



Contents lists available at ScienceDirect

Journal of Chromatography A

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



Multivariate analytical figures of merit as a metric for evaluation of quantitative measurements using comprehensive two-dimensional gas chromatography–mass spectrometry

Ali Eftekhari, Hadi Parastar*

Department of Chemistry, Sharif University of Technology, P.O. Box 11155-3516, Tehran, Iran

ARTICLE INFO

Article history:

Received 1 July 2016

Received in revised form 7 September 2016

Accepted 7 September 2016

Available online xxx

Keywords:

Figures of merit
Comprehensive two-dimensional gas chromatography
Chemometrics
MCR-ALS
PARAFAC

ABSTRACT

The present contribution is devoted to develop multivariate analytical figures of merit (AFOMs) as a new metric for evaluation of quantitative measurements using comprehensive two-dimensional gas chromatography–mass spectrometry (GC × GC–MS). In this regard, new definition of sensitivity (SEN) is extended to GC × GC–MS data and then, other multivariate AFOMs including analytical SEN (γ), selectivity (SEL) and limit of detection (LOD) are calculated. Also, two frequently used second- and third-order calibration algorithms of multivariate curve resolution-alternating least squares (MCR-ALS) as representative of multi-set methods and parallel factor analysis (PARAFAC) as representative of multi-way methods are discussed to exploit pure component profiles and to calculate multivariate AFOMs. Different GC × GC–MS data sets with different number of components along with various levels of artifacts are simulated and analyzed. Noise, elution time shifts in both chromatographic dimensions, peak overlap and interferences are considered as the main artifacts in this work. Additionally, a new strategy is developed to estimate the noise level using variance-covariance matrix of residuals which is very important to calculate multivariate AFOMs. Finally, determination of polycyclic aromatic hydrocarbons (PAHs) in aromatic fraction of heavy fuel oil (HFO) analyzed by GC × GC–MS is considered as real case to confirm applicability of the proposed metric in real samples. It should be pointed out that the proposed strategy in this work can be used for other types of comprehensive two-dimensional chromatographic (CTDC) techniques like comprehensive two dimensional liquid chromatography (LC × LC).

© 2016 Published by Elsevier B.V.

1. Introduction

Comprehensive two-dimensional chromatographic (CTDC) systems with multivariate detectors, such as mass spectrometer (i.e., GC × GC–MS and LC × LC–MS) have been emerged as a milestone in chromatographic history. This new insight in chromatography meets an increasing need for the analysis of complex samples due to its superior separation efficiency, peak capacity, and enhanced resolution [1]. In fact, using two independent columns with different polarities, inner diameters, and lengths have remarkably increased separation power in CTDC [2]. The interface between two columns, modulator, is a key feature in CTDC owing to transfer sample from the first column to the second column by taking into account Giddings's conservation rules [3]. Among different CTDC techniques, comprehensive two-dimensional gas chromatography

(GC × GC) has been proposed as a promising tool for the analysis of various complex sample matrices. On the other hand, qualitative and quantitative analysis using GC × GC has been comprehensively reviewed [4,5]. Different strategies have been proposed for quantitative analysis using CTDC which can be classified into conventional and multivariate approaches. Actually, quantitative analysis can be performed by using major selected modulation peaks in second chromatographic column. In conventional approaches, two or three major modulated peaks are selected, and summation of their areas are considered as a measure for quantification [6,7]. In spite of the validity of this strategy in simple sample matrices, but, there is no absolute guarantee to select suitable modulated peaks in complex mixtures due to the presence of different chromatographic artifacts and known and/or unknown interferences. Additionally, modulated peaks at tailings maybe lost because of low signal-to-noise ratio (S/N) [8,9]. Notwithstanding the challenges, it can be demonstrated that chemometric methods can be used to exploit pure analyte signal from a complex mixture (second-order advantage). In this regard, different multivariate resolution

* Corresponding author.

E-mail addresses: h.parastar@sharif.edu, h.parastar@gmail.com (H. Parastar).

methods have been proposed for GC \times GC data analysis [10–12]. Multivariate curve resolution-alternating least squares (MCR-ALS) and parallel factor analysis (PARAFAC) are two frequently used second- and third-order calibration algorithms for GC \times GC data [13]. In addition, n -way partial least squares (NPLS) and unfolded PLS (UPLS) combined with residual trilinearization (RTL) have been also reported for GC \times GC data analysis [14–17].

In modern analytical chemistry, defining numerical parameters which can help to evaluate the performance of a method or an instrument is extremely substantial. On this matter, analytical figures of merit (AFOMs) can be used for this purpose. Using AFOMs, it is possible to compare performance of different experimental procedures, optimization of a given methodology under the different experimental conditions, and development of official protocols for validation and analysis [18–21]. On the other hand, using multivariate methods to extract information from high-order data sets, such as GC \times GC systems makes it necessary to utilize multivariate calibration instead of univariate calibration. In this regard, defining multivariate AFOMs instead of univariate one is of prime importance. It is imperative to mention that over the past decade multivariate AFOMs have been developed for hyphenated chromatographic data [22,23]. However, to the best of our knowledge, multivariate AFOMs have not been developed for CTDC and especially GC \times GC. In the present contribution, multivariate AFOMs based on uncertainty propagation developed by A.C. Olivieri et al. [21] was extended to higher-order GC \times GC–MS data. The core of AFOMs definition was SEN and other AFOMs were calculated based on this figure. On this matter, MCR-ALS and PARAFAC as two frequently used high-order calibration algorithms for GC \times GC–MS were used to obtain pure component profiles (both chromatographic dimensions and mass spectra), to build calibration curves and to calculate SEN, and other AFOMs for target analytes. Additionally, a new strategy was developed to estimate the noise level using variance-covariance matrix of residuals which is very important to calculate multivariate AFOMs. For this purpose, various two-component systems with calibrated components in calibration set and interferences in test set were simulated with different levels of chromatographic artifacts and random shifts in two chromatographic columns. Finally, real GC \times GC–MS data sets were used to confirm the validity of the proposed strategy in this work.

2. Theory

2.1. Multivariate AFOMs

As it has been mentioned earlier, applying multivariate resolution algorithms to extract pure component profiles in second-order data sets (e.g., GC \times GC data set), make it necessary to utilize multivariate AFOMs rather than univariate ones. Among different higher-order calibration methods (i.e., second- and third-order calibration), multi-set and multi-way algorithms have been frequently used in analytical chemistry. Multi-set algorithms are based on decomposition of a two-way array to the pure profiles using different objective functions. As an instance, MCR-ALS [24,25] as a representative of multi-set methods decomposes the data matrix “ \mathbf{X} ” as follows:

$$\mathbf{X}_{\text{aug}} = \mathbf{C}_{\text{aug}}\mathbf{S}^T + \mathbf{E}_{\text{aug}} \quad (1)$$

where $\mathbf{X}_{\text{aug}}(I, J, K)$ is the augmented CTDC data matrix with I elution times in first chromatographic column, J elution times in second chromatographic column as rows and K spectral variables as columns. Also, $\mathbf{C}_{\text{aug}}(I, J, N)$ is the matrix of pure first and second chromatographic profiles, $\mathbf{S}(N, K)$ is matrix of pure spectral profiles, and N is the number of chemical constituents in the original

matrix (i.e., chemical rank). $\mathbf{E}_{\text{aug}}(I, J, K)$ is the error matrix contains the unmodeled part of data matrix.

On the other side, PARAFAC [26] as a representative of multi-way analysis algorithms decomposes a multi-way data array \mathbf{X} to three loadings as follows:

$$\mathbf{X} = \mathbf{C}\mathbf{B}\mathbf{S}^T + \mathbf{E} \quad (2)$$

where $\mathbf{X}(I \times J \times K)$ represents the original multi-way data set, $\mathbf{C}(I \times N)$ is the matrix of pure elution profiles in first chromatographic column, $\mathbf{B}(J \times N)$ is a matrix containing the pure elution profiles in second chromatographic column and $\mathbf{S}(K \times N)$ represents the matrix of pure spectral profiles. In addition, $\mathbf{E}(I \times J \times K)$ is the unmodeled part of original data array.

Actually, the core of definition of AFOM lies on SEN, which is also a key parameter in estimation of other AFOMs [27,28]. Moreover, calculation of multivariate AFOM is algorithm specific. Therefore, for multiset algorithms (e.g., MCR-ALS), SEN can be defined as follows [27]:

$$\text{SEN} = m_n \left[(\mathbf{S}^T \mathbf{S})_{nm}^{-1} \right]^{-1/2} \quad (3)$$

where n is the index for analyte of interest in a mixture, m_n is the slope of pseudounivariate calibration curve for this analyte (SEN(m)), \mathbf{S}^T is the resolved profile for target compound in non-augmented mode (i.e., spectral profiles).

For multiway algorithms (e.g., PARAFAC), SEN can be defined as follows [28]:

$$\text{SEN} = m_n [\text{nth row of } \left[(\mathbf{I} - \mathbf{Z}_{\text{unx}} \mathbf{Z}_{\text{unx}}^+) \mathbf{Z}_{\text{exp}} \right]^+]^{-1} \quad (4)$$

where n and m_n are as before, \mathbf{Z}_{unx} and \mathbf{Z}_{exp} are defined according to Eqs. (5) and (6):

$$\mathbf{Z}_{\text{exp}} = m_n (\mathbf{C}_{\text{exp}} \odot \mathbf{B}_{\text{exp}}) \quad (5)$$

$$\mathbf{Z}_{\text{unx}} = [\mathbf{c}_1 \otimes \mathbf{I}_b | \mathbf{I}_c \otimes \mathbf{b}_1 | \mathbf{c}_2 \otimes \mathbf{I}_b | \mathbf{I}_c \otimes \mathbf{b}_2 | \dots] \quad (6)$$

where \mathbf{B}_{exp} and \mathbf{C}_{exp} are respectively the pure elution profiles in first and second chromatographic columns for the desired constituents in the calibration set. The profiles $\mathbf{b}_1, \mathbf{b}_2$ and $\mathbf{c}_1, \mathbf{c}_2$ are the unexpected constituents in first and second elution modes, respectively. Furthermore, \mathbf{I}_b and \mathbf{I}_c are identity matrices with appropriate dimensions, of $J \times J$ and $K \times K$, respectively. The number 1, 2, ... shows the total number of unexpected constituents. The symbols \odot and \otimes indicate the Khatri-Rao and Kronecker products, respectively [29,30]. In this regard, other multivariate AFOMs, such as analytical sensitivity (γ), selectivity (SEL), limit of detection (LOD), and limit of quantification (LOQ) can be defined for each second-order calibration algorithm. As an instance, γ for is defined as follows:

$$\Upsilon_n = \text{SEN}_n / \sigma_{x_n} \quad (7)$$

where σ_{x_n} is uncertainty in instrumental signal, subscript n identifies a particular analyte of interest. The unit of γ is independent of measured signal which can be employed to compare different methodologies. In addition, SEL can be defined as a dimensionless ratio between SEN in the mixture and SEN when all the other sample constituents are absent [21] (Eq. (8)).

$$\text{SEL}_n = \text{SEN}_n(\text{in a mixture}) / \text{SEN}_n(\text{pure}) \quad (8)$$

In another form, SEL can be defined as follows:

$$\text{SEL}_n = \text{SEN}_n / m_n \quad (9)$$

where m_n is the slope of pseudounivariate calibration curve. As SEN is algorithm specific, therefore, the SEL also depends on the type of SOC algorithm (i.e., multi-set or multi-way). Afterwards, other important parameters that have to be reported as figures of merit

Download English Version:

<https://daneshyari.com/en/article/5135983>

Download Persian Version:

<https://daneshyari.com/article/5135983>

[Daneshyari.com](https://daneshyari.com)