Talanta 128 (2014) 450–459

Contents lists available at ScienceDirect

Talanta

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

Application of third-order multivariate calibration algorithms to the determination of carbaryl, naphthol and propoxur by kinetic spectroscopic measurements

Pablo Santa-Cruz, Alejandro García-Reiriz *

Departmento de Química Analítica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Instituto de Química Rosario (IQUIR-CONICET), Suipacha 531, Rosario, S2002LRK, Argentina

ARTICLE INFO

Article history:
Received 17 January 2014
Received in revised form
23 April 2014
Accepted 25 April 2014
Available online 6 May 2014

Keywords:
Fluorescence
Third-order data
Chemometrics

ABSTRACT

In the present work a new application of third-order multivariate calibration algorithms is presented, in order to quantify carbaryl, naphthol and propoxur using kinetic spectroscopic data. The time evolution of fluorescence data matrices was measured, in order to follow the alkaline hydrolysis of the pesticides mentioned above. This experimental system has the additional complexity that one of the analytes is the reaction product of another analyte, and this fact generates linear dependency problems between concentration profiles. The data were analyzed by three different methods: parallel factor analysis (PARAFAC), unfolded partial least-squares (U-PLS) and multi-dimensional partial least-squares (N-PLS); these last two methods were assisted with residual trilinearization (RTL) to model the presence of unexpected signals not included in the calibration step. The ability of the different algorithms to predict analyte concentrations was checked with validation samples. Samples with unexpected components, tiabendazole and carbendazim, were prepared and spiked water samples of a natural stream were used to check the recovered concentrations. The best results were obtained with U-PLS/RTL and N-PLS/RTL with an average of the limits of detection of 0.035 for carbaryl, 0.025 for naphthol and 0.090 for propoxur (mg L^{-1}), because these two methods are more flexible regarding the structure of the data.

1. Introduction

Owing to the current technological development, analytical instrumentation provides increasing possibilities to obtain large amounts of data in a system under study. This increase of information is obtained by simultaneously measuring various properties of the system, thus resulting in data of increasingly higher number of modes. In analytical calibrations, increasing the dimensionality of the data generates different benefits. First-order multivariate methods allow quantification of analytes even in the presence of interferences, but require that they be represented sufficiently in the calibration step, which involves a tedious task. It has the benefit of allowing one to diagnose a sample containing uncalibrated components as an outlier, because its spectrum is adjusted with significant error to the calibration model. This property is called the *first-order advantage* [1]. Despite this, it may be noted that although outlying samples are detected, the predictions are incorrect due to the presence of an uncalibrated component in the sample. Only second-order data (or higher, as the third-order data studied in this work) are appropriate for modeling the information from a sample containing uncalibrated components, obtaining correct predictions. This property is known as the *second-order advantage* [1]. Calibrations with higher-orders (third-order or higher) allow one to increase the selectivity and sensitivity of a technique [2]. Currently the bibliographic information on the subject is extensive, both theoretical and applied to methods of first- and second-order, but there are few examples in which third-order algorithms are applied to real systems.

The application of multivariate calibration models to data from third-order or greater is continuously growing and applied to various research fields [3–5]. The inclusion of an extra mode in the data, as previously mentioned, increases the selectivity and sensitivity of the analysis by the inclusion of additional information of the sample. In the case under study, this extra mode was obtained by measuring the time evolution of the fluorescence emission–excitation matrices (EEMs). There are only a few works in relation to the processing of time-dependent EEM data, used for the determination of one or more analytes present in complex matrices. Few examples of modeling third-order data are the oxidation of catecholamines [6], mixtures of folic acid or leucovorin with methotrexate [7–10], the alkaline hydrolysis of procaine [11], the alkaline hydrolysis of carbaryl [12,13], the derivatization of

* Corresponding author.
E-mail address: garciareiriz@iquir-conicet.gov.ar (A. García-Reiriz).

http://dx.doi.org/10.1016/j.talanta.2014.04.075
0039-9140/

malondialdehyde [14], and the photochemical degradation of several pesticides [15], polycyclic aromatic hydrocarbons [16] and folic acid and its two principal metabolites [17].

In this work a technique is developed and optimized for determining three analytes simultaneously by multivariate calibration methods of third order. Currently the increased production and use of chemical compounds have given rise to a growing concern about the effect that these compounds may have on terrestrial and aquatic ecosystems. Because of their chemical characteristics, pesticides are in most cases persistent pollutants that resist in varying degrees, photochemistry degradation, both chemical and biochemical, so their half-lives in the environment can be high [18–20]. The application of synthetic pesticides has been a routine practice in agriculture in the last fifty years. The indiscriminate use given in the past to these compounds has produced at present residues which are detected in the environment, with potential risks associated with public health [21]. Pesticides can be a component of urban waste water. The major source of water pollution with pesticides comes from agricultural practices, waste water from agricultural industries or waste water from pesticide manufacturing plants.

Pesticides that were studied in this work belong to the family of carbamates. Carbamates are synthetic organic substances formed by a nitrogen atom attached to a carbamic acid, which acts as a leaving group. This may have a neurotoxic effect in the corresponding dose, leading to death. Their main features are high toxicity, relatively low chemical stability and moderate accumulation in tissues. From the family of carbamates, carbaryl and propoxur were analyzed together with naphthol because any correct technique for the determination of carbaryl should include the latter, because it is the product of its hydrolysis. Naphthol is the major metabolite of chemical hydrolysis and degradation of carbaryl. It is a mutagenic compound that causes irritation of the respiratory system. Its presence indicates prior contamination of water samples with carbaryl.

There are various analytical methods, previously reported for the determination of pesticides based on high performance liquid chromatography (HPLC) [22,23]. However, although HPLC techniques are selective and sensitive, they have some disadvantages, as they require expensive equipment, sometimes toxic solvents and complicated pre-treatments of samples. Thus, here it is developed as an alternative kinetic spectroscopic technique for determining the analytes by chemometrics, presenting it as more accessible and less toxic for routine laboratories.

In this paper, we developed a fluorescent kinetic method, chemometrics-assisted for the determination of carbaryl, naphthol and propoxur in water samples. It is based on third-order data, obtained by measuring the time evolution of the EEMs of the alkaline hydrolysis of these pesticides for each sample and then analyzed with three different algorithms [24]: parallel factor analysis (PARAFAC), unfolded partial least-squares (U-PLS) and multi-dimensional PLS (N-PLS); these last two assisted with residual trilinearization (RTL).

2. Chemometric methods

2.1. Parallel factor analysis (PARAFAC)

In this multivariate method the matrices are called $\mathbf{X}_{c,i}$, corresponding to the data for calibration, and for each unknown sample there is an $\mathbf{X}_u$ array. These can be grouped to form an array $\underline{\mathbf{X}}$, if the original data is of third order (as in the case in study), the latter will have four ways. If the individual matrices are of size $J \times K \times L$, such as in the kinetic evolution of emission–excitation fluorescence matrices (where J is the number of emission wavelengths, K is the number of excitation wavelengths and L is the number of temporal data), the dimensions of $\underline{\mathbf{X}}$ are $(I+1)J \times K \times L$ (I is the number of calibration samples).

In the PARAFAC model a generic element $\underline{X}_{ijk}$ of the array $\underline{\mathbf{X}}$ is defined as [25]

$$\underline{X}_{\mathrm{ijkl}} = \sum_{n=1}^{N} a_{ni} b_{nj} c_{nk} d_{nl} + \underline{E}_{\mathrm{ijkl}} \quad (1)$$

where N is the total number of chemical components which produces the response or signal, $\underline{E}_{ijkl}$ is an element of the residual error $\underline{\mathbf{E}}$ (with the same dimensions as $\underline{\mathbf{X}}$), a_{ni}, b_{nj} c_{nk} and d_{nl} are elements of the column vectors $\mathbf{a}_n$, $\mathbf{b}_n$, $\mathbf{c}_n$ and $\mathbf{d}_n$ corresponding to: relative concentrations [$(I+1)1$], profiles of emission ($J \times 1$), excitation ($K \times 1$) and temporal ($L \times 1$) for each of the N components, respectively. The column vector $\mathbf{a}_n$ is stored in the $\mathbf{A}$ matrix of scores (containing the relative concentrations of the components), and the column vectors $\mathbf{b}_n$, $\mathbf{c}_n$ and $\mathbf{d}_n$ are stored in the matrices of loadings $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ (with columns normalized to unity). The structure of the model Eq. (1) is called quadrilinear. The decomposition of $\underline{\mathbf{X}}$ provides the loadings and the scores of the individual components of the whole mixture, whether they are chemically known or not, being the base of being able to achieve the second-order advantage. The decomposition is normally carried through an optimization scheme by alternating least-squares (least-squares alternating, ALS) [26,27].

To apply the PARAFAC model in multivariate calibration the following steps should be taken into account (1) set the number of factors which cause the response signal, (2) identify specific components from the information provided by the model, and (3) calibrate the model to obtain the absolute concentrations of a particular component in an unknown sample. The optimal number of factors can be estimated by several methods; in this work it was carried out by observing the reduction of least-squares residues of the PARAFAC fit. The number of components is taken as the number for which the residual is stabilized at a value close to the instrumental noise. The identification of the chemical components under investigation is achieved by comparing the three profiles estimated as emission, excitation and kinetic profiles of solutions of the analytes of interest. This is necessary because the components obtained by the decomposition of $\underline{\mathbf{X}}$ are ordered according to their contribution to the spectral variance and this order is not necessarily maintained when changing the unknown sample.

PARAFAC fits Eq. (1) using an ALS procedure, which requires initial values of the four matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$. This is usually accomplished using direct trilinear decomposition (DTLD) when the $\underline{\mathbf{X}}$ array is three-way or from vectors obtained by singular value decomposition (SVD) for four-way arrangements.

The absolute concentrations of analytes are obtained after calibration, since the decomposition of $\underline{\mathbf{X}}$ only provides relative values. This is done using the information of the composition of the calibration samples. Using Eq. (1), once the scores of the components were obtained, those of standard samples associated with a particular component are linearly related to the nominal concentrations of the analytes

$$[a_{n1}|a_{n1}|\cdots|a_{nI}] = k_{PARAFAC}\mathbf{y} \quad (2)$$

where n identifies a PARAFAC component, and $\mathbf{y}$ is a vector ($I \times 1$) containing the nominal concentrations of the analyte in the I calibration standards. The concentration of the analytes in the unknown sample is then estimated from

$$y_u = \mathrm{a}_{n(I+1)}/k_{\mathrm{PARAFAC}} \quad (3)$$

where $k_{PARAFAC}$ comes from Eq. (2).

For details of restrictions on the application of PARAFAC to kinetic-spectrophotometric data, see Refs. [28–32].

2.2. Partial least-squares (PLS)

PLS is a *soft model*, because the mathematical system proposed for the resolution and quantification of an analyte does not follow a hard physicochemical law. It is an algorithm developed for first-order data, but newer versions are already extending to higher-orders (N-PLS and U-PLS, see below).

A first-order system generates an **X** calibration matrix (size $I \times J$, I mixtures of known composition and J sensors in one mode). The PLS model equation for this type of system would be

$$\mathbf{X}_{I\times J} = \mathbf{T}_{I\times A}\mathbf{P}^{\mathrm{T}}_{A\times J} + \mathbf{E}_{I\times J} \tag{4}$$

To solve this equation an iterative algorithm is used, which performs compression of the information contained in the data matrix **X** into a smaller one of scores **T**, in order to minimize the effect of noise **E**. The matrix **P** of Eq. (4) is known as the matrix of PLS. The PLS algorithm also provides a matrix **W**, of identical size to **P**, which is known as the weight loading matrix.

The parameter A (number of latent variables) can be selected by techniques such as cross-validation, leaving a sample out, as suggested by Haaland and Thomas, and described in the literature [33].

The information of the concentration of analyte in the calibration samples is used in the calibration step, excluding unknown sample data, to obtain the vector **v** of regression coefficients, which is then used for prediction of unknown samples. During the prediction step, the concentration of each analyte in the unknown sample is quantified determining their scores $\mathbf{t}_u$ from its measured data $\mathbf{x}_u$ (a vector, since data being studied are first order). The concentrations of the different components of the unknown mixture are determined from the scores ($\mathbf{t}_u$) and the vector **v** obtained in the calibration using the following equation:

$$y_{\mathrm{u}} = \mathbf{t}_u^T\mathbf{v} \tag{5}$$

2.2.1. Unfolded partial least-squares (U-PLS)

This algorithm works similarly to that described for PLS, but unlike this it allows one to operate with data with orders higher than one. The concentration information was introduced in the calibration step again (excluding unknown sample data) in order to obtain two kinds of latent variables: the factors in the matrix **P** and the factors in the matrix **W**. These are estimated from the I calibration samples, which are formed by three-dimensional arrays **X** ($J \times K \times L$), and the vector **y** of calibration of concentrations ($I \times 1$, where I is the number of calibration samples).

However, U-PLS does not work with **X** arrangements stacked in tetradimensional arrangements, but it unfolded there, which is the vectorization and brings in a new matrix **Z**:

$$\mathbf{Z} = [\mathrm{vec}(\mathbf{X}_{JKL,1})|\mathrm{vec}(\mathbf{X}_{JKL,2})|\cdots|\mathrm{vec}(\mathbf{X}_{JKL,I})] \tag{6}$$

The symbol vec($\cdot$) denotes application of the vectorization operator, which converts the $J \times K \times L$ arrangements in $JKL \times 1$vectors. The objective of the vectorization is to obtain a second-order tensor **Z**, which is able to apply the PLS procedure. Then, the matrix **Z** ($JKL \times I$) is decomposed into the respective matrices **P** and **W** (both of identical size $JKL \times A$, where A is the number of latent factors) and the vector of the regression coefficients **v** ($A \times 1$) analogously to conventional PLS. If calibration is accurate, and there are no unexpected components, the analytics concentration of a species in an unknown sample can be predicted with the vector **v** using the following relation as in PLS:

$$y_{\mathrm{u}} = \mathbf{t}_u^T\mathbf{v} \tag{7}$$

where $\mathbf{t}_u$ is the score vector of the unknown sample, obtained in an appropriate projection of the unfolded data into the matrices of the calibration loadings (space of A latent factors):

$$\mathbf{t}_u = (\mathbf{W}^T\mathbf{P})^{-1}\mathbf{W}^T\mathrm{vec}(\mathbf{X}_{\bar{u}}) \tag{8}$$

If unexpected components are present in the sample, the scores obtained by Eq. (8) are inappropriate for the prediction of analyte concentration using Eq. (7), and residues in the PLS prediction stage [s_p, in Eq. (9) below] are significantly higher than the instrumental noise level.

$$s_p = ||\mathbf{e}_p||/(JKL-A)^{1/2} = ||\mathrm{vec}(\mathbf{X}_u) - (\mathbf{W}^T\mathbf{P})^{-1}\mathbf{W}^T\mathrm{vec}(\mathbf{X}_u)||/(JKL-A)^{1/2} \tag{9}$$

where $\|\cdot\|$ denotes the Euclidean norm. This situation can be solved with RTL as discussed below for N-PLS, since both U-PLS and N-PLS processed higher-order information, but on its way to work do not exploit the second-order advantage.

2.2.2. Multi-dimensional partial least-squares (N-PLS)

The N-PLS method can be applied to third-order data without unfold. It employs the I calibration data arrays with the original structure together with the vector **y** of standard concentrations (size $I \times 1$) to obtain the group loadings $\mathbf{W}^j$, $\mathbf{W}^k$ and $\mathbf{W}^l$ (with their respective sizes $J \times A$, $K \times A$ and $L \times A$, where A is the number of latent factors) as well as the regression coefficients **v** (size $A \times 1$). If the unknown sample does not contain any unexpected components, **v** can be used to estimate the analyte concentration according to

$$y_u = \mathbf{t}_u^T\mathbf{v} \tag{10}$$

where $\mathbf{t}_u$ is the vector of scores of the unknown sample obtained by the appropriate projection of the unknown data in the matrices of the calibration loadings. If in the unknown sample are presented unexpected components, the scores obtained are inadequate for prediction of analyte by Eq. (10). In this case, the residues from N-PLS phase prediction [s_p, see Eq. (11) below] are abnormally large compared with the typical instrument noise level

$$s_p = \|e_p\|/(JKL-A)^{1/2} = \|\mathrm{vec}(\mathbf{X}_u) - \mathrm{vec}(\hat{\mathbf{X}}_u)\|/(JKL-A)^{1/2} \tag{11}$$

where $\hat{\mathbf{X}}_u$ is the data array of third-order of the ($\mathbf{X}_u$) sample, rebuilt by the N-PLS model and $\|\cdot\|$ denotes the Euclidean norm. This again can be handled with an additional procedure called residual trilinearization.

2.2.3. Residual trilinearization (RTL)

The residual trilinearization models the effects of unexpected component of samples for correct concentration prediction analogous to RBL (Residual Bilinearization) only extends to one more mode. This analysis is based on the Tucker3 model of the unexpected component effects [7]. RTL minimizes the residual error calculated by fitting the data of the sample to the sum of relevant contributions:

$$\mathbf{X}_u = \mathrm{reshape}\{\mathbf{t}_u[\mathbf{W}^j|\otimes|\mathbf{W}^k)|\otimes|\mathbf{W}^l]\} + \mathrm{Tucker3}(\hat{\mathbf{X}}_u - \mathbf{X}_u) + \mathbf{E}_u \tag{12}$$

where 'reshape' indicates the transformation of a vector $JKL \times 1$ in a three-dimensional array of size $J \times K \times L$, and $|\otimes|$ indicates the operator Kathri-Rao [27]. It uses the weight loadings (or heavy loadings vectors) $\mathbf{W}^j$, $\mathbf{W}^k$ and $\mathbf{W}^l$ obtained with U-PLS or N-PLS; these are kept constant in the calibration values, and $\mathbf{t}_u$ is varied until the norm of $\|\mathbf{E}_u\|$ is minimized using the Gauss–Newton method [see eq. (12)]. Once done, the analyte concentration can be calculated by the equation corresponding to each particular algorithm [N-PLS Eq. (10)], introducing the last value of vector $\mathbf{t}_u$ found by RTL. Tucker3 model is constructed usually with N_{unx} components in each dimension [7], where N_{unx} indicates the number of unexpected components in the unknown sample. This last number can be adjusted, plotting s_u values as a function of

N_{unx}, until it stabilizes at a value comparable to instrumental noise. It should be noted that a single component unexpected in Tucker3 analysis of Eq. (12) provides the profiles corresponding to the three dimensions, but for additional unexpected components, the recovered profiles do not correspond to the real spectra (or temporal profile).

Finally, note that N-PLS/RTL or U-PLS/RTL can easily be extended to higher-order data, resulting in N-PLS or UPLS coupled to residual multilinearization (N-PLS/RML or U-PLS/RML), which surely in the near future will be increasingly applied to modern instruments, leading to increasingly high data complexity.

3. Experimental methods

3.1. Instrumental

The measurements of fluorescence matrices were performed on a Cary Eclipse fluorescence spectrophotometer of Varian, equipped with two Czerny–Turner monochromator and a xenon flash lamp connected to a microprocessor in a PC by a serial interface IEE 488 (GPIB) equipped with a thermostatic bath. The matrices were recorded in a quartz cell of 10 mm of path length. The readings were made in the range of 244 at 354 nm in excitation every 5 nm, and emission from 280 to 490 nm every 5 nm at a scan rate of 24,000 nm min^{-1}. It was used for 13 cycles with a time of 90 s for each matrix, thus having $43 \times 23 \times 13 = 12{,}857$ data points.

It was employed a thermostating bath to fix the temperature to 35 °C. The voltage of the photomultiplier detector was 600 V and the slit of excitation and emission monochromators was 5 nm for each one. This criterion was applied to all samples processed.

Absorbance measurements were performed on a Beckman DU 640 (Fulletron USA) using a quartz cell of 10 mm path length in the range from 240 to 360 nm.

3.2. Reagents and solutions

Solid drugs used were as follows: carbaryl (Sigma-Aldrich), propoxur (Sigma-Aldrich), naphthol (Sigma-Aldrich), carbendazim (Sigma-Aldrich), thiabendazole (Sigma-Aldrich), methanol (Merck, HPLC grade), sodium monohydrogen phosphate (Merck), sodium hydroxide (Merck) and hydrochloric acid (Merck).

Standards solutions of carbaryl (1120 mg L^{-1}), naphthol (1120 mg L^{-1}), and propoxur (1070 mg L^{-1}) were prepared in 10.00 ml volumetric flasks by dissolving an appropriate amount of the respective solid drug in methanol, ensuring its dissolution applying ultrasound for 5 min and completing with the same solvent. Solutions were prepared monthly since drugs were stable in the organic phase. Standard interferent pesticides, thiabendazole (100 mg L^{-1}) and carbendazim (200 mg L^{-1}), were prepared weighing the solid drug in a 10.00 mL volumetric flask; the solid was dissolved in methanol with ultrasound and then the volume was completed with the same solvent.

The stock solutions of carbaryl (10.00 mg L^{-1}), naphthol (5.00 mg L^{-1}), and propoxur (30.00 mg L^{-1}) were prepared taking aliquots of standards solutions and placed in flasks of 25.00 mL; then methanol was evaporated to dryness with the aid of a N_2 stream and carried to final volume with double distilled water. The stock solutions of individually interferent pesticides were prepared by taking aliquots of standards prepared, placed in flasks of 10.00 mL, the solvent was evaporated with N_2 and the final volume was completed with double distilled water to yield concentrations of 12.00 mg L^{-1} and 4.80 mg L^{-1}, for carbendazim and thiabendazole, respectively. All of these solutions were prepared weekly because analytes in aqueous phase are hydrolyzed slowly. All solutions were stored at 8 °C.

3.3. Samples of calibration, validation and spiked interferents

- A set of 15 calibration samples was prepared by triplicate with the three analytes together. Concentration levels were determined using a central compound design. The following concentration ranges were used: for carbaryl of 0.00–1.00 mg L^{-1}, for naphthol of 0.00–0.50 mg L^{-1} and for propoxur of 0.00–4.00 mg L^{-1}. The samples were prepared adding the analytes in appropriate amounts to 1.00 mL of phosphate buffer (0.1 M, pH 10.2) and double distilled water to bring the final volume of 3.00 mL.
- Validation samples were prepared identically to the calibration, but with different concentrations. Six samples were prepared by triplicate containing the analytes at random concentrations within the respective ranges of calibration.
- Carbendazim and thiabendazole were used as unexpected components to generate the non-modeled signal in the calibration step. The samples contained two levels of concentration of these interferents. Samples were prepared by triplicate on the cell placing appropriate volumes of the respective solutions of each analyte and the interferents, 1.00 mL phosphate buffer (0.1 M, pH 10.2) and double distilled water to bring the final volume to 3.00 mL.

3.4. Spiked samples of river water

The first step was filtering the samples using a microporous membrane of cellulose acetate 22 μm in diameter with the help of a vacuum pump. Then the absorbance spectrums in the wavelengths used for excitation of the calibration samples (244–354 nm) were measured. Therefore it was possible to choose an appropriate dilution in order to avoid the inner filter effect. They were diluted 1:10. The three analytes were added together to diluted samples in concentrations within the linear range. Fluorescence measurements were performed as described for the calibration samples.

3.5. Software

All routines used to perform the calculations described in this work were written in MATLAB [34]. Those for applying PARAFAC and N-PLS (without the second-order advantage, provided by RTL) are available on Internet, thanks to Bro [35]. U-PLS/RTL and N-PLS/RTL were developed by our group and successfully incorporated in a graphical interface (MVC2 or MVC3), the type already published for first-order multivariate calibration [36], available at http://www.chemometry.com/Index/Links%20and%20downloads/Programs.html. The routine to correct the Rayleigh scattering used was provided by Zepp [37].

4. Results

This system has the complexity of evolution over time, because carbaryl and propoxur in basic medium are hydrolyzed. This complexity can be turned into an advantage when the data is analyzed by third-order methods, which can discriminate the components of a mixture by their behavior in the three dimensions of the studied system (in our case excitation-emission-time). Fig. 1 and 2 shows the behavior of two samples, one with carbaryl and another with propoxur, both under alkaline conditions. Here it can appreciate how the signal of the carbaryl and propoxur species

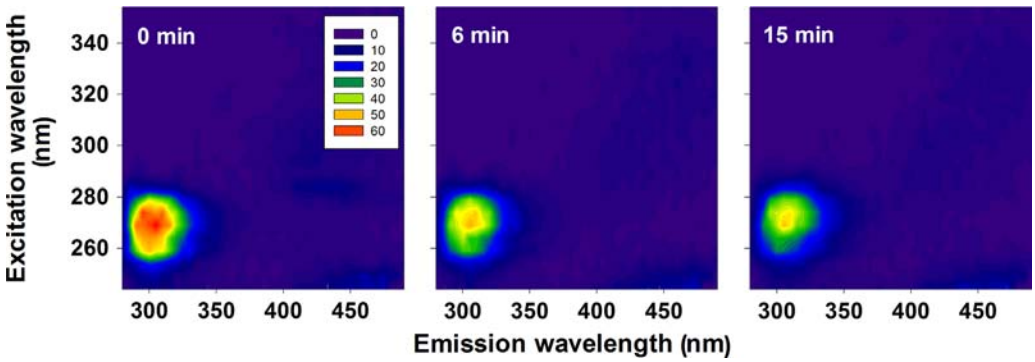

Fig. 1. Contour plots corresponding to EEMs versus time for the hydrolysis of a propoxur solution (3.00 mg L^{-1}) at pH 10.2. The legend shows different colored fluorescence intensities in UA. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

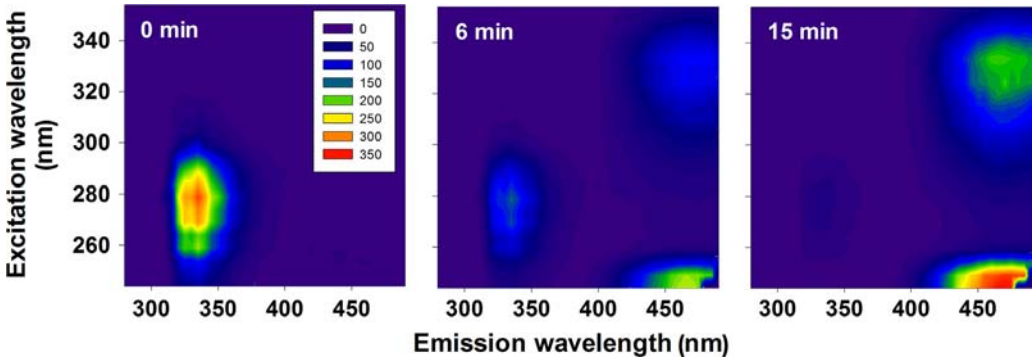

Fig. 2. Contour plots corresponding to EEMs versus time for the hydrolysis of a carbaryl solution (1.00 mg L^{-1}) at pH 10.2. The legend shows different colored fluorescence intensities in UA. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

decrease as alkaline hydrolysis proceeds. The carbaryl analyte has the fastest kinetic and it presents the particularity that its hydrolysis product is also fluorescent (naphthol); therefore, their samples show how the signal of carbaryl decreases to result in a stronger signal corresponding to naphthol.

Before preparing the calibration samples the system variables that influence the kinetics of reaction must be optimized, so the relationship between the reaction time and the speed of measurement should be taken into account. If time is extremely long, the technique is not practical. And if the reaction is too rapid compared with the time of measurement of one EEM, the measures will lose the intrinsic multilinearity of the data, thereby making it impossible for further processing by chemometric algorithms. In order to keep the multilinearity characteristic of data, the EEMs must be measured as quickly as possible, so that between measuring an emission spectrum and another, the reaction has not advanced significantly and thus each EEM reflects a sample instant. Therefore, making a risk-benefit ratio between the practicality of the technique, the scanning speed and increased fluorometer measurement noise by increasing the speed, set as the optimization target time in 15 min.

The influence of pH and temperature on kinetics of the alkaline hydrolysis as affect at fluorescent signal of the three analytes was studied. It can be concluded that the optimum operating pH is 10.2 and the optimum working temperature is 35 °C.

4.1. Preprocessing of data

The proposed chemometric algorithms require trilinear data. This means that the data matrix can be decomposed linearly on vectors of each dimension containing information about how the analyte behaves in this particular dimension. In our case a vector would be emission spectrum, other excitation spectrums, and another time behavior.

The data of the system under study has the problem in that besides having the signal of analytes studied and unexpected components, all samples have a Rayleigh signal. This signal does not respond to trilinear modeling. Therefore, before analyzing any data it was necessary to 'clean' this signal of samples. A routine written in MATLAB was applied to each sample to replace the data in the area affected by such dispersion on EEMs with their polynomial estimate [37].

4.2. Results of the multivariate calibration

4.2.1. Validation samples

Validation samples were analyzed by PARAFAC, N-PLS and U-PLS to corroborate the predictive ability of these three models. The samples were measured under optimum conditions and following the procedure described in the *Experimental Methods* section.

One of the main difficulties in this system is that one of the compounds of interest (naphthol) is an analyte and it is also a reaction product of one of the other analytes. This fact generates linear dependencies or collinearities between naphthol and carbaryl and complicates the analysis using the PARAFAC method. In previous works, this problem was solved using multiple linear regressions to the obtained scores [11,13].

As expected, the PARAFAC model showed certain difficulties in obtaining precise and accurate results. The number of components was established to be equal to 4 ($N=4$) due to the observation of PARAFAC error adjustment (16.43, 6.09, 4.54, 2.22, 2.07, errors in arbitrary fluorescence units (AFU) for one to five PARAFAC components respectively), thus it means that the system has four components with different behaviors (carbaryl, propoxur, original naphthol and naphthol produced by hydrolysis of carbaryl).

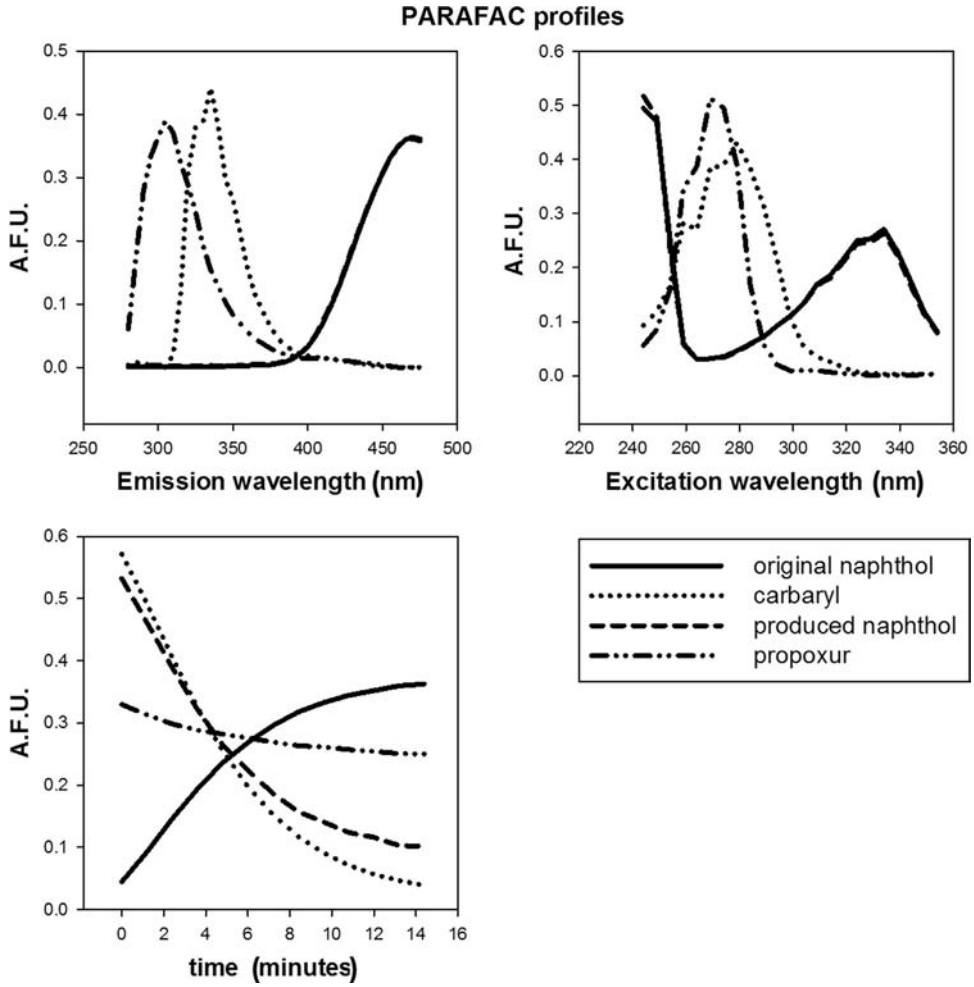

Fig. 3. Spectral profiles obtained by PARAFAC when processing a validation sample containing carbaryl (0.60 mg L^{-1}), naphthol (0.20 mg L^{-1}) and propoxur (1.75 mg L^{-1}). All profiles are normalized to unity.

It worked with the non-negativity restriction for all dimensions, i.e. the scores and loadings are not negative. The recovery of the excitation and emission profiles was successful, but not so for time profiles (Fig. 3).

The estimated concentrations of the validation samples by PARAFAC were not successful; REP% of 14.00, 8.50 and 5.40 for carbaryl, naphthol and propoxur, respectively, were obtained.

In Fig. 3 in the plot corresponding to the temporal profile one can see that the curve for the original naphthol (blue line) does not agree with its behavior. Since, as previously stated, it is not susceptible to hydrolysis, its kinetic curve should be constant. The large linear dependence of the system justifies the fact that PARAFAC does not provide good results, as expected. To solve this problem PARAFAC scores were processed for multiple linear regression (MLR) as was done in a previous work [11,13]; however, due to the presence of an extra analyte (propoxur), it does not succeed in this analysis.

Then N-PLS and U-PLS models were applied. The algorithms mentioned could provide better results because of the characteristics of the system. The number of latent variables was determined according to the criterion of Haaland, applying cross-validation technique as described above. This one reported that carbaryl predictions needed only two latent variables, two for naphthol predictions and three for propoxur predictions. The predictions of validation samples with the latent variables proposed were not correct again for carbaryl and naphthol, but they were correct for propoxur. This is due to the strong linear dependence mentioned above between the species carbaryl and naphthol. Therefore in order to obtain the optimal number of latent variables for all three analytes, the residual of predictions of the validation samples was set as a function of the number of latent variables to the N-U-PLS and U-PLS methods. In this way four latent variables were determined, which are enough to quantify naphthol, three to carbaryl and three to propoxur, by N-PLS and U-PLS.

The fact that the number of factors is greater than the number of analytes, in some cases, is because PLS takes into account additional factors of the signal of analytes that produce spectral variation, for example the reaction progress in the measurement time, temperature stabilization, dispersion of signals, non-linearities, among others. Therefore more factors are required to explain the variance in the data. The results of prediction are expressed in Table 1. Table 1 presents the root mean squared error of perdition (RMSEP) and the relative error of prediction (REP%) calculated; these showed very good results for N-PLS and U-PLS.

4.2.2. *Samples with interferents*

Some substances commonly present with the pesticides studied here may interfere with the determination of their concentration using the proposed methods. Many samples containing carbaryl, naphthol and propoxur may possess among of other pesticides in their composition. The possibility that the signal of fluorescence of these substances affect to the

Table 1
Results predicted by N-PLS and U-PLS for the validation samples in mg L^{-1}. CAR=Carbaryl; NAP=Naphthol; PRO=Propoxur.

Nominal			N-PLS predictions[a]			U-PLS predictions[a]		
CAR	**NAP**	**PRO**	**CAR**	**NAP**	**PRO**	**CAR**	**NAP**	**PRO**
0.00	0.35	1.50	0.024 (2)	0.35 (3)	1.54 (4)	0.008 (1)	0.34 (2)	1.53 (1)
0.30	0.00	2.60	0.31 (1)	0.01 (2)	2.66 (3)	0.31 (1)	0.06 (2)	2.60 (4)
0.45	0.15	0.00	0.45 (2)	0.15 (1)	0.05 (1)	0.48 (3)	0.14 (2)	0.089 (2)
0.60	0.20	1.75	0.61 (2)	0.19 (2)	1.74 (2)	0.62 (2)	0.19 (3)	1.77 (2)
0.65	0.45	1.40	0.66 (3)	0.46 (3)	1.39 (3)	0.68 (3)	0.43 (1)	1.41 (1)
0.75	0.30	3.10	0.76 (1)	0.28 (2)	3.13 (1)	0.76 (2)	0.28 (2)	3.20(2)
RMSEP			**0.01**	**0.01**	**0.04**	**0.02**	**0.02**	**0.05**
REP %			**1.90**	**2.60**	**1.40**	**2.80**	**3.70**	**1.60**

[a] Standard deviation in parentheses corresponds to the last significant figure. Average of three determinations.

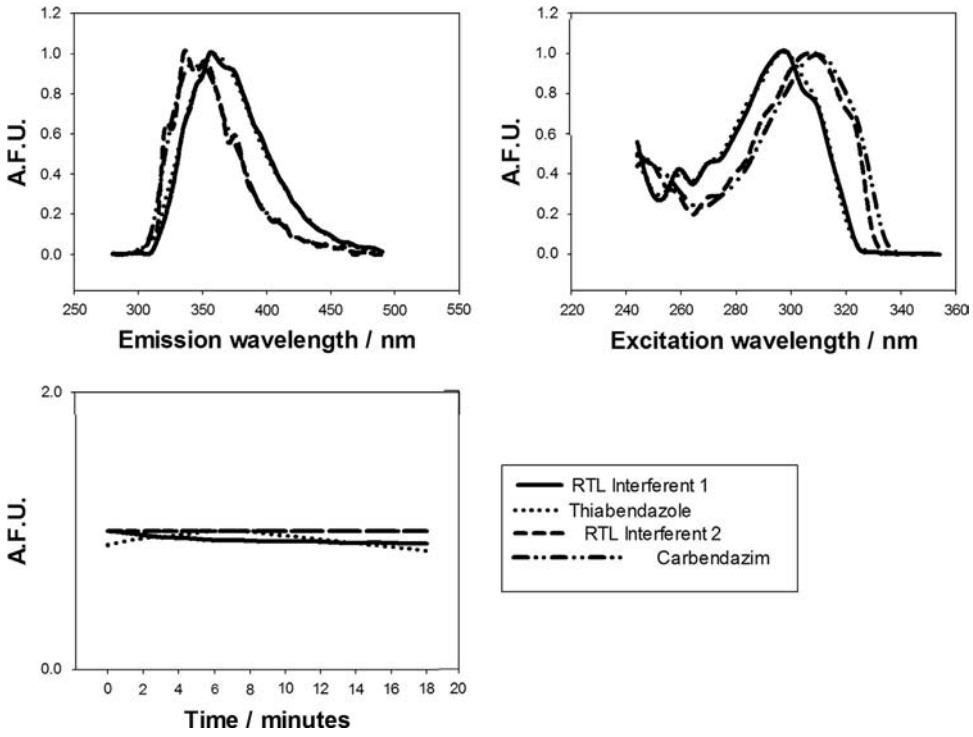

Fig. 4. Spectral profiles obtained by RTL when processing a sample with the three analytes together with carbendazim (2.00 mg L^{-1}) and thiabendazole (0.8 mg L^{-1}). All profiles are normalized to unity.

Table 2
Results predicted by N-PLS and U-PLS for samples with interferents in mg L^{-1}. CAR=Carbaryl; NAP=Naphthol; PRO=Propoxur; CBZ=Carbendazim; and TBZ=Thiabendazole.

Nominal					N-PLS predictions[a]			U-PLS predictions[a]		
CAR	**NAP**	**PRO**	**CBZ**	**TBZ**	**CAR**	**NAP**	**PRO**	**CAR**	**NAP**	**PRO**
0.20	0.20	3.10	2.00	0.80	0.21 (2)	0.19 (1)	3.20 (3)	0.25 (1)	0.25 (6)	3.2 (1)
0.45	0.35	1.80	2.00	0.80	0.46 (1)	0.33 (2)	1.83 (2)	0.52 (1)	0.32 (2)	1.80 (3)
0.20	0.20	3.10	2.80	1.12	0.21 (2)	0.18 (3)	3.0 (1)	0.25 (3)	0.21 (3)	2.90 (2)
0.45	0.35	1.80	2.80	1.12	0.45 (2)	0.34 (2)	1.90 (2)	0.50 (2)	0.31 (2)	1.90 (2)
RMSEP					**0.01**	**0.01**	**0.06**	**0.02**	**0.03**	**0.08**
REP %					**1.90**	**2.90**	**1.90**	**5.60**	**8.30**	**2.50**

[a] Standard deviation in parentheses corresponds to the last significant figure. Average of three determinations.

concentration predictions was examined. The interference substances studied were carbendazim and thiabendazole, at two levels of concentrations of each one.

Four samples by triplicate containing the analytes in aqueous solution, combined with amounts of carbendazim and thiabendazole, were prepared. Predictions concentrations were made for each analyte studied in all samples. The predictions were made using N-PLS and U-PLS assisted with RTL. Besides the latent variables required for each analyte, an extra variable for each interferent in the RTL procedure was necessary. The profiles obtained for these

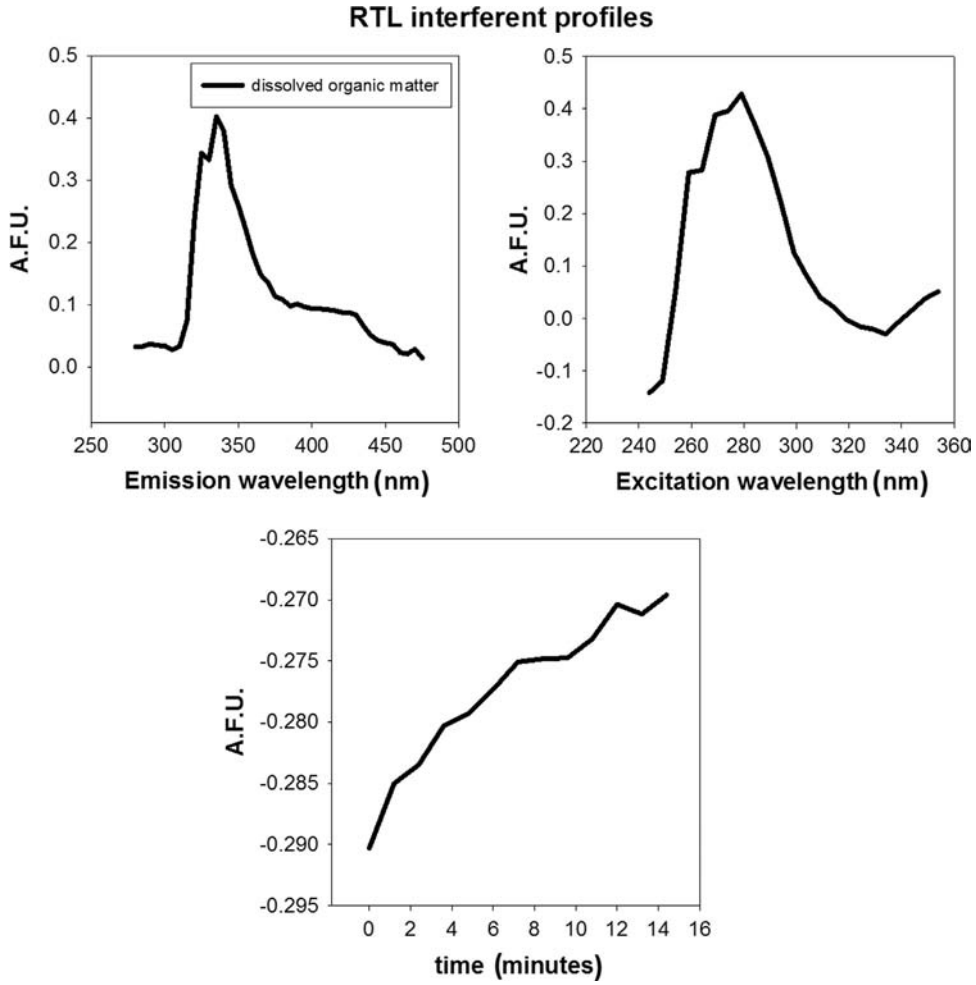

Fig. 5. Spectral profiles obtained with RTL when a river water sample added artificially with the analites is processed. All profiles are normalized to unity.

two interferents are shown in Fig. 4, which agree well with the experimental profiles of these pesticides, and the predictions of the samples in the presence of these interferents are provided in Table 2. Successful results were obtained taking into account the complexity of the samples. Comparing the values of RMSEP and REP%, it can be seen that the best results were obtained with N-PLS/RTL.

4.2.3. *Spiked river water samples*

Water samples collected from Ludueña Stream and Ibarlucea Channel, south of the province of Santa Fe, Argentina, were analyzed. These were filtered, their absorbance measured and then diluted according to the protocol described in the *Experimental Methods* section. Then their EEMs were measured in time by triplicate in order to make a blank, and they were subsequently predicted by the algorithms N-PLS and U-PLS assisted by RTL. Subsequently the samples were artificially added with carbaryl, propoxur and naphthol on the same samples, in order to make predictions in the presence of other components forming part of the complex signal of the matrix and background.

The five samples prepared were measured by triplicate following the procedure described; N-PLS and U-PLS were applied to predict them, assisted by RTL. RTL required only an extra latent variable to 'clean' the EEMs data from the native signal of water samples. Fig. 5 shows the excitation and emission profiles. It shows that they are consistent with the highest of fluorescence reported in the literature for dissolved organic matter of surfactants and/or whitening agents [38]. Although the temporal profile shows an increase in signal, taking into account the scale of the profile and that all samples to be placed in the fluorometer show a slight increase in signal due to stabilization with temperature, it can be concluded that the behavior over time of this interference is constant. The results shown in Table 3 are the average of the predictions obtained and the value of the standard deviations is in brackets, considering the presence of interferents as unexpected component. Again comparing the values of RMSEP and REP% for each analyte it can be concluded that better results were obtained processing samples by N-PLS/RTL.

4.3. *Figures of merit*

In the present work we calculated figures of merit for the three analytes, with N-PLS and U-PLS; these values are reported in Tables 4 and 5, respectively. Sensitivity (SEN) was calculated using equations found in the literature for third-order data [13]. Analytical sensitivity (γ_n) was calculated from the ratio between sensitivity and instrumental noise, expressed in units on mg^{-1} L. The instrumental noise considered for this case was 1.7 AFU. Conversely also analytical sensitivity (γ_n^{-1}) was calculated in units of mg L^{-1}. Finally, the limit of detection (LOD) was calculated in units of mg L^{-1} for each analyte.

For the samples with interferents and samples of river water added, which used RTL, it can be seen that merit figures showed an unfavorable change with respect to the validation samples. This is due to the complexity of the samples and the significant spectral

Table 3
Results predicted by N-PLS and U-PLS for spiked river water samples in mg L^{-1}. CAR=Carbaryl; NAP=Naphthol; and PRO=Propoxur.

Nominal			N-PLS predictions[a]			U-PLS predictions[a]		
CAR	**NAP**	**PRO**	**CAR**	**NAP**	**PRO**	**CAR**	**NAP**	**PRO**
0.00	0.00	0.00	0.020 (1)	0.019 (3)	0.024 (2)	0.035 (3)	0.029 (4)	0.042 (2)
0.55	0.30	0.00	0.50 (2)	0.33 (3)	0.002 (3)	0.50 (3)	0.36 (4)	0.10 (1)
0.60	0.00	1.30	0.59 (1)	0.02 (2)	1.30 (2)	0.58 (1)	0.001 (2)	1.00 (3)
0.00	0.45	3.00	0.01 (2)	0.47 (3)	2.90 (1)	0.02 (3)	0.50 (3)	2.80 (2)
0.70	0.20	2.40	0.69 (2)	0.23 (2)	2.30 (2)	0.59 (4)	0.28 (3)	2.30 (2)
0.40	0.35	1.70	0.38 (2)	0.37 (1)	1.70 (3)	0.39 (3)	0.38 (2)	1.70 (2)
RMSEP			**0.02**	**0.02**	**0.06**	**0.08**	**0.05**	**0.20**
REP %			**4.00**	**5.80**	**2.80**	**12.20**	**11.40**	**7.90**

[a] Standard deviation in parentheses corresponds to the last significant figure. Average of three determinations.

Table 4
Figures of merit for the N-PLS method.

Samples	Figures of merit	Carbaryl	Naphthol	Propoxur
Validation	SEN (U.A.F. mg^{-1} L)	2600	2500	250
	γ_n (mg^{-1} L)	1600	1400	130
	γ_n^{-1} (mg L^{-1})	6.25×10^{-4}	7.14×10^{-4}	7.69×10^{-4}
	LOD (mg L^{-1})	0.030	0.016	0.079
Interferences	SEN (U.A.F. mg^{-1} L)	2100	2000	200
	γ_n (mg^{-1} L)	1200	1300	65
	γ_n^{-1} (mg L^{-1})	8.33×10^{-4}	7.69×10^{-4}	0.014
	LOD (mg L^{-1})	0.035	0.025	0.090
River samples	SEN (U.A.F. mg^{-1} L)	2300	2100	130
	γ_n (mg^{-1} L)	1300	1200	70
	γ_n^{-1} (mg L^{-1})	8.33×10^{-4}	7.69×10^{-4}	0.014
	LOD (mg L^{-1})	0.035	0.025	0.090

Table 5
Figures of merit for the U-PLS method.

Samples	Figures of merit	Carbaryl	Naphthol	Propoxur
Validation	SEN (U.A.F. mg^{-1} L)	2500	2400	230
	γ_n (mg^{-1} L)	1600	1300	140
	γ_n^{-1} (mg L^{-1})	6.25×10^{-4}	7.69×10^{-3}	7.14×10^{-3}
	LOD (mg L^{-1})	0.031	0.018	0.082
Interferences	SEN (U.A.F. mg^{-1} L)	2100	2000	200
	γ_n (mg^{-1} L)	1150	1100	67
	γ_n^{-1} (mg L^{-1})	8.70×10^{-4}	9.09×10^{-4}	0.015
	LOD (mg L^{-1})	0.038	0.029	0.093
River samples	SEN (U.A.F. mg^{-1} L)	2200	2000	120
	γ_n (mg^{-1} L)	1400	1200	74
	γ_n^{-1} (mg L^{-1})	7.14×10^{-4}	8.33×10^{-4}	0.014
	LOD (mg L^{-1})	0.038	0.029	0.093

overlap. Despite this, the results were very good, demonstrating the effectiveness of RTL to 'clean' the signal not modeled in the calibration step. When the values obtained with U-PLS and N-PLS are compared in general, it can be concluded that the best results are always obtained with N-PLS.

5. Conclusions

The efficiency of algorithms of third-order multivariate calibration to determine three analytes together was corroborated. Using third-order data it was possible to overcome several drawbacks of extreme spectral overlap, increase the sensitivity and selectivity of the system and maintain the second-order advantage to be used when necessary. The results allowed us to propose the applied chemometric tools (N-PLS and U-PLS) as efficient alternatives for the quantification of pesticides in various aqueous samples.

In this paper, we were able to develop a simple, accurate, rapid, economical and environmentally friendly method for the simultaneous determination of carbaryl, naphthol and propoxur in water samples. As the analytes have fluorescent behavior, a method based on the kinetics of hydrolysis of these compounds in an alkaline medium with fluorescence measurements was developed. Third-order data was processed using different chemometric methods. The results showed that the models of N-PLS and U-PLS are more successful than other models such as PARAFAC, allowing quantification of the analytes even in the presence of severe spectral overlap and interferences not modeled, thanks to the application of the RTL procedure. The predictions in spiked river samples with N-PLS and U-PLS showed the success of these two methodologies for simultaneous quantification of the analytes.

Acknowledgments

The following institutions are gratefully acknowledged for financial support: Universidad Nacional de Rosario, CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas) and ANPCyT (Agencia Nacional de Promoción Científica y Tecnológica, PICT-2011-0033).

References

[1] K.S. Booksh, B.R. Kowalski, Anal. Chem. 66 (1994) 782A–791A.
[2] A.C. Olivieri, Anal. Chem. 80 (2008) 5713–5720.
[3] R. Bro, Crit. Rev. Anal. Chem. 36 (2006) 279–293.
[4] G.M. Escandar, N.M. Faber, H.C. Goicoechea, A. Muñoz de la Peña, A.C. Olivieri, R.J. Poppi, Anal. Chem. 26 (2007) 752–765.
[5] R. Bro, Anal. Chim. Acta 500 (2003) 185–194.
[6] R.P.H. Nikolajsen, K.S. Booksh, A.M. Hansen, R. Bro, Anal. Chim. Acta 475 (2003) 137–150.
[7] A.C. Olivieri, J.A. Arancibia, A. Muñoz de la Peña, I. Durán-Merás, A. Espinosa Mansilla, Anal. Chem. 76 (2004) 5657–5666.
[8] J.A. Arancibia, A.C. Olivieri, D. Bohoyo Gil, A. Espinosa Mansilla, I. Durán-Merás, A. Muñoz de la Peña, Chemom. Intell. Lab. Syst. 80 (2006) 77–86.
[9] A. Muñoz de la Peña, I. Durán Merás, A. Jiménez Girón, Anal. Bioanal. Chem. 385 (2006) 1289–1297.
[10] A. Muñoz de la Peña, I. Durán Merás, A. Jiménez Girón, H.C. Goicoechea, Talanta 72 (2007) 1261–1268.
[11] P.C. Damiani, I. Durán-Merás, A.G. García-Reiriz, A. Jiménez-Girón, A. Muñoz de la Peña, A.C. Olivieri, Anal. Chem. 79 (2007) 69496958.
[12] Z. Shao-Hua, W. Hai-Long, X. A-Lin, N. Jin-Fang, B. Jing-Chao, C. Chen-Bo, Y. Ru-Qin, Talanta 77 (2009) 1640–1646.
[13] R. Maggio, P. Damiani, A. Olivieri, Anal. Chim. Acta 677 (2) (2010) 97–107.
[14] A.G. García-Reiriz, P.C. Damiani, A.C. Olivieri, F. Cañada-Cañada, A. Muñoz de la Peña, Anal. Chem. 80 (2008) 7248–7256.
[15] M.L. Nahorniak, G.A. Cooper, Y.C. Kim, K.S. Booksh, Analyst 130 (2005) 85–93.
[16] Y.C. Kim, J.A. Jordan, M.L. Nahorniak, K.S. Booksh, Anal. Chem. 77 (2005) 7679–7686.

[17] A. Jiménez Girón, I. Durán-Merás, A. Espinosa-Mansilla, A. Muñoz de la Peña, F. Cañada-Cañada, A.C. Olivieri, Anal. Chim. Acta 622 (2008) 94–103.
[18] C.D.S. Tomin, The pesticide Manual: A World CompendiumCroydon, UK, 11th ed., British Crop. Protection Council (1997) 1997.
[19] H.M.G. Van der Werf, Agric. Ecosyst. Environ. 60 (1996) 81–96.
[20] L.A. Albert, 11th ed., British Crop. Protection Council, Croydon, UK, (1997), México D.F. (1998).
[21] E.R. Bandala, J.A. Octaviano, V. Albiter, L.G. Torres, Anal. Chem. 60 (1998) 177–182.
[22] R.M. Maggio, P.C. Damiani, A.C. Olivieri, Talanta 80 (2011) 1173–1180.
[23] Technical Support Center, Office of Ground Water and Drinking Water, United States Environmental Protection Agency, Method 531.2. Measurement of N-methylcarbamoyloximes and N-methylcarbamates in Water by Direct Aqueous Injection HPLC with Postcolumn Derivatization, Cincinatti, Ohio, 2001.
[24] Y.N. Ni, C.H. Huang, S. Kokot, Chemom. Intell. Lab. Syst. 71 (2) (2004) 177–193.
[25] S. Leurgans, R.T. Ross, Stat. Sci. 7 (1992) 289.
[26] R. Bro, Chemom. Intell. Lab. Syst. 38 (1997) 149–171.
[27] R. Bro, Multi-Way Analysis in the Food Industry (Doctoral thesis), Universidad de Amsterdam, Holanda, 1998.
[28] M. Azubel, F.M. Fernández, M.B. Tudino, O.E. Troccoli, Anal. Chim. Acta 398 (1999) 93–102.
[29] J.C.G. Esteves da Silva, C.J.S. Oliveira, Talanta 49 (1999) 889–898.
[30] S.R. Crouch, J. Coello, S. Maspoch, M. Porcel, Anal. Chim. Acta 424 (2000) 115–126.
[31] A. Espinosa-Mansilla, A. Muñoz de la Peña, H.C. Goicoechea, A.C. Olivieri, Appl. Spectrosc. 58 (2004) 83–90.
[32] M.A.B. Levi, I.S. Scarminio, R.J. Poppi, M.G. Trevisan, Talanta 62 (2004) 299–305.
[33] D.M. Haaland, E.V. Thomas, Anal. Chem. 60 (1988) 1193–1202.
[34] MATLAB 7.0. The MathWorks Inc., Natick, MA, 2000.
[35] ⟨http://www.models.kvl.dk/source/⟩.
[36] A.C. Olivieri, H.C. Goicoechea, F.A. Iñón, Chemom. Intell. Lab. Syst. 73 (2004) 189–197.
[37] R. Zepp, W.M. Sheldon, M.A. Moran, Mar. Chem. 89 (2004) 15–36.
[38] N. Hudson, A. Baker, D. Reynols, River Res. Appl. 23 (2007) 621–649.