concentration of 20 mg L^{-1} .

Fig 6 (a) Plot of the MCR-ALS predicted lipid concentrations versus actual lipid concentrations using the area correlation constraint. (b) MCR-ALS resolved elution profiles of three of the lipids in the mixture samples at four different concentrations and three replicates. (c) MCR-ALS resolved pure mass spectra of these three lipids allowing their identification (given in the left corner of the subplot).