Multicomponent Determination of Chlorinated Hydrocarbons Using a Reaction-Based Chemical Sensor. 2. Chemical Speciation Using Multivariate Curve Resolution

R. Tauler,[†] A. K. Smilde,[‡] J. M. Henshaw,[§] L. W. Burgess, and B. R. Kowalski[*]

Center for Process Analytical Chemistry, University of Washington, BG-10, Seattle, Washington 98195

A new multivariate curve resolution method that can extract analytical information from UV/visible spectroscopic data collected from a reaction-based chemical sensor is proposed. The method is demonstrated with the determination of mixtures of chlorinated hydrocarbons by estimating the kinetic and spectral profiles of the chemical species formed in the Fujiwara reaction. The three key aspects of the proposed method are (1) the initial estimation of the kinetic concentration profiles from evolving factor analysis; (2) the implementation of an alternating and constrained least squares method to optimize the determination of both the spectral and concentration profiles of the species present in the reaction, and (3) the development of a quantitative approach based on the simultaneous analysis of standards and unknowns for the determination of the initial concentration of the analytes in the mixtures.

Multivariate curve resolution has seen little application to the study of kinetic reactions despite its potential for the interpretation of the mechanisms of such reactions and their application in analytical chemistry.[1] The description and modeling of the evolution of a time-dependent chemical reaction is important for both theoretical and practical analytical reasons. Most of the kinetic-based analytical methods in the literature are of an univariate nature, i.e., only make use of one channel or wavelength in the analysis.[2,3] Even when multicomponent analysis is performed with these methods, the distinction between analytes (selectivity) is accomplished from univariate measurements based on the differences between rate constants and not on their global different spectroscopic responses. The methods described in the literature[4] that use a full-spectrum approach are chemical model-based methods (hard model approach), which means that the model has to be precisely determined and the parameters of that model estimated at the same time that the analytical information is obtained. The large number of parameters to estimate increases the possibility of uncertainties and artifacts in the numerical treatment.

In contrast to this situation, self-modeling curve resolution methods do not rely on the initial proposal of a specific kinetic model and do estimate directly the changes in concentration and unit responses of the species formed in the kinetic reaction without any previous assumptions. It is possible to determine what kind of kinetic model the reaction follows and the best approach to use that kinetic reaction for analytical purposes from the changes in concentration of the different species during the kinetic reaction.

Curve resolution decomposes a bilinear data matrix into the product of two simpler matrices which are related respectively to each one of the two orders (e.g., spectral and temporal) of the original data matrix. The goal of the curve resolution methods is the determination of those decompositions which have physical and chemical meaning.[5,6] The experiment involves periodically acquiring spectra during the progress of a chemical reaction. The spectra are placed in a matrix which is, hopefully, decomposed to yield one matrix containing the pure spectra of the individual reacting components and another containing concentration profiles as a function of time. Obtaining meaningful solutions from curve resolution decompositions requires some assumptions about the data such as nonnegative spectral intensities or closure.

However, in general, such treatments do not guarantee unique solutions because the rotational and intensity ambiguities inherent to factor analysis decompositions can still be present after applying the above-mentioned constraints.[7,8] As shown in the present work, these ambiguities can be partly overcome with the use of some special techniques. One of the most interesting which has not received much attention in kinetic analysis is evolving factor analysis (EFA).[9,10] EFA provides valuable information concerning the windows of existence and relative importance of the species formed in the unknown mixtures at any time during the kinetic reaction. When evolving factor analysis is applied to cases where selectivity for some component is present in any of the two orders, the determination of the concentration profile and spectroscopic response of such a component can be estimated at least qualitatively without any other additional requirement.

In addition, if different kinetic determinations are performed over different samples with different concentrations

† Present address: Department of Analytical Chemistry, University of Barcelona, Diagonal 647, Barcelona 08028, Spain.
‡ Present address: Laboratory for Analytical Sciences, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands.
§ Present address: ORS Environmental Systems, 7 Barnabas Rd., Marion, MA 02738.
(1) Mottola, H. A.; Perez-Bendito, D. *Anal. Chem.* 1992, *64*, 407R–428R
(2) Mottola, H. A.; *Kinteic Aspects of Analytical Chemistry*, John Wiley: New York, 1988.
(3) Perez-Bendito, D.; Silva, M. *Kinetic methods in Analytical Chemistry*, John Wiley: New York, 1988.
(4) Pardue, H. L. *Anal. Chim. Acta* 1989, *216*, 69–107.

(5) Hamilton, J. C.; Gemperline. P. J. *J. Chemom.* 1990, *4*, 1–14.
(6) Windig, W. *Chemom. Intell. Lab. Syst.* 1992, *16*, 1–16.
(7) Lawton, W. H.; Sylvestre, E. A. *Technometrics* 1971, *13*, 617–632.
(8) Borgen, O. S.; Kowalski, B. R. *Anal. Chim. Acta* 1985, *174*, 1–26.
(9) Gampp, H.; Maeder, M.; Meyer, Ch.; Zuberbuhler, A. D. *Talanta* 1985, *32*, 1133–1139.
(10) Gampp, H.; Maeder, M.; Meyer, Ch.; Zuberbuhler, A. D. *Talanta* 1986, *33*, 943–951.

of the analytes and at least one of the two orders is common between them (e.g., the same wavelength channels in the different spectra), the intensity ambiguities associated with the analysis of a single kinetic determination can be resolved.[11,12] Assuming that the same species in the different experiment has the same unit spectrum, the simultaneous analysis of the different experiments will give the relative amounts of the common species in the different mixtures analyzed. From these relative amounts, it is possible to estimate the initial concentration of the analytes in the samples if a set of standards is provided in the analysis. In the third paper of the present series,[13] a different approach is described that takes advantage of the full second order structure of the data[14,15] and makes possible the quantitation of the analytes in the presence of unknown and unmodeled interferents.

Multivariate reaction-based chemical sensors couple a selective chemical reaction with a multivariate analysis technique such as those based on spectroscopy. In the previous paper of this series[16] the Fujiwara reaction-based chemical sensor for the determination of mixtures of organohalides has been described.

In the Fujiwara reaction, halogenated hydrocarbons react with pyridine in basic medium giving different species, some of them characteristic of each analyte, and others common or very similar for different analytes.[17,18] The three analytes studied are 1,1,1-trichloroethane (TCA), trichloroethylene (TCE) and chloroform. Although the analytes themselves do not absorb in the spectral region under study, the product species of the Fujiwara reaction with these three analytes can be monitored using UV/visible spectroscopy.

However, the reaction is not selective for similar organohalides, and a considerable level of overlap and correlation is present in the two order responses of the different analytes, rendering traditional calibration unable to optimally analyze unknown mixtures of the analytes. In the present work, a recently developed multivariate curve resolution method[11,12] is used to determine the time-dependent concentration profiles and unit spectral responses of the species formed in the Fujiwara reaction. Resolving the pure species' spectral and temporal profiles enables the study of the reaction kinetics and the development of a method for the quantitation of the analytes in the mixtures.

EXPERIMENTAL SECTION

The experimental data used to illustrate the proposed multivariate curve resolution method was obtained using the experimental setup described in the Experimental Section of the first paper of the present series.[16] The selected experiments were those where the kinetic reaction was monitored in a thermostated spectrophotometric cuvette. The reagents and analytes were manually mixed, and the spectra were acquired at evenly spaced times using a diode array spectrophotometer
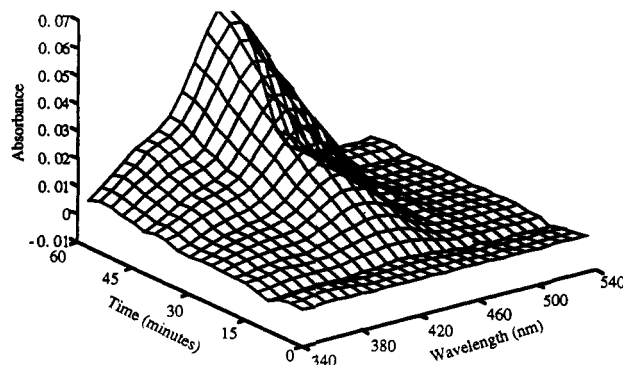
(11) Tauler, R.; Izquierdo-Ridorsa, A.; Casassas, E. Chemom. Intell. Lab. Syst. 1993, 18, 293–300.
(12) Tauler, R.; Fleming, S.; Kowalski, B. R. Anal. Chem. 1993, 65, 2040–2047.
(13) Smilde, A. K.; Tauler, R.; Henshaw, J. M.; Burgess, L. W.; Kowalski, B. R. Anal. Chem., third paper of the three in this issue.
(14) Sanchez, E.; Kowalski, B. R. J. Chemometrics 1988, 2, 265–280.
(15) Smilde, A. K. Chemom. Intell. Lab. Syst. 1992, 15, 143–157.
(16) Henshaw, J. M.; Burgess, L. W.; Booksh, K. S.; Kowalski, B. R. Anal. Chem. first paper of the three in this issue.
(17) Lugg, G. A. Anal.Chem. 1966, 38, 1532–1536.
(18) Reith, J. F.; van Ditmarsch, W. C.; de Ruiter, Th. Analyst, 1974, 99, 652–656.

Figure 1. TCA (66.9 ppm) response in the Fujiwara reaction. Absorption increases during the reaction.
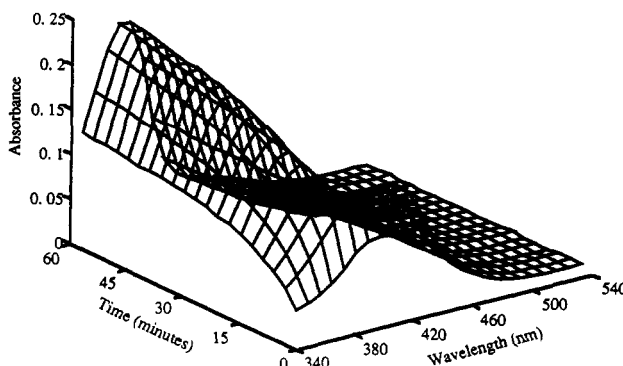

Figure 2. TCE (7.31 ppm.) absorption response in the Fujiwara reaction. Absortion increases and shifts to shorter wavelengths during the reaction.
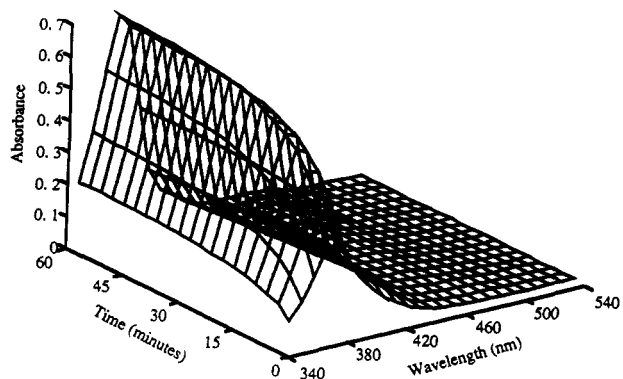

Figure 3. CHCl₃ (4.45 ppm) absorption response in the Fujiwara reaction. Absorption increases and shifts to shorter wavelengths during the reaction.

and stored in the computer. The temperature was held constant at 20 °C.

In Figures 1–3, examples of the spectra acquired during the analysis of the pure analytes 1,1,1-trichloroethane (TCA), trichloroethylene (TCE), and chloroform (CHCl₃) are respectively given. TCA shows a continuous growth of the same profile during the time of the experiment. TCE and CHCl₃ spectral profiles changed considerably in the first part of the experiment, suggesting that more than one chemical species is present in both systems. Observe also the presence of an isobestic point in the case of TCE. In Figure 4, an example of the spectra acquired in the analysis of a mixture of the three analytes is also given. In this figure, the contribution of TCA is at the higher wavelengths, whereas TCE and CHCl₃ are overlapped in the lower wavelength region.
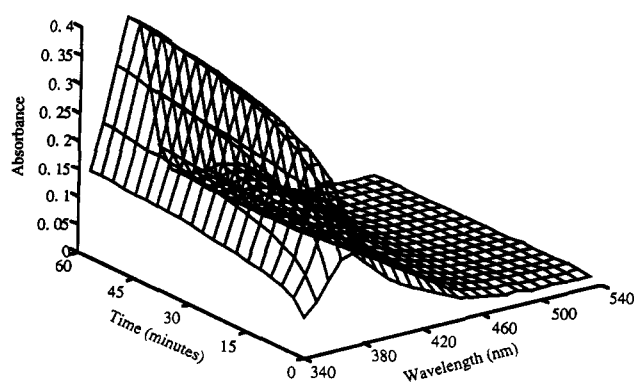
Figure 4. Absorption response of the mixture of TCA (66.9 ppm), TCE (2.93 ppm) and CHCl₃ (1.48 ppm).

In the following, it will be assumed that a plurality of spectra have been acquired during the analysis of each sample at evenly spaced time intervals until the kinetic reaction has reached a determined level. The spectral intensities are stored in a matrix **D** with one spectrum per row. The dimensions of **D** are the number of spectra acquired at different points in time during the reaction by the number of spectral channels (e.g., wavelengths). Each analyzed sample is characterized by a data matrix. The goals of the study are to determine the number of chemical species formed during the reaction, their pure spectra and concentration/time profiles, and the quantitation of the amount of analyte initially present in the sample. The term analyte will be used for the chemical compounds TCA, TCE, and chloroform to be determined in the analysis. The concentration of the analytes refers then to the initial concentration of these three compounds in the original samples before the reaction takes place. During the reaction, these analytes are consumed and new chemical species are formed. The term species refers to the actual different chemical species present at any time during the reaction. Not all species formed in the reaction can be detected spectroscopically; in particular, the species corresponding to the analytes themselves are not detected in the spectral range under study, and only some of the products of the reaction are detected. The total number of species in the system will be at least equal to or higher than the total number of analytes.

THEORY

Linear model. The linear model proposed is based on three fundamental assumptions which are validated during the analysis of the data. The eventual confirmation of such a model and of the underlying assumptions is derived from the final results obtained using the proposed multivariate curve resolution method.

Generalized Lambert–Beer's Law. The first important assumption relies on the matrix generalization of Beer's law. The spectroscopic data matrices obtained in the kinetic experiments are bilinear, such that

$$D = CB \qquad (1)$$

$$d_{IJ} = \sum c_{IN} b_{NJ} \qquad (2)$$

where **D** is the data matrix which contains the spectra acquired

at fixed time intervals during the reaction. **C** and **B** are respectively the matrices of the concentration profiles and of the unit pure spectra for the spectroscopically active chemical species involved in the reaction. The dimensions of these matrices are $D(I \times J)$, $C(I \times N)$ and $B(N \times J)$, where I is the number of spectra analyzed, J is the number of spectroscopic channels, and N is the number of chemical species in the mixtures. The goal of curve resolution methods is to obtain the actual **C** and **B** matrices from **D**.

Rank Analysis and Complexity of the Systems under Study. The second assumption is that the formation of the different chemical species in the pure analyte samples and in the mixture samples depends only on the initial concentration of the analyte. In addition, there must be no new species and no new interaction between the analytes, and consequently between the species, in the mixtures. To check these assumptions, the rank (pseudorank) of the data matrices obtained in the analysis of the pure analyte samples and of the data matrices obtained in the analysis of mixture analyte samples are calculated and compared in the following way:

(i) The rank of the pure analyte data matrices gives the number of spectroscopically active species given by that analyte in the kinetic reaction. As was mentioned before, the analytes do not absorb in the spectral range under study and will not contribute to the rank of the corresponding spectral data matrices. However, the rank associated with the global response of the analytes (TCA, TCE, CHCl₃) can be greater than 1, since they can produce more than one species in the chemical reaction.

(ii) The rank of the data matrices containing more than one analyte must be equal to or less (when some species are in common) than the sum of the ranks associated with the pure analyte responses present in the particular mixture, which is equal to the total number of spectroscopically active species identified in the analysis of the corresponding pure analyte samples.

(iii) The rank of the augmented data matrices formed by stacking the sample data matrices columnwise or rowwise (Figure 5) must be equal to or less than the total number of species identified in the analysis of the individual data matrices used to build the augmented data matrix.

When these three conditions are met, the pure spectral and kinetic concentration profile for every species in the different experiments is equal and allows the application of second-order methods.[13]

Kinetic Model. The kinetic experiments are assumed to be carried out under reproducible conditions. This assumption means that the concentration of the different species formed during the reaction of the analytes is always linearly related to the initial concentration of the analyte or parent molecule. For this assumption to be true, kinetic conditions of pseudofirst order hold.

The following kinetic model is proposed as the basis of the work. For TCA, the analyte gives irreversibly (because of the large excess of the other reagents) a single detectable product in the reaction (see below)

$$TCA + py + OH^- \rightarrow P_1 \qquad (3)$$
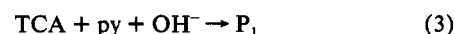
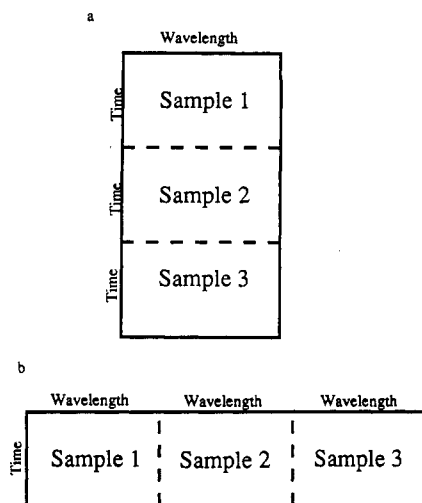TCA is the analyte and P₁ is the product; py is the pyridine

Figure 5. (a) Augmented data matrix formed by stacking columnwise (wavelength) three sample data matrices; (b) augmented data matrix formed by stacking rowwise (time) three sample data matrices.

reagent and OH^- is the base. Under excess concentration of pyridine and base, the reaction is of pseudofirst order. The concentration of TCA and of the product P_1 at any point in time during the reaction is directly proportional to the initial concentration of TCA. As the only species spectroscopically active in the wavelength region under study is the product P_1, measuring its concentration allows the determination of the initial concentration of TCA when compared to a calibration standard.

For TCE and $CHCl_3$, the situation is more complex because they give two products in a two-step reaction. In excess pyridine and base, the situation can be described as coupled first-order irreversible reactions

$$TCE + py + OH^- \rightarrow P_2 \rightarrow Q_1 \qquad (4)$$

$$CHCl_3 + py + OH^- \rightarrow P_3 \rightarrow Q_2 \qquad (5)$$
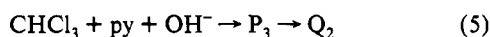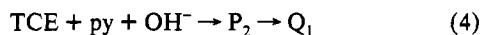
Solving the kinetic differential equations for the intermediate or transient products P_2 or P_3 gives the relation between the rate of change in concentration of them and the initial concentration of analyte, which is also linear[2,19,20] if the conditions of pseudofirst order prevail in the two stages of the reaction. The concentrations of the intermediate product and of the final product, species P, are linearly related to the initial concentration of the analyte. Since in the wavelength region under study, the only species which are spectroscopically active (see Experimental Section) are the intermediates and products of the reaction, monitoring the spectral changes will give the concentration changes of these species and provide the initial concentration of the analytes.

Multivariate Curve Resolution. Introduction. Data analysis can be carried out over one sample (one data matrix) or over different samples (several data matrices) simultaneously. Suppose there are K different samples at different initial conditions or starting concentrations of the analytes. In the analysis of each sample a data matrix D_k is obtained:

(19) Hughes, E. *J. Chem. Educ.* **1989**, *66*, 46.
(20) Schechter, I. *Anal. Chem.* **1991**, *63*, 1303-1307.

$$D_k = C_k B, \quad k = 1, 2, ..., K \qquad (6)$$

C_k is the matrix of the time/concentration profiles of the chemical species spectroscopically active formed during the analysis of sample k, and B is the matrix of the unit or pure spectra of these species. As the number of columns (wavelengths) is the same for all the D_k matrices, the analysis can be performed simultaneously over more than one data matrix, as follows:

$$D = \begin{bmatrix} D_1 \\ D_2 \\ ... \\ ... \\ D_K \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \\ ... \\ ... \\ C_K \end{bmatrix} B = CB \qquad (6)$$

The new augmented data matrix D is obtained by setting each of the sample data matrices D_k to be analyzed on top of the others, with the columns (wavelengths) in common and with a number of rows equal to the total number of acquired experimental spectra in the different sample analyses. In the case of the analysis of only one data matrix, this equation reduces to the usual case of the generalized Beer's law shown before. In the case of simultaneous analysis of different samples, the new augmented data matrix will be the product of an augmented concentration matrix times the unit spectra matrix. The augmented concentration matrix will contain the different submatrices C_k corresponding to the concentration of the species present in each of the data matrices D_k analyzed.

Selectivity and Evolving Factor Analysis. It is extremely important in multivariate curve resolution to detect and use the selective information present in the data to remove the rotational ambiguity. Evolving factor analysis[9,10] has been applied mostly to the study of spectroscopic experiments of multiequilibria systems[21,22,23] and to liquid chromatography with diode array detection.[24,25] The basic idea of this procedure is to provide an initial estimation of the concentration profiles by examining how the singular values evolve and change in magnitude the experiment (kinetic reaction in this case). From the results of evolving factor analysis, the selective regions in the data are usually detected, and the rotational ambiguities eventually solved.

Other similar methods have been proposed to detect and use the selectivity of the system such as local rank analysis,[26] window factor analysis,[27] fixed-size moving window evolving factor analysis,[28] the pure variable self-modeling mixture analysis method,[29] and the HELP method.[30] In the case of kinetic reaction-based data, it is better to take advantage of the intrinsic order in the time domain as in evolving factor analysis.

(21) Tauler, R.; Casassas, E. *J.Chemom.* **1988**, *3*, 151-161.
(22) Tauler, R.; Casassas, E. *Anal. Chim. Acta* **1989**, *223*, 257-268.
(23) Tauler, R.; Casassas, E.; Izquierdo-Ridorsa, A. *Anal. Chim. Acta* **1991**, *248*, 447-458.
(24) Maeder, M. *Anal. Chem.* **1987**, *59*, 527-530.
(25) Tauler, R.; Durand, G.; Barcelo, D. *Chromatographia* **1992**, *33*, 244-254.
(26) Geladi, P.; Wold, S. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 273-281.
(27) Malinowski, E. R. *J. Chemometrics* **1992**, *6*, 29-40.
(28) Keller, H. R.; Massart, D. L. *Anal. Chim. Acta* **1991**, *246*, 379-390.
(29) Windig, W.; Stephenson, D. A. *Anal. Chem.* **1992**, *64*, 2735.
(30) Kvalheim, O. M.; Liang, Y.-z. *Anal. Chem.* **1992**, *64*, 936-946, 946-956.

Constrained Alternating Least-Squares Optimization of the Concentration Profiles and Unit Spectra. Analysis of a Single Kinetic Experiment. From the results of evolving factor analysis, the windows or range of existence of each species and its concentration profile can be obtained. These concentration profiles are used as initial values in a constrained alternating least-squares optimization procedure. At each iteration of the optimization a new estimation of the matrix of spectra B and of the concentration profiles C is obtained successively using the two following equations:

$$B = C^+D^* \tag{7}$$

$$C = D^*B^+ \tag{8}$$

where the matrix C^+ is the pseudoinverse of the matrix C and the matrix B^+ is the pseudoinverse of the matrix B. Obviously, the selection of the right number of components in the calculation of D^* is essential. The use of this matrix instead of the experimental data matrix D will provide more stability in the calculations, since D^* is noise-filtered for the particular number of principal components.

In order to limit the number of possible solutions to these equations, the unit spectra are constrained to be nonnegative and the time/concentration profiles are constrained to be nonnegative[5–11,21–23,25] and of unimodal shape.[11,21–23,24,31] However, as was mentioned before, the solutions obtained from the analysis of a single sample data matrix are still ambiguous in their scale/intensity dimensions, and quantitation is not possible from them.

Simultaneous Analysis of Multiple Samples (Calibration and Unknowns). As a first step in the curve resolution method, the augmented matrix of the concentration profiles C is built up from the initial estimation of the $C_1, C_2, ..., C_K$ submatrices obtained by evolving factor analysis of the $D_1, D_2, ..., D_K$ submatrices, as described before. An initial estimation of the augmented C data matrix is then obtained simply by setting the estimation of the C_i matrices on top of each other in the same order as they are in D. The question about the presence or absence of a certain species in an experiment can be solved by looking at the pseudorank of the individual data matrices C_i and also by looking at the results of the individual alternating least-squares optimization. If the species spectra of two resolved species in different experiments are similar (fingerprint matching), then probably they correspond to the same species and so they will be postulated in the simultaneous analysis (common columns in matrix C and common rows in matrix B). Conversely, species with different spectra in different kinetic experiments correspond to different species in the simultaneous analysis. At each iteration of the alternating least-squares method a new estimation of the matrix of spectra B and of the concentration profiles C are obtained. The same procedure as before is used except now it is applied to the augmented matrix obtained by principal component analysis for the number of components D^*. In addition to the constraints present in the single data matrix analysis, in the simultaneous analysis of multiple data matrices, there are two more constraints to apply:

(i) common species have the same spectrum in all the kinetic experiments[11,12]

(ii) common species have concentration profiles with equal shape in all the kinetic experiments:

Let C^n the nth column of the augmented matrix C contain the current estimation of the concentration profiles of species n in the different K samples analyzed. This column can be folded in a matrix $C'(I,K)$, where I is the number of spectra acquired in the analysis of every sample (synchronization is assumed between experiments). Under the assumption of equal shape, this matrix can be decomposed as the product of two vectors:

$$C' = u'^*t^T \tag{9}$$

where $u'(I)$ is the unit concentration profile containing the common shape information of C', and $t^T(K)$ contains the relative amounts of this unit profile in C'. The unit profile u' is equal to the first score vector of C' and the relative analyte concentrations t are obtained from the first loading of C' scaled by its first singular value.

This method assumes that the concentration profiles of the same species have the same shape in the different experiments and consequently that the matrix C' is rank one and E is at the level of the experimental error. An easy way to test this assumption during the analysis is to compare the magnitude of the next singular values with the first singular value.

(iii) Zero concentration components: when a species is known not to be present in the specific kinetic experiment, then the concentration of such species in C is forced to be equal to zero.

It is important to note that the use of the three described constraints leads to the second-order advantage[14,15] and the solutions obtained with the proposed method are very close to those obtained by using the second-order resolution methods.[13]

Testing the Number of Chemical Species. If the number of species is uncertain, the complete iterative alternating least-squares optimization is performed for the most plausible number of components. At the end, it is considered that the right number is the one which gives a best fit to the experimental data matrix D (not to D^*) and physically meaningful solutions. This means that only those solutions with reasonable shapes for the unit spectra and concentration profiles are considered. The alternating least-squares procedure is repeated until convergence is achieved or until a predetermined number of cycles has occurred.

Quantitation of the Analytes. Once the concentration profiles have been determined using the proposed multivariate curve resolution method, the initial concentration of the analytes can be estimated. If the conditions previously described for linearity hold, the concentrations of the species formed at any point of the reaction are linearly related with the initial concentration of the analytes. Different methods of quantitation are tested and validated in the present work. The area under the concentration profile of a certain species (which can be determined from the sum of the concentration values of this species at the different measured times) will be then proportional to the initial concentration of the analyte which produced that species in the reaction.

(i) The simpler method of quantitation consists first of estimating the pure spectra of the different species from the

analysis of known pure analyte samples. Once they are determined, the concentration of the species present in any mixture of the analytes can be determined by least squares analysis from

$$C_{unk} = D_{unk}B^+ \tag{10}$$

where B^+ is the pseudoinverse of the unit spectra matrix B obtained in the analysis of the pure analyte samples, D_{unk} is the matrix of a new unknown sample, and C_{unk} is the matrix of the unknown kinetic concentration profiles.

The initial concentrations of the analytes in D_{unk} can then be determined from the ratio of the areas under the concentration profiles of the species in the unknowns compared with the areas under the concentration profiles of the corresponding species in the standards.

(ii) If several samples are analyzed simultaneously with the proposed procedure, the ratio between the areas of the concentration profiles for a particular species in the different samples will give the ratio between the initial concentrations of the analyte which produce this species in the reaction. If some of the samples are standards of known analyte concentration, the concentration of the analytes in unknown samples can be determined.

This method of quantitation is improved if the constraint of equal shape over the concentration profiles is applied. Under complete reproducible pseudo-first-order conditions, the intensity of the concentration profiles of the different species in the different experiments will only depend upon the initial concentration of the analyte. The relative amounts of the analytes in the different samples are directly recovered from the simultaneous constrained alternating least-squares optimization previously described.

RESULTS AND DISCUSSION

Rank Analysis. Results of rank analysis confirm that TCA only gives one spectroscopically active species, whereas TCE and CHCl3 give two spectroscopically active species. From the analysis of the pure analyte data matrices, a maximum number of five different species is detected. The rank of the augmented data matrices formed by several samples of the same analyte at different concentrations is equal to the rank of any of the single pure analyte data matrices, regardless of the order used to form the augmented matrix, the wavelength variables, or the time variables.

The pseudorank of the mixture analyte data matrices is equal to or less than the sum of the pseudoranks of the pure analyte data matrices. The rank of the augmented data matrices is also equal to or less than the sum of the ranks of the data matrices present in the augmented matrix. These results are in agreement with the initial hypothesis that the same species are formed in the mixtures and in the pure analyte samples (no new species are formed in the mixtures) and that the kinetic behavior of the different analytes is independent of the presence of the other analytes. No new interactions appear in the analysis of the mixtures. The assumption of equal shape in the time profiles for the different species is then also confirmed in the rank analysis of the augmented data matrices.
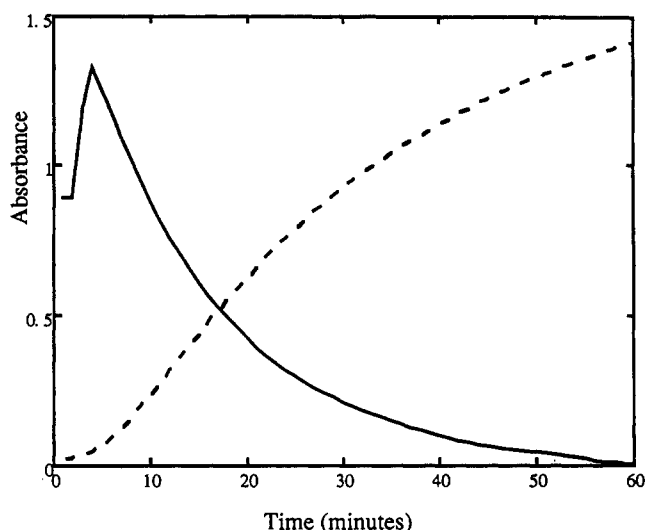
Figure 6. Concentration profiles of the two species obtained in the reaction of TCE initially estimated by evolving factor analysis (species 2, —; species 3, - - -).

Selectivity and Initial Estimates Using EFA. Evolving factor analysis of the different spectral data matrices provides the starting values for the concentration matrix in the alternating least-squares optimization. In Figure 6, the evolving factor analysis plot [9,10] of a pure TCE data matrix is shown for example. From this plot, an initial estimation of how the concentration of the two species change during the TCE reaction is obtained. Some selectivity for the two species given by TCE exists at the beginning and at the end of the reaction. Analogous observations can be made for CHCl3, although here the selectivity is lower. For TCA, only one spectroscopically active species is present and there is no rotational ambiguity in the curve resolution analysis for this analyte.

Identification of the Species from the Pure Analyte Sample Analysis. Figure 7 shows the resolved pure unit spectra of the species formed in the analysis of the pure analyte samples using the multivariate curve resolution method previously described. Several samples at different concentrations of the same analyte were analyzed simultaneously with the outlined procedure. High overlap is observed for the species spectra of TCE and CHCl3. Species 3 (TCE) and species 5 (CHCl3) have unit spectra with similar shape and could correspond to the same chemical species. The spectra of these two species are also close to the spectrum obtained by the degradation of pyridine at high base concentrations, in agreement with the mechanisms proposed for the Fujiwara reaction in the literature.[17,18] Species 3 and 5 gave concentration profiles with a different slope, showing that TCE and CHCl3 react with pyridine at different rates. The only significant differences between TCE and CHCl3 are in the spectra of the intermediates, species 2 and 4, respectively, indicating a possible way to differentiate and quantitate them in the mixtures containing both. TCA gives species 1 with a unique spectrum (around 480 nm) and a different time/concentration profile (slower reaction). Therefore quantitation of TCA in the mixtures is simpler. In part 3 of the present series,[13] a better treatment is proposed to measure the similarities between the time profiles and unit spectra of the different
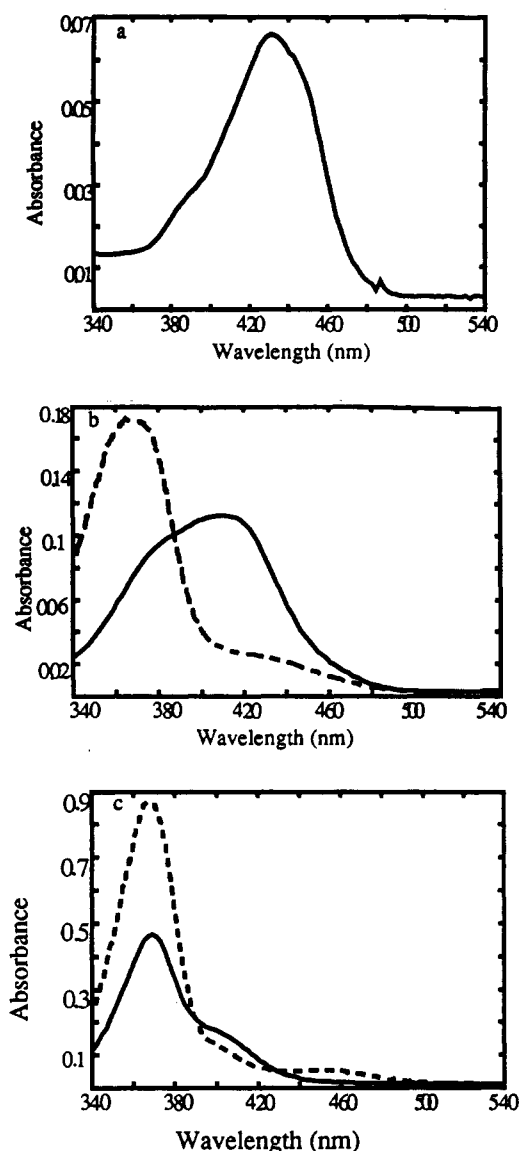
Figure 7. Unit spectra for the five detected species in the analysis of the pure analyte samples: (a) species 1 from TCA; (b) species 2 (—) and 3 (- - -) from TCE; (c) species 4 (—) and 5 (- - -) from CHCl₃.

Table 1. Results of Quantitation[a] Using the Pure Spectra

sample[b]	concn[c]	species[d]	area[e]	pred concn[f]	% error[g]
TCA	66.9	1	22.32	std	
TCA	133	1	45.98	138	3.6
TCA	267	1	100.5	301	13
TCA	267	1	87.29	262	2.0
TCE	7.31	2	22.43	7.43	1.6
		3	56.87	7.43	1.6
TCE	7.31	2	22.07	std	
		3	55.96		
TCE	29.3	2	75.89	24.7	15
		3	207.1	24.0	7.7
TCE	58.6	2	182.0	60.3	2.9
		3	451.5	59.0	0.6
TCE	1.46	2	3.340	1.11	24
		3	11.09	1.45	0.8
CHCl₃	1.48	4	7.150	1.48	0.1
		5	11.09	1.50	1.1
CHCl₃	4.45	4	21.53	std	
		5	32.98		
CHCl₃	7.42	4	41.80	8.65	16.6
		5	52.57	7.09	4.4
TCA	66.9	1	22.62	67.5	1.0
TCE	2.93	2	9.770	3.23	10
		3	20.63	2.69	8.0
CHCl₃	1.48	4	9.090	1.88	29
		5	14.35	1.94	31
TCE	2.93	2	11.57	3.83	31
		3	20.28	2.65	9.5
CHCl₃	1.48	4	22.42	4.63	4.0
		5	32.76	4.42	0.7
TCE	7.30	2	19.83	6.57	10
		3	47.44	6.19	15
CHCl₃	1.48	4	9.110	1.88	27
		5	13.13	1.77	19

[a] Results of the analysis of 15 samples using the pure species spectra. [b] Content of analyzed samples. [c] Concentration in ppm. [d] The species used for analyte quantitation (see Figure 7 for species identification). [e] Area under concentration profile (arbitrary units). [f] Estimated concentration of the analyte. [g] Percent error in the prediction.

species present in the system, as well as a global measurement of the similarities between the different analytes.

Quantitation of the Analytes Using the Pure Species Spectra. The results obtained in the analysis of the pure analyte samples confirm the initial assumptions made about the linear model for any of the species identified. Once the pure species spectra have been estimated from the pure analyte samples, quantitation in new samples is possible. Table 1 shows the results obtained by this procedure. For the quantitation of the pure analyte samples, the results are rather good, taking into account the experimental limitations[16] and that only one sample is taken as standard. The quantitation is not only possible in relative terms but also in absolute terms if the concentration of the analytes in some of the samples are known (standards). Quantitation using the pure species spectra is also done over four samples containing mixtures of the analytes. The results obtained are given in Table 1. The errors in prediction, although higher than in the analysis of pure analyte samples, are still reasonable when the high overlap between the TCE and CHCl₃ spectral and concentration profiles is considered.

The use of the pure species spectra for the quantitation of the analyte in the samples is a full-model approach; this means that all the data variance is explained by the given species spectra. This approach fails in the presence of unknown overlapping interferences because then the estimated concentrations of the analytes will be in error.

Quantitation by Simultaneous Analysis of Standards and Unknowns. A different way to perform the quantitation using multivariate curve resolution is by the simultaneous analysis of a set of standard and unknown samples. In Table 2, a summary of the results obtained when this approach is used is shown. Unknown mixtures and pure analyte samples are analyzed simultaneously. A total number of 15 data matrices are simultaneously analyzed. In order to have the absolute concentrations, three pure analyte samples are used as standards. The values obtained along with the estimation of the errors in prediction are shown in Table 2. The results obtained with the proposed method are comparable to the results obtained in the quantitation when the unit species spectra of the different species were used, which is reasonable because both approaches are full-model descriptive approaches and there are no interferents. In the mixtures of TCE and CHCl₃, worse results are obtained for the analyte contributing less to the signal (at lower concentration). This method has some advantages as previously described since the quantitation is achieved in one step and the concentration profiles can be constrained to be positive, unimodal, and of equal shape.

Table 2. Results of Quantitation[a] Using the Profile Shapes

sample[b]	concn	species[d]	rel concn[e]	pred concn[f]	% error[g]
TCA	66.9	1	11.96	std	
TCA	133	1	24.64	137.6	3.4
TCA	267	1	49.93	279.3	4.6
TCA	267	1	45.73	255.8	4.2
TCE	7.31	2	32.80	7.54	3.1
		3	72.70	7.47	2.2
TCE	7.31	2	31.82	std	
		3	71.14		
TCE	29.3	2	116.0	26.7	8.6
		3	246.8	25.3	13
TCE	58.6	2	270.6	62.2	6.1
		3	614.9	63.2	7.8
TCE	1.46	2	5.765	1.32	9.8
		3	12.47	1.28	13
CHCl$_3$	1.48	4	20.13	std	
		5	19.14		
CHCl$_3$	4.45	4	44.49	3.27	26
		5	59.22	4.58	2.9
CHCl$_3$	7.42	4	70.14	5.16	30
		5	92.89	7.18	3.2
TCA	66.9	1	10.99	61.5	8.1
TCE	2.93	2	14.57	3.35	14
		3	24.72	2.55	13
CHCl$_3$	1.48	4	14.04	1.03	30
		5	24.72	1.91	29
TCE	2.93	2	17.10	3.93	34
		3	0.029	****	****
CHCl$_3$	1.48	4	41.28	3.03	32
		5	57.76	4.47	0.9
TCE	7.30	2	30.89	7.09	2.9
		3	59.10	6.07	17
CHCl$_3$	1.48	4	15.14	1.11	25
		5	23.69	1.83	26

[a] Results of the analysis of 15 samples using the pure species spectra. [b] Content of analyzed samples. [c] Concentration in ppm. [d] The species used for analyte quantitation (see Figure 7 for species identification). [e] Estimated relative concentration of the analyte compared to the other samples. [f] Estimated absolute concentration of the analyte. [g] Percent error in the prediction.

Table 3. Results of Quantitation in the Presence of Unknown Interferences

sample[a]	concn[b]	species[c]	rel concn[d]	pred concn[e]	% error[f]
TCE	7.31	2	22.99	std	
TCE	7.31	2	19.00	6.04	17
(CHCl$_3$)	(1.48)				
TCE	7.31	2	17.98	std	
TCE	7.31	2	16.55	6.72	8.0
(TCA)	(66.9)				
CHCl$_3$	4.45	4	28.98	std	
CHCl$_3$	4.45	4	22.87	3.51	21
(TCE)	(2.93)				

[a] Content of analyzed samples. [b] Concentration in ppm. [c] The species used for analyte quantitation (see Figure 7 for species identification). [d] Estimated relative concentration of the analyte compared to the other sample. [e] Estimated absolute concentration of the analyte. [f] Percent error in the prediction.

Quantitation in the Presence of Unknown Interferents. Finally, the curve resolution method presented here is tested by the quantitation of a certain analyte in the presence of an unknown interferent. The augmented data matrix is formed now by two data matrices, the unknown mixture data matrix containing the analyte and the interferent and the pure analyte data matrix. In Table 3, the results of three determinations are given for different mixtures. In the first case, the determination of TCE in a mixture of TCE and TCA is given; TCA is considered the interferent. As the first species of TCE is formed at earlier times in the reaction and its spectrum is quite different from that of TCA, quantitation is accurate (Table 3). In the second case, the determination of TCE in a mixture of TCE and CHCl$_3$ (now the interferent being CHCl$_3$) is given. The quantitation of TCE is also rather good using species 2 because of its unique spectral characteristics (see Figure 7b). Quantitation would be much worse if species 3 were used for the TCE quantitation, because this species

overlaps highly with the species given by CHCl$_3$. The third case is the quantitation of CHCl$_3$ in the presence of TCE as interferent. The results are now worse because of the high overlap between the spectra of the two species given by CHCl$_3$ and also with the species 3 given by TCE.

CONCLUSIONS

Multivariate curve resolution has been shown to be a very useful technique for extracting qualitative and quantitative information from kinetic reactions. Multivariate curve resolution improves the selectivity of the analytical determinations obtained with reagent-based chemical sensors. In particular, the proposed multivariate curve resolution method allowed the identification and interpretation of the products (UV/visible spectroscopically active) given by TCA, TCE, and CHCl$_3$ in the Fujiwara reaction. The quantitation of the three analytes both in pure analyte samples and in their mixtures was possible in spite of the high overlap of their analytical signals. The method is proposed for the multicomponent determination of mixtures of chlorinated hydrocarbons using a reaction-based chemical sensor.

ACKNOWLEDGMENT