ELSEVIER

Chemometrics and Intelligent Laboratory Systems 55 (2001) 125–132

Chemometrics and intelligent laboratory systems

www.elsevier.com/locate/chemometrics

Modeling kinetic spectrophotometric data of aminophenol isomers by PARAFAC2

Jacqueline M. Dueñas Cueva, Adriana V. Rossi, Ronei J. Poppi *

Instituto de Química, Universidade Estadual de Campinas, CP 6154 CEP 13083-970 Campinas, Brazil

Received 8 May 2000; accepted 5 December 2000

Abstract

This paper describes an application of PARAFAC2 in the modeling of kinetic-spectrophotometric data. The data were obtained by spectrophotometric monitoring of the reaction between aminophenol isomers and $NaNO_2$ in acid solution. The data set of various samples produces a three-way data array. This reaction has the property that isomers in different proportions produce different kinetic profiles. Due this property, PARAFAC2 is suitable to model the system because it does not assume parallel proportional kinetic profiles. The results with PARAFAC2 were satisfactory, it being possible to recover the spectral and kinetic profiles, as well as the initial isomer concentrations with good accuracy.

Keywords: Multi-way; PARAFAC2; Kinetic spectrophotometric analysis

1. Introduction

Analytical methods that generate two-dimensional data have become common in analytical chemistry. Excitation–emission fluorescence [1], MS/MS spectrometry [2], HPLC with spectrophotometric diode array detection [3], acid–base titrations using emission spectrofluorimetry at different excitation wavelengths [4] and kinetic spectrophotometric analysis [5] are some examples of these methods. An important characteristic of these methods is the large quantity of information obtained for only one sample. The situation is more complex when data of various samples are compiled, producing a three-way array of data. The development of multi-way methods [6–8] allows the use of the intrinsic tri-linear structure of the data and makes it possible to solve many practical analytical problems.

The incorporation of rapid scan multichannel detectors (Diode Array detector) [9] in spectrophotometry allows fast full spectra acquisition in the ultraviolet/visible region. It improves the study of fast kinetic systems, increasing their range of applications. Thus, kinetic analysis is an area of increasing interest in analytical chemistry.

This kinetic spectrophotometric approach, together with chemometric methods, represents a powerful tool to solve analytical problems as, for instance, quantification and separation of structural isomers, specifically position isomers, such as, *o*-aminophenol, *m*-aminophenol e *p*-aminophenol.

* Corresponding author.
E-mail address: ronei@iqm.unicamp.br (R.J. Poppi).

0169-7439/01/$ - see front matter
PII: S0169-7439(00)00129-5

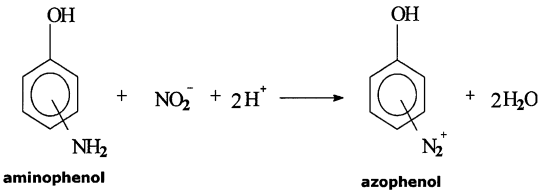

Fig. 1. Formation of azophenol from aminophenol and nitrite in acid medium.

These isomers react with NO_2^- by a reaction called diazotization [10], producing aryl diazonium salts that are more stable than aliphatic diazonium salts, as shown in Fig. 1. At pH values higher than 4, the reaction is slower and problems of coupling between the diazonium salt and aminophenol isomers can occur. The formation of the aryl diazonium salts in reactions with nitrite in acid medium (pH = 4) can be spectrophotometrically monitored and gives information about the initial reactant isomers, giving rise to kinetic based determinations.

Investigative kinetic measurements of the reaction of aminophenol isomer mixtures with nitrite in an acid medium were performed before starting the data treatment of this work. The mechanism does not seem to be trivial because changing the concentration of the reagents does not change the reaction rate in a direct way. Therefore, a multivariate approach must be employed in order to achieve the identification and the quantification of each isomer in the mixtures.

A data matrix is obtained by kinetic monitoring, in several wavelengths, of the reaction between aminophenol isomers and nitrite. The kinetic data of different isomer mixtures produces a three-way data array. In first mode are the absorbance measurements of the different samples; the second mode consists of absorbance measurements at different times and the third mode is absorbance measurements at several wavelengths. A representation of the three-way data generated is presented in Fig. 2.

There are several multi-way methods to process superior order data, such as Partial Least Squares (PLS three way) [11,12], Generalized Rank Annihilation method (GRAM) [13], Direct Trilinear Decomposition (TLD) [14], Tucker [15], PARAFAC1 [16,17] and PARAFAC2 [18–20].

In this work, PARAFAC1 was initially used, but the results were not satisfactory, probably due to the non-uniformity on the changes of reaction rate with the changes in reagent concentrations. This behavior can not be modeled by PARAFAC1 since it assumes constant profiles in all dimensions. In this case, a kinetic constant profile in the second mode is assumed. PARAFAC2 was used instead, because the method produces different profiles in the second mode and maintains constant profiles of the first and third mode, which seems to be quite suitable for the present system.

1.1. PARAFAC1

PARAFAC1 is a generalization of Principal Components Analysis (PCA) [21] used for multiway data modeling. A condition necessary for modeling of three-way data, represented by $\underline{\mathbf{X}}$, with PARAFAC1, is that it can be modeled with the Direct Trilinear Decomposition (TLD), given by the following equation:

$$\underline{\mathbf{X}} = \sum_{f=1}^{F} \mathbf{a}_f \otimes \mathbf{b}_f \otimes \mathbf{c}_f \qquad (1)$$

where $\mathbf{a}_f$, $\mathbf{b}_f$ and $\mathbf{c}_f$ are the fth columns of loadings matrices **A**, **B** and **C**, respectively.

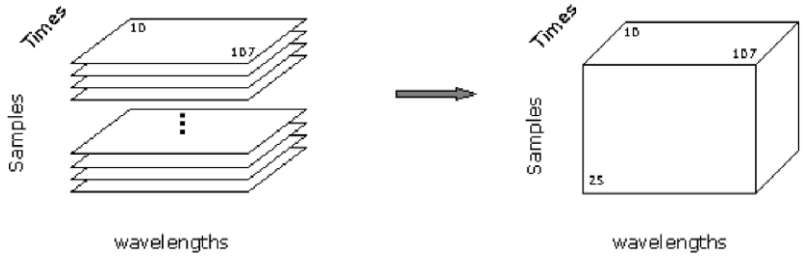

Fig. 2. Example of three-way data generated by reaction monitoring.

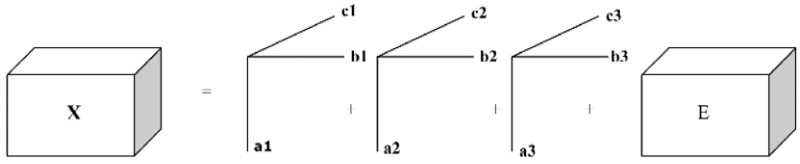

Fig. 3. Decomposition of three-way data array for PARAFAC1 using three components.

For all matrix representations, the following correspondence is used: lowercase is used for scalars, lowercase bold for vectors, uppercase bold for two-way matrices and underlined uppercase bold for a third order tensor. The transpose of a matrix is represented by superscript "T". The symbol ⊗ represents the Kronecker product, and the symbol $|\otimes|$ represents the Khatri–Rao product.

A graphic representation of the decomposition with PARAFAC1 is show in Fig. 3. In this figure, the vectors **a1**, **a2** and **a3** form the matrix loading **A**, the vectors **b1**, **b2** and **b3** form the matrix loading **B** and the vectors **c1**, **c2** and **c3** form the matrix loading **C**.

The tensor $\underline{\mathbf{X}}$ rearranged as a matrix with dimension $I \times JK$ can be modeled by following expression:

$$\mathbf{X} = \mathbf{A}(\mathbf{C}\,|\otimes|\,\mathbf{B})^{\mathrm{T}} \tag{2}$$

A slab, $\mathbf{X}_k$, can be calculated by:

$$\mathbf{X}_k = \mathbf{A}\mathbf{D}_k\mathbf{B}^{\mathrm{T}} \tag{3}$$

Where $\mathbf{X}_k$ is a slab of dimension $I \times J$ and $\mathbf{D}_k$ is the diagonal formed by elements of the kth row of **C**.

1.2. PARAFAC2

In some cases, the PARAFAC1 model can not be applied because the three-way data structure is not trilinear. The loss of trilinearity can be produced by sampling problems, physical properties or chemical interactions. Also, it can not be used when the variables of the different slabs are not comparable. Some problems in the use of PARAFAC1 can be solved by the use of PARAFAC2. This method was developed to overcome the difficulty in use PARAFAC1 in cases in which a violation of the assumption of parallel proportional profiles occurs.

The model of PARAFAC2 [22] is given by the following mathematical expression:

$$\mathbf{X}_k = \mathbf{A}\mathbf{D}_k(\mathbf{P}_k\mathbf{B})^{\mathrm{T}} \quad , k = 1, \ldots, K \tag{4}$$

where: $\mathbf{X}_k$ is a slab of data $(I \times J)$ and K is the number of slabs. **A** are the scores for the first mode loadings, $\mathbf{D}_k$ is a diagonal matrix that holds the kth row of **C** in its diagonal, and **C** is the third mode loadings. **B** is an $F \times F$ matrix and $\mathbf{P}_k$ is a $J \times F$ orthogonal matrix, where F is the number of components.

Eq. (4) can be rewritten as:

$$\mathbf{X}_k = \mathbf{A}\mathbf{D}_k\mathbf{B}_k^{\mathrm{T}} \tag{5}$$

and:

$$\mathbf{B}_k = \mathbf{P}_k\mathbf{B} \tag{6}$$

Eq. (6) calculates the kth second mode loadings, where $\mathbf{B}_k$ has dimensions $J \times F$.

A graphical representation of $\underline{\mathbf{X}}$ decomposed by PARAFAC2 is shown in Fig. 4. In this figure, the vectors **a1**, **a2** and **a3** form the matrix loading **A**, and the vectors **c1**, **c2** and **c3** form the matrix loading **C**.

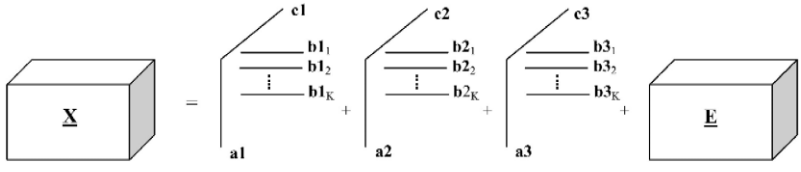

Fig. 4. Decomposition of three-way data by PARAFAC2 using three components.

Table 1
Samples and their initial concentration used in the PARAFAC2 model for simultaneous determination of *o*-, *m*- and *p*-aminophenol

Sample	*o*-Aminophenol (μmol l^{-1})	*m*-Aminophenol (μmol l^{-1})	*p*-Aminophenol (μmol l^{-1})
1	91.6	0.0	0.0
2	0.0	146.6	0.0
3	0.0	0.0	22.9
4	45.8	110.0	11.5
5	45.8	110.0	22.9
6	45.8	110.0	34.4
7	45.8	146.6	11.5
8	45.8	146.6	22.9
9	45.8	146.6	34.4
10	45.8	183.3	11.5
11	45.8	183.3	22.9
12	45.8	183.3	34.4
13	91.6	110.0	11.5
14	91.6	110.0	22.9
15	91.6	110.0	34.4
16	91.6	146.6	11.5
17	91.6	146.6	22.9
18	91.6	146.6	34.4
19	91.6	183.3	11.5
20	91.6	183.3	22.9
21	91.6	183.3	34.4
22	137.5	110.0	11.5
23	137.5	110.0	22.9
24	137.5	110.0	34.4
25	137.5	146.6	22.9

The vectors loadings $\mathbf{b1}_1$, $\mathbf{b1}_2$ and $\mathbf{b1}_3$ form the matrix loading $\mathbf{B}_1$. In a similar way, the matrices loadings $\mathbf{B}_2$, $\mathbf{B}_3, \ldots, \mathbf{B}_K$ are formed.

2. Experimental

2.1. Reagents

P.A. grade *o*- (Carlo Erba), *m*- and *p*-aminophenol (British Drug Houses), sodium nitrite (Mallinckrodt), citric acid (Quimibrás), boric acid (Merck) and trisodium orthophosphate (Ecibra) were used. Milli Q water was used in all preparations.

2.2. Experimental design

Stock solutions of the *ortho* and *meta* isomers with concentrations of 4.582×10^{-3} mol l^{-1} and of the *para* isomer with a concentration of 1.145×10^{-3} mol l^{-1}, were prepared in 0.06 mol l^{-1} hydrochloric acid. From this stock, ternary and pure samples were prepared in a buffer solution (mixture of 0.310 mol l^{-1} boric acid, 0.078 mol l^{-1} citric acid and 0.045 mol l^{-1} trisodium orthophosphate), to maintain pH = 4. A set of 25 samples with different concentrations of isomers was prepared. The final concentrations at the reaction cell, to be mixed with 0.013 mol l^{-1} sodium nitrite are shown in Table 1.

2.3. Injection simultaneous kit

A home-made system [23], shown in Fig. 5, was used for simultaneously injecting equal volumes of reactant solutions into the reaction cell. This simultaneous injection system has a support of nylon (a) where the glass syringes are fixed (b) for injection of reagents. Glass bottles (c) serve as reservoirs of the reagents. Each bottle is connected to a three-way Teflon valve (d). The syringe pistons are fixed in a support (e) that passes through an axle (f), impelled

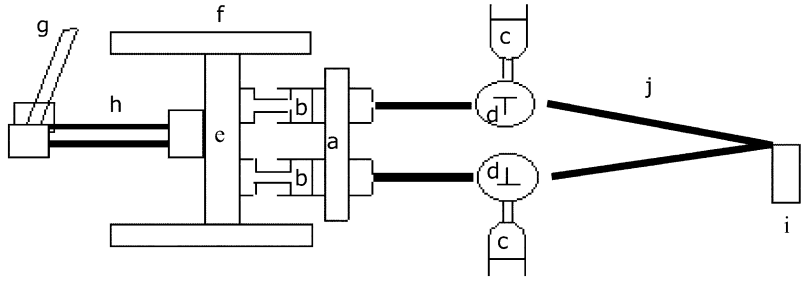

Fig. 5. Simultaneous injection system.

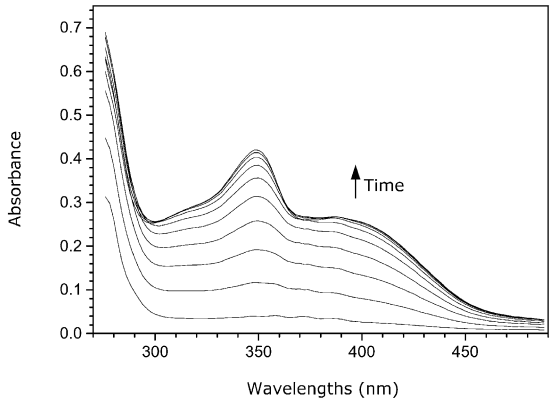

Fig. 6. Illustration of a typical data set obtained in the monitoring of the aminophenol reaction with $NaNO_2$, in the conditions studied in this work.

by a manual rod (g) located in piston (h). The manual rod controls the flow that fills the reaction cell (i), through polyethylene tubes (j).

2.4. Kinetic data acquisition

The reactions were performed at 25 ± 0.1°C, with solutions previously maintained at this temperature for 1 h. Equal volumes of isomer solution and $NaNO_2$ solution were added using the system for simultaneous injection previously described. The spectra were scanned from 276 to 488 nm, with 2 nm of resolution, at intervals of 3 s from an initial time of 1 s until 28 s of reaction, using a HP8496 diode array spectrophotometer and a thermostated reaction cell (1 cm optical path) with a Peltier system. The blank was prepared by mixing equal volumes of the buffer solution and sodium nitrite.

Fig. 6 shows a data set typically obtained in the monitoring of the reaction of aminophenol with $NaNO_2$ at pH = 4.

2.5. Computer programs

The experimental data were processed in programs written in Matlab 5.0. The PARAFAC2 program was developed by Rasmus Bro and is available at http://www.models.kvl.dk/source/.

3. Results and discussion

A set of 25 sample mixture of *o*-, *m* and *p*-aminophenol isomers was used to construct the model. The kinetic data collected was arranged in a three-way array with dimensions of $107 \times 10 \times 25$ (wavelengths, times and samples, respectively).

Several types of data pre-processing such as centering in the first mode (wavelengths) or in the second mode (times) were studied and compared with no centering results. Centering in more than one mode at a time was not accomplished because it can destroy the multilinear behavior of the data. It was verified, using three components in the calculations, a decrease in explained variance after centering and no advantages to the model were introduced with this procedure. Based on these results, it was decided do not use any pre-processing.

Restrictions of non-negativity in the first (wavelengths) and the third mode (samples) were applied, since only positive values of absorbance and concentration are meaningful. The iterative process stops when the relative difference in fit between two successive iterations is lower than 1×10^{-5}.

The number of components used in the construction of the model was based on the calculations of the residual tensor $\underline{\mathbf{E}}$ shown in Fig. 4. This tensor is obtained from the difference between the experimental data and the data predicted by the model. Once this tensor is obtained, a sum of squares of the all elements is performed. The results are presented in Fig. 7. In this figure, it is possible to see that three is an adequate number of components to be used in the construction of the PARAFAC2 model, because the sum of squared residuals is practically zero using

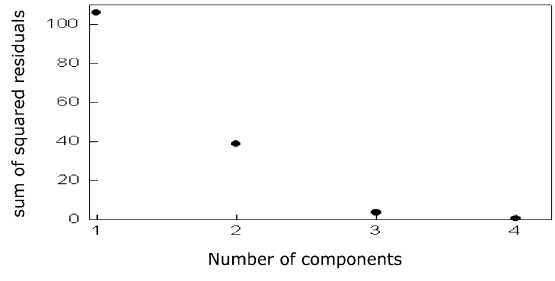

Fig. 7. Evaluation of residual tensor $\underline{\mathbf{E}}$ according to the number of components applied in PARAFAC2.

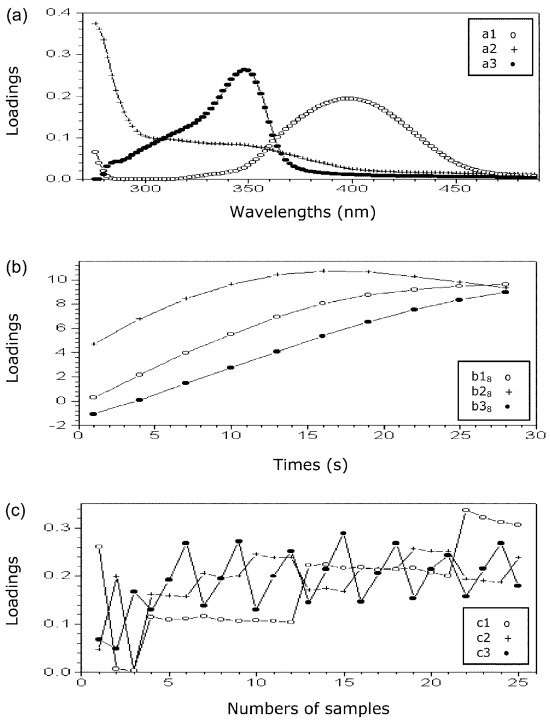

Fig. 8. (a) Spectral, (b) kinetics and (c) relative concentration profiles obtained by PARAFAC2 model for the isomers: (○) *o*-azophenol, (+) *m*-azophenol and (●) *p*-azophenol.

three or four components. This result agrees with the initial values expected, since there are three different isomers being analyzed.

The PARAFAC2 model obtained with three components decomposes the tensor $\underline{\mathbf{X}}$ in a loading matrix of spectral profile, 25 loading matrices of kinetic profiles and a loading matrix of relative concentrations of the three isomers. Fig. 8a shows a graphic of the loadings corresponding to the first dimension of the three-way tensor $\underline{\mathbf{X}}$. This matrix is formed by three loading vectors, **a1**, **a2** and **a3**. Each ones furnishes the spectral profile of each reaction product. Fig. 8b illustrates the matrix loading of the kinetic profile for one sample (8th sample). This matrix is formed by three loading vectors $\mathbf{b1}_8$, $\mathbf{b2}_8$ and $\mathbf{b3}_8$ and represents the kinetic profiles of each isomer for the 8th sample. Fig. 8c is the plot of the loading matrix of the third dimension. This matrix is formed by loading vectors **c1**, **c2** and **c3**, and supplies information on the relative concentrations of each isomer in each sample.

A linear regression can be performed between the relative concentrations C_R supplied by loadings of the third dimension and the experimental concentrations C_T of isomer in the samples. This relation can be used to estimate the isomer concentrations, it being possible to quantify each isomer in the mixture. The results for prediction of isomer concentrations are shown in Fig. 9.

The RMSEP (root mean square error of prediction) can be used to estimate the error in the concentration prediction, and was calculated by Eq. (12):

$$\mathrm{RMSEP} = \sqrt{\frac{\Sigma(C_T - C_E)^2}{N}} \qquad (12)$$

where: C_T is the experimental concentration, C_E is the estimated concentration, and N is the number of samples.

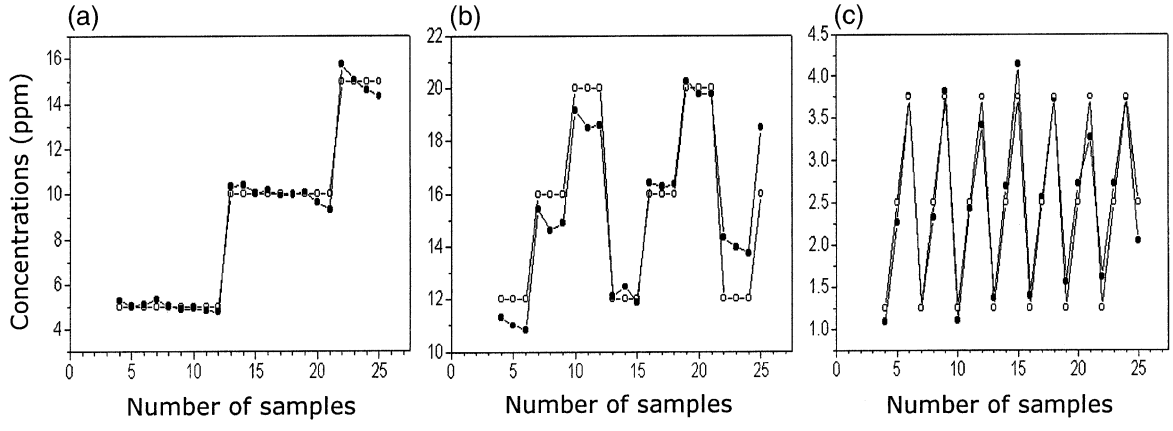

Fig. 9. (○) Experimental and (●) estimated concentrations for (a) *o*-azophenol, (b) *m*-azophenol, and (c) *p*-azophenol.

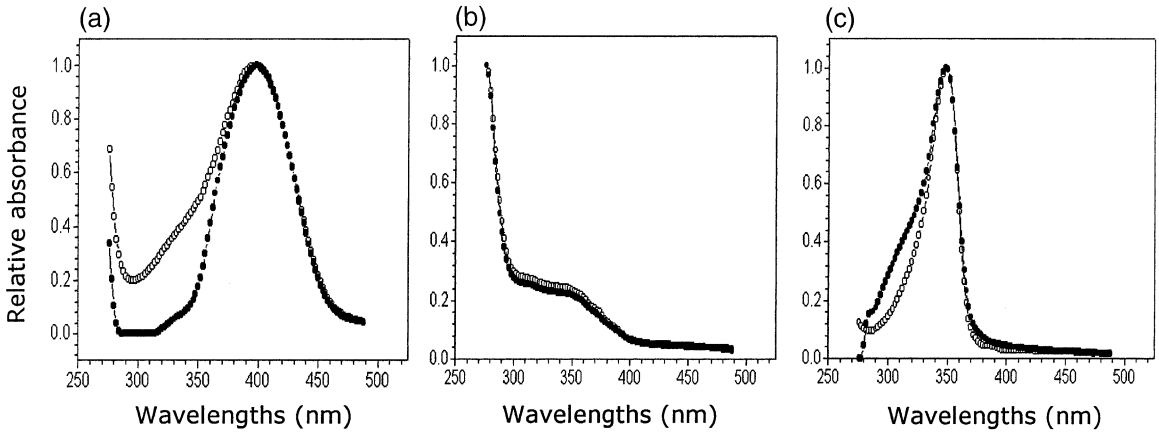

Fig. 10. (○) Experimental and (●) estimated spectral profile by PARAFAC2 for (a) *o*-azophenol, (b) *m*-azophenol and (c) *p*-azophenol.

The RMSEP values were 3.05 μmol l^{-1} for the *ortho* isomer; 10.76 μmol l^{-1} for the *meta* isomer and 2.17 μmol l^{-1} for the *para* isomer. According to the results presented, the PARAFAC2 model was able to predict the isomer concentrations with acceptable accuracy.

The normalized spectra of the first dimension loading and the normalized experimental spectra for each isomer's reaction product are presented in Fig. 10. The experimental spectra were obtained by monitoring the reaction of a single isomer. It can be observed that the experimental spectra are similar to the spectra estimated by PARAFAC2. The spectral range with more differences between the normalized experimental spectra and that predicted by the model is from 276 to 360 nm where a significant superposition occurs. In addition, the maximum absorption of the spectra obtained by PARAFAC2 for all isomers agrees with the experimental value. The wavelength of maximum absorption for the *o*-azophenol isomer is 398 nm, for the *m*-azophenol isomer, it is 276 nm, and for the *p*-azophenol isomer, it is 349 nm.

As result of the PARAFAC2 data modeling, 25 kinetic loading matrices were obtained, corresponding to the 25 different samples prepared (variables of the second dimension). The kinetic result comparison between the PARAFAC2 value and the experimental, can be made easily for *ortho* isomer, because it presents a spectral region (above 400 nm) without overlap. According to the results, a good agreement was observed between the experimental kinetic curves at 410 nm and the PARAFAC2 model for *ortho* isomer, as can be seen in Fig. 11a and b for samples numbers 4 and 6. This figure also presents the kinetic curve obtained by PARAFAC1 modeling calculated at the same conditions as PARAFAC2. This PARAFAC1 kinetic curve shows significant differences with the experimental data.

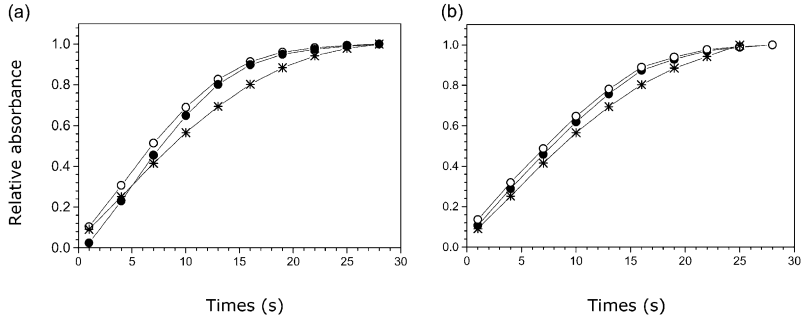

Fig. 11. Kinetic profiles at 410 nm for *o*-aminophenol using (○) experimental data, (●) PARAFAC2 model and (*) PARAFAC1 model. (a) sample 4 and (b) sample 6.

4. Conclusion

In this paper, it was demonstrated that, using PARAFAC2, it was possible to model kinetic data obtained by the reaction between aminophenol isomers and nitrite. The spectral profiles recovered by the model for each reaction product present a good similarity with the experimental spectra obtained in the reaction of samples containing only one isomer.

The concentrations estimated by the calibration model developed present a good correlation with the true concentrations, making it possible to quantify isomers mixtures with acceptable accuracy.

The model provides kinetic curves for *o*-aminophenol, in good agreement with the kinetic curve obtained by monitoring the single *ortho* isomer reaction.

The PARAFAC2 can be applied to model intricate reaction systems, that can not be modeled by other chemometric methods, such as PARAFAC1, which presume a common kinetic profile for each reaction product in all samples.

References

[1] N. Ho, D. Christian, R. Davidson, Anal. Chem. 50 (1978) 1108–1113.
[2] G. Johson, R. Yost, Anal. Chem. 57 (1985) 758A–760A.
[3] P. Hindmarch, K. Kavianpour, R.G. Brereton, Analyst 122 (1997) 871–877.
[4] R. Tauler, I. Marqués, E. Casassas, J. Chemom. 12 (1998) 55–75.
[5] R. Tauler, K. Smilde, M. Henshaw, B. Burgess, R. Kowalski, Anal. Chem. 66 (1994) 3337–3344.
[6] P. Geladi, Chemom. Intell. Lab. Syst. 7 (1989) 11–30.
[7] A.K. Smilde, Chemom. Intell. Lab. Syst. 15 (1992) 143–157.
[8] H.A.L. Kiers, Psychometrika 56 (1991) 449–470.
[9] M. Blanco, J. Gene, H. Ituriaga, S. Maspach, J. Riba, Talanta 34 (1987) 987–993.
[10] A. Carey, J. Sundberg, Advanced Organic Chemistry, Plenum, New York, 1990, 588.
[11] L. Stahle, Chemom. Intell. Lab. Syst. 7 (1989) 95–100.
[12] R. Bro, J. Chemom. 10 (1996) 47–61.
[13] E. Sanchez, R. Kowalski, J. Chemom. 2 (1988) 265–280.
[14] E. Sanchez, R. Kowalski, J. Chemom. 4 (1990) 29–46.
[15] L.R. Tucker, Psychometrika 31 (1966) 279–311.
[16] B. Mitchell, D. Burdick, Chem. Intell. Lab. Syst. 20 (1993) 149–161.
[17] J. Martins, M.M. Sena, R.J. Poppi, F.B.T. Pessine, Appl. Spectrosc. 53 (1999) 510–522.
[18] L. Kiers, R. Bro, T. Berge, J. Chemom. 13 (1999) 275–294.
[19] R. Bro, A. Andersson, L. Kiers, J. Chemom. 13 (1999) 295–309.
[20] F. Ten Berge, L. Kiers, J. de Leeuw, Psychometrika 61 (1996) 123–132.
[21] H. Martens, T. Naes, Multivariate Calibration, Wiley, Chichester, 1989, 12.
[22] N. Cliff, Psychometrika 31 (1966) 33–42.
[23] A.V. Rossi, E. Ferreira, 20th Brazilian Chemical Society Meeting 1997. Proceedings of the Conference, 24–27 May 1997, Poços de Caldas, Brazil, 1997, ED47.