Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 204 (2018) 354–361

Contents lists available at ScienceDirect

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy

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

Random initialisation of the excitation-emission matrix fluorescence spectral variables in constraint fashion for subsequent multivariate curve resolution alternating least square analysis on a peculiarly designed calibration set: Simultaneous sensing of nine polycyclic aromatic hydrocarbons in water samples

Riham El Kurdi [a], Keshav Kumar [b,*], Digambara Patra [a,*]

[a] Department of Chemistry, American University of Beirut, Beirut, Lebanon
[b] Institute for Wine Analysis and Beverage Research, Hochschule Geisenheim University, Geisenheim 65366, Germany

ARTICLE INFO

Article history:
Received 6 February 2018
Accepted 18 June 2018
Available online 19 June 2018

Keywords:
Polycyclic aromatic hydrocarbons
MCR-ALS
EEMF: Random initialisation
Water samples
Chemometrics

ABSTRACT

Polycyclic aromatic hydrocarbons (PAHs) are carcinogenic and mutagenic in nature therefore their sensing in water sample is an important analytical task. In the present work, a novel approach that is based on the random initialisation of the excitation-emission matrix fluorescence (EEMF) spectral variables in constraint fashion for subsequent multivariate curve resolution alternating least Square (MCR-ALS) analysis is introduced for simultaneously sensing the complex dilute aqueous mixture of PAHs. The usefulness of the proposed analytical approach is successfully demonstrated by applying it intentionally on a calibration set that is peculiar in many senses. The peculiarity mainly arises because the designed (i) the calibration set consist of nine PAHS having significant spectral overlap, (ii) the concentration of each PAH in different samples are kept constant and (iii) any two samples differ only in the presence and absence of the PAHs. The proposed approach is found to make precise and accurate estimation of each of the nine PAHs without involving any pre-separation. In summary, the proposed approach provides a simple and cost-effective procedure for simultaneous sensing of several PAHs in water samples. The proposed approach could be very useful in developing countries.

1. Introduction

Polycyclic aromatic hydrocarbons (PAHs) are ubiquities in nature and well-known environmental pollutants [1–6]. PAHs are carcinogenic and mutagenic as a result their presence in water samples beyond a permissible limit is a serious health concern [5]. Industrial effluents, oil spills, run-off from sewage and leakage from vehicle oil tank are some of the common reason that brings PAHs in water samples [2, 7–9]. In order to ensure the quality of water samples, monitoring of PAHs is an unavoidable important analytical task. PAHs are usually fluorescent in nature that provides the necessary ground to use fluorescence based analytical procedure to achieve a simple, sensitive and quantification of the PAHs in water samples [5, 6, 8]. Excitation-emission matrix fluorescence (EEMF) [7–11] can be used to capture the fluorescence signatures of all the fluorophores present in a multifluorophoric mixture. EEMF spectroscopy can provide a user friendly procedure mainly because of the following two reasons. (i) All the available fluorimeters have appropriate software that allows the automatic fluorescence data acquisition, (ii) these software also enables auto removal of both first and second order Rayleigh scattering, and (iii) EEMF data sets are truly trilinear in nature that makes it compatible with most of the chemometric techniques [11]. In the context of EEMF spectroscopy, trilinearity comes naturally and it is so because each fluorophore has a unique emission and excitation spectra and shape of which is invariant to the changes in excitation and emission wavelength, respectively [11–14].

Multivariate curve resolution alternating least square is a curve resolution technique [8, 11, 12, 14–23]. It can deconvolute the overlapping spectral profiles without requiring any a priori information related to the peak maxima and spectral bandwidth. MCR-ALS has second order advantages i.e. it can analyse the components even in the presence of an unknown interferences [8, 11, 12, 14–23]. MCR-ALS can be used to analyse the EEMF data sets provided (i) the data is suitably arranged and (ii) the data strictly obey the Beer-Lambert law i.e. intensity is directly proportional to concentration [8, 11, 12, 14]. Mathematically, MCR-ALS decomposes a data set X of dimension I × J (sample × variable) as a product of two matrices C and S of dimensions I × F (sample

---

* Corresponding authors.
E-mail addresses: keshavkumar29@gmail.com, (K. Kumar), dp03@aub.edu.lb (D. Patra).

https://doi.org/10.1016/j.saa.2018.06.067
1386-1425/

Table 1
Concentration of all the nine polycyclic aromatic hydrocarbons (PAHs) in all the 28 samples of calibration set. Sign (×) indicates absence of the PAHs in the mixtures.

Sample	Fluoranthene μg/mL	Dibenz(*ah*)anthracene μg/mL	Benz(*b*)fluoranthene μg/mL	Benz(*ghi*)perylene μg/mL	Chrysene μg/mL	Pyrene μg/mL	Ideno(1,2,3-*cd*)pyrene μg/mL	Anthracene μg/mL	Benz(*a*)pyrene μg/mL
1	0.2023	0.2784	0.2523	0.2763	0.2283	0.2023	0.2763	0.1782	0.2524
2	×	×	×	0.2763	0.2283	0.2023	0.2763	0.1782	0.2524
3	0.2023	×	×	×	0.2283	0.2023	0.2763	0.1782	×
4	0.2023	0.2784	×	×	×	0.2023	0.2763	0.1782	0.2524
5	0.2023	0.2784	0.2523	×	×	×	0.2763	0.1782	0.2524
6	0.2023	0.2784	0.2523	0.2763	×	×	×	0.1782	×
7	0.2023	0.2784	0.2523	0.2763	0.2283	×	×	×	×
8	×	0.2784	0.2523	0.2763	0.2283	0.2023	×	×	×
9	0.2023	×	0.2523	0.2763	×	×	0.2763	0.1782	×
10	×	×	×	×	×	×	0.2763	0.1782	0.2524
11	0.2023	0.2784	×	×	×	×	×	×	×
12	×	×	0.2523	0.2763	×	×	×	×	×
13	×	×	×	×	0.2283	0.2023	×	×	×
14	×	×	×	×	×	×	0.2763	0.1782	×
15	0.2023	×	×	×	×	×	×	×	×
16	×	0.2784	×	×	×	×	×	×	×
17	×	×	0.2523	×	×	×	×	×	×
18	×	×	×	0.2763	×	×	×	×	×
19	×	×	×	×	0.2283	×	×	×	×
20	×	×	×	×	×	0.2023	×	×	×
21	×	×	×	×	×	×	0.2763	×	×
22	×	×	×	×	×	×	×	0.1782	×
23	×	×	×	×	×	×	×	×	0.2524
24	0.2023	0.2784	0.2523	×	×	0.2023	0.2763	0.1782	×
25	0.2023	0.2784	0.2523	×	×	×	×	0.1782	×
26	0.2023	×	×	×	0.2283	×	0.2763	0.1782	×
27	×	×	×	0.2763	0.2283	0.2023	×	×	×
28	×	×	×	×	×	0.2023	0.2763	0.1782	0.2524

× Factor) and J × F (Variable × Factor). The matrix C contains the concentration related information and S contains the spectral related information for each of the F factors. E is the residual matrix of dimension I × J (sample × variable). The decomposition can be summarised using Eq. (1).

$$\mathrm{X} = \mathrm{CS}^{\mathrm{T}} + \mathrm{E} \tag{1}$$

MCR-ALS algorithm solves the above equation by carrying a series of iterations between the given below Eqs. (2) and (3).

$$\mathrm{C} = \mathrm{XS}\left(\mathrm{S}^{\mathrm{T}}\mathrm{S}\right)^{-1} \tag{2}$$

$$\mathrm{S}^{\mathrm{T}} = \left(\mathrm{C}^{\mathrm{T}}\mathrm{C}\right)^{-1}\mathrm{C}^{\mathrm{T}}\,\mathrm{X} \tag{3}$$

The iteration is terminated once the Frobenious norm of residual matrix E is minimised.

$$E = \left\|X - CS^{T}\right\|^{2} \tag{4}$$

The success of MCR-ALS analysis is often bottlenecked by the fact how C and S matrices are initialised for the subsequent iterations. The conventional approach involves the initialisation of the concentration matrix C with evolve factor analysis (EFA) approach [24–27] and the initialisation of spectral matrix (S) using pure variable approach [16, 17, 27, 28]. Both the approaches have their limitations, for example EFA can only be used if the data sets belong to sequential process. The real life samples mostly belong to non-sequential process that limits the application of EFA approach. The pure variable approach such as SIMPLISMA algorithm [16, 17, 27, 28] can be successfully used to initialise the spectral variables provided the data set have pure variable for each factors. With real life samples, it is seldom possible to have pure variables for each of the factors present in the data sets. Another approach, which is recently introduced, involves the random initialisation of the variables in a space spanned by minima and maxima of each variable [27]. This approach does not require the need of data must come from sequential process and there must be pure variable for each factor. Therefore, it can be realised that *MCR-ALS analysis assisted with random initialisation of the spectral variable is most suitable for the analysis of complex multifluorophoric mixtures having significant spectral overlap and belonging to non-sequential process*.

The objective of the present work is (i) to provide a simple and cost – effective analytical procedure and (ii) more specifically to evaluate the efficiency of analytical approach based on the random initialisation of EEMF spectral variables for subsequent MCR-ALS analysis towards the simultaneous analysis of several PAHs having significant spectral overlap in water samples without pre-separation. To carry out the present work, aqueous complex mixtures of nine PAHs fluoranthene, dibenz(*ah*)anthracene, benz(*b*)fluoranthene, benz(*ghi*)perylene, chrysene, pyrene, ideno(123,*cd*-)pyrene, anthracene, benz(*a*)pyrene have been

Table 2
Concentration of all the nine polycyclic aromatic hydrocarbons (PAHs) in all the 6 samples of validation set. Sign (×) indicates absence of the PAHs in the mixtures.

Sample	Fluoranthene μg/mL	Dibenz(*ah*)anthracene μg/mL	Benz(*b*)fluoranthene μg/mL	Benz(*ghi*)perylene μg/mL	Chrysene μg/mL	Pyrene μg/mL	Ideno(1,2,3-*cd*)pyrene μg/mL	Anthracene μg/mL	Benz(*a*)pyrene μg/mL
1	0.2023	0.2784	0.2523	0.2763	0.2283	0.2023	0.2763	0.1782	0.2524
2	0.2023	0.2784	0.2523	0.2763	0.2283	×	×	×	×
3	0.2023	×	0.2523	0.2763	0.2283	0.2023	0.2763	×	×
4	0.2023	0.2784	0.2523	×	×	×	0.2763	0.1782	0.2524
5	0.2023	0.2784	0.2523	×	0.2283	0.2023	0.2763	×	×
6	×	×	×	×	0.2283	0.2023	×	0.1782	0.2524