Multivariate Curve Resolution Methods in Imaging Spectroscopy: Influence of Extraction Methods and Instrumental Perturbations

L. Duponchel,* W. Elmi-Rayaleh, C. Ruckebusch, and J. P. Huvenne

Laboratoire de Spectrochimie Infrarouge et Raman, LASIR, CNRS UMR 8516, Bât. C5, Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq Cedex, France

Imaging spectroscopy is becoming a key field of analytical chemistry. In the face of more and more complex samples, we actually need accurate microscopic insight. Nowadays, the methods used to produce concentration maps of the pure compounds from spectral data sets are based on the classical univariate approach although multivariate approaches are sometimes investigated. But in any case, the analytical quality of the chemical images thus provided cannot be discussed since no reference methods are at our disposal. Thus the proposed research focuses on the application of multivariate methods such as Orthogonal Projection Approach (OPA), SIMPLE-to-use Self-modeling Mixture Analysis (SIMPLISMA), Multivariate Curve Resolution − Alternating Least Squares (MCR-ALS), and Positive Matrix Factorization (PMF) for imaging spectroscopy. A systematic and quantitative characterization of the accuracy of spectra and images extraction is investigated on mid-infrared spectral data sets. Of special interest is the influence of instrumental perturbations such as noise and spectral shift on the extraction ability to access the algorithm's robustness.

INTRODUCTION

Improvements in analytical instrumentation allow collecting more and more data for shorter and shorter acquisition times.[1] Multivariate data consisting of a very large number of variables are often acquired particularly in spectroscopy. Furthermore, instrumental coupling introduces a multiway character to the data.[2−6] Although such data are more complex, the ability to solve a specific analytical problem is significantly increased, and advanced chemometrics tools have to be developed to extract knowledge on chemical information. Among the numerous fields of chemometrics, curve resolution methods have become one of the most important since 10 years.[1] This craze is explained by the potential of these methods which can extract simultaneously pure concentration and spectral profiles from the spectral data matrix with no prior knowledge. These methods are thus found in numerous direct applications of spectroscopy like FT-MIR spectroscopy,[7−9] FT-NIR spectroscopy,[10−12] Raman spectroscopy,[13−15] UV−vis spectroscopy,[16−18] circular dichroism spectroscopy,[19−21] mass spectrometry,[22−24] NMR spectroscopy,[25−27] fluorescence spectroscopy,[28−30] electron spectroscopy,[31] ... or in indirect applications where spectroscopies are used for detection purposes such as in liquid[32] and gas[33] chromatography or flow injection analysis.[34] In fact all these analytical techniques and curve resolution methods are generally used for equilibria or kinetic surveys i.e., for time scale experiments. Actually few articles[31,35,36] deal with multivariate curve resolution methods applied to spatial scale experiments as imaging spectroscopy. In general, many researchers only apply univariate data treatments[37−40] to produce spectroscopic images. As there is no reference method to verify the quality of the produced distribution maps, the univariate methodology is never criticized.

The objective of our study is to characterize the most accurate multivariate curve resolution methods for the extraction of the images and the spectra of the pure products from spectral data according to different instrumental perturbations.

GENERAL PRINCIPLE OF IMAGING SPECTROSCOPY

In imaging spectroscopy, a spectrometer is coupled with a microscope. A systematic spectral analysis (e.g. spectral mapping) is carried out with a fixed step size shift over a large sample area defined by a width (x) and a length (y) expressed in pixels. So, each pixel is a microzone of analysis for which a characteristic spectrum is provided (λ variables per spectrum). The spectral analysis is then characterized by the acquisition of $x \times y$ spectra generating the experimental data cube. The state of the art of data analysis in imaging spectroscopy is the following. A characteristic spectral range of a compound of interest is selected from spectral data.[38] Then, the integration of the area under this peak over all spectra gives an estimation of the concentration. By associating a color range to the pseudoconcentration range, a chemical image of the distribution of the considered pure product is obtained.

Unfortunately, this procedure of image extraction is very debatable and has many drawbacks. The industrial or natural samples are very complex. Effectively, they are often composed of a high number of compounds which are not all known. In this way, finding a characteristic spectral zone is impossible. The result is an overestimation of the concentrations providing erroneous spectroscopic images. Furthermore, the current use of microspectrometric data does

---

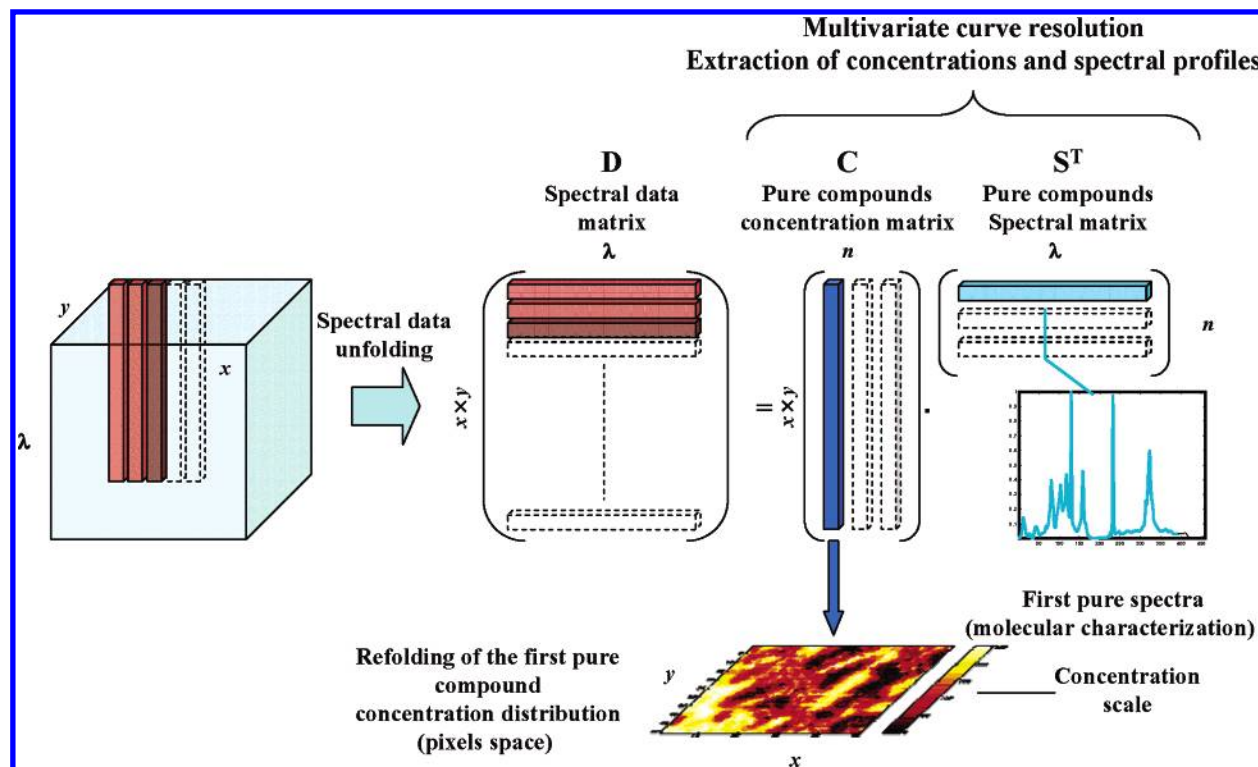* Corresponding author e-mail: ludovic.duponchel@univ-lille1.fr.

Figure 1. Multivariate curve resolution methods applied to imaging spectroscopy.

not permit to extract concentration maps of the products which are not suspected to be present in the analyzed sample. The pure spectra are never extracted although this point is essential for molecular characterization. To resume, the lack of spectral data exploration decreases the expertise level of such imaging techniques.

MULTIVARIATE CURVE RESOLUTION APPLIED TO IMAGING SPECTROSCOPY

Multivariate curve resolution methods permit, with no prior knowledge, to extract simultaneously the spectra of the pure products and their corresponding concentration maps from the experimental data matrix (Figure 1). Unlike the univariate method, there is neither the need to select a characteristic spectral zone nor the need to know all the constituents of the mixtures. To apply multivariate curve resolution methods in imaging spectroscopy, the extraction procedure is the following. We start unfolding the experimental data cube into the spectral data matrix $\mathbf{D}$. This first unfolding procedure is necessary due to the two-way character of the spectral data set despite its three dimensions ($x \times y \times \lambda$). Then a classical line by line unfolding method of the spectral data matrix $\mathbf{D}$ is used in our study. The x first spectra corresponding to the x first pixels of the analysis area are placed along the x first lines of the matrix $\mathbf{D}$ and so on. Afterward, it is necessary to estimate the global rank of $\mathbf{D}$ i.e., the number of pure species (n) present in the whole data set $\mathbf{D}$ from statistical tools.[41-43] Finally, curve resolution methods are performed to extract simultaneously the pure compounds concentration matrix $\mathbf{C}$ (n columns and $x \times y$ rows) and the pure compounds spectral matrix $\mathbf{S^t}$ (n rows and λ columns). Pure spectra are used for identification purposes, whereas concentrations are refolded to recover the pixel space giving the distribution maps for each pure compounds in the analyzed sample.

TESTED MULTIVARIATE CURVE RESOLUTION METHODS

Few articles deal with systematic comparative studies of the extraction accuracy of multivariate curve resolution algorithms in imaging spectroscopy.[35] In this work, we compare the resolution methods currently applied. In a first step, simple methods such as PCA,[42] SIMPLISMA,[44] or OPA[45] produce pure spectral or concentration profiles. In a second step, these profiles are used as initial estimates to be refined by MCR-ALS. The only constraint implemented here is non-negativity on spectra and concentrations during the MCR-ALS optimization. It is in fact one of the only applicable constraint in imaging spectroscopy, local rank being another one. As in other related works, PMF is performed alone.[46]

All calculations are developed with MATLAB v6.1 (The Mathworks, Natick, MA). Matlab homemade codes are used to perform PCA, OPA, and PMF. The SIMPLISMA Matlab toolbox is also used (Courtesy of W. Windig, Eigenvector Research, Inc. 830 Wapato Lake Road Manson, WA 98831). MCR-ALS is performed with the freely available Matlab toolbox developed by R. Tauler (http://www.ub.es/gesq/mcr/mcr.htm, Group of Solution Equilibria and Chemometrics, Analytical Chemistry Department, University of Barcelona).

Let us consider now a spectral data matrix $\mathbf{D}$ with m rows corresponding to m samples and n columns corresponding to n different wavelengths. The aim of curve resolution is to produce a bilinear factorization of $\mathbf{D}$ into a product of two matrices hopefully having a physical sense

$$\mathbf{D} = \mathbf{C} \cdot \mathbf{S^t} \qquad (1)$$

with $\mathbf{C}$ and $\mathbf{S}$, the concentration and the pure spectra matrices, respectively. Various methodologies to provide this factorization are now explained.

CURVE RESOLUTION METHODS IN IMAGING SPECTROSCOPY

J. Chem. Inf. Comput. Sci., Vol. 43, No. 6, 2003 **2059**

PCA. Principal components analysis (PCA) is a common form of factor analysis.[42] This method is widely used in the chemometrics community.[47] Results are generally provided through eigenvector analysis of the correlation matrix. Beyond the definition given by eq 1, the spectral data matrix $\mathbf{D}$ can also be defined in terms of singular value decomposition as

$$\mathbf{D} = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V^t} = \bar{\mathbf{U}} \cdot \bar{\mathbf{S}} \cdot \bar{\mathbf{V}}^t + \mathbf{E} \qquad (2)$$

where $\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$ are the first r columns (so-called rank of $\mathbf{D}$) of the matrices $\mathbf{U}$ and $\mathbf{V}$. The matrices $\mathbf{U}$ and $\mathbf{V}$ are calculated from eigenvalue − eigenvector analyses of the matrices $\mathbf{D} \cdot \mathbf{D^t}$ and $\mathbf{D^t} \cdot \mathbf{D}$. There are respectively related to the scores of the objects and to the loadings of the variables. It can be shown that the error matrix term $\mathbf{E}$ in eq 2 estimates $\mathbf{D}$ in the least-squares-sense, giving the lowest possible value for eq 3.

$$\sum_i \sum_j e_{ij}^2 = \sum_i \sum_j (d_{ij} - \sum_k c_{ik} s_{kj})^2 \qquad (3)$$

After this decomposition, the matrices $\bar{\mathbf{U}} \cdot \bar{\mathbf{S}}$ and $\bar{\mathbf{V}}^t$ can be used as estimations of the matrices $\mathbf{C}$ and $\mathbf{S^t}$, respectively.

SIMPLISMA. The SIMPLE-to-use Self-modeling Mixture Analysis (SIMPLISMA) is one of the first self-modeling curve resolution methods to be used in spectroscopy.[44,48,49] Unlike methods based on factor analysis, SIMPLISMA relies on basic statistical tools, which are the mean and the standard deviation. In fact, the key point of SIMPLISMA is the selection of so-called pure variables or objects from the data matrix $\mathbf{D}$. We focus here on the selection of variables for the algorithm description. A pure variable is a variable to which only one compound of the mixtures contributes. The first purity $p_j^{(1)}$ of a variable j is defined by eq 4

$$p_j^{(1)} = \frac{\sigma_j}{(\mu_j + \alpha)} \qquad \text{for } j = 1,...,n \qquad (4)$$

with

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^m (d_{i,j} - \mu_j)^2}{m}} \quad \text{and} \quad \mu_j = \frac{\sum_{i=1}^m d_{i,j}}{m} \quad \text{for } j = 1,...,n$$

The offset parameter α, user defined, avoids attributing a high purity value to a variable with a low mean absorbance. So the purest variable is the variable for which $p_j^{(1)}$ is the higher. In the next step, it is necessary to subtract from the data matrix the contribution of the first pure compound corresponding to the first selected variable. For this purpose, a weighting parameter $w_j^{(2)}$ is used to reduce the influence of variables which are correlated with the first selected one. More details on the weighting parameter can be retrieved elsewhere.[43] The second purity $p_j^{(2)}$ of a variable j is defined by eq 5

$$p_j^{(2)} = w_j^{(2)} \frac{\sigma_j}{(\mu_j + \alpha)} \text{ for } j = 1,...,n \qquad (5)$$

The purest variable is the variable for which $p_j^{(2)}$ is the higher. By iterating this procedure, the kth purest variable can be selected in the same way by eq 6.

$$p_j^{(k)} = w_j^{(k)} \frac{\sigma_j}{(\mu_j + \alpha)} \text{ for } j = 1,...,n \qquad (6)$$

The total number of selected variables (the total number of pure species in a mixture) is reached when $p_j^{(k)}$ seems to follow a noisy shape. According to the subjective character of this evaluation, statistical tests have been developed.[43] In the final step, the resolution of the spectral data matrix $\mathbf{D}$ is performed. The concentration matrix $\mathbf{C}$ is obtained by merging the k columns of $\mathbf{D}$ corresponding to the k purest variables previously selected. The pure spectra matrix $\mathbf{S}$ is then retrieved using eq 1.

OPA. Like SIMPLISMA, Orthogonal Projection Approach (OPA)[45,50,51] looks for the purest variables or the purest spectra in the spectral data matrix $\mathbf{D}$. Working on the rows i.e., spectra, OPA is denoted OPA(spec) here. The main idea of OPA(spec) is that the purest spectra are mutually more dissimilar than the corresponding mixture spectra.[47] A dissimilarity index k_i is used to select the purest spectrum. It is defined as the determinant of the dispersion matrix $\mathbf{Y_i} \cdot \mathbf{Y_i^t}$. $\mathbf{Y_i}$ is constituted of the mean spectrum $\mathbf{d_{mean}}$ of the data matrix $\mathbf{D}$ and $\mathbf{d_i}$ the ith sample spectrum. Given m different values of i, m different matrices $\mathbf{Y_i}$ of size $(2 \times n)$. Then the purest spectra will have the highest k_i value. From a more geometrical point of view, k_i correspond to the vectorial product of $\mathbf{d_i}$ and $\mathbf{d_{mean}}$ estimating indirectly their correlation (eq 7)

$$k_i = \det(\mathbf{Y_i} \cdot \mathbf{Y_i^t}) = [\|\mathbf{d_{mean}}\| \cdot \|\mathbf{d_i}\| \cdot \sin(\alpha_i)]^2 \text{ for } i = 1,...m \qquad (7)$$

where $\|\mathbf{d_{mean}}\|$ and $\|\mathbf{d_i}\|$ represent the norms of the two vectors $\mathbf{d_{mean}}$ and $\mathbf{d_i}$, respectively, and α_i is the angle between them. To select the second purest spectrum, m dissimilarities values are now calculated by replacing the mean spectra $\mathbf{d_{mean}}$ by the first purest previously selected spectrum in $\mathbf{Y_i}$. For each new iteration, the previously selected pure spectra are merged to $\mathbf{Y_i}$. The total number of spectra can be estimated visually or statistically.[43] An estimation of the pure spectral matrix $\mathbf{S}$ is then obtained by merging the pure spectra previously selected. Complementary, working on the columns of $\mathbf{D}$, OPA(var) aims at finding pure variables rather than pure spectra, so providing an estimation of the pure concentration matrix $\mathbf{C}$.

PMF. The Positive Matrix Factorization (PMF) is generally referred to within air pollution research community as receptor modeling.[52−56] Like the other methods, the spectral data matrix $\mathbf{D}$ is factorized into a product of two other ones.[46] This algorithm has been applied both to 2-way and 3-way data. In fact, the 3-way PMF model is the PARAFAC one. The more interesting PMF features are the use of error estimation of the experimental data matrix $\mathbf{D}$ and the implementation of non-negativity constraints during the matrix decomposition. Given $\mathbf{D}$ and the rank of the factorization r, PMF solves a bilinear factorization problem (eq 1). A F function of the matrix $\mathbf{E}$ is then to be minimized as a function of matrices $\mathbf{C}$ and $\mathbf{S}$ (eq 8)

2060 *J. Chem. Inf. Comput. Sci., Vol. 43, No. 6, 2003*

DUPONCHEL ET AL.

$$F(\mathbf{E}) = \sum_{i,j} \frac{(\mathbf{D} - \mathbf{C} \cdot \mathbf{S^t})_{i,j}^2}{\sigma_{ij}^2} = \sum_i \sum_j \left(\frac{e_{ij}}{\sigma_{ij}}\right)^2 \qquad (8)$$

where σ_{ij} is an estimate of the uncertainty of the ith variable measured in the jth sample. The task is then to minimize F by constraining $\mathbf{C}$ and $\mathbf{S}$ to be non-negative. This optimization is obtained by the alterning least-squares method.[57]

MCR-ALS. As the methods previously presented, multivariate curve resolution-alterning least squares (MCR-ALS)[58,59] try to solve a bilinear factorization of $\mathbf{D}$ expressed by eq 1. In a first step, initial estimates of either pure concentration matrix $\mathbf{C}$ or pure spectral matrix $\mathbf{S}$ have to be obtained from methods such as PCA, OPA, or SIMPLISMA giving respectively $\mathbf{C_{ini}}$ or $\mathbf{S_{ini}}$. In a second step, MCR-ALS algorithm is performed and thus considered as a refinement method. Starting for example with $\mathbf{S_{ini}}$, the concentration matrix $\mathbf{C_{est}}$ is calculated by least squares according to eq 1.

$$\mathbf{C_{est}} = (\mathbf{D} \cdot \mathbf{S_{ini}}) \cdot (\mathbf{S_{ini}^t} \cdot \mathbf{S_{ini}})^{-1} \qquad (9)$$

To reduce the size of the solution space, constraints are then applied on $\mathbf{C_{est}}$ to obtain $\mathbf{C_{cor}}$. Hopefully, this corrected matrix contains the solution chemically relevant. The most usual constraints are the non-negativity (only positive profiles), the unimodality (unimodal shape), and the closure (constant total concentration).[60]

Then $\mathbf{S_{est}}$, a least-squares estimation of $\mathbf{S}$, **is** calculated according to eq 1.

$$\mathbf{S_{est}^t} = (\mathbf{C_{cor}^t} \cdot \mathbf{C_{cor}})^{-1} \cdot (\mathbf{C_{cor}^t} \cdot \mathbf{D}) \qquad (10)$$

The matrices $\mathbf{S_{est}}$ and $\mathbf{C_{est}}$ are then iteratively refined though alternating least squares. During MCR-ALS optimization procedure, the error matrix $\mathbf{E}$ and the residual sum of squares *RSS* are computed.

$$\mathbf{E} = \mathbf{D} - \mathbf{C} \cdot \mathbf{S^t} \; RSS = \sum_i \sum_j e_{ij}^2 \qquad (11)$$

When the *RSS* error reaches the threshold fixed by the user according to the noise level, the system is considered to have converged. The extraction of the pure concentration and pure spectral profiles is achieved.

THE DATA SET

For a quantitative estimation of the extraction ability, it is necessary to test the algorithms on synthetic data sets. Effectively, it is the only way to have reference values in order to validate the extractions. Moreover, another advantage of this synthetic approach is that it is possible to implement different instrumental perturbation levels. The spectral data matrix (experimental data cube) is obtained by a linear combination of six mid infrared reference spectra (six polymers) and their associated reference concentrations maps (Figure 2). To obtain a general insight of the extraction ability, it is possible to choose reference data coming from any spectroscopic technique. Nevertheless the mid-infrared spectroscopy was selected first for the good representativeness of vibrational spectroscopy in imaging and second for the relative complexity of mid-infrared spectra where many overlaps are observed. The six polymers, polyamide 6,

polyester terephthalate, poly(vinylidene fluoride), cellophane, poly(acrylonitrile), and polyetherurethane, are selected from the Hummel Polymer Sample Ftir Library (ThermoNicolet). Their spectra cover the classical mid-infrared range from 455 to 3996 cm^{-1} with a resolution of 7.714 cm^{-1} i.e., 460 wavenumbers. Figure 2 presents the reference spectra of the six components. As one can see from the spectra of the equimolar mixture, it is hardly possible to define a selective spectral range for each of the six pure compounds spectra. The spectral overlapping is undoubtedly a critical aspect to overcome in imaging spectroscopy.

Concerning the data pretreatment, all the six pure reference spectra are scaled between 0 and 1 absorbance unit to avoid variance differences between them. Furthermore, all the six pure concentration maps (50 pixels height, 50 pixels width) are scaled between 0 and 256 arbitrary unit for the same reason. On concentration maps, a color scale is used to express the concentration level i.e., white or black for higher or lower concentrations, respectively. Thus intermediate colors correspond to intermediate concentration levels.

At the same time, increasing levels of white noise (x% normally distributed random noise) are added to the experimental data cube to get closer to the real spectral conditions. For 10% of noise added, we can observe from Figure 2 that the initial spectral information is really perturbed.

In a second part of the work, the influence of a spectral shift on the extraction accuracy of the algorithms is studied. We performed a uniformly distributed random selection of spectra from the experimental data cube on which a z wavenumber shift is applied. In this way, the first third of spectral data set is moved forward z wavenumbers, while the second third is moved backward. The last third is left unchanged.

In the study, a dissimilarity index *dis* is used to establish a quantitative characterization of the extraction ability in spectral and image domain (eq 12)[61]

$$dis(\mathbf{x_{Ext}}, \mathbf{x_{Ref}}) = \sqrt{1 - r^2(\mathbf{x_{Ext}}, \mathbf{x_{Ref}})} \qquad (12)$$

with $\mathbf{x_{Ext}}$ and $\mathbf{x_{Ref}}$ being the extracted and reference profiles, respectively ($\mathbf{x}$ is a spectrum or an image of a pure compound). Thus, we compare the extracted spectrum and image with the reference ones according to their correlations r. We chose to evaluate the extraction accuracy of algorithms with a dissimilarity index because direct comparisons of $\mathbf{x_{Ext}}$ and $\mathbf{x_{Ref}}$ are not allowed due to the intensity ambiguity in bilinear factorizations. Indeed, the amplitude of the concentration matrix $\mathbf{C}$ can be increased by a factor of k and be decreased in its corresponding spectral matrix $\mathbf{S}$ by a factor of k^{-1} (eq 13).[62]

$$\mathbf{D} = \mathbf{C} \cdot \mathbf{S^t} = (\mathbf{C} \cdot k) \cdot (k^{-1} \cdot \mathbf{S^t}) \qquad (13)$$

The less is the dissimilarity, the more is the capability of extraction of the considered algorithm for the given compound. Finally, the mean dissimilarity over the six extracted compounds is used to characterize a given algorithm thus having a global insight of the results.

RESULTS AND DISCUSSION

We propose here to work on a synthetic data set to get a quantitative comparison of the various resolution methods.
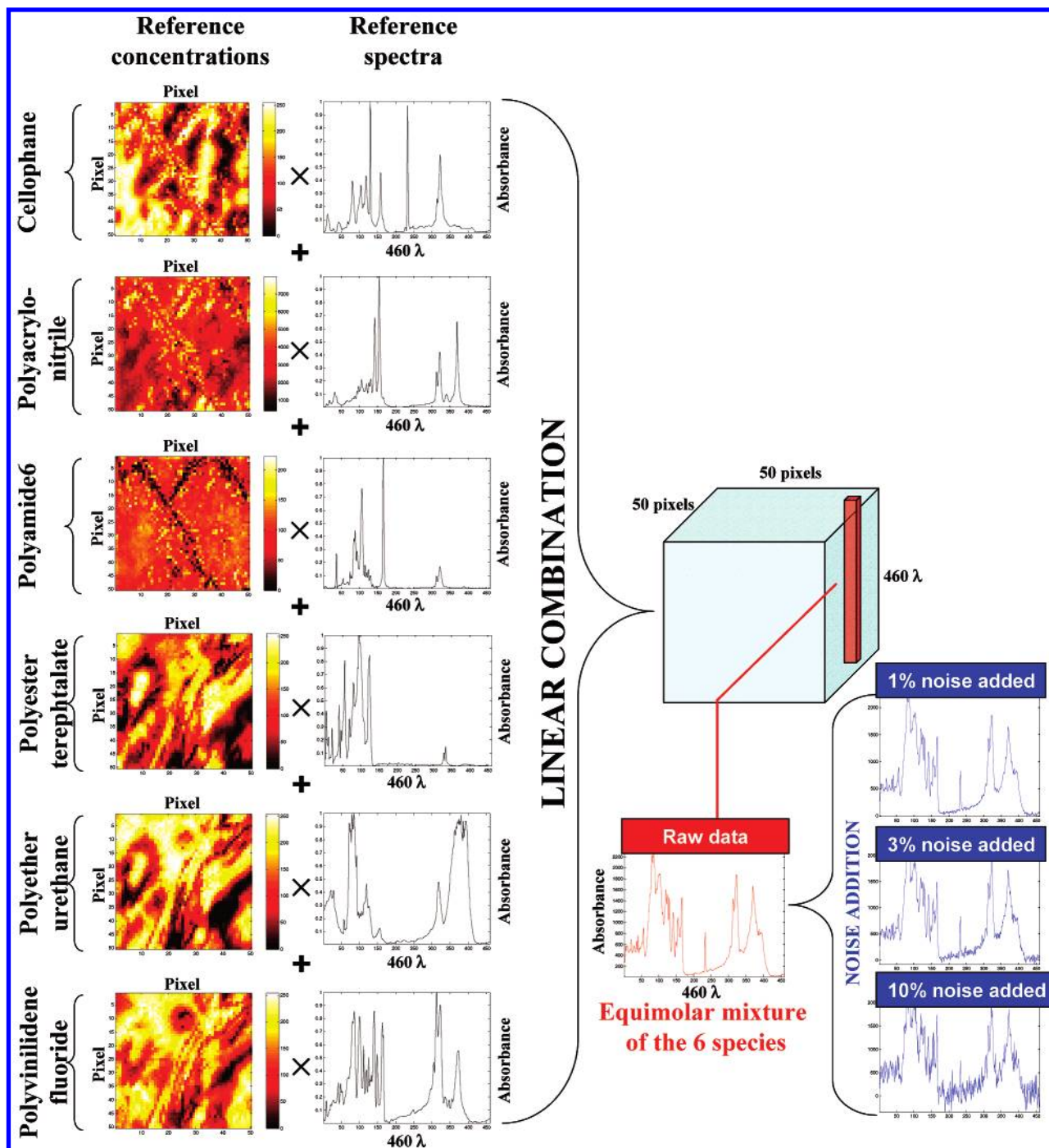
Figure 2. Construction of the data set.

It is then possible, first, to study systematically the extraction ability of each method and, second, to evaluate the robustness of the algorithms to experimental perturbations. It should be noticed that the purpose of the work is not to evaluate different statistical tools for the estimation of the total number of pure species as it can be found elsewhere.[43] Consequently, whatever the curve resolution method (simple or refined one) applied to the experimental data cube, six spectra and their corresponding concentration distributions are systematically extracted. After the extractions, the method is characterized by the mean dissimilarity.

Image and Spectral Extractions with Simple Methods. In the first part, only simple methods such as SIMPLISMA, OPA, or PCA are studied according to instrumental perturba-

tion. First, the signal-to-noise ratio is of highest importance in all imaging spectroscopy. Indeed, the observed signal-to-noise ratio is often low because the number of spectra (equal to the number of analyzed pixels) to be acquired is high and the total acquisition time is limited for the analytical purposes. Second, the influence of wavelength shift is studied. Although modern spectroscopic instrumentation leads to very high wavelength repeatability, this kind of instrumental perturbation can give an idea of the algorithms robustness.

Figure 3a,b shows the influence of noise on the extraction ability of spectra and images for simple algorithms. The noise level covers the range from 1% to 10% corresponding to what is usually considered as good and very bad spectral
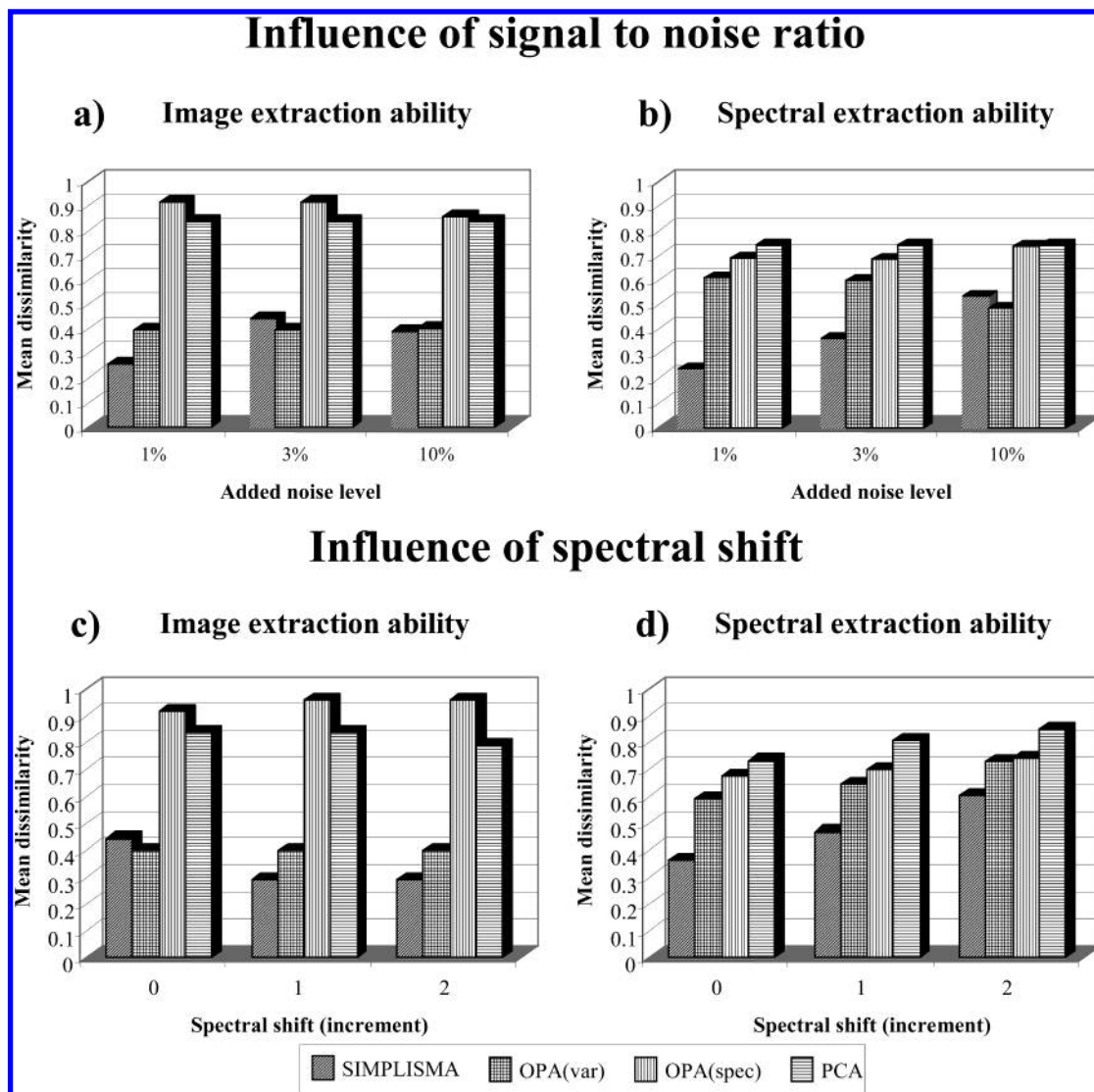
Figure 3. Influence of instrumental perturbations on extraction accuracy of simple methods.

quality, respectively. Figure 3c,d shows the influence of spectral shift for the same algorithms. As the same behaviors are observed for the other noise level data sets, the results of the spectral shift (from 0 to 2 increments) are only presented for the 3% noise data set. From a general point of view, we observe on all these figures very different results depending on the applied algorithm. A first result is that the lowest dissimilarity is always obtained with SIMPLISMA and OPA(var) methods. It corresponds to the best image and spectral extractions. On the contrary, OPA(spec) and PCA methods show the worst extraction results even with the low-perturbed conditions (no shift, 1% added noise). According to Figure 3a,b, increasing noise level reduces effectively the extraction ability of the algorithms. Nevertheless, a surprising decrease of OPA(var) spectral dissimilarity at 10% noise level is observed. In fact, we consider that there is no significant difference between high noise levels for OPA-(var) extractions since the dissimilarity is always very high i.e., the correlation is very low. SIMPLISMA seems to have good extraction ability when the noise level is low. OPA-(var) seems to be less sensitive to noise at high perturbation level, retrieving the extraction accuracy of SIMPLISMA. In Figure 3c, an interesting behavior is observed but left unexplained until now: algorithms appear insensitive to

spectral shift for image extraction. In Figure 3d, the increase of perturbation i.e., the spectral shift, effectively induces a higher dissimilarity for the spectral extraction.

At this point, we can try to explain why the worth extraction results are obtained with PCA or OPA(spec) and the best ones with OPA(var) whatever the perturbation level. By definition, PCA extracts perpendicular vectors which are considered in this application as pure spectra. Table 1 shows angles between pairs of pure reference spectra obtained from their corresponding scalar product. The perpendicular property is not at all observed because angles are often far from 90°. Consequently, the factors are often far from the real pure spectra. Concerning OPA(spec) and OPA(var) a deeper analysis is needed. In fact, the only difference comes from the direction of the selection in the data set. OPA(spec) selects the purest spectra from the experimental data matrix i.e., 2500 mixture spectra. Thus the six selected spectra should correspond to the six pure compound spectra. On the other hand, OPA(var) selects the purest variable from the experimental data matrix i.e., 460 wavenumbers for our application. Then the selected variables should correspond to pure compound concentrations. Table 2a,b focuses on the 1% noise data set. They present the selected objects for the two OPA methods as well as their agreement with the

Table 1. Angles Calculated between Pure Reference Spectra[a]

	cellophane	poly(acrylonitrile)	polyamide 6	polyester terephthalate	polyetherurethane	poly(vinylidene fluoride)
cellophane	0	57	63	63	63	47
poly(acrylonitrile)		0	72	74	61	48
polyamide 6			0	59	71	56
polyester terephthalate				0	63	57
polyetherurethane					0	49
poly(vinylidene fluoride)						0

[a] Expressed in degrees.

Table 2. Evaluation of the Purity of Selected Information with (a) OPA(spec) and (b) OPA(var)

(a) OPA(spec)

purest selected spectrum	Reference Concentration (Arbitrary Unit)					
	cellophane	poly(acrylonitrile)	polyamide 6	polyester terephthalate	polyetherurethane	poly(vinylidene fluoride)
#2037	255	136	196	23	6	86
#1663	116	94	81	212	255	255
#450	255	192	127	69	226	113
#2198	33	129	115	19	73	233
#1978	127	201	200	196	151	66
#2340	118	32	133	84	233	229

(b) OPA(var)

purest selected variable	Correlation between Selected Variables and Reference Concentrations (R^2)					
	cellophane	poly(acrylonitrile)	polyamide 6	polyester terephthalate	polyetherurethane	poly(vinylidene fluoride)
#233	0.99	0.01	0.08	0.00	0.01	0.02
#85	0.04	0.01	0.01	0.48	0.90	0.58
#155	0.01	0.86	0.10	0.06	0.14	0.05
#166	0.12	0.03	0.44	0.02	0.21	0.57
#95	0.02	0.02	0.01	0.86	0.76	0.22
#141	0.01	0.16	0.03	0.06	0.44	0.78

reference profiles. We should remember that the concentration ranges are between 0 and 255 arbitrary units. For OPA-(spec), a selected spectrum can be considered as pure if the reference concentration of only one compound is high and the five other ones are relatively low. It is not exactly the case, as noticed in Table 2a, because the selected spectra are often mixtures of at least two major products. For OPA-(var), a selected variable (wavenumber) can be considered as pure if the absorbance values corresponding to each spectra are correlated with the reference concentration of one compound only. The purity information given by OPA(var) seems to be more important as shown in Table 2b. We can thus conclude that OPA(var) produces better profiles than OPA(spec). Nevertheless, OPA(var) does not always give the best results. It is the analytical application that will dictate the search direction for pure objects i.e., across rows or columns of the experimental data set. If there is pure spectra in the data set, then OPA(spec) can also produce good initial profiles.

Image and Spectral Extractions with Refined Methods. In this second part of the work, MCR-ALS is used as a refined method. The aim is to establish if it is possible to increase the extraction accuracy for images and pure spectra production obtained from previous results. In imaging spectroscopy, it is the first attempt to refine PCA or SIMPLISMA solutions through MCR-ALS algorithm. On the contrary, OPA is currently associated with MCR-ALS. We also evaluate PMF alone as a refined extraction method.

For this purpose, Figure 4 shows the influence of experimental perturbations on extractions obtained with refined methods. Comparing these results with Figure 3, the dissimilarity is roughly reduced by a factor of two whatever the experimental conditions and the algorithms are. It should be noticed that this improvement is in fact even more important, the dissimilarity index being a function of the correlation. In fact, the extraction is improved by the non-negativity constraints of the concentration and spectral profiles imposed during the MCR-ALS optimization. Furthermore, alternating least squares used in MCR-ALS procedure lead to a noise filtering effect reducing drastically the mean dissimilarity.[63] Again OPA(spec) and PCA are the less accurate, despite refinement with MCR-ALS. In general, we notice that SIMPLISMA/MCR-ALS, OPA(var)/MCR-ALS, and PMF can give good extractions. The good results of the PMF method are explained by the implementation of the non-negativity constraint as well. Nevertheless, the use of PMF for such spectroscopic applications is limited because the calculations are relatively slow with regard to the other methods. From this work, we can summarize that MCR-ALS makes good extractions as long as the initial profiles are close to the reference ones. But more precisely, if we compare the capacities of these algorithms for an increasing noise level, OPA(var)/MCR-ALS is the most effective. It keeps a relatively stable mean dissimilarity despite an increasing data perturbation. For 10% added noise the mean dissimilarity of SIMPLISMA/MCR-ALS is in fact twice the OPA(var)/MCR-ALS one. According to a previous work,[45] OPA seems to perform better than SIMPLISMA when
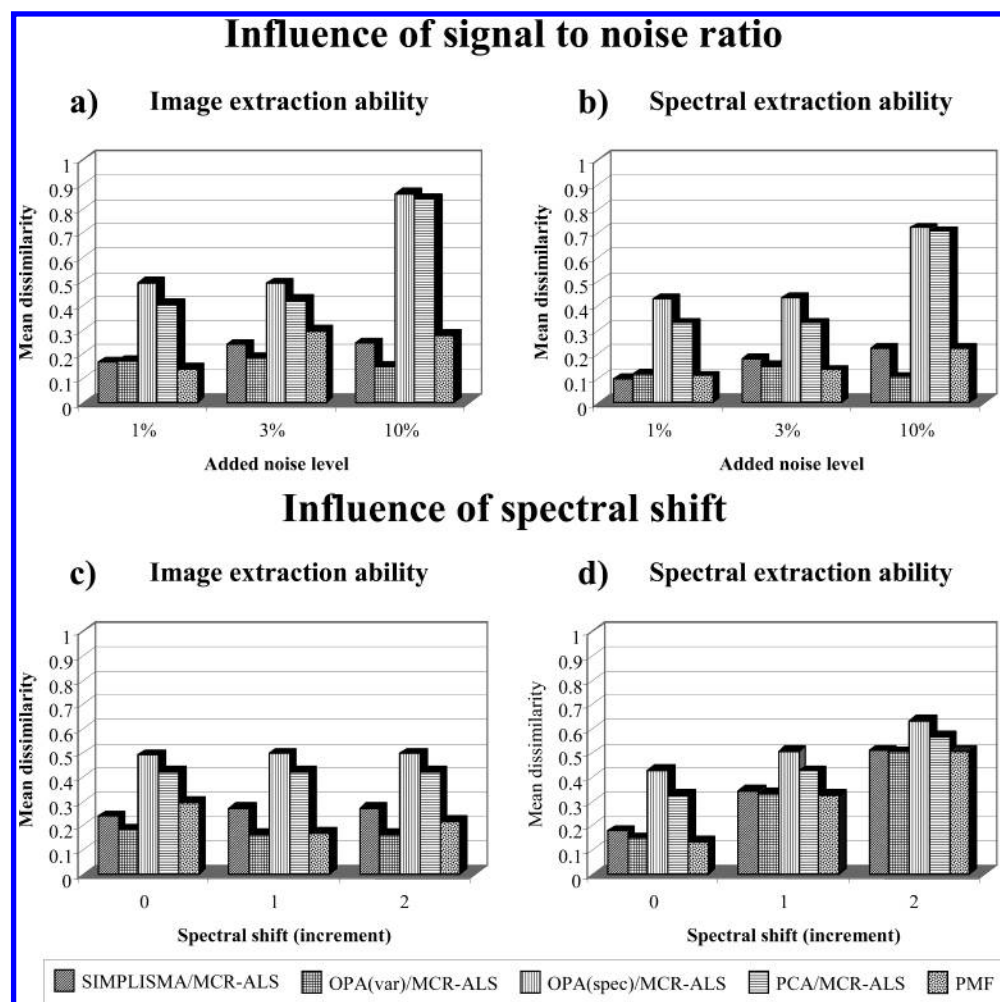
Figure 4. Influence of instrumental perturbations on extraction accuracy of refined methods.

information is difficult to extract. It is important to use MCR-ALS algorithm after SIMPLISMA to increase the image and spectral extraction accuracy. Thus, simple methods such as SIMPLISMA or PCA should be considered as first insight methods only giving initial estimates of the matrices **C** and **S** to be refined by MCR-ALS.[64,65,66] The influence of spectral shift is not discussed here because we can observe the same behavior as in Figure 3.

Visual Insight of the Extraction Accuracy. The results of the refined methods for one of the six extracted compounds i.e., the polyester terephthalate with 1% and 10% of added noise are proposed in Figure 5. Effectively, it is not easy to appreciate the link between the dissimilarity index and the quality of the extractions (spectra and images). The polyester terephthalate is chosen because it is the most difficult to extract for this application. It is also for this compound that the higher dissimilarity differences between the algorithms are observed. At 1% noise level, SIM-PLISMA/MCR-ALS and OPA(var)/MCR-ALS have good spectral extraction accuracy with the lower dissimilarity index i.e., 0.04 and 0.09, respectively. From the observation of OPA(var)/MCR-ALS and PMF extractions, we observe that the dissimilarity index is well suited for the comparison of extraction methods. Indeed, the observed spectral differences are rather small between the two methods, while the dissimilarity difference is important i.e., 0.09 and 0.21. As noticed previously, it is rather difficult to characterize visually the quality of OPA(var)/MCR-ALS and PMF images, while

associated dissimilarities are really different i.e., 0.50 and 0.38, respectively. In general, when the dissimilarity index is high we observe artifacts attributed to spectral contributions of other pure compounds of the mixtures marked with red triangles. For example at 1% noise level, OPA(spec)/MCR-ALS presents seven artifacts coming from the poly(vinylidene fluoride) (486 cm^{-1}, 530 cm^{-1}, 609 cm^{-1}, 795 cm^{-1}, 879 cm^{-1}, 1180 cm^{-1}, 1404 cm^{-1}), four contributions of the polyamide 6 for PCA/MCR-ALS (1543 cm^{-1}, 1643 cm^{-1}, 3070 cm^{-1}, 3300 cm^{-1}), and three contributions of the polyetherurethane (1535 cm^{-1}, 3324 cm^{-1}) and the poly-(acrylonitrile) (2244 cm^{-1}) for PMF. In this situation, the identification of pure compounds becomes impossible. Thus, it is very important to choose the methods from the observation of the dissimilarity index. Comparing now the spectral extractions for 1% and 10% of noise, the high stability of OPA(var)/MCR-ALS is noticed. The reference spectrum of the polyester terephthalate is always very well retrieved. For SIMPLISMA/MCR-ALS the noise filtering effect is less effective showing eight artifacts coming from the polyetherurethane unobserved at the 1% noise level (772 cm^{-1}, 818 cm^{-1}, 1535 cm^{-1}, 1597 cm^{-1}, 2800 cm^{-1}, 3047 cm^{-1}, 3124 cm^{-1}, 3324 cm^{-1}). This example attests the sensitivity of SIMPLISMA/MCR-ALS to high noise level. For the extractions of the concentration map, increasing the noise level induces overestimations (blue ellipses) or underestimations (green ellipses) of the polyester terephthalate concentration. We have then a slanted vision because it
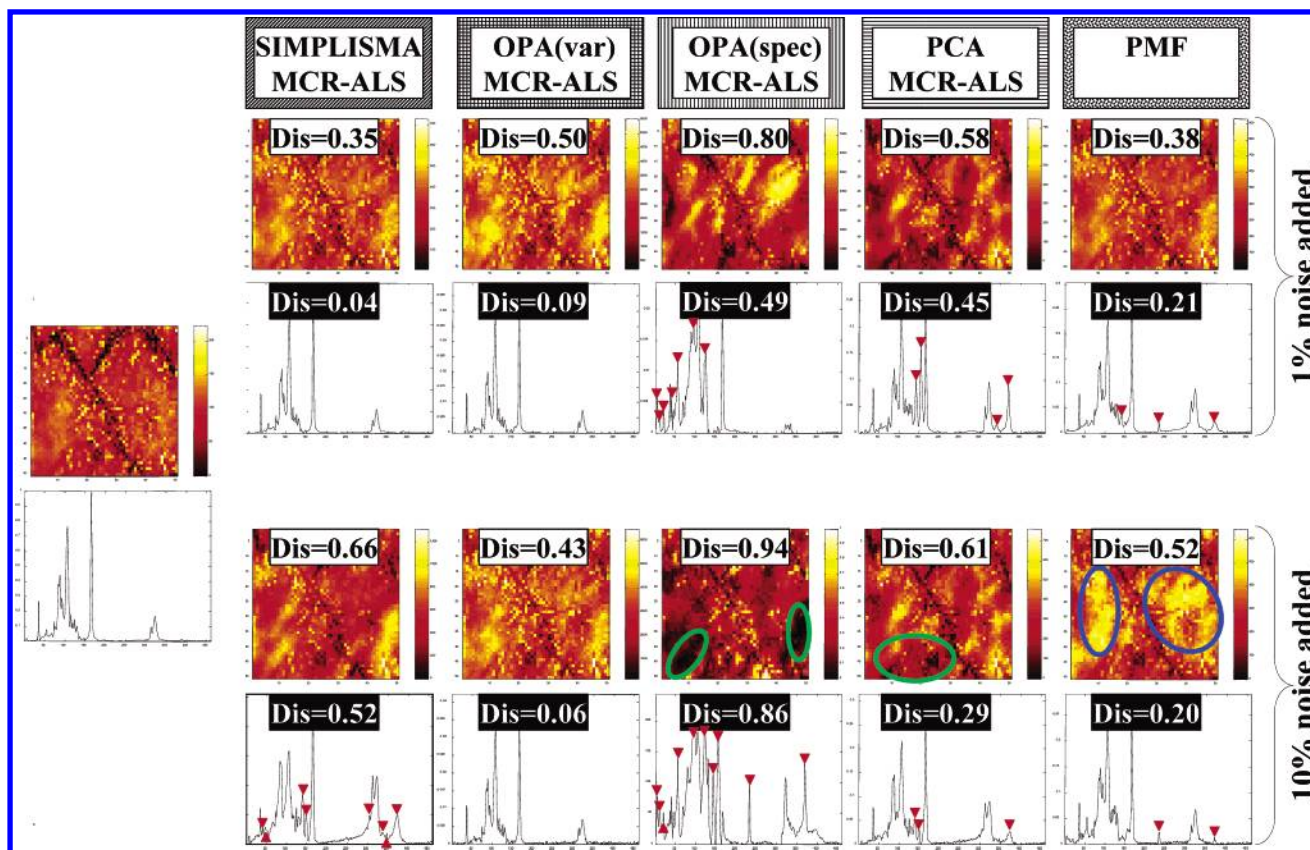
Figure 5. Visual insight into the relation between the dissimilarity index and the extraction accuracy.

indicates the presence of polyester terephthalate where there is not a presence. The opposite might also be observed. This point is particularly important because all the methods can produce concentration maps, but, as it is seen, they can be far from the analytical reality.

CONCLUSION

We proposed here a systematic and comparative evaluation of the multivariate curve resolution methods. It appeared necessary since very different extraction behaviors were observed even at low instrumental perturbation levels. When spectral extraction accuracy was low i.e., the mean dissimilarity was very high, extracted pure spectral profiles were perturbed by artifacts coming from other pure compounds. For analytical purposes, the molecular identification became sometimes difficult. In the same way, bad extraction accuracy has induced over- or underestimation observed on concentration maps. The use of synthetic data was also necessary to access algorithm accuracy because no reference data were at our disposal and second to control the perturbation level. The dissimilarity index was suitable to access the image and spectral extraction ability and have a quantitative comparison.

In general, we noticed that simple methods had to be refined by MCR-ALS in order to obtain good image and spectral extractions of the pure compounds. SIMPLISMA, PCA, and OPA methods should be used as first data exploration insight providing initial estimates only. We have also observed that MCR-ALS algorithm led to a noise filtering behavior. More precisely, OPA(var)/MCR-ALS and SIMPLISMA/MCR-ALS algorithms were the most accurate for the extraction of the pure spectra and images in infrared microspectrometry with strong spectral overlaps. Moreover,

OPA/MCR-ALS was less sensitive to the signal-to-noise ratio and the spectral drift when the perturbation level was high. In view of these results OPA/MCR-ALS or SIMPLISMA/MCR-ALS algorithms will probably perform well on other imaging spectroscopy devices. Since the spectral data matrix were first unfolded in this work, it will be possible to generalize conclusions to others with two way spectroscopic data sets. Future works will focus on the influence of the spectral matrix unfolding on the extraction results.[67] Improvements should also be possible by implementing other constraints such as local rank during MCR-ALS optimization[62,68] to restrict the solution space.

REFERENCES AND NOTES

(1) Lavine, B. K. Chemometrics. *Anal. Chem.* **2000**, *72*, 91R−97R.

(2) Bezemer, E.; Rutan, S. Study of the hydrolysis of a sulfonylurea herbicide using liquid chromatography with diode array detection and mass spectrometry by three-way multivariate curve resolution-alternating least squares. *Anal. Chem.* **2001**, *73*, 4403−4409.

(3) Jiji, R. D.; Booksh, K. S. Mitigation of Rayleigh and Raman spectral interferences in multiway calibration of excitation−emission matrix fluorescence spectra. *Anal. Chem.* **2000**, *72*, 718−725.

(4) Zampronio, C. G.; Gurden, S. P.; Moraes, L. A. B.; Eberlin, M. N.; Smilde, A. K.; Poppi, R. J. Direct sampling tandem mass spectrometry (MS/MS) and multiway calibration for isomer quantitation. *Analyst* **2002**, *127*, 1054−1060.

(5) Bylund, D.; Danielsson, R.; Malmquist, G.; Markides, K. E. Chromatographic alignment by warping and dynamic programming as a pre-processing tool for PARAFAC modelling of liquid chromatography−mass spectrometry data. *J. Chromatogr. A* **2002**, *961*, 237−244.

(6) Barbieri, P.; Adami, G.; Piselli, S.; Gemiti, F.; Reisenhofer, E. A three-way principal factor analysis for assessing the time variability of freshwaters related to a municipal water supply. *Chemom. Intell. Lab. Syst.* **2002**, *62*, 89−100.

(7) Ozaki, Y.; Olinga, A.; Siesler, H. W. Comparison of various chemometric evaluation approaches for on-line FT-NIR transmission

and FT-MIR/ATR spectroscopic data of methyl methacrylate solution polymerization. *Anal. Chim. Acta* **2002**, *452*, 265−276.

(8) De Braekeleer, K.; De Juan, A.; Cuesta Sanchez, F.; Hailey, P. A.; Sharp, D. C. A.; Dunn, P.; Massart, D. L. Determination of the End Point of a Chemical Synthesis Process Using On-Line Measured Mid-Infrared Spectra. *Appl. Spectrosc.* **2000**, *54*, 601−607.

(9) Leger, M. N.; Wentzell P. D. Dynamic Monte Carlo self-modeling curve resolution method for multicomponent mixtures. *Chemom. Intell. Lab. Syst.* **2002**, *62*, 171−188.

(10) Sasicacute, S.; Kita, Y.; Furukawa, T.; Watari, M.; Siesler, H. W.; Ozaki Y. Monitoring the melt-extrusion transesterification of ethylene-vinyl acetate copolymer by self-modeling curve resolution analysis of on-line near-infrared spectra. *The Analyst* **2000**, *125*, 2315−2321.

(11) Ozaki, Y.; Ai, S.; Jiang, J. H.; Siesler H. W. Self-modeling curve resolution analysis of on-line vibrational spectra of polymerisation and transesterification. *Macromolecular Symposia* **2002**, *184*, 229−248.

(12) Windig, W. Spectral data files for self-modeling curve resolution with examples using the Simplisma approach. *Chemom. Intell. Lab. Syst.* **1997**, *36*, 3−16.

(13) Ai, S.; Ozaki, Y.; Kleimann, M.; Siesler H. W. On the ambiguity of self-modeling curve resolution: orthogonal projection approach analysis of the on-line Fourier transform-Raman spectra of styrene/1,3-butadiene block-copolymerization. *Anal. Chim. Acta* **2002**, *460*, 73−83.

(14) Nicholson, M. A.; Aust, J. F.; Booksh, K. S.; Bell, W. C.; Myrick, M. L. Kinetic and spectroscopic profiles of three pyridine complexes at a silver electrode using surface-enhanced Raman scattering and evolving factor analysis. *Vib. Spectrosc.* **2000**, *24*, 157−163.

(15) Andrew, J.; Hancewicz T. Rapid Analysis of Raman Image Data Using Two-Way Multivariate Curve Resolution. *Appl. Spectrosc.* **1998**, *52*, 797−807.

(16) Vives, M.; Tauler, R.; Gargallo, R. J. Study of the influence of metal ions on tRNAPhe thermal unfolding equilibria by UV spectroscopy and multivariate curve resolution. *J. Inorg. Bioch.* **2002**, *89*, 115−122.

(17) Bijlsma, S.; Boelens, H. F. M.; Hoefsloot, H. C. J.; Smilde, A. K. Constrained least squares methods for estimating reaction rate constants from spectroscopic data. *J. Chemom.* **2002**, *16*, 28−40.

(18) Bijlsma, S.; Boelens, H. F. M.; Hoefsloot, H. C. J.; Smilde, A. K. Estimating reaction rate constants: comparison between traditional curve fitting and curve resolution. *Anal. Chim. Acta* **2000**, *419*, 197−207.

(19) Jaumot, J.; Escaja, N.; Gargallo, R.; González, C.; Pedroso, E.; Tauler R. Multivariate curve resolution: a powerful tool for the analysis of conformational transitions in nucleic acids. *Nucleic Acids Res.* **2002**, *30*, 92.

(20) Gargallo, R.; Vives, M.; Tauler, R.; Eritja, R. Protonation studies and multivariate curve resolution on oligodeoxynucleotides carrying the mutagenic base 2-aminopurine. *Biophys. J.* **2001**, *81*, 2886−2896.

(21) Mendieta, J.; Díaz-Cruz, M. S.; Monjonell, A.; Tauler, R.; Esteban, M. Complexation of cadmium by the C-terminal hexapeptide Lys-Cys-Thr-Cys-Cys-Ala from mouse metallothionein: study by differential pulse polarography and circular dichroism spectroscopy with multivariate curve resolution analysis. *Anal. Chim. Acta* **1999**, *390*, 15−25.

(22) Zampronio, C. G.; Moraes, L. A. B.; Eberlin, M. N.; Poppi, R. J. Multivariate curve resolution applied to MS/MS data obtained from isomeric mixtures. *Anal. Chim. Acta* **2001**, *446*, 493−500.

(23) Salau, J. S.; Honing, M.; Tauler, R.; Barceló, D. Resolution and quantitative determination of coeluted pesticide mixtures in liquid chromatography-thermospray mass spectrometry by multivariate curve resolution. *J. Chromatogr. A* **1998**, *795*, 3−12.

(24) Brakstad, F. The feasibility of latent variables applied to GC-MS data. *Chemom. Intell. Lab. Syst.* **1995**, *29*, 157−176.

(25) Pedersen, H.; Bro, R.; Engelsen, S. Towards Rapid and Unique Curve Resolution of Low-Field NMR Relaxation Data: Trilinear SLICING versus Two-Dimensional Curve Fitting. *J. Magn. Reson.* **2002**, *157*, 141−155.

(26) Bezemer, E.; Rutan, S. Resolution of overlapped NMR spectra by two-way multivariate curve resolution alternating least squares with imbedded kinetic fitting. *Anal. Chim. Acta* **2002**, *459*, 277−289.

(27) Ute, K.; Niimi, R.; Matsunaga, M.; Hatada, K.; Kitayama, T. On-Line SEC-NMR Analysis of the Stereocomplex of Uniform Isotactic and Uniform Syndiotactic Poly(methyl methacrylate)s. *Macromol. Chem. Phys.* **2001**, *202*, 3081−3086.

(28) Saurina, J.; Leal, C.; Compao, R.; Granados, M.; Prat, M. D.; Tauler, R. Estimation of figures of merit using univariate statistics for quantitative second-order multivariate curve resolution. *Anal. Chim. Acta*, **2001**, *432*, 241−251.

(29) Vives, M.; Gargallo, R.; Tauler, R. Three-way multivariate curve resolution applied to speciation of acid−base and thermal unfolding transitions of an alternating polynucleotide. *Biopolymers* **2001**, *59*, 477−488.

(30) Saurina, J.; Leal, C.; Compañó, R.; Granados, M.; Tauler, R.; Prat, M. D. Determination of triphenyltin in sea-water by excitation−emission matrix fluorescence and multivariate curve resolution. *Anal. Chim. Acta* **2000**, *409*, 237−245.

(31) Artyushkova, K.; Fulghum, J. E. Identification of chemical components in XPS spectra and images using multivariate statistical analysis methods. *J. Electron Spectrosc. Relat. Phenom.* **2001**, *121*, 33−55.

(32) Lavine, B. K.; Ritter, J. P.; Voigtman, E. Multivariate curve resolution in liquid chromatography−resolving two-way multi-component data using a Varimax extended rotation. *Microchem. J.* **2002**, *72*, 163−178.

(33) Statheropoulos, M.; Mikedi, K. PCA-ContVarDia: an improvement of the PCA-VarDia technique for curve resolution in GC-MS and TG-MS analysis. A*nal. Chim. Acta* **2001**, *446*, 351−368.

(34) Saurina, J.; Hernandez-Cassou, S.; Izquierdo-Ridorsa, A.; Tauler, R. pH-Gradient spectrophotometric data files from flow-injection and continuous flow systems for two- and three-way data analysis. *Chemom. Intell. Lab. Syst.* **2000**, *50*, 263−271.

(35) Wang, J.-H.; Hopke, P. K.; Hancewicz, T. M.; Zhang, S. L. Application of modified alternating least squares regression to spectroscopic image analysis. *Anal. Chim. Acta* **2003**, *476*, 93−109.

(36) Batonneau, Y.; Laurens, J.; Merlin, J. C.; Brémard, C. Self-modeling mixture analysis of Raman microspectrometric investigations of dust emitted by lead and zinc smelters. *Anal. Chim. Acta* **2002**, *446*, 23−37.

(37) Wetzel, D. L.; Williams, G. P. Synchrotron infrared microspectroscopy of retinal layers. *Vibr. Spectrosc.* **2002**, *30*, 101−109.

(38) Lamontagne, J.; Durrieu, F.; Planche, J. P.; Mouillet, V.; Kister J. Direct and continuous methodological approach to study the aging of fossil organic material by infrared microspectrometry imaging: application to polymer modified bitumen. *Anal. Chim. Acta* **2001**, *444*, 241−250.

(39) Lasch, P.; Boese, M.; Pacifico, A.; Diem, M. FT-IR spectroscopic investigations of single cells on the subcellular level. *Vibr. Spectrosc.* **2002**, *28*, 147−157.

(40) Tramini, P.; Bonnet, B.; Sabatier, R.; Maury, L. A method of age estimation using Raman microspectrometry imaging of the human dentin. *Forensic Sci. Int.* **2001**, *118*, 1−9.

(41) Gampp, H.; Maeder, M.; Meyer, C. J.; Zuberbuhler, A. D. Quantification of a known component in an unknown mixture. *Anal. Chim. Acta* **1987**, *193*, 287−293.

(42) Malinowski, E. R. In *Factor Analysis in Chemistry*; Wiley-Interscience Ed.: 2002; p 86.

(43) Gourvénec, S.; Massart, D. L.; Rutledge, D. N. Determination of the number of components during mixture analysis using the Durbin-Watson criterion in the Orthogonal Projection Approach and in the SIMPLe-to-use Interactive Self-modelling Mixture Analysis approach. *Chemom. Intell. Lab. Syst.* **2002**, *61*, 51−61.

(44) Windig, W.; Lippert, J. L.; Robbins, M. J.; Kresinske, K. R.; Twist, J. P.; Snyder, A. P. Interactive self-modeling multivariate analysis. *Chemom. Intell. Lab. Syst.* **1990**, *9*, 7−30.

(45) Sanchez, F. C.; Toft, J.; Van den Bogeart, B.; Massart, D. L. Orthogonal Projection Approach Applied to Peak Purity Assessment. *Anal. Chem.* **1996**, *68*, 79−85.

(46) Paatero, P.; Tapper, U. Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* **1994**, *5*, 111−126.

(47) Vandeginste, B. G. M.; Massart, D. L.; Buydens, L. M. C.; De Jong, S.; Lewi, P. J.; Smeyers-Verbeke, J. In *Handbook of Chemometrics and Qualimetrics*; Elsevier Edition: 1998; p 519.

(48) Snyder, A. P.; Windig, W.; Toth, J. P. Interactive self-modeling multivariate analysis of thermolysis mass spectra. *Chemom. Intell. Lab. Syst.* **1991**, *11*, 149−160.

(49) Windig, W.; Heckler, C. E.; Agblevor, F. A.; Evans, R. J. Self-modeling mixture analysis of categorized pyrolysis mass spectral data with the SIMPLISMA approach. *Chemom. Intell. Lab. Syst.* **1992**, *14*, 195−207.

(50) Sanchez, F. C.; Vandeginste, B. G. M.; Hancewicz, T. M.; Massart, D. L. Resolution of Complex Liquid Chromatography-Fourier Transform Infrared Spectroscopy Data. *Anal. Chem.* **1997**, *69*, 1477−1484.

(51) De Braekeleer, K.; Massart, D. L. Evaluation of the orthogonal projection approach (OPA) and the SIMPLISMA approach on the Windig standard spectral data sets. C*hemom. Intell. Lab. Syst.* **1997**, *39*, 127−141.

(52) Henry, R. C. Multivariate receptor modelscurrent practice and future trends. *Chemom. Intell. Lab. Syst.* **2002**, *60*, 43−48.

(53) Juntto, S.; Paatero, P.; Tapper, U.; Jarvinen O. Analysis of Daily Precipitation Data by Positive Matrix Factorization. *Environmetrics* **1994**, *5*, 127−144.

(54) Antilla, P.; Paatero, P.; Tapper, U.; Jarvinen, O. Source identification of bulk wet deposition in Finland by positive matrix factorization. *Atmos. Environ.* **1995**, *29*, 1705−1718.

CURVE RESOLUTION METHODS IN IMAGING SPECTROSCOPY

J. Chem. Inf. Comput. Sci., Vol. 43, No. 6, 2003 **2067**

(55) Polissar, A. V.; Hopke, P. K.; Malm, W. C.; Sisler, J. F. The ratio of aerosol optical absorption coefficients to sulfur concentrations, as an indicator of smoke from forest fires when sampling in polar regions. *Atmos. Environ.* **1996**, *30*, 1147−1157.

(56) Miller, S. L.; Anderson, M. J.; Daly, E. P.; Milford, J. B. Source apportionment of exposures to volatile organic compounds. I. Evaluation of receptor models using simulated exposure data. *Atmos. Environ.* **2002**, *36*, 3629−3641.

(57) Paatero, P.; Tapper, U. Analysis of different modes of factor analysis as least squares problems. *Chemom. Intell. Lab. Syst.* **1993**, *18*, 183−194.

(58) Tauler, R.; Barceló, D. Multivariate curve resolution applied to liquid chromatography−diode array detection. *TrAC Trends Anal. Chem.* **1993**, *12*, 319−327.

(59) Tauler, R.; Kowalski, B.; Fleming, S. Multivariate curve resolution applied to spectral data from multiple runs of an industrial process. *Anal. Chem.* **1993**, *65*, 2040−2047.

(60) Tauler, R. Calculation of maximum and minimum band boundaries of feasible solutions for species profiles obtained by multivariate curve resolution. *J. Chemom.* **2001**, *15*, 627−646.

(61) De Juan, A.; Rutan, S. R.; Tauler, R.; Massart, D. L. Comparison between the direct trilinear decomposition and the multivariate curve resolution-alternating least squares methods for the resolution of three-way data sets. *Chemom. Intell. Lab. Syst.* **1998**, *40*, 19−32.

(62) Tauler, R.; Smilde, A.; Kowalski, B. Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution. *J. Chemom.* **1995**, *9*, 31−58.

(63) Forina, M.; Drava, G.; Armanino, C.; Boggia, R.; Lanteri, S.; Leardi, R.; Corti, P.; Conti P.; Giangiacomo, R.; Galliena, C. Transfer of calibration function in near-infrared spectroscopy. *Chemom. Intell. Lab. Syst.* **1995**, *27*, 189−203.

(64) Gargallo, R.; Tauler, R.; Izquierdo-Ridorsa, A. Application of a Multivariate Curve Resolution Procedure to the Analysis of Second-Order Melting Data of Synthetic and Natural Polynucleotides. *Anal. Chem.* **1997**, *69*, 1785−1792.

(65) Vives, M.; Gargallo, R.; Tauler, R. Study of the Intercalation Equilibrium between the Polynucleotide Poly(adenylic)-Poly(uridylic) Acid and the Ethidium Bromide Dye by Means of Multivariate Curve Resolution and the Multivariate Extension of the Continuous Variation and Mole Ratio Methods. *Anal. Chem.* **1999**, *71*, 4328−4337.

(66) De Braeleleer, K.; Massart, D. L. Evaluation of the orthogonal projection approach (OPA) and the SIMPLISMA approach on the Windig standard spectral data sets. *Chemom. Intell. Lab. Syst.* **1997**, *39*, 127−141.

(67) Navea, S.; De Juan, A.; Tauler R. Three-way data analysis applied to multispectroscopic monitoring of protein folding. *Anal. Chim. Acta* **2001**, *446*, 185−195.

(68) Manne, R. On the resolution problem in hyphenated chromatography. *Chemom. Intell. Lab. Syst.* **1995**, *27*, 89−94.